Albert Heuberger
Günter Elst
Randolf Hanke
*Editors*

# Microelectronic Systems

## Circuits, Systems and Applications

Springer

Microelectronic Systems

Albert Heuberger • Günter Elst • Randolf Hanke
*Editors*

# Microelectronic Systems

## Circuits, Systems and Applications

Springer

*Editors*

Prof. Dr.-Ing. Albert Heuberger
Fraunhofer-Institut für Integrierte
Schaltungen IIS
Am Wolfsmantel 33
91058 Erlangen
Germany
albert.heuberger@iis.fraunhofer.de

Prof. Dr. Günter Elst
Fraunhofer-Institut für Integrierte
Schaltungen IIS
Zeunerstraße 38
01069 Dresden
Germany
guenter.elst@eas.iis.fraunhofer.de

Prof. Dr. Randolf Hanke
Fraunhofer-Institut für Integrierte
Schaltungen IIS
Dr.-Mack-Straße 81
90762 Fürth
Germany
randolf.hanke@iis.fraunhofer.de

*Editorial Coordination*
Janina Heppner, M.A.
Dr. Karlheinz Kirsch
Fraunhofer IIS
Am Wolfsmantel 33
91058 Erlangen, Germany

ISBN 978-3-642-23070-7 e-ISBN 978-3-642-23071-4
DOI 10.1007/978-3-642-23071-4
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011940678

# Preface to the Festschrift for Prof. Dr.-Ing. Heinz Gerhäuser

This book is dedicated to Prof. Dr. Heinz Gerhäuser on the occasion of his retirement both from the position of Executive Director of the Fraunhofer Institute for Integrated Circuits IIS and from the Endowed Chair of Information Technologies with a Focus on Communication Electronics (LIKE) at Friedrich-Alexander University, Erlangen-Nuremberg. The contributions to this Festschrift were written by Fraunhofer IIS staff and external project team members in grateful appreciation of his lifetime academic achievements and his inspiring and much-valued leadership of Fraunhofer IIS.

Born in Munich in 1946, Heinz Gerhäuser began his education in 1960 with an apprenticeship as an electrician, which in 1965 led to a technical college entrance qualification. Besides his practical abilities, an interest in scientific research and engineering was already emerging. Accordingly, Heinz Gerhäuser went on to pursue an academic education in electrical engineering, studying at the Georg Simon Ohm Polytechnic in Nuremberg from 1965 to 1968 and subsequently at Friedrich-Alexander University, where he obtained his Master's degree in 1973.

His main research interest was in audio coding and its hardware implementation. In the course of his doctoral research, which he completed in 1980, he created one of the first digital signal processing systems for real-time encoding of audio signals. After conducting postdoctoral research as a visiting scientist at the IBM San José Research Laboratory in 1980–1981, he returned to Germany to establish the Liaison Office for Research and Technology Transfer at Friedrich-Alexander University. Completed in 1984, this work laid the foundations for a company called Zentrum für Mikroelektronik. The company was later incorporated into the Fraunhofer-Gesellschaft as a working group on integrated circuits, which in turn evolved into an institute in its own right and ultimately became the Fraunhofer Institute for Integrated Circuits IIS.

Having devoted himself wholeheartedly to applied research, Heinz Gerhäuser was made director of Fraunhofer IIS in Erlangen in 1993. Additionally, in 1999, he was appointed to the Endowed Chair of Information Technologies at Friedrich-Alexander University. Under his leadership, Fraunhofer IIS grew to become the

largest of Germany's 60 Fraunhofer Institutes, a position it retains to this day, currently employing over 730 staff.

Heinz Gerhäuser's vision and entrepreneurial spirit have made Fraunhofer IIS one of the most successful and renowned German research institutions. Early on, he understood the challenges and potential of digital broadcasting. Moreover, he realized that the implementation of digital systems depended on data compression and audio coding, and it was to these areas that he turned his scientific attention. As a member of the core project team, Heinz Gerhäuser contributed greatly to the development of the mp3 format, which was to become a worldwide success. He was also instrumental in founding the AudioLabs facility, which brings together scientists from various countries and is run in conjunction with Friedrich-Alexander University. Its mission is to help Erlangen maintain its leading position in audio coding research for years to come. Outside this field, Heinz Gerhäuser has provided particular impetus to work on communications and non-destructive testing, which constitute the principal research themes of the institute's Nuremberg and Fürth branches, respectively. At the same time, he has taken a personal interest in medical engineering and technologies geared towards today's aging societies.

As well as being the author of a large number of publications, Heinz Gerhäuser has been involved in numerous patented inventions. He helped Fraunhofer IIS spin off several companies, including IZT and Coding Technologies. In addition, he contributed to the creation of the Fraunhofer Working Group on Electronic Media Technology AEMT in Ilmenau, which started out as a branch of Fraunhofer IIS before becoming the Fraunhofer Institute for Digital Media Technology IDMT in 2004.

For his entrepreneurial and scientific commitment, Heinz Gerhäuser has received many honors, including the Officer's Cross of the Order of Merit of the Federal Republic of Germany, the Bavarian State Medal of Merit for Outstanding Services to the Bavarian Economy and the Bavarian Order of Merit. Furthermore, he has been inducted into the Consumer Electronics Hall of Fame in recognition of his part in developing and popularizing the mp3 format and has been awarded an honorary professorship by Vladimir State University, Russia.

Heinz Gerhäuser offered valued experience and advice in his long-time role as spokesman and chair of the Fraunhofer Group for Microelectronics as well as in his capacity as a member of the Science and Technology Advisory Panel of the Bavarian State Government and as a judge for the "Zukunftspreis" innovation award conferred by the German President.

On behalf of everybody at Fraunhofer IIS, the editors would like to express their heartfelt gratitude to Heinz Gerhäuser. His invaluable commitment and outstanding leadership are very much appreciated. May these qualities be emulated by many and serve as an inspiration for future leaders. Ever ready to give helpful guidance, Heinz Gerhäuser has also supported editors Albert Heuberger and Randolf Hanke in their scientific careers. Similarly, he provided their fellow editor Günter Elst with beneficial conditions, allowing him the scope to establish Fraunhofer IIS's Dresden-based Design Automation Division EAS as a regional center of excellence in microelectronics. For all this we are deeply indebted. We would also like to take the opportunity to thank all contributors to this Festschrift.

We wish Heinz Gerhäuser the very best of luck for the future, exciting and fruitful new projects as well as good health and a happy life with his family.

Erlangen, October 2011

Albert Heuberger
Günter Elst
Randolf Hanke

# Contents

**Part IV Audio and Multimedia**

# List of Contributors

**Daniel Arndt** Fraunhofer IIS, Projektgruppe Drahtlose Verteilsysteme DVT, Helmholtzplatz 2, 98693 Ilmenau, Germany, daniel.arndt@iis.fraunhofer.de

**Christian Arzt** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, christian.arzt@iis.fraunhofer.de

**Stefan Aschenbrenner** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, stefan.aschenbrenner@iis.fraunhofer.de

**Ingo Bauscher** Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany, ingo.bauscher@iis.fraunhofer.de

**Stefan Bayer** International Audio Laboratories Erlangen, Am Wolfsmantel 33, 91058 Erlangen, Germany, stefan.bayer@audiolabs-erlangen.de

**Ashweeni Beeharee** University College London, Malet Place, London WC1E 6BT, United Kingdom, A.Beeharee@cs.ucl.ac.uk

**Tobias Bergen** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, tobias.bergen@iis.fraunhofer.de

**Philipp Berglez** TeleConsult Austria GmbH, Schwarzbauerweg 3, 8043 Graz, Austria, pberglez@teleconsult-austria.at

**Javier Gutiérrez Boronat** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, javier.gutierrez@iis.fraunhofer.de

**Marco Breiling** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, marco.breiling@iis.fraunhofer.de

**Jörg Bretschneider** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, joerg.bretschneider@eas.iis.fraunhofer.de

**Christoph Clauß** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, christoph.clauss@eas.iis.fraunhofer.de

**Giovanni Del Galdo** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, giovanni.delgaldo@iis.fraunhofer.de

**Manfred Dietrich** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, manfred.dietrich@eas.iis.fraunhofer.de

**Sascha Disch** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, sascha.disch@iis.fraunhofer.de

**Ulrich Donath** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, ulrich.donath@eas.iis.fraunhofer.de

**Markus Eberhorn** Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany, markus.eberhorn@iis.fraunhofer.de

**Ernst Eberlein** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, ernst.eberlein@iis.fraunhofer.de

**Andreas Eidloth** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, andreas.eidloth@iis.fraunhofer.de

**Günter Elst** Fraunhofer IIS/EAS, Zeunerstr. 38, 01069 Dresden, Germany, guenter.elst@eas.iis.fraunhofer.de

**Olaf Enge-Rosenblatt** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, olaf.enge@eas.iis.fraunhofer.de

**Andreas Ernst** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, andreas.ernst@iis.fraunhofer.de

**Jürgen Ernst** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, juergen.ernst@iis.fraunhofer.de

**Nikolaus Färber** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, nikolaus.faerber@iis.fraunhofer.de

**Sven Feilner** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, sven.feilner@iis.fraunhofer.de

**Daniel Fischer** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, daniel.fischer@iis.fraunhofer.de

**Felix Fleischmann** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, felix.fleischmann@iis.fraunhofer.de

**Christian Forster** Chair of Information Technologies, University of Erlangen-Nürnberg, Am Wolfsmantel 33, 91058 Erlangen, Germany, forster@like.eei.uni-erlangen.de

**Siegfried Foessel** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, siegfried.foessel@iis.fraunhofer.de

**Norbert Franke** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, norbert.franke@iis.fraunhofer.de

**Johannes Freitag** Alfred-Wegener-Institut für Polar- und Meeresforschung, 27568 Bremerhaven, Germany, johannes.freitag@awi.de

**Guillaume Fuchs** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, guillaume.fuchs@iis.fraunhofer.de

**Theobald Fuchs** Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany, theobald.fuchs@iis.fraunhofer.de

**Ralf Geiger** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, ralf.geiger@iis.fraunhofer.de

**Robért Glein** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, robert.glein@iis.fraunhofer.de

**Guillermo Grau** Indra, Barcelona, Spain, ggrau@indra.es

**Günther Greiner** Chair of Computer Graphics, University of Erlangen-Nürnberg, Am Wolfsmantel 33, 91058 Erlangen, Germany, greiner@informatik.uni-erlangen.de

**Bernhard Grill** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, bernhard.grill@iis.fraunhofer.de

**Roland Gruber** Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany, roland.gruber@iis.fraunhofer.de

**Michael Hackner** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, michael.hackner@iis.fraunhofer.de

**Stephan Haimerl** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, stephan.haimerl@iis.fraunhofer.de

**Randolf Hanke** Fraunhofer IIS/EZRT, Dr.-Mack Str. 81, 90762 Fürth, Germany, randolf.hanke@iis.fraunhofer.de

**Erik Haßlmeyer** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, erik.hasslmeyer@iis.fraunhofer.de

**Johann Hauer** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, johann.hauer@iis.fraunhofer.de

**Andy Heinig** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, andy.heinig@eas.iis.fraunhofer.de

**Christian Helmrich** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, christian.helmrich@iis.fraunhofer.de

**Jürgen Herre** International Audio Laboratories Erlangen, and Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, juergen.herre@audiolabs-erlangen.de

**Albert Heuberger** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, albert.heuberger@iis.fraunhofer.de

**Thomas Heyn** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, thomas.heyn@iis.fraunhofer.de

**Johannes Hilpert** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, johannes.hilpert@iis.fraunhofer.de

**Alexander Hofmann** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, alexander.hofmann@iis.fraunhofer.de

**Christian Hofmann** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, christian.hofmann@iis.fraunhofer.de

**Hans-Peter Hohe** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, hans-peter.hohe@iis.fraunhofer.de

**Alexander Ihlow** Fraunhofer IIS, Projektgruppe Drahtlose Verteilsysteme DVT, Helmholtzplatz 2, 98693 Ilmenau, Germany, alexander.ihlow@iis.fraunhofer.de

**Simon Jablonski** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, simon.jablonski@iis.fraunhofer.de

**Alexander Jaschke** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, alexander.jaschke@iis.fraunhofer.de

**Stephan Junger** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, stephan.junger@iis.fraunhofer.de

**Markus Kallinger** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, markus.kallinger@iis.fraunhofer.de

**André Kaup** Chair of Multimedia Communications and Signal Processing, University of Erlangen-Nürnberg, Cauerstraße 7, 91058 Erlangen, Germany, kaup@lnt.de

**Joachim Keinert** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, joachim.keinert@iis.fraunhofer.de

**Uwe Knoechel** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, uwe.knoechel@eas.iis.fraunhofer.de

**Bernd Kollmannthaler** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, bernd.kollmannthaler@iis.fraunhofer.de

**Christian Küblbeck**† Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany

**Fabian Küch** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, fabian.kuech@iis.fraunhofer.de

**Achim Kuntz** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, achim.kuntz@iis.fraunhofer.de

**Jérémie Lecomte** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, jeremie.lecomte@iis.fraunhofer.de

**Sebastian Lempert** Center for Intelligent Objects ZIO and Fraunhofer IIS, Dr.-Mack-Str. 81, 90762 Fürth, Germany, sebastian.lempert@iis.fraunhofer.de

**María Luis Valero** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, maria.luis.valero@iis.fraunhofer.de

**Manfred Lutzky** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, manfred.lutzky@iis.fraunhofer.de

**Kristin Majetta** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, kristin.majetta@eas.iis.fraunhofer.de

**Fritz Meier** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, fritz.meier@iis.fraunhofer.de

**Steffen Meyer** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, steffen.meyer@iis.fraunhofer.de

**Heinrich Milosiu** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, heinrich.milosiu@iis.fraunhofer.de

**Stefan Mödl** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, stefan.moedl@iis.fraunhofer.de

**Cécile Mongrédien** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, cecile.mongredien@iis.fraunhofer.de

**Christian Münzenmayer** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, christian.muenzenmayer@iis.fraunhofer.de

**Markus Multrus** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, markus.multrus@iis.fraunhofer.de

**Frederik Nagel** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, frederik.nagel@iis.fraunhofer.de

**Harald Neubauer** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, harald.neubauer@iis.fraunhofer.de

**Max Neuendorf** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, max.neuendorf@iis.fraunhofer.de

**Bernhard Niemann** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, bernhard.niemann@iis.fraunhofer.de

**Thomas Obenaus** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, thomas.obenaus@eas.iis.fraunhofer.de

**Matthias Oberst** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, matthias.oberst@iis.fraunhofer.de

**Steven Oeckl** Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany, steven.oeckl@iis.fraunhofer.de

**Frank Oehler** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, frank.oehler@iis.fraunhofer.de

**Matthias Overbeck** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, matthias.overbeck@iis.fraunhofer.de

**Volker Peters** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, volker.peters@iis.fraunhofer.de

**Alexander Pflaum** Center for Intelligent Objects ZIO and Fraunhofer IIS, Dr.-Mack-Str. 81, 90762 Fürth, Germany, alexander.pflaum@iis.fraunhofer.de

**Jan Plogsties** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, jan.plogsties@iis.fraunhofer.de

**Alexander E. Popugaev** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, alexander.popugaev@iis.fraunhofer.de

**Holger Priwitzer** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, holger.priwitzer@eas.iis.fraunhofer.de

**Sven Reitz** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, sven.reitz@eas.iis.fraunhofer.de

**Nikolaus Rettelbach** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, nikolaus.rettelbach@iis.fraunhofer.de

**Julien Robilliard** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, julien.robilliard@iis.fraunhofer.de

**Günter Rohmer** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, guenter.rohmer@iis.fraunhofer.de

**Alexander Rügamer** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, alexander.ruegamer@iis.fraunhofer.de

**Tobias Ruf** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, tobias.ruf@iis.fraunhofer.de

**Martin Rulsch** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, martin.rulsch@iis.fraunhofer.de

**Josef Sauerer** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, Josef.sauerer@iis.fraunhofer.de

**Timo Schlarb** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, timo.schlarb@iis.fraunhofer.de

**Peter Schmitt** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, peter.schmitt@iis.fraunhofer.de

**Peter Schneider** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, peter.schneider@eas.iis.fraunhofer.de

**Markus Schnell** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, markus.schnell@iis.fraunhofer.de

**Michael Schöberl** Chair of Multimedia Communications and Signal Processing, University of Erlangen-Nürnberg, Cauerstraße 7, 91058 Erlangen, Germany, schoeberl@lnt.de, michael.schoeberl@iis-extern.fraunhofer.de

**Werner Schön** Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany, werner.schoen@iis.fraunhofer.de

**Oliver Scholz** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, oliver.scholz@iis.fraunhofer.de

**Mario Schühler** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, mario.schuehler@iis.fraunhofer.de

**Lars Seifert** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, lars.seifert@iis.fraunhofer.de

**Jochen Seitz** Chair of Information Technology, University of Erlangen-Nürnberg, Am Wolfsmantel 33, 91058 Erlangen, Germany, seitz@like.eei.uni-erlangen.de

**Andreas Silzle** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, andreas.silzle@iis.fraunhofer.de

**Markus Stahl-Offergeld** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, markus.stahl-offergeld@iis.fraunhofer.de

**Dirk Steckhan** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany

**Jörn Stolle** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, joern.stolle@eas.iis.fraunhofer.de

**Matthias Struck** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, matthias.struck@iis.fraunhofer.de

**Daniel Tantinger** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, daniel.tantinger@iis.fraunhofer.de

**Jörn Thielecke** Chair of Information Technology, University of Erlangen-Nürnberg, Am Wolfsmantel 33, 91058 Erlangen, Germany, thielecke@like.eei.uni-erlangen.de

**Oliver Thiergart** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, oliver.thiergart@iis.fraunhofer.de

**Sabino Titomanlio** M.B.I. S.r.l., via F. Squartini 7, Loc. Ospedaletto, 56121 Pisa, Italy, stitomanlio@mbigroup.it

**Wladimir Tschekalinskij** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, wladimir.tschekalinskij@iis.fraunhofer.de

**Norman Uhlmann** Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany, norman.uhlmann@iis.fraunhofer.de

**Franz Uhrmann** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, franz.uhrmann@iis.fraunhofer.de

**Aharon Vargas** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, aharon.vargas@iis.fraunhofer.de

**Thorsten Vaupel** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, thorsten.vaupel@iis.fraunhofer.de

**Nanko Verwaal** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, nanko.verwaal@iis.fraunhofer.de

**Virginia Voland** Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany, virginia.voland@iis.fraunhofer.de

**Matthias Völker** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, matthias.voelker@iis.fraunhofer.de

**Thomas von der Grün** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany thomas.vondergruen@iis.fraunhofer.de

**Rainer Wansch** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, rainer.wansch@iis.fraunhofer.de

**Elmar Wasle** TeleConsult Austria GmbH, Graz, Austria, ewasle@teleconsult-austria.at

**Norbert Weber** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, norbert.weber@iis.fraunhofer.de

**Christian Weigand** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, christian.weigand@iis.fraunhofer.de

**Thomas Wenzel** Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany, thomas.wenzel@iis.fraunhofer.de

**Andreas Wilde** Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany, andreas.wilde@eas.iis.fraunhofer.de

**Stephan Wilde** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, stephan.wilde@iis.fraunhofer.de

**Nicolas Witt** Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany, nicolas.witt@iis.fraunhofer.de

**Thomas Wittenberg** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, thomas.wittenberg@iis.fraunhofer.de

**Daniel Wolf** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, daniel.wolf@iis.fraunhofer.de

**Thorsten Zerfaß** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, thorsten.zerfass@iis.fraunhofer.de

**Jinghua Zhong** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, jinghua.zhong@iis.fraunhofer.de

**Haiyan Zhou** Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, haiyan.zhou@iis.fraunhofer.de

# Part I
# Circuits

# Advanced Circuit Design and Design Automation for Electronics and Heterogeneous Systems

**Günter Elst, Peter Schneider, Josef Sauerer, Andreas Wilde, and Manfred Dietrich**

**Abstract** The ongoing miniaturization as well as advances in packaging technology paves the way to more powerful, intelligent and complex electronic systems, which are used in many applications. Furthermore new ideas to apply modern manufacturing technology enable new sensor concepts. This paper summarizes the work of the Fraunhofer IIS in the field of design of electronic systems. Current challenges of design for manufacturability and reliability as well as design of highly complex systems are discussed and directions are given for exploitation of the advantages in manufacturing technology.

## 1 Introduction

Microelectronics is the most important driver for innovations in numerous industrial sectors. Thanks to advances in modern manufacturing technology highly complex and intelligent systems can be built at low cost. Together with the worldwide trend towards new electronic products such as smart mobile devices, autonomous intelligent systems etc. this leads to increasing pervasion of nearly all technical systems with electronics.

However, microelectronic products must feature rich functionality and high performance as well as high quality, dependability and reliability in order to be competitive on the market. An essential precondition for competitive products is a well-suited design process.

Currently the most important challenges in design of electronic and heterogeneous systems are

Günter Elst (✉)
Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany
guenter.elst@eas.iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_1, © Springer 2011

- the reduction of the energy consumption, which is necessary for environmental as well as for functional reasons,
- the influences of the manufacturing process and operating conditions on the electrical behavior, which must be known, accounted for in the design and minimized,
- the high complexity and heterogeneity of the systems, which yields larger and more comprehensive models for simulation.

The Fraunhofer Institute of Integrated Circuits IIS meets these challenges by developing and applying new methods for the design automation of electronic and heterogeneous systems. The elaborated models, methods, and tools enable a fast implementation of product specifications into circuits or complete systems. They complement commercial tools and application specific design flows.

## 2 Technology-Aware Design: Manufacturability and Reliability

During the last decades feature sizes on chips kept decreasing which yielded increasing on-chip complexity at more or less constant die sizes. At the same time many new packaging and assembly concepts evolved causing a huge variety of realization possibilities.

Variations in the manufacturing processes for nanoelectronics lead to considerable variations of the devices' electrical behaviour and hence the complete circuits. Knowing the impact in terms of parameter variations in the corresponding behavioral models it is possible to minimize these influences, e.g. with an adequate circuitry and tolerance optimization during the design process. The impact of process variations on power dissipation and time delay in digital circuits was investigated and appropriate design methods were developed [1].

The influence of device aging on circuit behavior can be explored in simulations if equivalent aging models of degradation mechanisms are available and the local stress at each device is determined. As an important effect hot carrier lifetime degradation was investigated and corresponding aging models are used in an aging simulation flow. Furthermore, the aging models are used to compute safe operating area (SOA) diagrams that allow the determination of operational regimes which might cause premature degradation. These analysis methods support design verification to ensure robustness and reliability of the produced circuit [2].

Recently, one prominent development in assembly technology was the formation of stacks of thinned dies, which are connected electrically by so called through silicon vias (TSVs). TSVs allow for very short interconnect lengths compared to bond wires. However, due to the increased density of active elements in the 3-D stack new issues of signal integrity and thermal management occur. Another serious problem is the generation of thermo-mechanical stress which has an immediate influence on the circuit behavior and reliability. Furthermore, testing of the very complex stacked systems is crucial task for the future. For further details see [3].

## 3 Design of Highly Complex and Heterogeneous Systems

The increased complexity of current electronics system arising from continuing device miniaturization, new integration concepts and a heterogeneous system environment demand improved design methods, tools and languages.

Fraunhofer IIS/EAS has made significant contributions to the standardization of the SystemC AMS language for modeling and efficient simulation of mixed signal systems. Following the SystemC philosophy, the SystemC AMS extensions focus on abstract modeling to permit overall system-level simulations of complex hardware software systems in "realtime" application scenarios [4]. System components or parts of circuits can be modeled at various abstraction levels which enables detailed examinations of the electrical behavior of a partial circuit embedded in less detailed system model. The possibility of flexible hardware/software co-simulation as well as the possibility to integrate non-electrical models meet the requirements of industrial designers and are key selling points of SystemC AMS [5].

The object oriented modeling language Modelica was designed for description and simulation of complete multi-physical systems. Beside the physical model of a machine describing electrical, mechanical, hydraulic or other effects and interactions the control algorithms can be specified. More details are described by Clauss et al. [6].

The mathematical model of a system description with Modelica and also with SystemC AMS is a so-called hybrid system. The solution of such systems demands algorithms to solve the equations of the time-continuous and time-discrete parts. The current work aims at algorithms and tools for computation of models with structural changes during operation [6, 7].

In the future new concepts for fault tolerant systems covering redundant hardware setups as well as intelligent fault detection and reconfiguration approaches will be developed, which are based on multiple realization of critical sub-functions, extension of the system function with redundant parts, and an efficient online diagnostics.

## 4 Circuits and Systems for More-than-Moore Applications

Continuous research in semiconductor technologies to further follow Moore's law for smaller and smaller feature sizes brings along new materials in semiconductor processes, smaller feature sizes, new devices and higher digital integration density. These achievements cross-fertilize More-than-Moore applications by opening new possibilities for heterogeneous systems.

One example is nanostructured optical filters. As width and spacing of metal structures in today's CMOS processes are below the wavelength of visible light these layers, normally only used for electrical signal and power distribution, can be used as optical devices in front of optical sensors like photo diodes. Hence, high speed polarization cameras become feasible without any additional processing steps

opening new applications in numerous areas like driver assistance or traffic control systems or quality control. Spectral filters built this way profit from enhanced transmission caused by surface Plasmon effects.

There are also structures on standard CMOS processes which can be used as sensors. The properties like sensitivity or offset of such devices sometimes are not as good as those of sensors made in dedicated technologies. However, these sensors can easily be combined with comprehensive analog and digital signal processing to improve performance. With this approach novel solutions for intelligent sensor nodes become possible employing self calibration or self monitoring which is especially important for safety critical systems. Hall-Sensors are a prominent candidate for CMOS compatible sensors but due to the low lateral dimensions of CMOS technologies vertical Hall Sensors suffered from poor performance. But operating them in an intelligent way and using comprehensive offset and sensitivity control, 3-D Hall-Sensors become possible opening up new solutions for robust, multidimensional position sensing.

Analog-to-digital converters are key devices for mixed signal applications. But classical circuit topologies suffer from low supply voltages and high device variability coming along with advanced technologies. New ADC topologies and circuit approaches employing "digital assisted analog" help to overcome these limitations.

Wireless communication is another important working area. Autonomous systems will be necessary in numerous future applications where low power dissipation and high bandwidth efficiency are important. Research projects cover extremely low power always-on wake-up receiver, cognitive radio architecture for bandwidth efficient multi-standard receiver and multidirectional antennas with low form-factors.

## References

1. Dietrich M, Eichler U, Haase J (2010) Digital Statistical Analysis Using VHDL – Impact of Variations on Timing and Power Using Gate-Level Monte Carlo Simulation. Proc Design, Automation & Test in Europe (DATE), Dresden, March 8–12, pp 1007–1010
2. Jancke R, Ellmers Ch, Gaertner R (2009) Design of Reliable Circuits by Determination of SOA Borders as Part of the Degradation Analysis. Semiconductor Conference Dresden (SCD), April 29–30
3. Schneider P, Elst G (2008) Design for 3-D Integration. In Handbook of 3-D Integration, 2 vols, Technology and Applications of 3-D Integrated Circuits. Ramm P, Bower Ch, Garrow Ph (eds), Wiley-VCH publishing, ISBN 3527320342
4. Open SystemC Initiative (OSCI) AMS Working Group (2010) Standard SystemC AMS Extensions Language Reference Manual. http://www.systemc.org/downloads/standards/ams10/
5. Open SystemC Initiative (OSCI) AMS Working Group (2008) An Introduction to Modeling Embedded Analog/Mixed-Signal Systems using SystemC AMS Extensions. 7th Symposium on Electronic System-Level Design with SystemC, Anaheim, California, USA
6. Clauß C, Donath U, Enge-Rosenblatt O, Majetta K (2011) Modelica. A modeling language for heterogeneous systems. In A. Heuberger, G. Elst, R. Hanke (eds.) Microelectronic Systems – Circuits, Systems & Applications, Springer, Heidelberg

7. Enge-Rosenblatt O, Bastian J, Clauß C, Schwarz P (2007) Numerical Simulation of Continuous Systems with Structural Dynamics. 6th EUROSIM Congress on Modelling and Simulation, Ljubljana, Slovenia, Proc on CDROM (EUROSIM, Simulation News Europe 18 (2008), No 2, pp 24–32, ISSN 0929-2268)

# Nanostructured Optical Filters in CMOS for Multispectral, Polarization and Image Sensors

**Norbert Weber, Jürgen Ernst, Stephan Junger, Harald Neubauer, Wladimir Tschekalinskij, and Nanko Verwaal**

**Abstract** Sub-wavelength gratings and hole arrays in metal films are applicable for polarization and spectral selective sensors, respectively. We demonstrate the fabrication of wire grid polarizers using standard complementary metal–oxide semiconductor (CMOS) processes. Extraordinary optical transmission of hole arrays was achieved by using the dedicated layer of a modified CMOS process. The structures were simulated using the finite-difference time-domain (FDTD) method and fabricated using the work flow of integrated circuits. A high-speed polarization image sensor with a pixel size of 6 μm was designed and demonstrated, and multispectral sensing was implemented using nanostructures with different spectral filter performances on a single chip.

## 1 Introduction

The optical properties of nanostructures in metal films have been intensively studied by many groups [1–3] and can be used for sensing and imaging devices. Sub-wavelength gratings and hole arrays are applicable for polarization and spectral selective sensors, respectively. Conventional optical sensors require additional optical elements for the wavelength- and polarization selective detection of light, i.e. dedicated filter layers deposited on chip or external filters. In this paper we demonstrate a concept where the filtering is performed by optical nanostructures which are fabricated directly within the metal layers of a CMOS semiconductor process. Together with CMOS photodiodes implemented below the filtering nanostructures, novel integrated devices for polarization imaging and multispectral sensing are generated on a single chip. In this way hybrid elements are generated that combine nanostructured filters and photo diodes, offering additional functionality ("More than Moore").

Norbert Weber (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
norbert.weber@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_2, © Springer 2011

## 2 Methodology

The aim of the work presented in this paper is to generate nanostructures during the CMOS process without additional post-processing. Advanced CMOS processes with gate-lengths of 180 nm and below and image sensor processes allow the fabrication of structures with critical dimensions of 100–200 nm for metal layers. Typically these metal layers (aluminum in most cases) are used for electrical interconnections only. If the applicable design rules of the process are considered, the metal layers available by default can be patterned in order to achieve optical functionality. For improved optical performance dedicated "optical" layers without any electrical constraints like ampacity and electromigration are advantageous. Tailored optical filters can be achieved by appropriate design of the lateral shape of the nanostructured layers at a fixed layer thickness, and the layout of the patterns is defined using the typical work flow when designing integrated circuits.

The sub-wavelength structures used in this work are wire grid polarizers and hole arrays, fabricated in standard and enhanced CMOS processes, respectively (Fig. 1).

The optical properties of these nanostructures were simulated using the finite-difference time-domain (FDTD) method, using the software package MEEP [4] from MIT as well as the commercial software package OptiFDTD [5]. FDTD is based on a discrete decomposition of the Maxwell equations into spatial and temporal steps, using numerical techniques to calculate the evolution of the electromagnetic fields in time. Provided that precise material models and adequate spatial resolutions are used for simulation, FDTD yields reliable results matching numerous experiments.

Figure 2 shows an example of an FDTD simulation. The simulated structure is an array of holes with diameter 200 nm and period 320 nm, patterned in a 200 nm thick aluminum layer which is embedded in silicon dioxide, a typical material system of a CMOS process.

The peak wavelength of approximately 570 nm (green) is due to a resonant transmission based on surface plasmon effects ("enhanced optical transmission") and depends mainly on the period of the array, while the filter bandwidth is influenced significantly by the diameter of the holes. The minimum at 480 nm can be interpreted as Rayleigh–Wood anomaly.

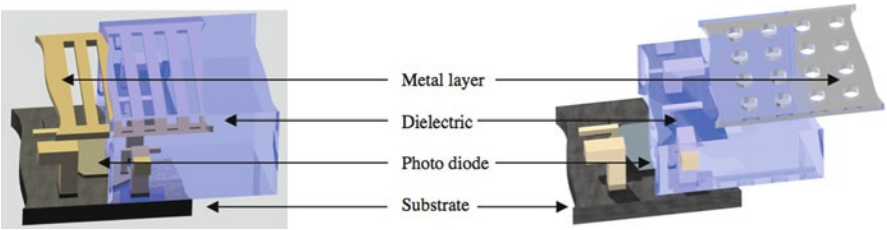



**Fig. 1** CMOS photo diodes with sub-wavelength grating using metal layers of standard CMOS (*left*) and with sub-wavelength hole array in dedicated metal layers (*right*)

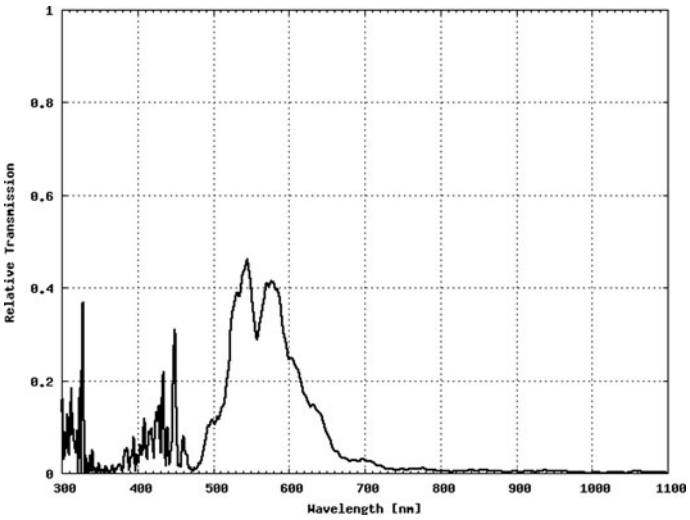



**Fig. 2** Simulated transmission of a sub-wavelength hole array in 200 nm thick aluminum layer

## 3 Novel Sensor Types

### *3.1 Polarization Image Sensor*

Sub-wavelength gratings have been used as wire grid polarizers (WGP) for a long time for radio frequencies. The recent progress of semiconductor processing enables their fabrication for infrared and visible light, too. This concept can be applied to single polarization sensors, and also to pixel arrays where different orientations of the gratings allow the detection of polarized light for polarization imaging as described in our previous work [6]. Figure 3 shows the arrangement of such a pixel and polarizer array in a comparable manner as the well-known Bayer matrix. In this case, a metapixel consists of 4 sub-pixels with a reference and three grid-covered pixels with an orientation of 0°, 45° and 90°.

In order to demonstrate this concept, a dedicated polarization image sensor was designed and fabricated using an unmodified CMOS image sensor process (UMC 0.18 µm CIS). In a test chip, an array of 560 × 254 pixels with a size of 6 µm each was covered by an array of polarization filters. The resolution was chosen, so that the imager fits on one 5 × 5 mm multi-project-wafer (MPW) tile. The architecture and electronics of the readout was designed to make a resolution of up to 4000 × 3000 pixel possible. The imager has an analog correlated double sampling circuit with a differential output. To suppress reset noise a pinned (4T) pixel was chosen. After illumination first the reset signal of the floating diffusion is sampled. Then the accumulated charge is transferred to the readout node and this signal is sampled. Through subtraction of these two values the reset noise is eliminated. The difference is converted to a differential signal at this stage. Afterwards the signal is

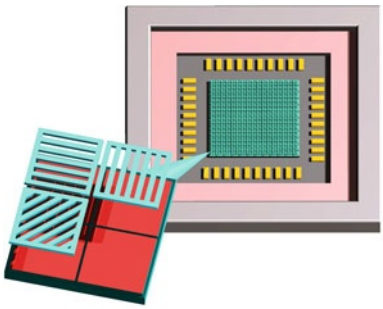


**Fig. 3** Pattern of the nanowire polarization filter array used for a CMOS polarization image sensor

passed through the output buffers to the camera. The sensor has 8 differential analog outputs. Each output has a throughput of 40 MPix/s, which enables high speed imaging with this sensor. 72 columns are handled by one output. To suppress block artifacts a special method for the readout was implemented. Columns and rows can be addressed individually, so parts of the sensor can be read out individually. This can be further used to increase sensor speed. The smallest period of the grating allowed by the design rules of the process was 400 nm, yielding a good polarization selectivity in the near infrared above 800 nm. In Fig. 4, the measurement of a single grid-covered photodiode is shown when illuminated with TE- and TM-polarized light, respectively. For wavelengths beyond 800 nm the transmission for TE polar-

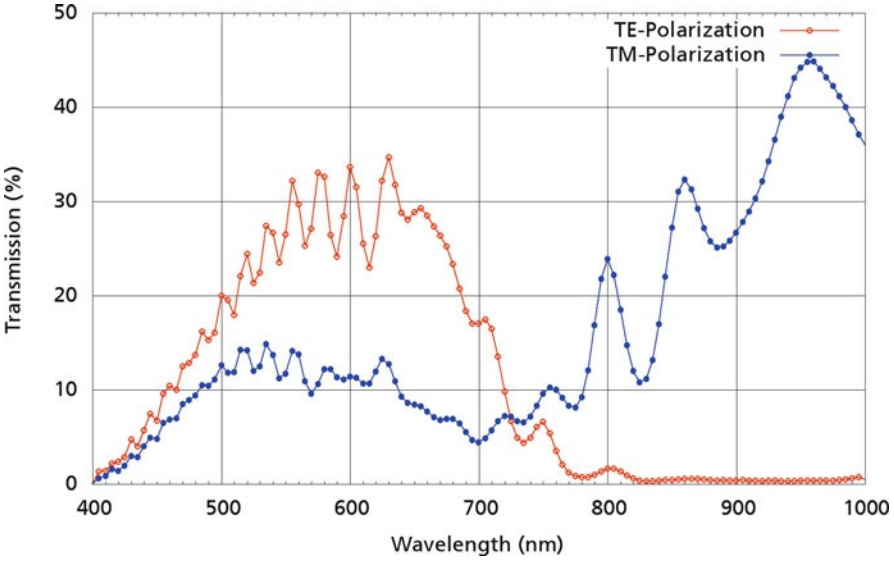



**Fig. 4** Measured spectral transmission of a grating with a period of 400 nm for TE- and TM-polarized light

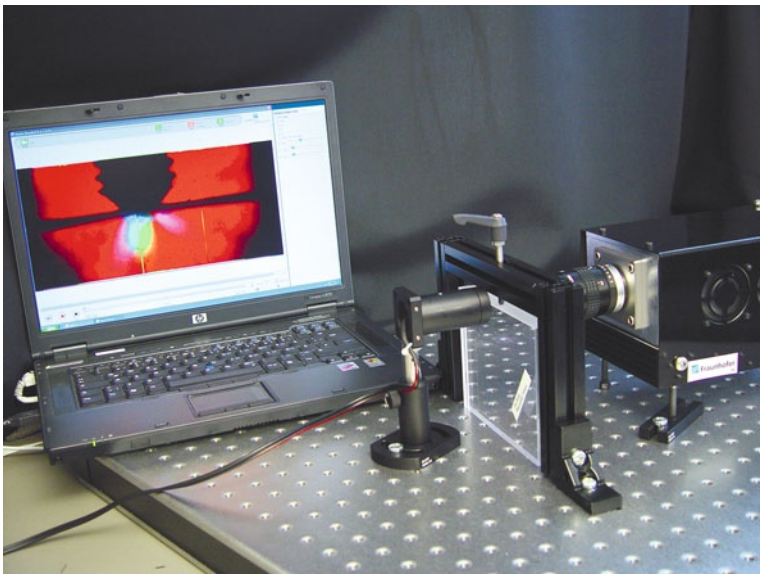


**Fig. 5** Polarization camera analyzing stress birefringence of a polymer slab

ized light is very low, while TM polarized light has an increasing transmission in the near infrared. This measured relative transmission is in good agreement with the simulation results.

The polarization image sensor (according to Fig. 3) was built into a prototype camera setup which is able to be operated at high frame rates of up to 1000 frames per second.

In a test arrangement, the stress birefringence of a polymer slab can be visualized, where the calculated Stokes parameters are shown as false colors (Fig. 5). A new generation of a compact camera system for industrial applications is currently under development (Fig. 6).


**Fig. 6** Novel compact camera system for industrial polarization imaging

### *3.2 Multispectral Sensor*

For polarization sensing with grating polarizers a standard CMOS process can be used if near infrared illumination is acceptable. The generation of color filters based on sub-wavelength hole arrays, however, requires process extensions with dedicated "optical" layers in order to achieve a good filter performance. Several test chips with varying nanostructures for color and multispectral sensing were fabricated by LFoundry GmbH, Landshut, Germany, using an extended 150 nm CMOS process (LF150). Diffraction is observed at small incident angles, dependent on the period

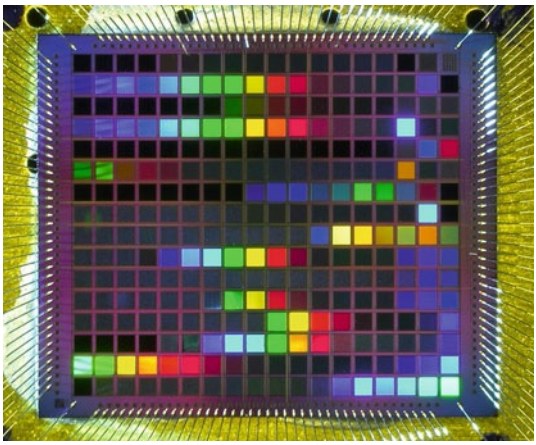


**Fig. 7** Diffractive effects of different nanostructures on a test chip with white light illumination

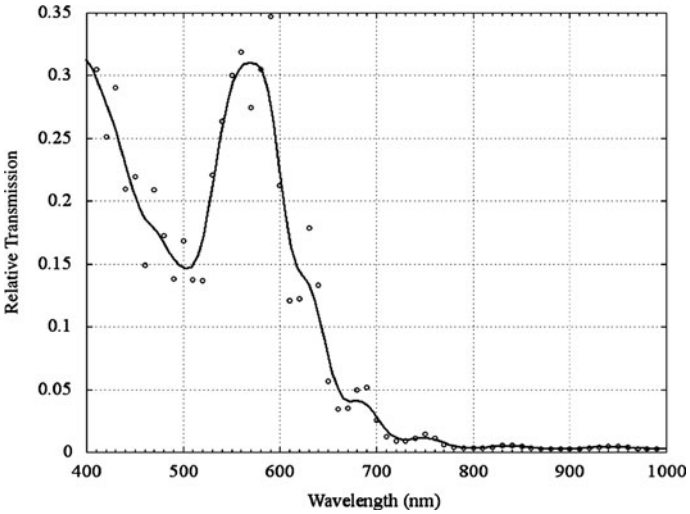



**Fig. 8** Measured spectral transmission of a hole array (hole diameter 200 nm, period of the array 320 nm)

of the structures as expected (Fig. 7). The size of each photo diode is $300 \times 300\ \mu m^2$, the total size of the chip approximately $8 \times 7\ mm^2$.

The spectral responsivity of a reference photo diode and the photo diodes covered by the nanostructures was measured using a white light source and a monochromator. The light of the output fiber was collimated first and focused onto the photo diodes using an achromatic lens. In order to eliminate the performance of the CMOS photo diode, the spectral transmission of the nanostructures was calculated as the ra-

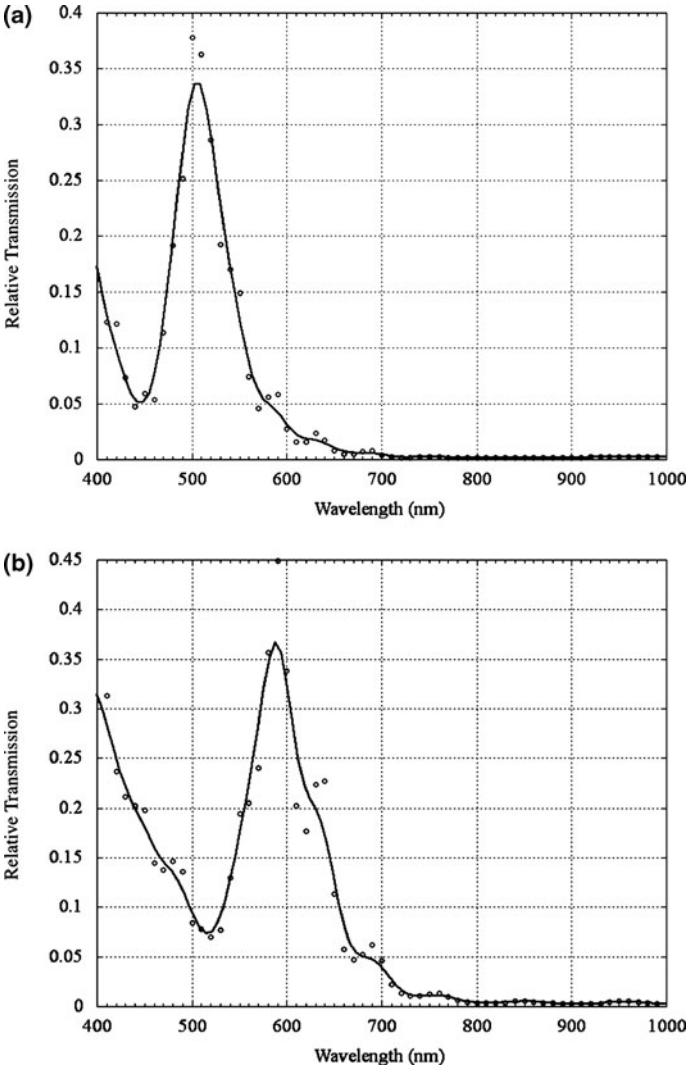



**Fig. 9** Relative transmission of (**a**) a *green* band pass filter (510 nm) with a peak transmission of 35%, (**b**) a *yellow* band pass filter (590 nm) with a peak transmission of 35%

tio of the measured spectral responsivity of the device under test and the reference photo diode.

Figure 8 shows the measured spectral transmission of a hole array with the same geometry as used for the simulation given in Fig. 2. The peak wavelength of 580 nm is in agreement with simulation, the minimum at approximately 490 nm is not as distinct but still observable.

A further improvement of the band pass characteristics can be achieved if a second optical layer is used. Optimizing both layers offers significant degrees of design freedom but requires extensive simulations. The result of this optimization procedure is shown in Fig. 9a, where the filter bandwidth is reduced without lowering the peak transmission of approximately 35%.

The possibility of tailoring the wavelength of the band pass filter is an important feature for implementing many photo diodes with different spectral sensitivities on a single chip for on-chip color and multispectral sensors. In Fig. 9b, the measurement of a two-layer nanostructure designed for a different wavelength is shown to illustrate this possibility.

## 4 Conclusions

We demonstrated the feasibility to fabricate wire grid polarizers using standard complementary metal–oxide semiconductor (CMOS) processes, where the existing aluminum layers, intended for electrical interconnects only, are used for patterning the nanostructures. Regarding hole arrays with extraordinary optical transmission, a band pass filter performance of the plasmonic structures with peak transmissions of 35% was achieved by using the dedicated layer of an enhanced CMOS process. The tailoring of the transmission spectra by varying hole size and period of the hole array was evident. All metallic structures were simulated using the finite-difference time-domain (FDTD) method and layouted using the work flow of integrated circuit design. The spectral response of uncovered photodiodes and photodiodes including nanostructured layers were measured and the relative transmission of the on-chip filters was calculated for different polarization states of the incident light. Based on these results, a high-speed polarization image sensor with a pixel size of 6 $\mu$m was designed, fabricated, and demonstrated using a camera system developed for polarization imaging. Multispectral sensing was implemented on several test chips using arrays of photodiodes and nanostructures with different spectral pass band characteristics.

## Acknowledgements

This work is supported by the funding program "Zentrales In-novationsprogramm Mittelstand (ZIM)" of the German Federal Ministry of Econo-mics and Technology, project "NanoOptoSensors" (KF2316703). The cooperation with Landshut Silicon Foundry GmbH (LFoundry) is gratefully acknowledged.

## References

1. Ahn S-W et al (2005) Fabrication of a 50 nm half-pitch wire grid polarizer using nanoimprint lithography. Nanotechnology 16:1874–1877
2. Catrysse PB, Wandell BA (2003) Integrated color pixels in 0.18 μm complementary metal oxide semiconductor technology. J. Opt. Soc. Am. A 22:2293
3. Ebbesen TW, Lezec HJ, Ghaemi HF, Thio T, Wolff PA (1998) Extraordinary optical transmission through sub-wavelength hole arrays. Nature (London) 391:667
4. Oskooi AF, Roundy D, Ibanescu M, Bermel P, Joannopoulos JD, Johnson SG (2010) MEEP: A flexible free-software package for electromagnetic simulations by the FDTD method. Computer Physics Communications 181:687–702
5. OptiFDTD by Optiwave Corporation, Ottawa, Ontario, Canada
6. Junger S, Tschekalinskij W, Verwaal N, Weber N (2010) Polarization- and wavelength-sensitive sub-wavelength structures fabricated in the metal layers of deep submicron CMOS processes. Proc. SPIE 7712:7712 0F

# Electronic Design Automation for Implementation of 3-D Integrated Systems

**Uwe Knoechel, Andy Heinig, Jörn Stolle, Sven Reitz, and Andreas Wilde**

**Abstract** The technology of vertical integration using Through Silicon Vias (TSVs) now is mature for commercial products with a smaller form factor, better performance, less power consumption and lower cost. This paper addresses two challenges faced with the design using vertical integration. First, methods for the characterization of the physical behavior of the new interconnect structures are described. Second, since issues of reliability and thermal management become more important and difficult and the design space is drastically larger, new early stage design tools are needed. A new floorplanning flow is introduced which supports the cost and performance optimized implementation of digital systems in a stack.

## 1 Introduction

After several years of research, the technologies for 3-D integration of electronic systems are ready for commercial applications. They enable the dense integration of different dies (e.g. analog, CMOS, sensors) in a single package.

First 3-D products were image sensors integrated with high performance processors [1]. Today products are mainly based on the reuse of existing designs. Interposers provide the redistribution of signals and interconnections between individual dies of a stack as shown in Fig. 1. This is sometimes called 2.5D design.

Through silicon vias (TSV) enable full 3-D-integration by direct interconnection between dies without interposers. This offers a variety of possibilities for new system concepts:

- higher performance, due to reduced interconnect length,
- lower power consumption by using dedicated silicon technologies for each subsystem,

Uwe Knoechel (✉)
Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany
uwe.knoechel@eas.iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_3, 

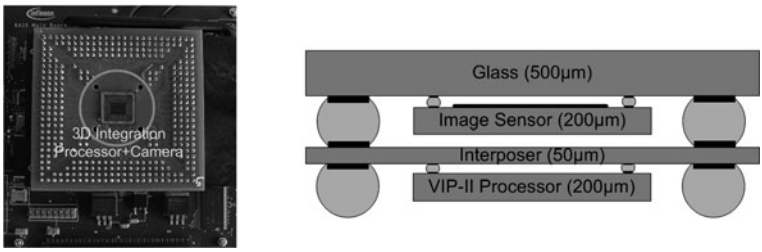



**Fig. 1** Interposer based integration of image sensor and processor [1]

- reduction of manufacturing costs for large SoC or multi-processor setups [2], whereat the yield is reduced with increasing die size
- and small form factors for applications in smart systems technology [3].

Assuming, that the design starts from scratch, the design space is large. Options are e.g. the number of dies and their individual silicon technologies; interconnect technologies and the order of dies in stack; and the partitioning and placement of functional blocks. Any choice of these options has an impact on performance, yield, reliability, testability, and production costs of a product.

Design tools made for 2-D support the design of the individual dies but cannot handle problems of stacking, such as electrical behavior of through silicon vias, heat transfer in a stack, or electromagnetic coupling between different layers. Therefore, 3-D design demands new EDA solutions that support detailed analyses on physical level and provide a modeling strategy how to consider those effects on system level (see Sect. 2). Based on knowledge and models of the multi-physical dependencies in a stack, a cost and thermal aware 3-D floorplanning algorithm is introduced in Sect. 3.

## 2 Physical Level Analyses in 3-D Design

Due to the dense interconnections between the stacked layers and increasing operating frequencies, parasitics like skin and proximity effect, cross talk or signal delays may influence the electrical function of the entire 3-D system significantly. Especially the RF behavior of the TSVs, (Fig. 2) has to be taken into account.

Furthermore, the very close placement of functional blocks may affect function and lifetime, e.g. due to thermal and thermo-mechanical effects. Power losses result in hot spots and require sufficient thermal management. The heating of the stack as well as different processing temperatures may cause additional mechanical stress in the system which may change the behavior of mechanical sensing elements (e.g. pressure sensors) and may cause reliability issues [4].

Regarding these aspects, a system design for 3-D is a multi-criteria optimization task with a huge amount of design parameters. Equivalent network or behavioral

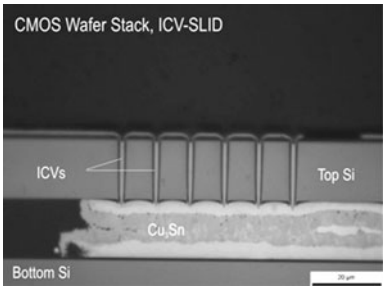



**Fig. 2** 3-D Interconnect structure with through silicon vias (source: Fraunhofer EMFT)

models, design guidelines as well as new modeling approaches and efficient simulation algorithms are necessary for this system design support. Further explanations will be focused on the RF and thermal behavior of the TSVs as the main part of this multi-criteria and multi-physics problem in 3-D system design.

## *2.1 Electrical Characterization of TSVs*

Electromagnetic field solvers, e.g. the full wave simulator Microwave Studio (MWS) from CST, can be used to analyze the influence of interconnect structures (including TSVs) on the behavior of particular subsystems or signal paths. Results of the simulation, which solves the underlying Maxwell equations with the finite integration technique (FIT), are scattering parameters ($S$-parameters) describing exactly the electrical behavior of the structure in the RF domain. The first part of Fig. 3 shows the RF simulation model (MWS) for a TSV with connection pads, silicon oxide layer, and silicon substrate for a preferably detailed copy of the real structure. Results are $E$- and $H$-field distributions as well as $s$-parameters considering skin and proximity effect in the RF domain (Fig. 3, second part).

The obtained $S$-parameters of the MWS-simulation are the basis for matching and optimizing equivalent network models of the investigated structures. By adapting these parametric RLCG network models (Fig. 3, right), the electrical behavior of the TSVs can be represented appropriately and used in circuit simulation, e.g. in Spectre or SPICE. The entire flow of the optimization is realized within our in-house tool MOSCITO Simon. Based on this approach, models for typical interconnect-TSV structures can be generated efficiently.

In addition to the ohmic losses ($R_{\mathrm{TSV}}$) in the depicted RLCG network model, the capacities $C_{\mathrm{OX}}$ (silicon oxide layer) and $C_{\mathrm{SI}}$ (silicon substrate) are also important. They particularly affect the signal delay. Therefore, besides the geometry of the TSVs, oxide thickness, and substrate material (in terms of electrical conductivity) must also be considered.

As dedicated libraries they are included into the system design flow and used for electrical simulation w.r.t. parasitic effects like skin and proximity effect.

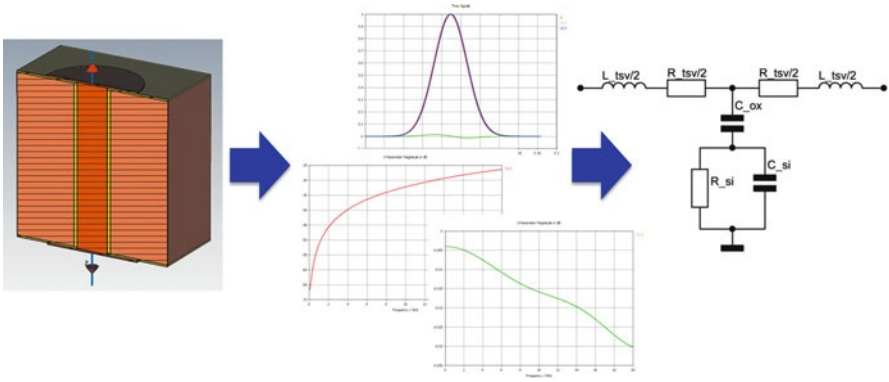



**Fig. 3** CST MWS-model of TSVs with S-parameters, time signals, and RLC-network model

## *2.2 Thermal Characterization of TSVs*

Due to the high packaging density, power dissipating blocks and temperature-sensitive devices might be located very close together. The resulting hotspots may influence the system behavior significantly. Thermal gradients or temperature changes in the system may also have a large impact on the mechanical reliability of the system. Therefore, thermal management is very important for the design of entire 3-D systems.

Figure 4 left depicts a half model of a TSV with ICV-SLID connection to the lower die as well as a transmission line on top. In the right part of the same figure the thermal influence of this TSV structure with a diameter of 10 µm at the proximity of a power device is shown. The temperature of the hot spot is reduced

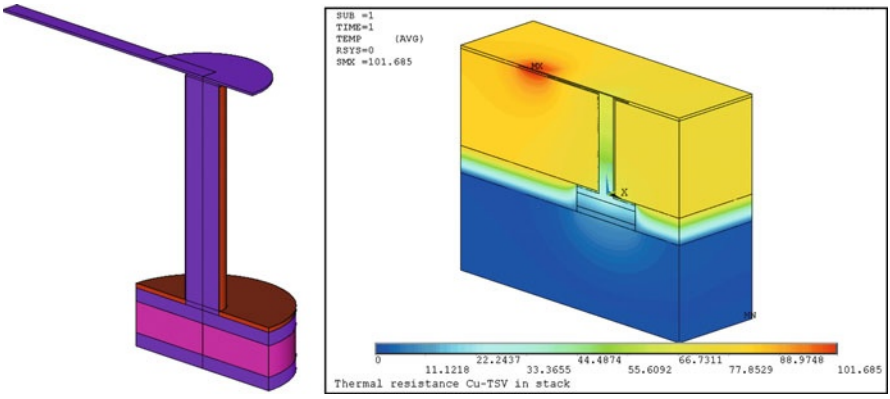



**Fig. 4** *Left*: Half model of TSV with ICV-SLID connection to the lower die as well as transmission line on top – *right*: Temperature distribution of a 3-D stack with power device using this TSV structure with different diameters ($d = 10\,\mu$m shown)

by about 13 degrees to 101 °C by increasing the diameter of the TSV from 5 to 10 µm. The higher the diameter the better the heat flux to lower dies of the stack. Further parameters which have an impact on heat removal and are investigated in parameter studies are the thickness of the oxide between die material and TSV, the distance between TSV and power device as well as width of the transmission line and used materials. Current 3-D systems have hundreds up to thousands of TSVs. Using dedicated pattern of TSVs the heat flow in a 3-D stack can be improved significantly.

## 3 Implementation of Digital Systems in 3-D

While traditional 2-D-designs are mainly optimized for minimum area, in 3-D various interdependencies between performance, multi-physical effects, and costs must be considered. The number and placement of TSVs plays an important role:

- TSVs consume significant die area, which increases wafer costs (at 45 nm node one small TSV takes the area of about 16 NAND cells and an additional keep off zone).
- TSVs can reduce the interconnect length which increases circuit performance.
- TSVs transfer heat within a stack.

Figure 5 depicts a first approach for a design flow based on a 3-D floorplanning algorithm. The number of layers which are efficient for a design and location of the

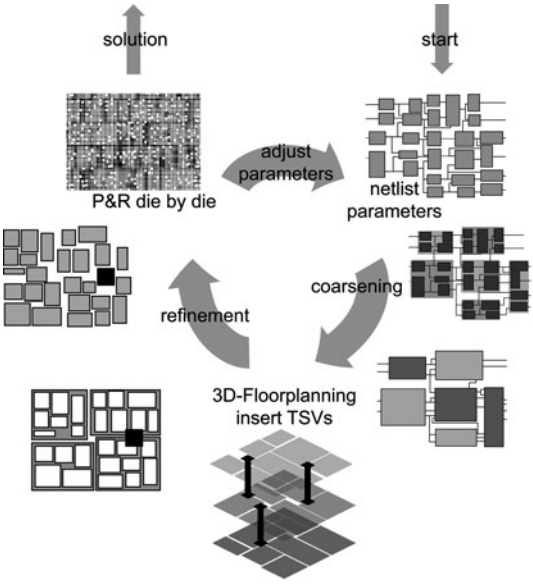



**Fig. 5** Netlist to Layout Flow for 3-D System

functional blocks in a stack are determined during the cycle. After TSV insertion the place and route (P&R) is done by 2-D P&R tools. This cycle of design steps can be repeated for optimization. The design steps are outlined in the following sections.

### *3.1 Coarsening of Gate Level Netlist*

The input of the flow is a digital gate level netlist as commonly used in 2-D design. Because the number of cells at gate level exceeds the capacity of a floorplanner, the cells are partitioned in a coarsening step into so called supercells.

Because TSVs use significant die area, the number of TSVs should be reduced. Each interconnection between two supercells can result in a TSV if those are placed at different dies. Therefore the coarsening step encourages the goal of reducing the number of TSVs by minimizing the number of cuts in interconnects between supercells. In this way, the supercells are created independently from the hierarchy of a design. As a further constraint, the total area of the supercells should be in the same range. Such types of partitioning problems are known as NP-complete. A good heuristic strategy for solving this problem is given by the hMetis algorithm [5] that is used in a modified way.

### *3.2 3-D Floorplanning Algorithm*

The floorplanning step is the major part of the proposed design flow. It places the supercells to the individual dies and within the dies to non overlapping positions. To handle this mapping from supercells to positions at dies, an efficient data structure is needed. Some work was done for developing data structures for the floorplanning of 2-D chips. In the 3-D domain these structures can be extended to either 2.5D structures or real 3-D structures [6].

As shown in Fig. 6, stochastic optimization was selected to compute floorplan candidates. The cost function plays a central role for the evaluation of floorplanning results. It has to consider electrical and multi-physical properties which influence the system performance as well as overall cost of the system. The definition of the cost function and weighting between the mentioned factors is influenced by requirements of the target application and improvement of stacking technologies. Especially the formulation of cost related to thermal and electromagnetic aspects is an ongoing research task that is supported by multi-physic analyzes as outlined in Sect. 2.

Different stochastic optimization techniques such as threshold accepting algorithm, simulated annealing, or great deluge algorithm are used for floorplanning. In [7] we discovered that the threshold accepting algorithm delivers the best results for production cost optimized floorplanning. The algorithm needs operations to generate a new configuration from a current one. Only few changes should be made between two subsequent configurations. Six operations are used in this implementation:

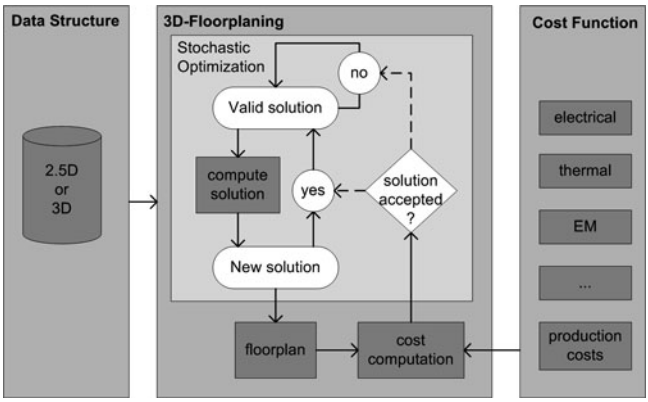



**Fig. 6** Floorplanning of 3-D-systems with stochastic optimization

- Move the position of a supercell at the same die,
- Swap the position of two supercells at the same die,
- Change the aspect ratio from one supercell,
- Move one supercell to another die,
- Swap two supercells between different dies,
- Move the point of origin.

At first a start configuration is chosen randomly. Following, in every optimization step a neighboring configuration is generated by selecting and executing one of those operations on the current solution. This new solution is accepted with a given probability, regarding the evaluation of cost function. The algorithm ends, if a stop criterion is fulfilled, e.g. maximum number of iterations is reached or if almost all of the neighboring solutions are discarded.

## *3.3 Via Insertion*

After the floorplanning of the supercells is done the TSVs are inserted into the floorplan. Because the supercells contain a lot of smaller standard cells, the TSVs are placed into it with priority. The additional area for the TSVs was already considered in the floorplanning step. If all possible TSV positions are already occupied, a search in the neighborhood is done.

## *3.4 Place and Route*

As a last step, place and route of the single dies is done with a standard P&R tool. Therefore, the original netlist is partitioned into individual netlists for each die. Furthermore, scripts are generated which place the TSVs to their determined positions.

The standard cells within the supercells are placed first to the center of the supercell. Later their position is optimized by the P&R tool. Further timing constraints for each die as calculated by the 3-D floorplanner are set by a script. After these preparations, the standard 2-D P&R can place the cells to their best positions considering timing and design rules. The fixed positions of the TSVs guarantee that corresponding TSVs are aligned on their dies. After completion for all dies, the 3-D implementation of the initial gate level netlist is done.

### *3.5 Example*

Figure 7 depicts the 3-D implementation of a high-scalable VLIW-processor designed with the proposed flow. In a 2-D-system the processor was placed at a die with 700 000 $\mu m^2$ and total costs of 100 units. With the proposed method, the same system can be implemented as a stack of three dies with a total area of 643 000 $\mu m^2$. As it is shown in Fig. 2, the cost optimal system consists of dies with different size. The dies cost $35 + 35 + 15$ units. With additional 10 units for TSV processing and stacking, the total manufacturing costs are with 95 units below the 2-D implementation. The cost advantages will growth with design size. Furthermore the timing was improved. The longest path in the 2-D-system has a delay of 2.9 ps whereas in 3-D the delay was reduced by 0.2 ps to 2.7 ps.

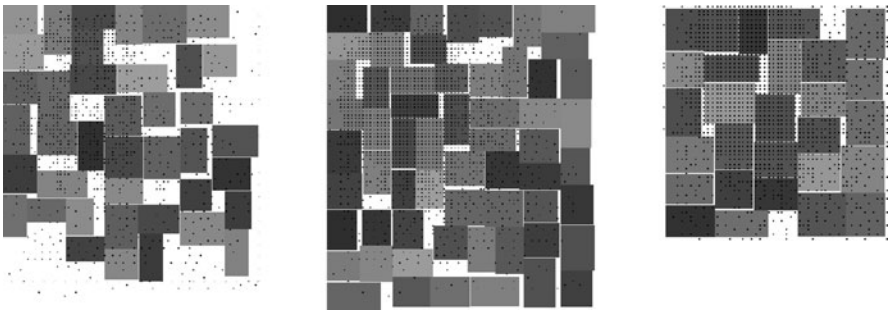


**Fig. 7** Supercell placement of a VLIW processor on three dies (*bottom*, *middle*, *top*)

## 4 Conclusion and Outlook

3-D integration opens up a large design space. Today most 3-D products are based on existing designs that are stacked and connected with interposer layers. Enhanced 3-D integration, with trough silicon vias, requests a common design of all individual dies. Therefore, new design tools are under development. The presented flow

for cost-efficient implementation of digital systems in a stack shows a first concept of a 3-D EDA solution. Beside the manufacturing costs, the system optimization must concern performance issues related to multi-physic effects in a stack. Therefore links between multi-physical analyzes and circuit design tools are needed. The standardization of an open data format for the description of 3-D designs could be an important step for building up a powerful 3-D design environment.

Further research is done to enhance the capability of floorplanning algorithms to support the complexity of future designs, as well as to provide support for stacks that contain digital, analog, RF and sensor components.

## References

1. Limansyah I, Wolf MJ, Klumpp A, Zoschke K, Wieland R, Klein M, Oppermann H, Nebrich L, Heinig A, Pechlaner A, Reichl H, Weber W (2009) 3d image sensor sip with tsv silicon interposer, ECTC 2009, San Diego
2. Karnik T (2010) 3d architectures and cad tools, D43-D Workshop, Lausanne
3. Schneider P, Elst G (2008) Modeling approaches and design methods for 3-D system design. In Garrou P (ed) Handbook of 3-D integration, vol 2, Technology and applications of 3-D integrated circuits, Wiley-VCH, Weinheim, pp 529–574
4. Elst G, Schneider P, Ramm P (2006) Modeling and Simulation of Parasitic Effects in Stacked Silicon. Proc. 2006 MRS Fall Meeting, November 27–December 1, Boston
5. Karypis G, Aggarwal R, Kumar V, Shekhar S (1997) Multilevel hypergraph partitioning: Applications in vlsi domain, 34th Design and Automation Conference, pp 526–529
6. Schneider P, Heinig A, Fischbach R, Lienig J, Reitz S, Stolle J, Wilde A (2010) Integration of multi physics modeling of 3-D stacks into modern 3-D data structures. 2nd IEEE International 3-D System Integration Conference, 3-DIC
7. Heinig A, Schuster T (2011) A 3d stack cost optimization floorplanning and routing approach, DATE Workshop for 3-D Integration, Grenoble

# Analog to Digital Converters for Mixed Signal ASICs and SOCs

**Johann Hauer, Stefan Mödl, Harald Neubauer, Matthias Oberst, Matthias Völker, and Haiyan Zhou**

**Abstract** Analog to Digital Converters (ADCs), key functions for mixed signal ASICs and SOCs stay in the main focus of our work as enabling technology for many different system applications. New requirements of system as well as of technology aspects call for new concepts for analog to digital conversion like direct sensor signal conversion or digital assisted analog functions. Four examples from current research and development projects are depicted in this article.

## 1 Introduction

The future of digital signal processing and data computing is closely linked to the performance of CMOS based processor implementation which is still driven by transistor size reduction and increase in gate density. This is popularly known as Moore's law, referring to a famous postulate of Gordon Moore, cofounder of Intel, who forecasted in the 1970s, that silicon based integrated circuits would have a great future. One of his expectations was, that the number of components (transistors) per given area will double every one or two years due to miniaturization. Nowadays the "International Technology Roadmap for Semiconductors" (ITRS-Roadmap) has overtaken the role of predicting the future of semiconductor process development. For CMOS based leading edge technologies, the ITRS-forecast in 2010 was from 24 nm processes in 2011 to 8 nm in 2022. That the race is still ongoing and even beating the forecast was demonstrated by Intel in June 2011 announcing its 22 nm process ready for production.

A tremendous increase of digital processing power can be foreseen. However the real world is analog and interfacing with the analog world requires analog signal processing, which is not as easy implemented on high density low voltage nanome-

Johann Hauer (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
johann.hauer@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_4, © Springer 2011

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | ...2022 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ASIC Gate length [nm] | 38 | 32 | 29 | 27 | 24 | 22 | 18 | 17 | 15 | 14 | 8.1 |
| ASIC MTransistors/mm$^2$ | 3.57 | 4.49 | 5.66 | 7.14 | 8.99 | 11.3 | 14.3 | 18.0 | 22.6 | 28.5 | 114 |

**Fig. 1** Extract from the ITRS-Roadmap

ter processes. Analog to Digital Converters (ADCs) and Digital to Analog Converters (DACs) are the key elements for connecting digital processing systems to sensors and actors in the real world. The resolution and bandwidth requirements for these interfaces strongly depend on the applications. Most of the interfaces are proprietary and the specifications are tailored to the actual needs of the processing system. Nevertheless some general tendencies can be identified. Typical resolutions of measurement systems for industrial applications range from 10 to 14 bit. Bandwidth requirements range from quasi static measurements of environmental conditions like temperature over many medium bandwidth implementations to very high frequency applications for communication purposes. Application areas are industrial control, audio and video signal applications, medical ultrasonic and Xray imaging and high frequency communication devices.

## 2 Pipeline ADC with Partial Amplifier Sharing

Pipeline ADCs are preferred architectures for high-speed (10–250 MS/s) data conversion at medium to high resolution (8–14 bits). They are employed in a variety of applications such as communication, imaging and fast control systems. Within this architecture, residue amplifiers are known to dominate power dissipation due to the simultaneous demand for low noise, high-speed, and precise linear amplification. This is especially true for the amplifiers in the first few-stages of the pipeline, which have the greatest impact on the ADC overall performance.

Several techniques have been developed to reduce the power dissipation [1]. One consists in reducing the power dissipation of the operational amplifiers. This can be achieved by exploiting stage scaling techniques, where switched capacitor (SC) circuits are determined by noise requirements in each stage. It has been proven that thermal noise contribution of a given stage is reduced by the gain of the previous stages, allowing the reduction of the capacitors size of that stage. Another technique is based on amplifier sharing between adjacent ADC stages working in opposite clock phases. This allows a theoretical reduction of half the number of amplifiers. However, amplifier sharing technique introduces some drawbacks since additional switches are needed in the implementation affecting the stage settling time. The complexity of the pipeline stage is increased and more clock signals with different phases are needed for the switches of the ADC. This is especially critical in the first stages of the pipeline. Therefore also a hybrid architectural version was analyzed

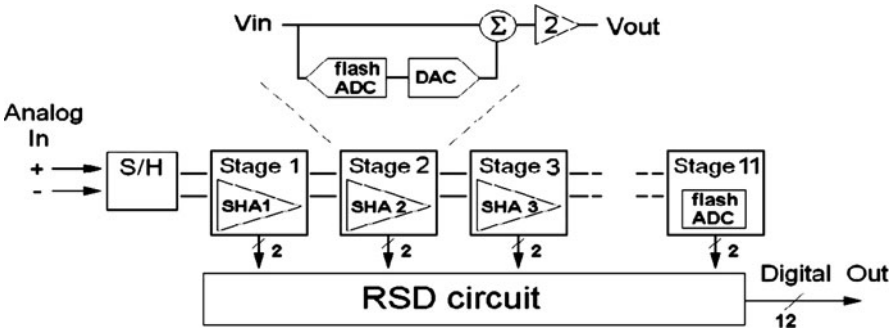



**Fig. 2** Pipeline ADC with conventional scaling

with two conventional stages at the beginning of the pipeline and amplifier sharing for the following stages.

Figure 2 shows a block diagram of a pipeline with conventional scaling. Each stage consists of the same architecture. The improved pipeline with full amplifier sharing is depicted in Fig. 3. For a partial amplifier sharing ADC the two first stages are designed as in conventional pipeline ADCs, each one with its own amplifier while the rest of stages take advantage of amplifier sharing technique. In this way, some of the aforementioned drawbacks, specially the settling time for first stages are minimized. In order to evaluate the performance of these topologies, a comparison of three equivalent converters has been done: a conventional type, one using full amplifier sharing and another one with partial amplifier sharing. The results are based on layout extracted simulations which include parasitic influences.

Simulated power reduction was 40% for the fully shared and 33% for the partially shared type compared to conventional pipeline. For the fully shared pipeline the SNDR drops from 68.1 dB to 66.6 dB, compared to the conventional type. A 20% area reduction could be achieved for the fully shared type compared to conventional pipeline. For the partially shared type, SNDR was simulated to be 68.7 dB. The area is practically the same as for the fully shared type. In the meantime different pipeline architectures are integrated and successfully tested.

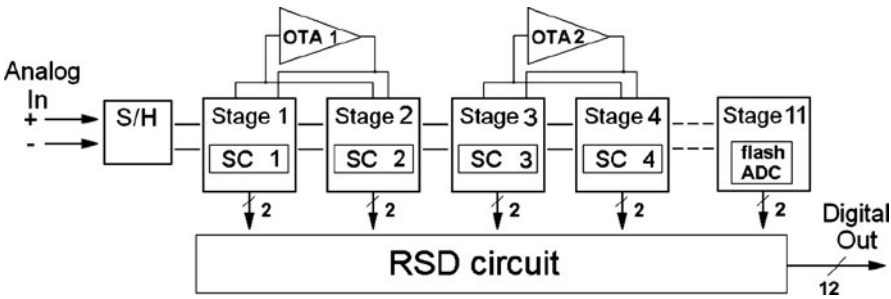



**Fig. 3** Pipeline ADC with amplifier sharing

## 3 Low Power Cyclic 12 Bit@1 MS/s ADC for Industrial Control Application

Position and rotations sensors are used in many industrial applications. Measurement principles are based on optical or capacitive sensing with nonius coded rings and plates. Primarily, this results in a sine signal with a given relation between sine phase and position. There is also a fixed relation between the frequency of the sine signal and the rotation speed. For most applications the signal frequency is between zero and up to 410 kHz, which relates to idleness or a rotation speed of 24 000 rotations/minute for a 1024 resolution encoder. Since in many cases the distance between sensor and processing unit could be quite long, digital data transmission is preferred. For many safety applications like cranes, lifts or robots the sensor acquisition, processing and data transfer must be designed for high reliability and safety levels related to regulations for functional safe product design like Safety Integrity Level (SIL), which are defined by international norm regulations IEC61508/IEC61511.

For industrial position and rotation sensors a low power, low area ADC with 12 Bit resolution @ 1 MS/s was developed. Since the final product should be miniaturized a high level of integration was needed. In the final ASIC two of the ADCs were integrated for a two channel sine-cosine measurement principle. The second part of the ASIC consists of the digital signal processing for a SIL-Level 2 approved data transmission. For the ADC the most critical requirements concerned power consumption and chip area (both as low as possible) and an extended ambient tem-

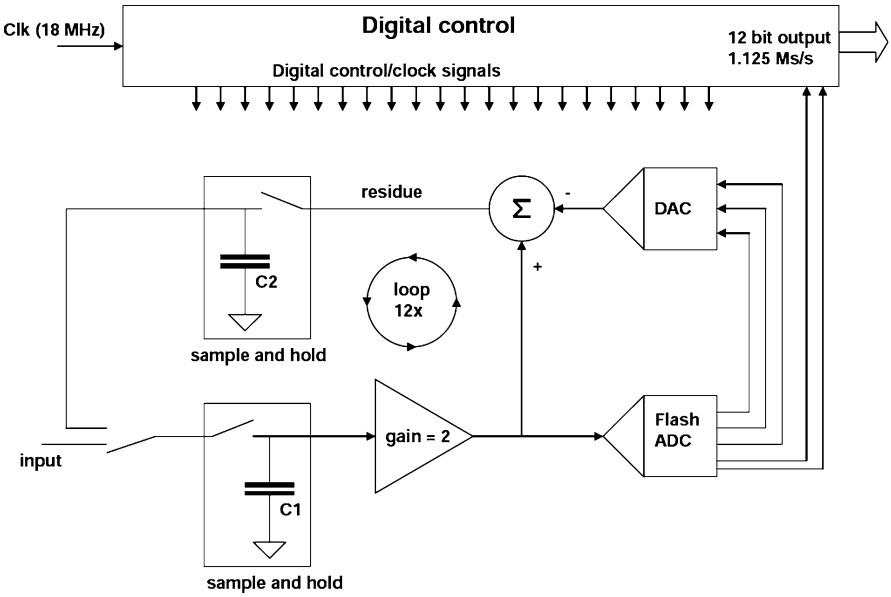



**Fig. 4** Block diagram of a cyclic ADC

perature range of −40 up to 125 °C. A block diagram of the cyclic ADC is shown in Fig. 4.

The ADC works with a clock of 18 MHz derived from the 72 MHz system clock of the ASIC. The input signal is sampled to the input capacitor C1. For the calculation of the MSB (First Bit) the gain of the amplifier is set to "1" and the 1.5 Bit Flash ADC makes a decision. The DAC calculates the corresponding analog value which is subtracted from the input value. The residue is stored in capacitor C2 and transferred to C1 for the next cycle. So for the calculation of each bit one clock cycle would be necessary. However, the demand for bit decision accuracy is much higher for the first few bits. The decision accuracy can be influenced either by increasing the circuit current or by allowing a longer signal settling time. At the actual implementation the timing for the first four bits consists of two clock cycles each. Therefore a full conversion cycle needs 16 clock cycles instead of 12. However it was possible to reduce the necessary power consumption to almost half of the value necessary for a full speed implementation. For further power reduction it would be possible to adjust the bias current for the gain amplifier and the comparators for every cycle. But this would lead to a very complex control structure and a quite unpredictable behaviour over temperature and process variation. The selected approach is a good trade off between power reduction and complexity and it yields a high reliable functionality. The ADC was simulated for a broad range of process corners and a chip temperature range of −40 °C to 150 °C. Sensitivity analysis was done by Monte Carlo simulations. Finally the ADC was implemented in a 350 nm CMOS technology of austriamicrosystems AG, Unterpremstätten, Austria. A measured power spectral density (PSD) plot is shown in Fig. 5. The input signal is normalized to 0 dB. The 3rd harmonic limits the calculated SNDR value to 66 dB.

The ADC macro occupies an area of 0.3 $mm^2$ and consumes 3 mW from a 3.3 Volt supply. Extensive evaluation showed that the ADC achieves 10.8 effective numbers of bits (ENOBs) over temperature und supply voltage range. The ASIC is in production and used for many industrial applications.

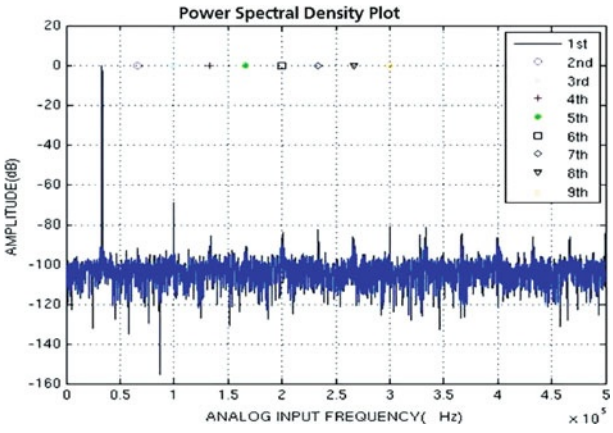



**Fig. 5** Power spectral density plot of the output of the Cyclic ADC

# 4 CT $\Delta\Sigma$-ADC for Direct Conversion of CdZnTe Detector Arrays

Cadmium–Zinc–Tellurium (CdZnTe) detectors are promising candidates for room temperature operated radiation detection sensor arrays [2]. Readout circuits for CdZnTe detector arrays measure charges caused by electron–hole pair generation due to ionizing events. These circuits are mainly based on semi-Gaussian shaping of the input signal followed by peak height detection and timing analysis in the analog domain. The semi-Gaussian shaping has been developed over many years and has been deployed in many detector circuits. A typical detector channel as used in current developments is depicted in Fig. 6. The charge pulse emitted by the sensor is integrated by a charge sensitive amplifier and filtered by shaping filters. Peak sampling or timing analysis can be done on the shaped pulses depending on the selected shaping time.

Figure 7 depicts the described signal processing chain. The analog shaping filters in front of the sampling circuit are required to avoid aliasing during the sampling process and to suppress high frequency noise. Relaxed requirements in terms of anti-alias filtering would offer the possibility to move pulse shaping, peak detection and peak sampling into the digital domain. Continuous-time (CT) delta-sigma ADCs meet these requirements very well:

- The shaping filters can be moved to the digital domain due to the implicit anti-alias filtering of these converters.
- The resolution scales with the data rate which makes it a perfect solution for high data rate timing analysis as well as high resolution peak sampling.
- The digital processing can be easily shared between several pixels.

A continuous-time delta-sigma CT $\Delta\Sigma$-ADC was especially designed for direct analog to digital conversion of CdZnTe detector arrays. Figure 8 shows a block diagram of the channel architecture of the new approach.

The continuous time ADC is the key element of the new channel architecture. A resolution of 12 bit is required which results in a minimum SNR of 74 dB. Due to the non-linearity correction in the digital signal processing the THD can be relaxed

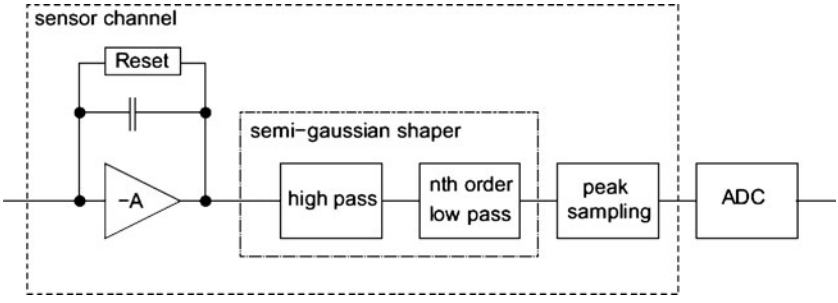



**Fig. 6** Conventional analog radiation detector channel

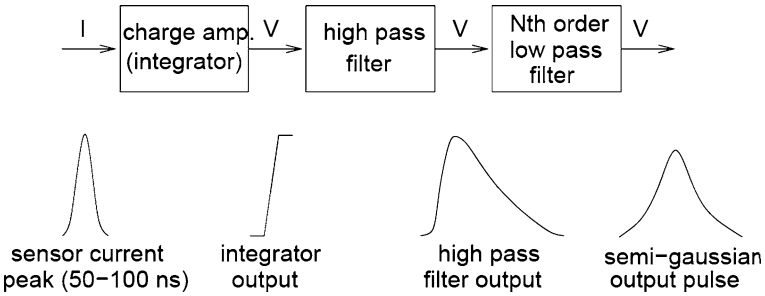



**Fig. 7** Signal processing chain in the radiation detector channel

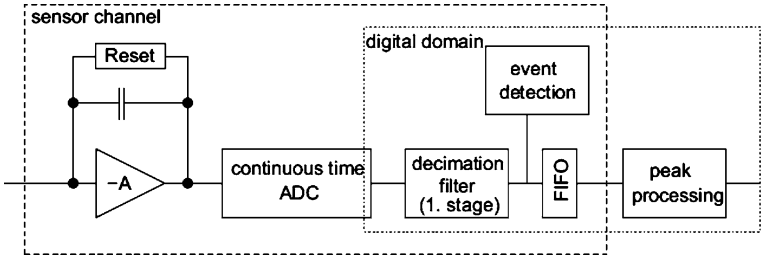



**Fig. 8** ADC-based radiation detector channel

to −50 dB. A signal band of 625 kHz is selected to allow a shaping time of 1500 ns. Shorter shaping times can be achieved with reduced resolution. The generation and distribution of very low jittered clocks on large readout ICs is a challenging task. A multi-bit implementation with non return-to-zero feedback pulses offers a high tolerance against clock jitter. It also allows a second order loop in combination with an oversampling ratio of 50 which reduces the number of operational amplifiers and the requirement for very high gain-bandwidth as well.

The modulator block diagram is illustrated in Fig. 9. Two RC integrators are implemented for the 2nd order loop filter. The 4 bit quantizer is realized as a flash ADC. Three current-mode DACs (IDACs) feed back the quantizer output to the inputs of the respective building blocks. IDAC1 and IDAC2 use cascoded current sources in a complementary structure. The IDAC cells deliver positive and negative output currents for the summing nodes. For the quantizer also a direct feedback is foreseen to compensate for excessive loop delay. A fixed delay of a half clock period is introduced to support signal dependent ADC decision. In order to avoid the complexity of summing two voltage signals at the input stage of the quantizer, the summing operation is shifted to the reference stage. The resulting function exhibits the same signal transfer function as achieved by summing at the input nodes.

Architectural block level simulation and optimization were done with HDL-Models. Finally, a transient simulation for the overall $\Delta\Sigma$ modulator was performed on circuit level. The clock was set to 50 MHz with a duty cycle of 50%. A sine signal with 0.6 V amplitude (−4.4 dBFS) and 500 kHz frequency was applied. The out-

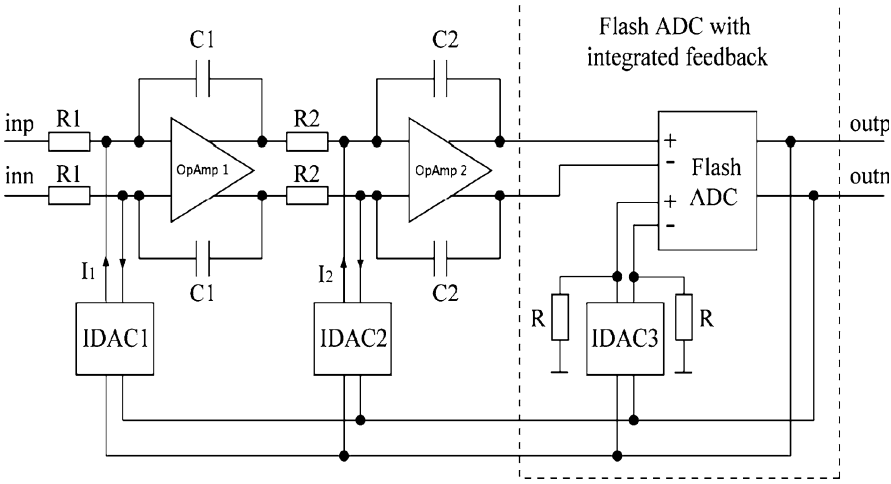



**Fig. 9** Block diagram of the multi bit continuous time delta sigma modulator

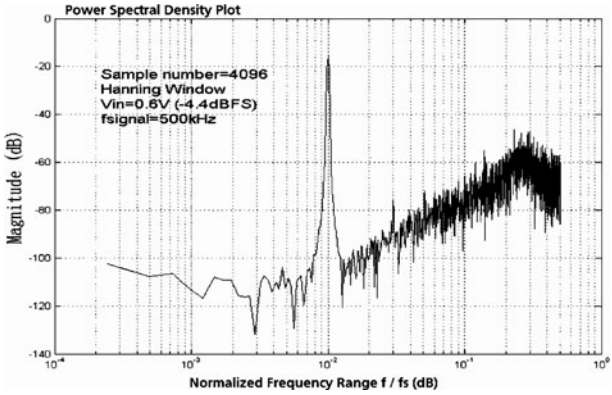



**Fig. 10** Simulated power spectrum density for the CT $\Delta\Sigma$ modulator

put data of the modulator were exported to Matlab to compute the spectrum using a 4096-point Hanning windowed FFT. The spectrum of the output data is depicted in Fig. 10. Signal to noise Ratio (SNR) was simulated for the whole input signal range. A maximum SNR of 78.3 dB was achieved at 0.7 V input amplitude (−3 dBFS). The calculated Dynamic Range (DR) is 81 dB.

## 5 Fluxgate Sensor ASIC for Space Applications

The advanced concept of direct sensor signal conversion was used for the implementation of a new control and measurement system of a fluxgate based magnetometer

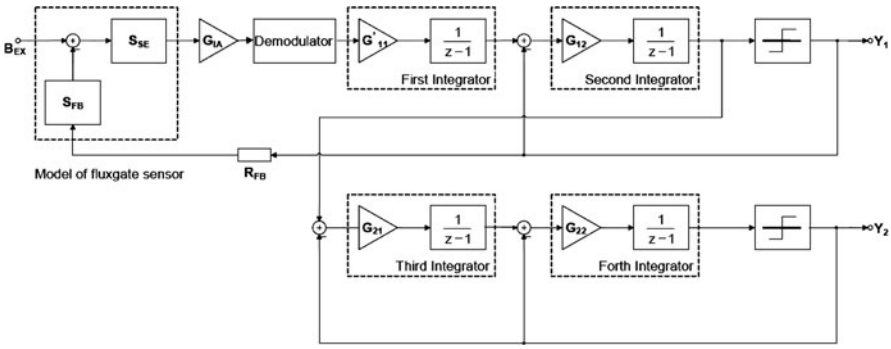



**Fig. 11** Block diagram of the 2-2$\Delta\Sigma$ modulator for fluxgate sensors

for space missions to monitor earth or planetary magnetic fields [3]. Fluxgate magnetometers consist of a set of coils producing a magnetic field which counteracts the external magnetic flux. So detecting the point at which external and internal fluxes are equal and measuring the current necessary to produce the internal flux gives a high sensitive result for the external magnetic field.

A cascaded 2-2$\Delta\Sigma$ modulator is used for analog to digital signal conversion. The sensor is connected to the input of the first modulator over a preamplifier und an asynchronous demodulator. As shown in Fig. 11 the fluxgate sensor is part of the feedback loop of the first 2$^{nd}$ order modulator, keeping the flux in the sensor in an equilibrium. An increased signal to noise ratio is achieved by the following 2$^{nd}$ order modulator. Four channels of the 4$^{th}$ order $\Delta\Sigma$-modulator were implemented

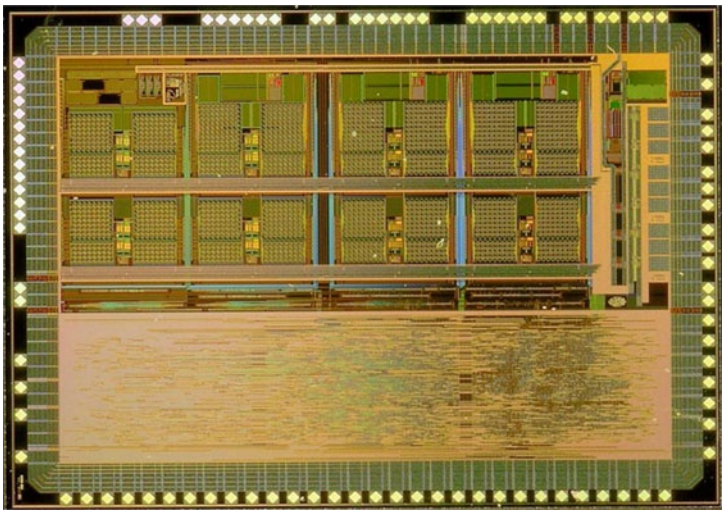


**Fig. 12** Layout of the multichannel ADC for spaceborn fluxgate sensors

on a 350 nm CMOS technology for a three axis magnetometer and a general house keeping channel for voltage, current and temperature measurements on the satellite. The modulators are combined with complex digital filter functions, which allow an online trade off selection of resolution and bandwidth. Figure 12 shows a photo of the measurement circuit. The devices were successfully tested in the laboratory, evaluated with different fluxgate sensors and qualified for space missions. The measurement of the house keeping channel showed an SNDR of 92 dB and a DR of 98 dB for a 10 Hz input signal. Using a fluxgate sensor under controlled conditions a minimum noise figure of 3 pT/sqrt(Hz) at 1 Hz was measured. Confirmed radiation tolerance up to 300 krad, low power consumption of 13 mW for each fluxgate channel and a total chip area of 20 $mm^2$ for all analog and digital functions make the ASIC very attractive for space applications. The fluxgate measurement instrument is selected for the Magnetospheric Multiscale Mission (MMS) with four satellites planned to be launched in 2014 by the NASA.

## 6 Summary

For many mixed signal systems the ADC presents the most critical requirements and demands considerable effort for the system implementation. In this paper four different ADC implementations are presented, which, due to their requirements, employ different architectures. Future research concentrates on sophisticated combinations of known architectures and on new designs for nanotechnologies.

## Acknowledgements

We thank Jose Angel Diaz Madrid, Dr. Gines Domenech Asensi and Professor Dr. Ramon Ruiz Merino from the Universidad Politecnica de Cartagena, Spain for the cooperation on the development of Pipeline ADCs. The work on readout circuits for CdZnTe was conducted in close cooperation with Dr. Jose Benloch from the IFIC, Barcelona and Dr. Jose-Manuel Perez from the CIEMAT, Madrid, Spain. The work was supported by the EU-FP6 project "*Mammography with molecular imaging, (MAMMI)*", LSHC-CT-2006-037555. The work on the fluxgate sensor ASIC was done in close cooperation with Dr. Werner Magnes and his group from the Institut für Weltraumforschung IWF, Graz, Austria. The project was funded by the European Space Agency, Science Payload and Advanced Concepts Office under contract no. 18391/04/NL/HB.

## References

1. Diaz-Madrid JA, Neubauer H et al (2009) Power reduction of a 12-bit 40 MS/s pipeline ADC exploiting partial amplifier sharing, Design Automation & Test in Europe

2. Völker M, Zhou H, Hauer J (2010) A Multibit Continous-Time Delta-SIGMA-ADC for Direct Conversion of CdZnTe Detector Arrays, Dresdner Arbeitstagung Schaltungs- und Systementwurf
3. Magnes W, Oberst M et al (2008) Highly integrated front-end electronics for spaceborn fluxgate sensor. IOP-Journal for Measurement Science and Technology, vol 19

# Robust Position Measurement Systems Based on Integrated 3-D Magnetic Field Sensors

**Hans-Peter Hohe, Michael Hackner, Markus Stahl-Offergeld, Volker Peters, and Josef Sauerer**

**Abstract** Today's position measurement systems are restricted to systems based on lateral Hall sensors and, hence, suffer from low robustness concerning temperature dependency and disturbing external magnetic fields. With the application of 3-D Hall sensors new principles in signal conditioning can be applied. Position values are then calculated from the complete flux vector and not only from a single component of the magnetic field. Thus it is possible either to perform multi axis position measurement or to use gradient approaches to come to more robust systems. The drawback of these approaches is that the complexity of system design increases. So it is necessary to set up new development methodologies for such multidimensional systems. In this article an approach to design such robust multidimensional position sensors systems is proposed.

## 1 Introduction

The basis of the new robust and/or multidimensional position sensor systems is a low cost integrated 3-D magnetic field sensor without magnetic active material in the signal path. These sensors can be realized on standard CMOS processes and can therefore be combined with signal conditioning circuitry around. Such a 3-D magnetic field sensor called magnetic pixel cell is shown in Fig. 1.

Even though the sensors for the three orthogonal magnetic field components can not be placed exactly at the same location on the chip, you can think of this 3-D sensor as a punctual sensor in almost all applications due to two reasons:

- the size of the sensor ($150 \times 150\,\mu m^2$) is small with respect to the size of most used permanent magnets,

Hans-Peter Hohe (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
hans-peter.hohe@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_5, © Springer 2011

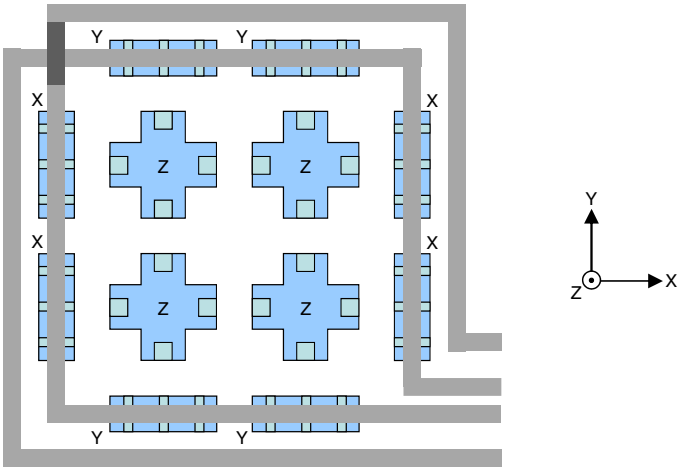



**Fig. 1** 3-D magnetic field sensor 'pixel cell'

- the gravity centers for the four hall elements needed for each component of the magnetic field vector are identical.

Only in a few specialized applications the size of the sensor and the location of the sensors for the three magnetic field components have to be considered.

Auxiliary to the basic 3-D measurement functionality each pixel cell has an integrated coil that is able to expose all three sensors for the magnetic field components with an additional magnetic field dependent on the current through the coil and hence well known in magnitude. This can be used to

- implement a regulator for a constant sensitivity of the sensors (no drift over temperature, no ageing, no stress effects, . . . ),
- do an electrical test of the sensors after fabrication (low testing costs),
- implement a self test of the sensors for safety critical applications (SIL).

This basic 3-D magnetic field sensor is ready to use for the two main robustness principles.

## 2 Robustness Principles for Applications

### *2.1 Relative or Angle Principle*

Using the relative or angle principle, one gets rid of temperature effects no matter where they come from.

Imagine a permanent magnet surrounded by its magnetic field. A magnetic field sensor is placed somewhere in the magnetic field at a certain location. If the temperature of the permanent magnet (or the temperature of the sensor chip) increases

the measured strength of a single magnetic field component in the sensor will decrease. This is due to the negative temperature coefficient of the magnetization of the permanent magnet (or the negative temperature coefficient of the sensitivity of the hall elements which are inside the sensor). A position calculated from this single component magnetic field value will be influenced by temperature changes of the magnet or the sensor. Temperature changes of the sensor can be measured by appropriate means on the chip and be corrected by activation of the sensitivity regulator using the integrated coil. But temperature changes of the permanent magnet (which is placed on the physically separated moving part) can not be corrected and lead to an error for example in the derived position value.

Repeating the same thought experiment and looking at the direction of the magnetic field lines at the position of the sensors (and not at the magnetic field magnitude) then there is no change. The direction of the magnetic field lines remain the same even if the temperature of magnet and/or sensor changes dramatically.

If it is possible to measure more than one component of the magnetic field vector, one can deduce a final position value from a quotient between two different magnetic field measurement values representing different components of the magnetic field. Doing so eliminates temperature dependent effects especially coming from the permanent magnet. Calculating an angle out of the same two magnetic field values has the same effect on temperature effects. The decision between quotient and angle calculation is only a matter of better linearization of the final position.

## *2.2 Difference or Gradient Principle*

The gradient principle can be used to reduce the influence of external homogenous disturbing fields (e.g. coming from magnetic field sources far away).

Now, imagine a small magnet which has a rather big distance to the measuring sensor. Such a system will be strongly influenced even by the earth magnetic field. To eliminate the influence of the earth magnetic field it is recommended to use two magnetic field sensors for the same magnetic field component in a certain distance (several millimeters) on the same chip. By using the difference (differential quotient) to calculate the position value it is possible to eliminate the constant part of the chosen field component and therefore for example the influence of the earth magnetic field.

Unfortunately this principle completely eliminates disturbing external fields only if they are constant over space. This is only fulfilled for the earth magnetic field. But all other local disturbing elements can be reduced in their impact on the position value.

## 3 Robust Position Measurement System

To implement a robust position measurement system using both robustness principles at the same time a sensor chip with two pixel cells at a certain distance was developed. The setup of this system is shown in Fig. 2.

A cylindric magnet is used moving in $x$-direction above the sensor chip. Two pixel cells (3-D magnetic field sensors) are located on the chip. Moving the magnet over the full distance of $+/-10$ mm results in the measurement values shown in Fig. 3.

Because of the geometry of the system, the $y$-components of the two pixel cells are always zero. The following values are simulated and shown in the diagram:

- '$Bx$' the average value of the $x$-components of both pixel cells
- '$Bz$' the average value of the $z$-components of both pixel cells
- 'd$Bxnx$' the difference value of the $x$-components of both pixel cells
- 'd$Bznx$' the difference value of the $z$-components of both pixel cells.

In the following sections two different calculation methods of the position value are described. The first uses only the 'angle' robustness principle and the second uses the 'angle' and the 'gradient' robustness principles. The formulas are shown in Fig. 4.

The difference between the two calculation schemes is that the second one is robust against external disturbing field and the first one is not. This is due to the fact, that the second one is based on gradient values and not on absolute (average) values like the first one.

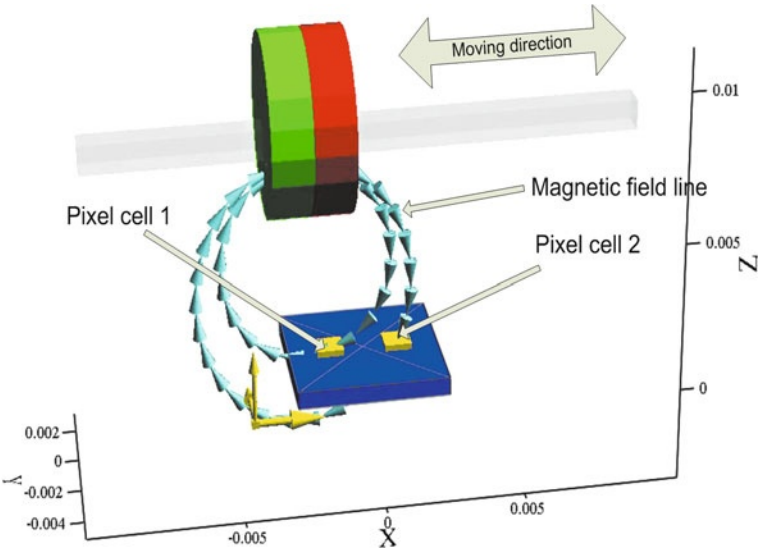



**Fig. 2** Position measurement system using 'difference' and 'relative' principle

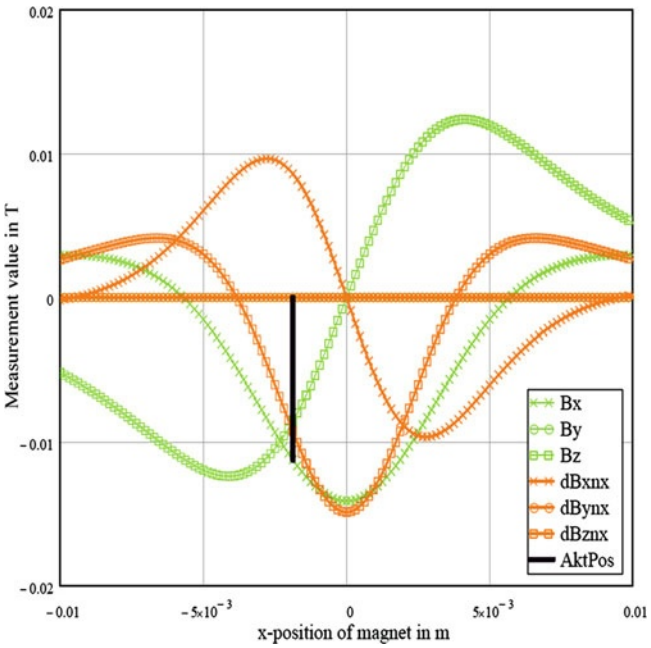



**Fig. 3** Magnetic field values over the full movement range

'*angle*'

$$\alpha = \arctan 2\left(\underbrace{Bx(Pixel2)+Bx(Pixel1)}_{\approx average(Bx)}, \underbrace{Bz(Pixel2)+Bz(Pixel1)}_{\approx average(Bz)}\right)$$

'*angle*' *and* '*gradient*'

$$\alpha = \arctan 2\left(\underbrace{Bx(Pixel2)-Bx(Pixel1)}_{\approx gradient(Bx)}, \underbrace{Bz(Pixel2)-Bz(Pixel1)}_{\approx gradient(Bz)}\right)$$

**Fig. 4** Formula to calculate the angle and gradient

The final step to get a position measurement system is the linearization. A nonlinear transfer function is used to derive the position out of the calculated angle value. In most cases this transfer function is done via look-up table and linear interpolation between the base points. Doing so leads to the following characteristic curve (Fig. 5) with the associated deviations shown in Fig. 6.

The remaining error mainly results out of the linear interpolation between the set points of the linearization table.

Blue curves show the behaviour of the ‘angle’ based system whereas the magenta curves represent the behaviour of the ‘angle and gradient’ based system. Fig-

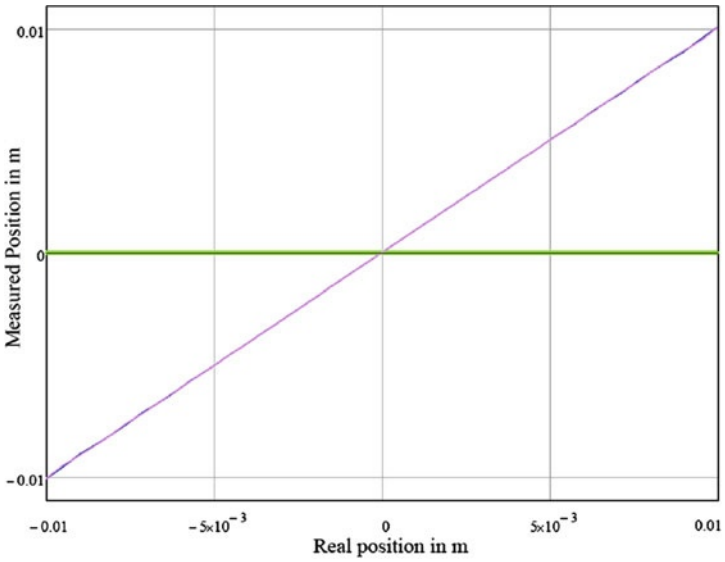



**Fig. 5** Final characteristic curve of the position measurement system

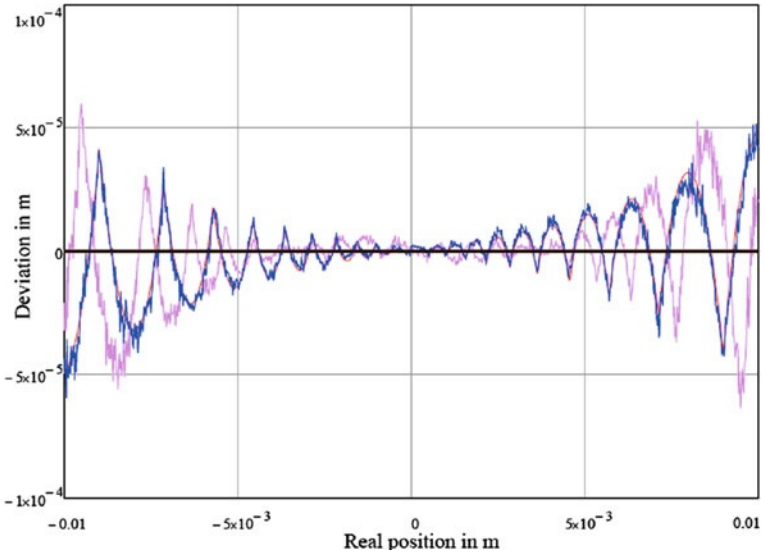



**Fig. 6** Remaining position error

ure 6 proves that the two calculation schemes ‘angle’ and ‘angle with gradient’ have nearly the same basic resolution and linearity. So a gradient based system is not necessarily worse than an absolute value based system.

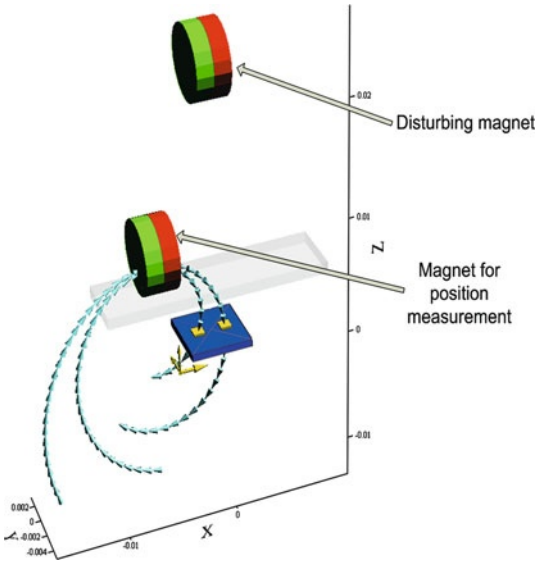



**Fig. 7** Position measurement system with a disturbing magnet

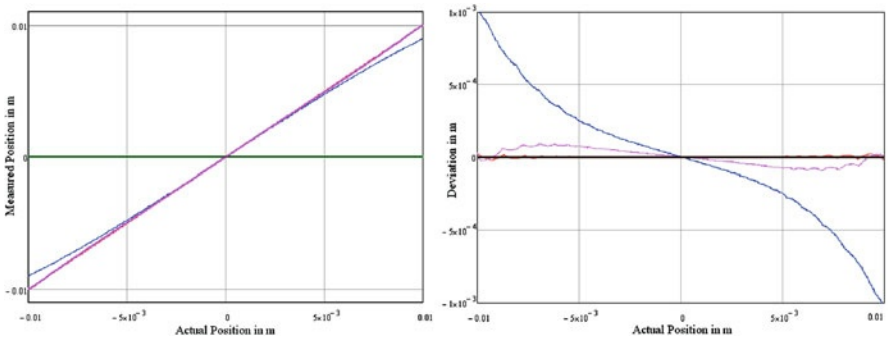



**Fig. 8** Characteristic curve of the disturbed position sensor for both calculation methods

To show the robustness against external disturbing fields a simulation with a second magnet with identical dimensions but shifted by 15 mm in positive $z$-direction is done. Figure 7 shows the configuration.

If the magnet for position measurement is moved in $x$-direction and the disturbing magnet is fixed to the shown location 15 mm above this results in the characteristic curve shown in Fig. 8 left side.

Again the blue curve represents the 'angle' based system – magenta curve represents the 'angle' and 'gradient' based system. It can be seen, that the disturbing magnet produces an error of factor 10 higher in the 'angle' based system. So the 'gradient' based system is a factor of 10 better even in conditions where non homogenous disturbing magnetic fields are applied to the system.

If there are homogenous disturbing fields (e.g. earth magnetic field) in one or more direction the ‘angle’ based system will be influenced whereas the ‘angle and gradient’ based system won’t be influenced at all.

## 4 Conclusion

Using the new 3-D Hall sensor technology will increase costs only a little with respect to common lateral magnetic field sensor based technologies but has big advantages in robustness. Temperature dependency of the permanent magnet in a position measurement system is no longer an issue. External disturbing fields are less critical than in absolute values based systems. This opens a lot of new applications for such 3-D based systems in the future. Even multidimensional position sensors will be possible with this approach.

## References

1. Kuhrt F, Lippmann HJ (1968) Hallgeneratoren – Eigenschaften und Anwendungen. Springer, Heidelberg
2. Melexis (2009) MLX 90333 – Triaxis 3-D-Joystick position sensor. Data sheet, Melexis
3. Hackner M, Hohe HP (2007) Offset-centering – a way towards higher resolution of Hall-sensor-signals. SENSOR conference 2007, 13th international conference, pp 253–258, ISBN 978-3-9810993-1-7
4. Hohe HP, Sauerer J (2001) Hall-Element. Patent DE 19,954,360
5. Ernst R, Hackner M, Hohe HP (2002) Realizing Highly Symmetric Vertical Hall Sensor Elements on a Standard CMOS Process. European Conference on Solid-State Transducers 16, Prague, pp 329–330, ISBN 80-01-02576-4
6. Allen PE, Hollberg DR (1987) CMOS Analog Circuit Design. Oxford University Press, Oxford, New York, ISBN 0-19-510720-9
7. Janisch J (2009) Magnetische Winkelencoder – Funktionsprinzip und Anwendung. Hall-Sensor Seminar, AMA

# Design of Multi-Dimensional Magnetic Position Sensor Systems Using the Example of an Inverse Pendulum

**Thomas Obenaus, Andreas Wilde, Holger Priwitzer, Jörg Bretschneider, and Olaf Enge-Rosenblatt**

**Abstract** The exact determination of displacement and rotation of moving parts is a frequent task in automotive and industrial automation. Systems consisting of permanent magnets and magnetic flux sensors provide the advantage of a precise, robust, non-contact, i.e. wear-free measurement method. Magnetic sensors based on HallinOne® technology enable the design of systems which make use of the entire (3-D) set of magnetic flux density for the determination of up to six relative translational and rotational position coordinates. However, design of corresponding multi-axis magnetic encoder applications poses a quite complex task for the engineer. In this contribution, we explain the design process for multi-axis magnetic position sensing systems and present an inverse pendulum solution as an example.

## 1 Introduction

Determination of relative displacement and rotation of moving parts is a frequent task in automotive, industrial automation, and other applications. Examples are gas pedal in cars, the steering column switch of trucks, or the task to determine the angle of rotation in robotized automation. Position measurement systems with sliding or switching contacts need a direct connection between the moving and the fixed part. Due to degeneration and oxidation, the life time of the whole system is shortened significantly. In contrast to that, systems consisting of permanent magnets and magnetic field sensors offer the advantage of a precise, robust, contactless, and wear-free measurement method. Such systems have replaced contact-based methods in wide areas. In comparison to optical methods they are much cheaper and significantly less sensitive to rough environments and dirt.

Sensors being able to measure the magnetic field in one direction are currently state of the art. However, there are many applications, even in the automotive sec-

Thomas Obenaus (✉)
Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany
thomas.obenaus@eas.iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_6, © Springer 2011

tor, with a degree of freedom higher than one. There is a considerable demand for solutions being able to measure more than one – up to all six – spatial degrees of freedom of a component in a contactless and sufficiently exact manner. Such systems have to meet high standards of reliability and have to be sufficiently robust with respect to magnetic interference fields. Therefore, differential measurement of rotational and translational movements is desired. The Fraunhofer Institute for Integrated Circuits (IIS) has done intensive research in the field of integrated hall sensor systems since 1994, resulting in HallinOne®, a low cost and CMOS based 3-D Hall sensor technology which opens new possibilities for the development of multi-axis position sensor systems [3].

This paper introduces a spatial position measurement system based on HallinOne® technology and integrated into a mechatronic demonstrator. The demonstrator is a planar SCARA-robot which balances a spatial inverse pendulum. The base of the pendulum rod is located right above the cardan joint at the Tool Center Point (TCP) of the robot. The system was built to demonstrate the concept of model-based design of mechatronic systems including the necessary nonlinear control as well as to show a magnetic field based multi-axis measurement system which measures both cardan angles contactless and with high precision.

## 2 Magnetic Position Sensor Systems

A magnetic position sensor system consists of a magnetic field sensor and a relatively moveable magnetic field generator, which is a permanent magnet in most cases (Fig. 1).

The simplest applications of these systems are used to detect the presence of certain components. For instance the states “open” or “closed” of a mobile phone with an extendable keyboard are identified. In this case a threshold-based method can be used. More challenging are applications where a continuous measurement of a position coordinate is needed. These so called linear or rotary encoders are part of

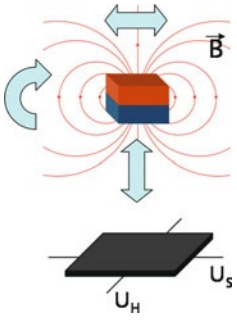



**Fig. 1** Principle view: Magnetic Position Sensor System with Hall element (1D-Sensor) and a relatively moveable magnetic field source

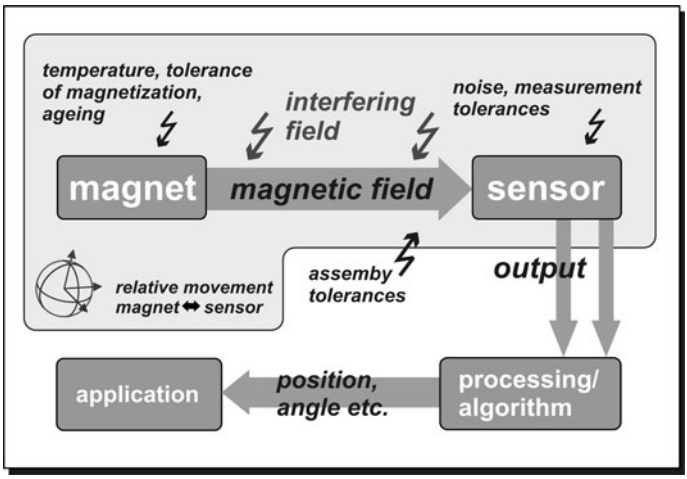



**Fig. 2** Data flow and influencing variables of a magnetic position sensor system

many mechatronic systems. The idea is to measure the magnetic field produced by the encoder magnet and to map the values to position coordinates afterwards.

The following error sources have to be considered in magnet based position sensor systems, since they might influence the accuracy and resolution of the system (Fig. 2):

- Explicit temperature dependency of the magnetic field strength of magnetic materials
- Deviations of the real magnetization from the desired magnetization
- Influence of stray fields distorting the expected field
- Assembly tolerances of magnets and sensors
- Offset and temperature dependency of the magnet sensors
- Noise caused by A/D-converter and signal processing components

There are different approaches available to map the measured magnetic field data to a position coordinate. The method to choose depends on the requirements of correctness and robustness of the position sensor system. In the following several methods are discussed.

One approach is based on using absolute values of the measured field to compute position coordinates. Therefore, it is necessary to find a configuration of the magnet and sensor that leads to a direct correlation between the measured value and the desired position. Here the computational effort is reduced to the calculation of a scaling factor, but this method is susceptible to all problems mentioned above and therefore more suited for uncritical applications.

Using a gradient-based method it is possible to increase the robustness of the measurement [4]. In this case the differences of values gained from magnetic field sensors which are slightly spatially separated are used. This way, homogeneous external magnet fields are inherently suppressed. In many applications inference fields

are caused by sources which are significantly farther away from the sensor than the magnet, thus they are supposed to be approximately homogeneous and their influence will be reduced by the gradient-based method significantly.

If several field components are needed for an application, there is the possibility to use the quotient of components of the magnetic vector to compute the angle between the field vector and the magnetic field sensor. Thereby, temperature depending error sources (e.g. magnetic field strength and sensor sensitivity) are inherently compensated. If separate sensors or Hall elements are used for each component, a local displacement will add an additional error. In contrast, HallinOne® sensors are able to measure precisely in 3-D due to their symmetric arrangement and coupling of Hall elements. Caused by the high integration-level of the HallinOne® technology, it is possible to realize 3-D applications with just one sensor IC. Besides space and product costs this reduces the complexity for control and signal processing and therefore appropriate error sources dramatically. Using several identical sensor components on the same IC, fail-safe systems are much easier to realize [5].

## 3 Design Support

Multi-axis magnetic position sensor systems offer many new, compact, robust, and cost-effective solutions for position measuring tasks which up to now have only been realized with lots of effort or not at all. Compared to the common magnetic position sensor system which is limited to one single axis, the design of applications with multi-axis measuring of the magnetic flux density obviously poses an explicit complexity step [1]. Problems result from the additional spatial movement possibilities which are measured simultaneously where each one combines and integrates the known error sources (Fig. 2).

For the application designer, methods of model-based design are preferable and helpful, in particular combined simulations and suitable visualizations (Fig. 3). The direct use of these methods in every single phase of the design process – conception, construction/verification, and optimization – saves design costs and expensive prototyping. Thereby, the accuracy of the used models grows in every phase.

During the concept phase, acceptable constellations for possible arrangements of magnet and sensor as well as magnet geometries are chosen out of first ideas. It has to be investigated if the measurement task is in principle solvable. In this phase, the exactness of the simulation is less important. However, the possibility to outline ideas and to get elementary statements of feasibility of the measuring system quite fast is crucial. Therefore, analytical models of simple design forms for calculating the magnetic field and simple behavior models of the sensor are used. Influences of temperature variations, adjustment errors, or inference fields are initially neglected. By means of visualization and processing of the simulation results it can be decided, whether an enough inhomogeneous and therefore locally distinguishable magnetic field is created within the traversing range of the sensor.

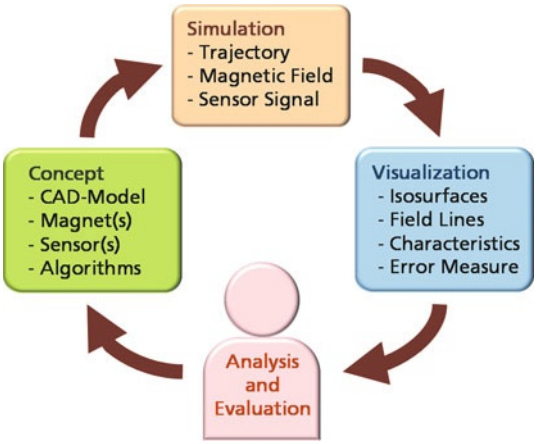



**Fig. 3** Model-based design cycle for magnetic position sensor systems. The cycle will be passed several times during prototype and product development. Only single components or even parameters will be varied

After noticing the general feasibility of one or more designs, typical error sources are stepwise considered within the simulation in the construction phase. Doing so, the measurement system can be analyzed according to its requirements (resolution, robustness, etc.) under realistic conditions. Especially for complicated geometries and error sources precise models are needed. These can be obtained by using numerical methods like the finite element method (FEM).

The main task of the optimization phase is to improve a verified and in the context of the requirements – functional system, with regard to costs and other needs.

Methods for visualization, adapted to the particular development step, are of vital importance for the evaluation and analysis of the modeled measurement system. Therefore, intermediate results of the simulation are as important as the analysis of the simulated output variables of the whole position measurement system.

During the conception phase, data views representing the results of single steps of the simulation are rather interesting. For example, the visualization of the components of the magnetic field which are theoretically measurable with the specific sensor helps to decide whether the concept should be developed further or not. During the construction phase, other visualizations are needed that display the violations of error thresholds. Exemplarily Fig. 4 shows the digital output signal of a Hallin-One® sensor that moves around the magnet on a spherical shell distance. The red dots are used to show areas wherein the sensor's A/D converter over-/understeers since the analog signal is too strong/too weak and therefore an explicit computation of the position is not possible. This information can only be gained by taking the sensor model into account. Thereby, the need of the combination of magnetic field and sensor simulation is illustrated. Due to its complexity, a theoretical realizable decoding algorithm can not always be realized due to the restrictions of the converting system (e.g. microcontroller). This needs to be considered during the design process as well and has to be verified by simulation.

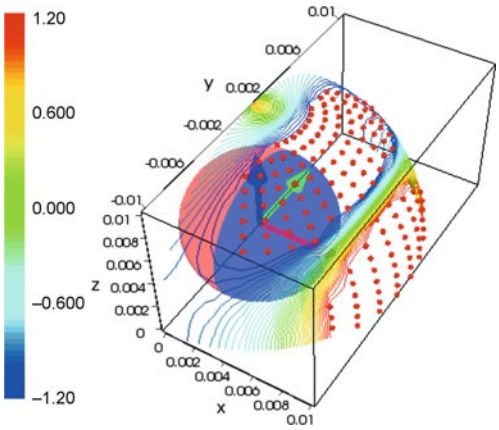



**Fig. 4** Hall signal of a magnetic field component with over-/understeering of the A/D converter

## 4 Inverse Pendulum

An inverse pendulum was designed [6] as a demonstrator for the methodology of model-based design. Application and benefits of contactless magnetic field sensors based on the HallinOne® technology developed at IIS for design of a robust system are illustrated and compared to common sensor systems.

### *4.1 System View*

Inverse pendulums exist in many places and in different designs. The limitation on balancing along a straight line (one degree of freedom), the usage of linear motors, or the use of actuators free of play make those systems easy to handle. Such inverse pendulums are also processed in many textbooks and dissertations [2, 7]. Ignoring the restrictions mentioned above, a robot (SCARA) whose behavior and design is inspired by the human arm was built. The SCARA is able to emulate the balancing of a freestanding rod on the palm of a hand. That is why the demonstrator presented in this article differs in its complexity in comparison to simple demonstrators known from literature.

According to the SCARA architecture, the robot consists of one forearm and one upper arm. Shoulder and elbow are designed as joints with one rotational degree of freedom. Two servo motors above the joints allow a specific movement of the arms and thus the manipulation of the TCP in the $XY$ plane. Using the angles $\theta_1$ and $\theta_2$ between shoulder, upper arm, and forearm (Fig. 5) it is possible to calculate the absolute position of the TCP. The embedded PC intended for the implementation of the control, including the needed AD-/DA converter cards, as well as the servo amplifier for the two driving systems are placed on the inside of the construction.

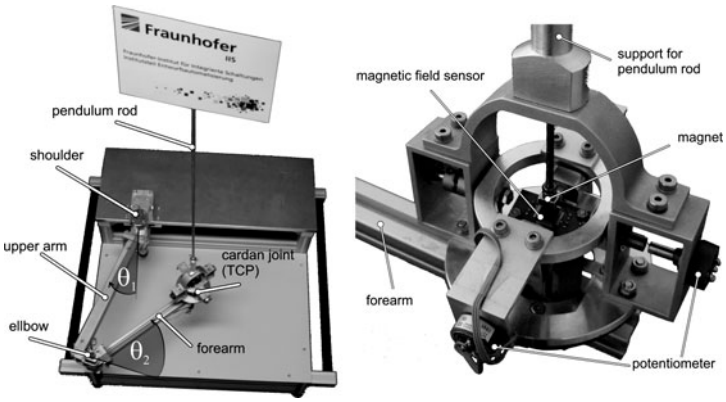



**Fig. 5** *Left*: Construction of the pendulum robot. *Right*: TCP of the robot as cardan joint

For the constructive implementation of the support for the pendulum rod within the TCP it is necessary to allow the pendulum to fall in all directions and to measure the inclination angle compared to its neutral position. The chosen cardan connection of the pendulum rod with the forearm of the robot is shown in Fig. 5 (right). The magnetic field sensor HATA-3-1 was placed in the center of the joint. Additional potentiometers at the end of both axes of the joints are used for the redundant measurement of the cardan angles.

The design of the nonlinear control for balancing the pendulum rod was done using Matlab/Simulink®, the add-on Stateflow® was used to implement the control system logic, and the xPC Target® was applied to establish the linkage between the control unit with the actuating elements and the sensor system.

## *4.2 Magnetic Sensor within the Pendulum*

To create a homogeneous magnetic field, one permanent magnet was installed above and one below the sensor. The magnets move according to the displacement of the pendulum rod, following a surface of a sphere around the sensor. The cardan angles are easy to compute by taking the quotient of the measured $Z$- and $X$- and corresponding $Y$-components of the magnetic field. This task is done by a microcontroller which converts the sensor data into the cardan angles and sends these values via RS232 interface to the PC hosting the control. An important fact for using the sensor within the present application is that there are constructive inaccuracies that have to be considered. Such unavoidable variations (e.g. the exact alignment of the sensor board) are corrected by software. Furthermore, before initiating the control, the neutral position of the pendulum (vertical pendulum rod) and the position of the robot arm which is needed to achieve the required position, have to be aligned. Figure 6 shows an example of a complete recording of the cardan angle $\psi_1$ and $\psi_2$

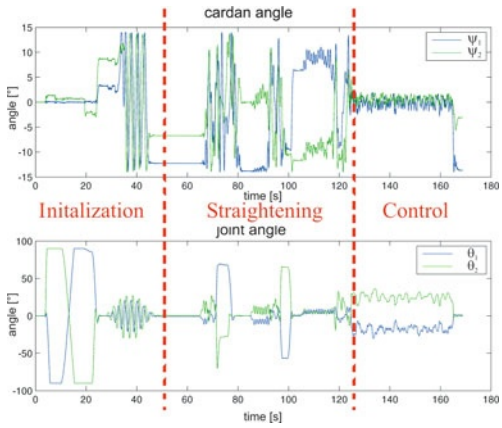



**Fig. 6** Measurement of cardan and joint angles during operation

as well as the joint angle of the robot arm $\theta_1$ and $\theta_2$ during the initialization, the straightening up of the pendulum, and the following change into controlled operation.

## 5 Summary

The article presents the use of spatial magnetic field based position sensor systems for contactless and very exact measurement of the position of moveable components. A spatial inverse pendulum was used as demonstrator.

The principle of magnetic position sensors (consisting of a sensor chip with Hall elements and a thereto moveable magnet) is based on the measurement of the magnetic field followed by a mapping to the position. The design of such systems is a complex task and therefore not effectively practicable without simulation and visualization. In this article a convenient design cycle had been introduced.

Exemplarily, the magnetic field sensor is used in a set-up of an inverse pendulum with two degrees of freedom. Both angles are simultaneously measured using only one sensor. With this demonstrator the model-based design of a system – from the dimensioning of the components to the control unit design, up to automatic code generation – is presented and tested under real conditions.

## References

1. Bretschneider J, Wilde A, Schneider P, Hohe H (2010) Design of multi-dimensional magnetic position sensor systems based on HallinOne® technology. IEEE Int. Symposium on Industrial Electronics, Bari, Italy, July 4–7, paper BD-017736.

2. Buhl M (2008) Sättigende strukturvariable Regelungen. Dissertation, TU München
3. Hackner M, Ernst R, Hohe H, Stahl-Offergeld M, Schlag U (2007) Vertikaler Hall-Sensor. Deutsches Patent DE102006017910A1
4. Hackner M, Hohe H, Sauerer J (2008) xD-Hallsensorik zur mehrdimensionalen Positionsmessung. In Gesch H (ed) Landshuter Symposium für Mikrosystemtechnik, p 93
5. Hohe H, Sauerer J (2008) Robuste Positionsmessung in Hydraulik- und Pneumatikzylindern. Mechatronik 11–12:60–63
6. Knoll C (2008) Entwurf und Realisierung eines roboterbasierten Pendelaufbaus. Praktikumsbericht, TU Dresden/Fraunhofer IIS/EAS
7. Nijmeijer H, van der Schaft AJ (1990) Nonlinear Dynamical Control Systems. Springer

# Sub-10-μA 868-MHz Wake-Up Receiver ASIC for In-Door Localisation and Geofencing Applications

**Heinrich Milosiu, Fritz Meier, Frank Oehler, and Alexander Pflaum**

**Abstract** Modern wireless receivers tend to decrease power consumption. This section shows the benefits of a sub-10 μA wake-up receiver circuit used for wireless geofencing and localisation applications with very low maintenance. The proposed wake-up receiver is a wireless receiver which continuously scans the radio channel for certain wake-up messages. Having received and decoded such a message containing additional data, the wake-up receiver triggers different actions in smart objects. The wake-up receiver consumes only 7.5 Microwatts and is suitable for mobile battery-operated or solar-cell driven smart objects. It is shown how the wake-up receiver can be used to implement a wireless in-door geofencing system. This application profits the most from both the very low deterministic current consumption and the short reaction time below 500 ms.

## 1 Ultra Low Current Wake-Up Receiver

In wireless sensor networks wake-up receivers considerably increase the service life and decrease the response time. Due to long lasting operation the battery-driven sensor nodes can now be run with low maintenance effort.

### *Principle Operation*

The presented wake-up receiver ASIC operates as continuous radio monitoring receiver at 868 MHz with a current consumption down to 2.4 μA at 3.0 volts supply [1]. A chip micrograph is shown in Fig. 1. The wake-up receiver ASIC is fabricated in a low-cost 0.18 μm CMOS technology.

Heinrich Milosiu (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
heinrich.milosiu@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_7, © Springer 2011

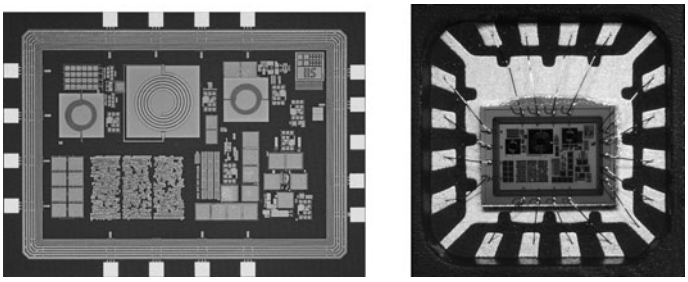


**Fig. 1** Chip micrograph of the wake-up receiver ASIC (2.2 mm × 1.7 mm) and assembled in a QFN16 package (5.0 mm × 5.0 mm)

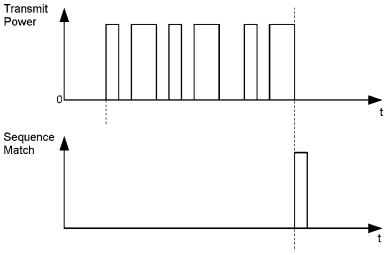



**Fig. 2** Wake-up sequence transmission and decoding scheme

The wake-up receiver can be used as an additional wireless receiver in a standard transceiver module while other components in the smart object module are suspended to sleep mode. It performs continuous listening to radio signals and can detect two types of transmitted wake-up sequences. Binary data reception is also provided. An additional crystal oscillator provides a 32 768 kHz clock signal and contributes an additional current consumption of less than 1 μA.

The wake-up sequence is received and decoded within 32 ms typically. Thus, a digital control signal is generated (Fig. 2). This can switch on a standard transceiver for wireless data exchange at higher data rates. Moreover, a local sensor or actuator can be activated. The wake-up receiver stays tuned and enables a persistent RF monitoring with low detection time. It also contains a forward error correction (FEC) decoder and is able to correct bit errors. Besides, quality-of-service data for the radio channel is also generated.

The wake-up receiver consists of a superheterodyne receiver whose analogue output signal is discretised and processed in the digital correlator (Fig. 3). If the received sequence information matches the reference sequence, a digital control signal is generated. The implemented receiver contains two separate correlators scanning two wake-up sequences (codes A and B). The wake-up sequence recognition is fault-tolerant. In practice, positive recognition of code A or B is ensured even if five out of thirty-one received bits within the wake-up sequence are false. Moreover, an 8-bit data decoder supplies continuous reception of binary data independent of wake-up sequence transmission.

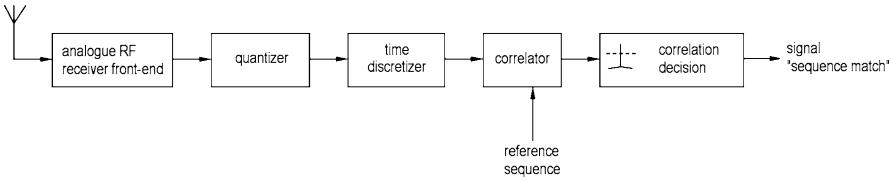



**Fig. 3** Block diagram of wake-up receiver

Measured sensitivity amounts to −60 dBm at the 868 MHz band. Assuming a transmit power level of 10 mW, a radio distance of 50 m can be reached. In-door experiments show a radio range up to 100 m with 200 mW transmit power using two monopole antennas. Driving a transmitter with 500 mW output power, a distance of 600 m is possible on a line-of-sight path.

## *Current Consumption*

The user can choose among a set of current consumption and data rate values during the receiver operation. This novelty offers many opportunities to the user. Besides, the wake-up reaction time can be adjusted according to the application requirements. In Table 1 a typical reaction time of 61 ms is reported at a current consumption of 7.6 μA. Thus, the continuous operating time amounts to 3 years using a CR2032 lithium button cell with 210 mAh charge.

Even 2.4 μA is possible at 485 ms reaction time. Thereby, continuous battery operation up to 9 years is feasible. This very low current can be delivered by many energy harvesting units such as solar cells, thermoelectric generators or microvibration generators.

The current consumption is not affected by reception of data or wake-up signals. Therefore, the user can predetermine the battery operating time by choosing the desired data rate. This is important for the wireless system design especially with a large amount of battery-driven nodes at low maintenance.

**Table 1** Selected operating modes for measured current consumption at 3 volts and data rate

| Current consumption | Data rate | Duration of wake-up sequence | Operating time with CR2032 |
|---|---|---|---|
| 2.4 μA | 64 bps | 485 ms | 9 years |
| 4.6 μA | 256 bps | 121 ms | 5 years |
| 7.6 μA | 512 bps | 61 ms | 3 years |
| 13.6 μA | 1024 bps | 30 ms | 21 months |
| 25.5 μA | 2048 bps | 15 ms | 11 months |
| 49.4 μA | 4096 bps | 7.6 ms | 6 months |

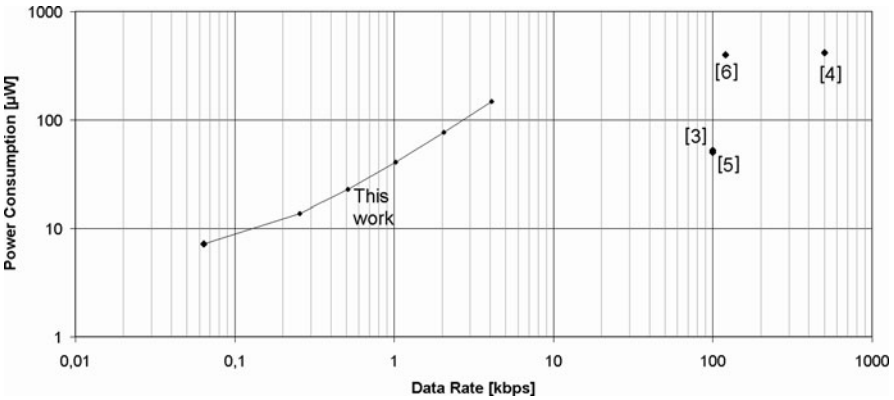



**Fig. 4** Comparison with state of the art receivers

Several wake-up schemes open up for the user. Utilizing the two wake-up sequences (named code A and code B), two separate control signals are generated from the wake-up receiver. Thus, two classes of smart objects can be addressed very easily. For example, code A addresses all objects in reception area and code B is dedicated to a certain class of objects depending on attributes or features.

Another wake-up method combines the reception of data packets with the original wake-up sequence. Therefore, an ID-based selection of smart objects can be done. Especially wireless networks with a large number of smart objects can benefit. Consequently, a minimum traffic on the wireless channel as well as minimum power consumption of the objects is obtained. ID-based selective wake-up prevents unnecessary activation of objects, ensures availability and increases life-time. In addition, a low response time for wireless inquiries facilitate new implementations of wireless network topologies for smart objects. Furthermore, selection of groups for wake-up can also be implemented.

Comparing state of the art integrated receiver publications (Fig. 4) in the submilliwatt range, the scalability of current consumption versus data rate illustrates a novelty. Moreover, the power consumption can be chosen less than 150 Microwatts down to 7.5 Microwatts. This value is seven times lower than in [3] and [5].

## 2 Beacon Sensitive Operation

Primarily, the ultra low current wake-up receiver is able to control an electrical component with higher power consumption initiated by a remote command via the radio channel. The origin of a transmitted wake-up sequence can either be an individual user, another smart object or a periodically transmitting radio beacon.

Using the built-in selective wake-up method, the beacon ID serves as object selector. Only certain objects scanning the radio channel for their desired beacon

ID can be activated. Consequently, several object classes can be addressed. This is a quite simple method to set up a singular fenced area.

Besides, the capability of the wake-up receiver to receive and decode data packets attached to the original wake-up sequence forms another method of beacon sensitive operation. Then, the data packet contains beacon information that may consist of a beacon ID and a time-stamp, for example. A small processing unit handles and interprets the received beacon information. Thus, context-aware activities of the smart object are established.

## *Beacon-Based Localization Method*

Applying a 3-stage transmission scheme (see Fig. 5), an ultra low current localisation method with room accuracy can be established.

The transmission scheme consists of three sub-messages containing 31 bits for a wake-up sequence (C), 24 bits for a room number (R) and 8 bits for the power level (P):

$$\text{CCCCRRRP}|_{\text{low}} \quad \text{CCCCRRRP}|_{\text{med}} \quad \text{CCCCRRRP}|_{\text{high}}$$

Each sub-message comprises 63 bits, the total 3-stage message amounts to a length of 189 bits. The bit sequences for room number and power level may contain redundancy information, e.g. FEC coded data. Depending on the current mode (4.6 μA, 7.6 μA or 13.6 μA) of the wake-up receivers, the duration of the 189 bits message is 738 ms, 369 ms or 185 ms respectively. If 185 ms beacon message duration and 13.6 μA current consumption is assumed, up to five beacon messages per second can be transmitted. The beacon module consists of a microcontroller, an OOK wireless transmitter, an wire antenna and a CR2450 (550 mAh) button cell. The maximum operating time for the beacon module is 7 years assuming a transmit period of 5 minutes and a transmitter current of 10 mA (the MCU is driven with a 32 kHz clock and consumes 2 μA). The receiver is operating in the 13.6 μA mode and provides a maximum operating time of 4 years using a CR2450 button cell. In-door measurement results utilizing the proposed wake-up receiver lead to the following proposal of the three power levels:

Consequently, a wake-up receiver recognizes all three sub-messages (low, medium, high) when it is put in the same room as the beacon transmits. In the adjacent

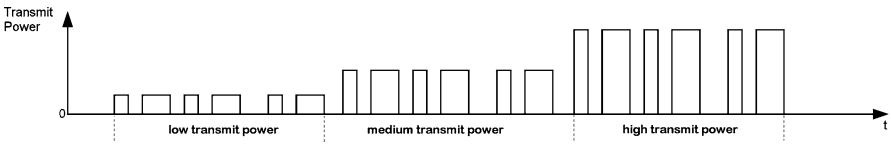



**Fig. 5** Transmission scheme for beacon

**Table 2** Choice of transmit power levels for the beacon

| Range | Power class | Transmit power | Description |
|---|---|---|---|
| <3 m | low | −4 dBm | same room |
| 3 . . . 12 m | medium | +7 Bm | adjacent room |
| >12 m | high | +14 dBm | two rooms next |

room only the reception of two sub-messages (medium, high) succeeds. Two rooms next, only the reception of the single sub-message (high) may work.

Depending on the reception of the sub-messages, a MCU can estimate the distance from the beacon.

## 3 Measurement Results

In order to prove the proposed multi-level beacon method, several in-door measurements have been carried out in an office building. The center floor is 2 metres wide. The packet reception depends on the distance between a mobile test receiver (wake-up receiver ASIC) and a stationary transmitting beacon (standard TX module with 4-ASK packet). Figure 6 shows the 2-D map of receive strength in four levels (0, 1, 2, 3). The beacon ("TX") is placed top right in the map.

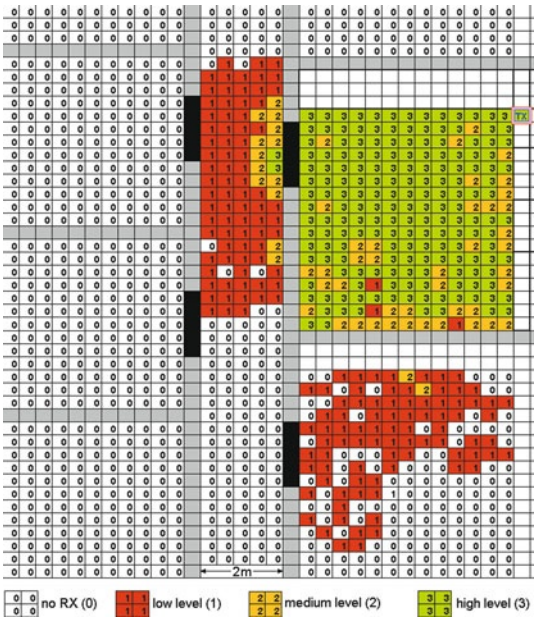



**Fig. 6** Packet Reception Map

Assigning reception levels 2 and 3 to the statement "receiver in the same room as beacon", correct assignments to room number can be done with 98% probability. Levels 1 or 0 do not contribute to a room decision. The receiver current consumption can be set to 7.6 μA at 3.0 V according to the total packet length of the beacon 369 ms. The periodical transmission of the beacon telegram can occur every 5 minutes. Thus, an average current consumption of the beacon module can be decreased to 50 μA at 4.0 V. This corresponds with a battery operating time of about 2 years using a 1000 mAh lithium accumulator.

## 4 Geofencing Application in a Research Department Using a Wake-Up Receiver

Deriving the context from the current position is a business driver. It is not only important to know where the objects are located right now, but together with other knowledge about the object business processes can be optimized. An emerging application geofencing is asset management of measurement equipment in R&D companies [2].

Assuming a single floor department that is observed in a geofencing application. The department consists of 15 office rooms. A first fence F1 comprises the exit gangway containing the exit door to outside. Objects in this area are likely to disappear from the observed building. A second fence F2 encloses a particular room with a higher security level. Precious goods such as high performance mobile measurement equipment are stored there. A third fence F3 encloses a lab for maintenance and calibration of the objects. Assuming this setup of fences, objects can be localised if they are in the storage room (F2), in maintenance (F3), or in the exit gangway (F1). Thus, many scenarios for a large number of smart objects used in the department are covered. Involving object attributes, the current position and the position before (including the time of entry or departure of the former fence), the context for every object can be concluded. Each fence contains a radio beacon that is transmitting an individual sequence periodically. Such a sequence can consist of a wake-up sequence, a beacon ID, and optionally a time stamp. All smart objects include an ultra low current wake-up receiver that decodes the incoming radio messages.

A benefit of the proposed indoor geofencing approach is the low number of beacons compared to a high number of observed objects. Each smart object can decide autonomously if its predetermined area matches the recognized area. Consequently, two fence decisions open up at first: match or mismatch. In case of match the smart object does not need to initiate any action or can transmit an OK message towards the beacons occasionally. Then, the beacons should be designed as transceiver for bidirectional radio communication. In case of mismatch the smart object can change its operation mode and starts collecting beacon information and stores it together with a time stamp. Otherwise, the object can retransmit the received beacon ID and its object ID in order to communicate its mismatch state to a backbone IT system.

Mismatched objects can even build up an adhoc multihop network with other smart objects that can be waken up to overcome possible radio barriers or interferers. Using the ultra low current wake-up receiver, a smart object can determine different kinds of events and together with context information it can trigger the designated activity. Fence crossing and touching events can be observed by the objects themselves.

Only from this fence crossing events a lot of information like movement profiles, actual localisation with room or fence accuracy, neighbourhood relations and probability of disappearance of objects can be derived when this information is stored in the smart object or a backend system.

Moreover, indication of user activity operating with objects e.g. intrusion of non-authorized personnel can be observed.

## *Operating Modes of Smart Objects in Geofencing Applications*

Operating modes of the smart object in a geofencing application can be:

- Idle mode: continuous listening and monitoring the RF channel (only wake-up receiver is running). This may be the default mode. A successful recognition of a fence crossing event selects one of the modes as follows.
- Tracking mode: receiving and storing beacon data in history memory within the smart object. Whether at a predetermined time or accidentally, the actual fence data is stored. Periodical memorizing, e.g. every minute, can provide a complete plot of tracked positions of the smart object on a fence scale. A posteriori decision can be made in order to attain the context of individual objects and of groups. Thus, a consistent long-term documentation of object positions or object activities is obtainable.
- Pursue mode: transmitting of received beacon data after leaving or enter- ing fences. Time critical applications with low reaction time (e.g. super vision of goods for security guards) need object information that is up to date. Immediate intervention by the user is possible.
- Anti-theft mode: repetitive transmitting of received beacon data in short intervals, e.g. every 10 seconds, in order to prevent theft or disappearance of a smart object. Intervening personnel can encircle or focus on certain fenced areas during non-authorized displacement. Real-time tracking of the object can be carried out.

If an inquiry of the status of an object is unsuccessful, the wireless connection to the smart object is likely to be broken. Reasons for the wireless communication failure can be non-authorized removal or damage of the wireless communication module or an empty battery. Consequently, the last fence information can be drawn from the IT backbone as a basis for a personal intervention for maintenance. The context-aware selection of the operation mode depends on the fence attributes. The smart object can perform this decision autonomously. This decision algorithm is determined by the user before roll-out or during operation. In some cases, the decision

could be made by a supervisor gathering the fence information of the monitored objects in order to get an entire position profile and a full system status.

### *Application Examples of Geofencing Systems*

An interesting application of the proposed geofencing method is enforced maintenance. Devices that need maintenance and are not located in the calibration and maintenance room (fence F3) should inform the IT system. Besides, assurance of escape routes in buildings can also benefit from a geofencing system. If objects stay in the corridor longer than a predetermined time (e.g. 1 hour) an optical indication at the smart objects should signalize a storage irregularity.

Asset Combination Management can be implemented with a geofencing system where the linkage of objects can be ensured and monitored.

Furthermore, wireless geofencing systems based on wake-up receivers can assist the booth constructors at business fairs. Defining the destination of fair boxes and locating the current position in an exhibition hall are typical geofencing methods to increase efficiency. The proposed geofencing system can provide an occupancy diagram showing the actual context of assets that can be "free for use", "used" or "reserved" including the current location. This diagram can be shown on a screen in the lab or can be accessed via a web browser. Many novel services can arise from that technology.

## 5 Outlook

The ASIC of the wake-up receiver is available and is currently combined with different complementary technologies, like active RFID-tags, nodes of a wireless sensor networks or energy harvesting modules.

Wireless technologies offer many advantages building up geofencing applications, to derive fence action events from the radio channel, persistent radio channel monitoring is mandatory. This is the key feature of the proposed wake-up receiver. If the geofencing application has some of the following additional requirements or boundaries, a wake-up receiver enhanced smart object is an adequate solution to consider:

- Energy restricted and long-term operation: due to the ultra low current consumption of 2.4 μA at 3.0 volts. A module can operate with a single battery up to 9 years (CR2032 battery).
- Energy harvesting supplies are suitable alternatives for battery supplies and enable autonomous long-term operation with minimum maintenance effort.
- Very short reaction time: fence action events should be detected with a minimum delay. The wake-up receiver provides a reaction time of 30 ms typically.

- Deterministic performance: reaction time, energy consumption and fence actions are highly deterministic. The wake-up receiver performs every time and under any circumstances with the same current consumption.
- Different kinds of fence crossing events: The beacon transmitter for the wake-up sequence can be optimized for different applications (e.g. par ticular radiation patterns of transmit antenna to realize space diversity).
- Small form factor: the wake-up receiver in QFN16 package has a size of 25 $mm^2$ and can therefore be easily integrated into other communication devices. Due to the very low power consumption, the power supply can be implemented much smaller.
- Expandable and flexible: new geofences can be added and removed to a site at any time.
- Innovation: the very low current consumption for wireless receivers in the UHF band offers a huge number of new applications and services.

Geofencing applications can be realized with the wake-up receiver as stand alone receiver. An integration in active RFID product (as replacement for today's built-in wake-on-radio mechanisms) or in wireless sensor networks can help to further optimize these products.

## References

1. Milosiu H, Meier F, Pflaum A, Oehler F, Preiß P (2011) A Novel Concept for a Long Lifetime Wireless Geofencing System With an Integrated Sub-10 µA Wake-Up Receiver. In 7th European Workshop on Smart Objects: Systems, Technologies and Applications (RFID-SysTech 2011, Dresden)
2. Milosiu H, Meier F (2009) Wake-Up-Receiver für drahtlose Sensornetzwerke in Logistik-Anwendungen. EEEfCOM 2009, Ulm
3. Pletcher NM, Gambini S, Rabaey J (2009) A 52 µW Wake-Up Receiver With 72 dBm Sensitivity Using an Uncertain-IF Architecture. IEEE Journal of Solid-State Circuits 44:269–280
4. Drago S, Leenaerts DMW, Sebastiano F, Breems LJ, Makinwa KAA, Nauta B (2010) A 2.4 GHz 830 pJ/bit duty-cycled wake-up receiver with −82 dBm sensitivity for crystal-less wireless sensor nodes. Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 224–225
5. Huang X, Rampu S, Wang X, Dolmans G, de Groot H (2010) A 2.4 GHz/915 MHz 51 µW wake-up receiver with offset and noise suppression. Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 222–223
6. Bohorquez JL, Chandrakasan AP, Dawson JL (2009) A 350 µW CMOS MSK Transmitter and 400 µW OOK Super-Regenerative Receiver for Medical Implant Communications. IEEE Journal of Solid-State Circuits 44:1248-1259

# Multi-Band GNSS Antenna

**Alexander E. Popugaev and Rainer Wansch**

**Abstract** A new low cost antenna for high-precision applications in Global Navigation Satellite Systems (GNSS) is described. The antenna has been designed to operate for all existing satellite navigation frequencies in L-band including GPS, GLONASS, Galileo and Compass (1164–1300 MHz and 1559–1610 MHz). The right-hand circularly polarized radiating structure consists of a broadband matched square patch with a fourpoint feeding. The antenna exhibits a 10 dB beamwidth of more than 150°, an antenna gain of more than 3 dBic and an excellent phase center stability. The desired radiation pattern roll-off and antenna gain are obtained using an electrically small radiating element with parasitic elements arranged around it. The dimensions of the developed antenna are a diameter of 146 mm and a height of 31 mm.

## 1 Introduction

Many navigation applications e.g. for civil engineering, agriculture and land surveying require high precision and integrity. There is a need for antennas working in all existing satellite navigation frequencies in L-band including GPS, GLONASS, Galileo and Compass (1164–1300 MHz and 1559–1610 MHz) to achieve the desired accuracy (Fig. 1). The antenna should have a known and a stable phase center, a pure right hand circular polarization (RHCP), good multi-path rejection and should show an antenna gain not less than 3 dBic with a 10 dB beamwidth of more than 150°. Such an antenna was developed at the Fraunhofer IIS and is described in this paper. In comparison to other commercially available antennas for high-precision GNSS applications it is a low cost product and should be especially attractive to customers.

Alexander E. Popugaev (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
alexander.popugaev@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_8, © Springer 2011

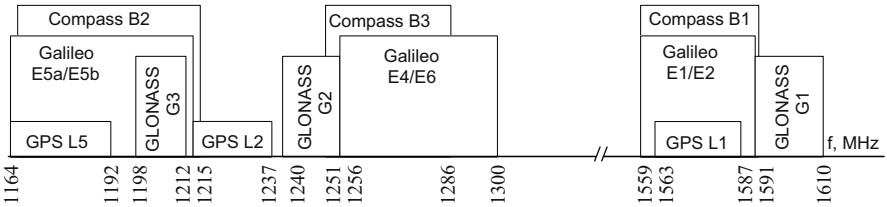



**Fig. 1** GNSS frequency plan

## 2 Antenna Concept

### *2.1 Radiator*

There are two possibilities to design a multi-band antenna. The first one is to use a separate radiator for each band. In this case stacked patch antennas come into consideration. As a result of using the stacked construction, the antenna would be either relatively large or very expensive (for example when using ceramic as a substrate). Stacked patches have a natural vertical offset between phase centres in the operating bands. The second approach is to make the antenna broadband. In this case there is no natural phase center offset. But the typical problem of conventional antennas is the fact, that the higher the frequency, the narrower the beam-width will be, and it is very difficult or almost impossible to achieve the required antenna gain and beam-width at all GNSS frequencies.

The problem was solved using the antenna concept described in [1,2]. The central idea is to use a broadband matched electrically small radiating element with parasitic elements arranged around it. Figure 2 shows the proposed antenna.

In order to achieve a good polarization purity and an omnidirectional radiation pattern, the radiator (50 mm square and 25 mm high) is fed in four points. The radiator concept allows a stable construction and a broadband matching. A ground-plane of 143 mm diameter is the smallest possible solution in order to achieve a good antenna gain and an effective multipath rejection.

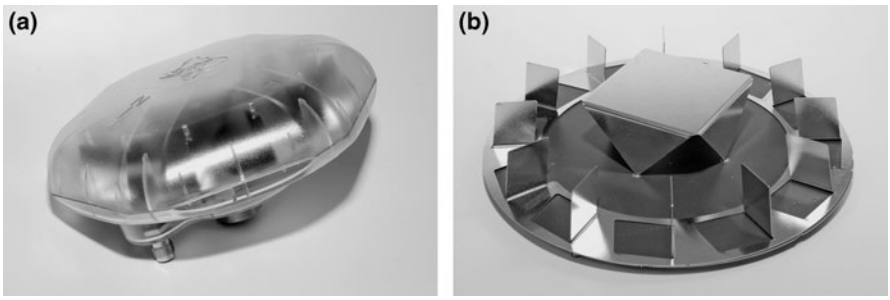



**Fig. 2** GNSS antenna

### *2.2 Feed Network*

A first version of the feed network of the GNSS antenna is shown in Fig. 3a [1]. The circuit consists of two main parts: a matching network and a phase shifter. The broadband matching was realized using a transformer and two open stubs for each of the four outputs of the radiator. The phase shifter drives it with nearly equal amplitudes and phases of appr. 0°, −90°, −180° and −270° in order to obtain a pure RHCP.

This passive antenna showed a very good performance and should be extended with a low noise amplifier (LNA). But there was not enough place for the LNA unless a substrate with a considerably high permittivity and/or multi layer structures were used (high fabrication costs). We developed a novel miniaturization technique [3] and solved this problem. The initial layout was meandered using a raster consisting of ring shape segments [3]. The result is presented in Fig. 3b [4]. The new design is very compact (20% of the conventional layout) and enables a successful integration of a LNA. Even delay lines for a double band filter (1.16–1.30 GHz and 1.56–1.61 GHz) could be integrated into the meandered layout.

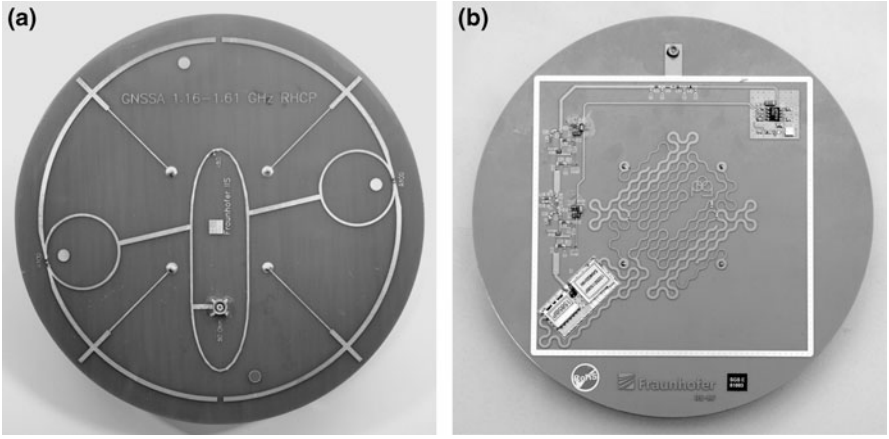



**Fig. 3** Feed network of the GNSS antenna

## 3 Measurement Results

### *3.1 Radiation Pattern*

Figure 4 shows the measured RHCP and LHCP patterns of the antenna at 1.16 GHz and 1.61 GHz. The antenna gain is 3.5 dBic and is nearly constant for all GNSS frequencies. The plots are normalized to this value. They show a wide angle beamwidth

of 180° at 1.16 GHz and 150° at 1.61 GHz. It can be seen, that the cross polarization discrimination (the difference between RHCP and LHCP) is larger than 15 dB for both operating bands. The gain variation at an elevation angle of 15° (Theta = 75°) is less than 2 dB.

## *3.2 Phase Center Variation*

High-precision GNSS applications require an antenna with a stable and preferably angle and frequency independent phase center. The phase center variation (PCV) should not exceed some millimeters.

It is physically impossible to build a circularly polarized antenna with a perfectly stable phase center because the phase diagram of such an antenna at an elevation of zero is an Archimedean spiral with an undefinable center. Nevertheless, a point

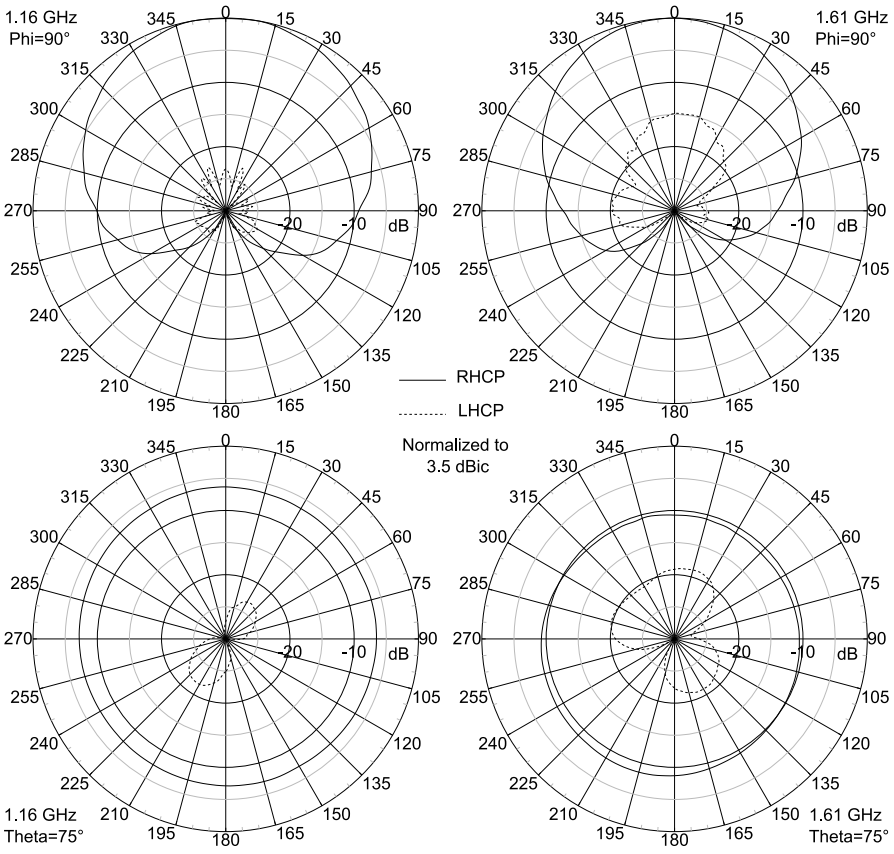



**Fig. 4** Measured radiation pattern, normalized to 3.5 dBic

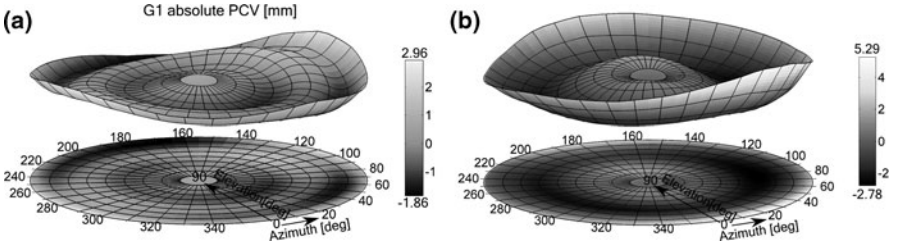



**Fig. 5** Measured Phase Center Variations of the GNSS antenna (University of Bonn)

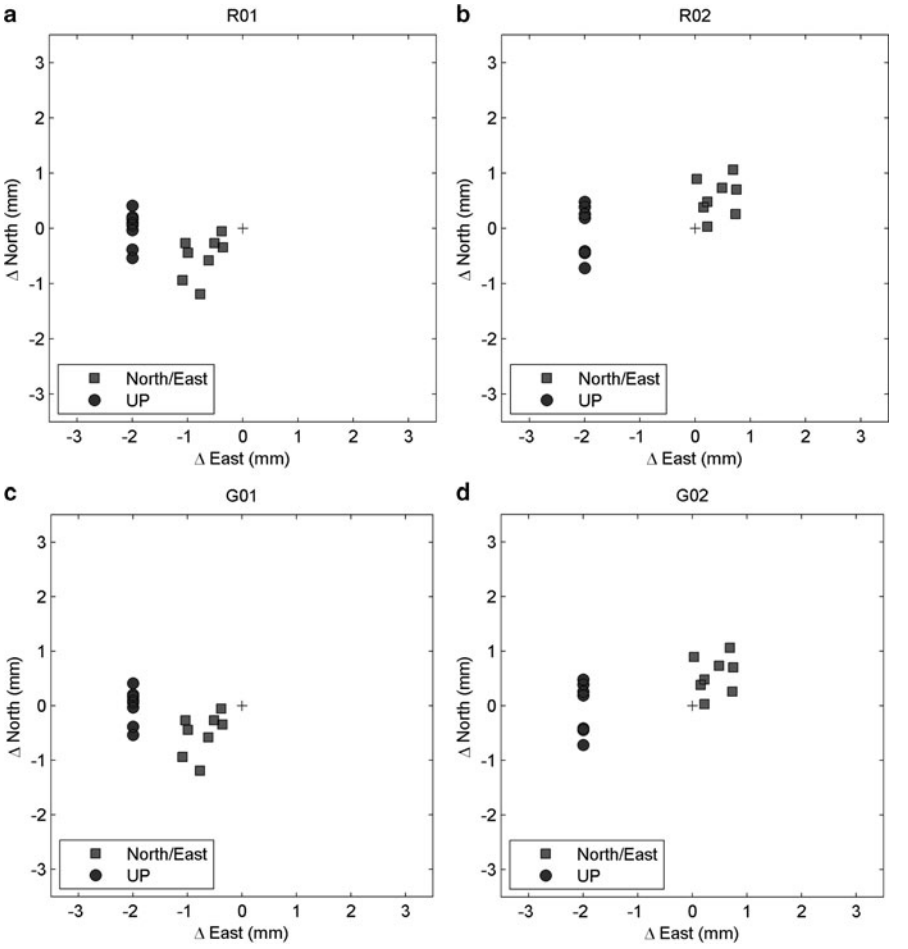



**Fig. 6** Repeatability of the PCVs for 8 GNSS antennas (Geo++): GLONASS G1(R01) and GLONASS G2(R02), GPS L1(G01) and GPS L2(G02)

can be found at which the angular dependency is minimal. It is the average phase center.

The GNSS antenna was calibrated at University of Bonn, Germany (measurements in an anechoic chamber) and at Geo++ in Garbsen, Germany (robot measurements using real GPS and GLONASS signals). Measured PCVs are presented in Fig. 5.

It can be seen, that the position of the average phase center in $z$-direction in the complete GNSS frequency band is stable up to $0.5 \cdot (62.00 - 58.53)\,\text{mm} = 1.735\,\text{mm}$. In the case of elevation ($x$–$y$-plane) the phase center variates even less: $0.5 \cdot (1.56 + 0.78)\,\text{mm} = 1.17\,\text{mm}$.

Very important is the repeatability of the PCVs for an individual antenna. The differences for 8 tested GNSS antennas are smaller than 1.2 mm (Fig. 6). The very small values of measured PCVs and their good repeatability guarantee a very high accuracy of coordinate measurements.

## 4 Conclusion

A novel and patent-pending antenna is presented which is designed to receive signals from all existing navigation systems in L-band. The antenna shows an omnidirectional hemispherical radiation pattern with a good circular polarization, a high multipath rejection and an excellent stable phase center for all operating bands. The desired beam shape and antenna gain are obtained using an electrically small radiating element with parasitic elements arranged around it. Measurements prove that the antenna is completely suitable for high-precision GNSS applications.

The antenna is licensed and is manufactured and sold by navXperience GmbH in Berlin.

## References

1. Popugaev AE, Wansch R, Urquijo SF (2007) A Novel High Performance Antenna for GNSS Applications, Antennas and Propagation, 2007. EuCAP 2007. The Second European Conference on, pp 1–5
2. WO/2008/092592 Antenna Apparatus for Transmitting and Receiving Electromagnetic Signals
3. Popugaev AE, Wansch R (2009) A Novel Miniaturization Technique in Microstrip Feed Network Design. 3rd European Conference on Antennas and Propagation (EuCAP 2009), Proceedings, CD-ROM: 23–27 March 2009, Berlin, Germany, VDE-Verlag, pp 2309–2313
4. Popugaev AE, Wansch R (2011) 3G+C: Hochpräzisions-Navigationsantenne, VDVmagazin – Zeitschrift für Geodäsie und Geoinformatik, 62:92–95

# Reconfigurable RF Receiver Front-End for Cognitive Radio

**Mario Schühler, Alexander Jaschke, and Alexander E. Popugaev**

**Abstract** A cognitive-radio (CR) radio-frequency (RF) receiver front-end architecture is proposed that allows for aggregation of spectral whites spaces. The architecture is able to exploit white spaces that are possibly spread over a wide bandwidth. Key components of the front-end are identified, having considerable impact on the performance of a CR system. Example implementations of those are discussed, including a dual-polarised dual-band antenna, a broadband low-noise amplifier LNA, and a tuneable pre-selection filter.

## 1 Introduction

More and more wireless communication systems are emerging today. To take advantage therefrom, highly flexible multiband user equipments are required. On the other hand, frequency spectrum has become a rare ressource – overcrowded frequency bands decrease the quality of service [7]. The regulation in several European countries is currently attributing the UHF band, also called *digital dividend*, TV white spaces (TVWS) at the frequency range from 470 MHz to 792 MHz, LTE–800 frequency band for radio mobile communications from 790 MHz to 862 MHz and a joint license frequency band to operators from 2.50 GHz to 2.69 GHz, known as LTE–2600. To efficiently exploit these frequency regions, cognitive radio approaches become more and more important. Cognitive radio is a new paradigm for wireless communication, where either the network or the wireless node itself changes certain transmission or reception parameters to efficiently execute its tasks [5]. To meet the requirements resulting therefrom, the signal processing is moved as far as possible to the digital part. The analogue part, however, has to be tailored to the needs of a CR system as well, including the implementation of tune-

Mario Schühler (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
mario.schuehler@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_9, © Springer 2011

able pre-selection filters, for instance, or of broadband or multiband antennas and LNAs to cover the bandwidth focussed on for CR operations.

Before CR systems are able to perform, several challenges, researchers and operators are faced with, have to be solved. As part of the 7th Framework Program, two projects funded by the European Community deal with these challenges. While SACRA [3] considers spectrum and energy efficiency of CR systems, QoSMOS [2] looks at quality-of-service aspects. Energy efficiency and flexibility in the use of radio spectrum are important research topics, serving as a basis for the development of future wireless communications technologies.

The present work deals with RF part of a CR receiver. It divides into two main parts. Section 2 focusses on a reconfigurable RF front-end architecture, providing spectrum aggregation capability, as a major objective in QoSMOS. Several key components are identified, here. The design of these key components is discussed in Sect. 3 by examples, including a dual-band dual-polarised antenna, a broadband LNA, and a tuneable pre-selection filter. Issues during the design process are discussed to give an impression on the challenges to be solved.

## 2 RF Receiver Capable of Spectrum Aggregation

CR front-ends are supposed to cover multiple frequency bands to allow for opportunistic and dynamic use of spectral white spaces. An architecture consisting of parallel signal paths, each of them is tailored to a single frequency band, is a straightforward approach that is pursued in SACRA, for instance [4]. Figure 1a illustrates the idea by example of a receiver front-end. The signal received by the antenna is amplified by an LNA. By means of a switch, the respective frequency region or channel can be selected, within which the information is expected. The chosen channel is down-converted to an intermediate frequency (IF) and over an anti-aliasing filter (AAF) fed to an analogue-digital converter (ADC), working in subsampling mode. Depending on the link budget, one or more amplifying stages can optionally be included as well as variable attenuators.

To reduce the efforts following from a switching architecture, a tuneable architecture can be used, as portrayed by example in Fig. 1b for the receiving case. After the antenna and the LNA a tuneable bandpass follows, serving as pre-selection filter and basically substituting the switching between several channels. The selected channel is down-converted to the IF band and, again, is fed to an ADC after passing an AAF. As for the switching architecture, one or more amplifying stages and variable attenuators can optionally be included in the front-end.

In QoSMOS, an approach is pursued that even extends the idea of a tuneable front-end [1]. An objective of QoSMOS is to use fragmented spectra, i.e., to simultaneously exploit various channels being possible spread over a wide bandwidth. The technique called *spectrum aggregation* allows for a high throughput, even though the bandwidth required is unavailable in a whole but available fragmented.

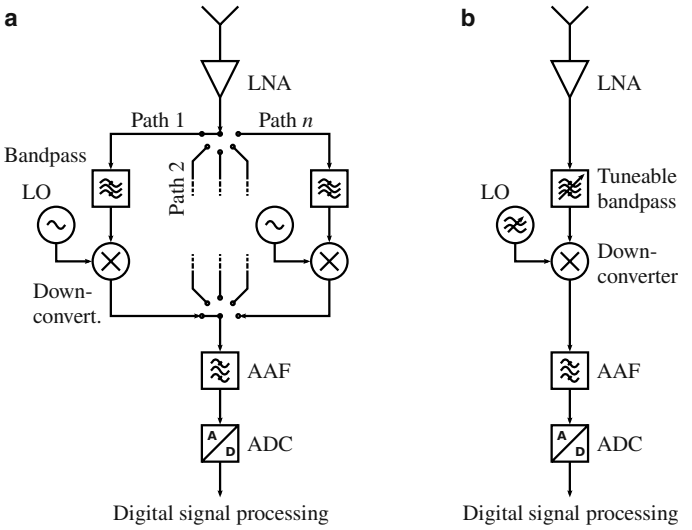



**Fig. 1** Schematic view of a RF receiver front-end with (**a**) switchable frequency channels and (**b**) tuneable pre-selection filters

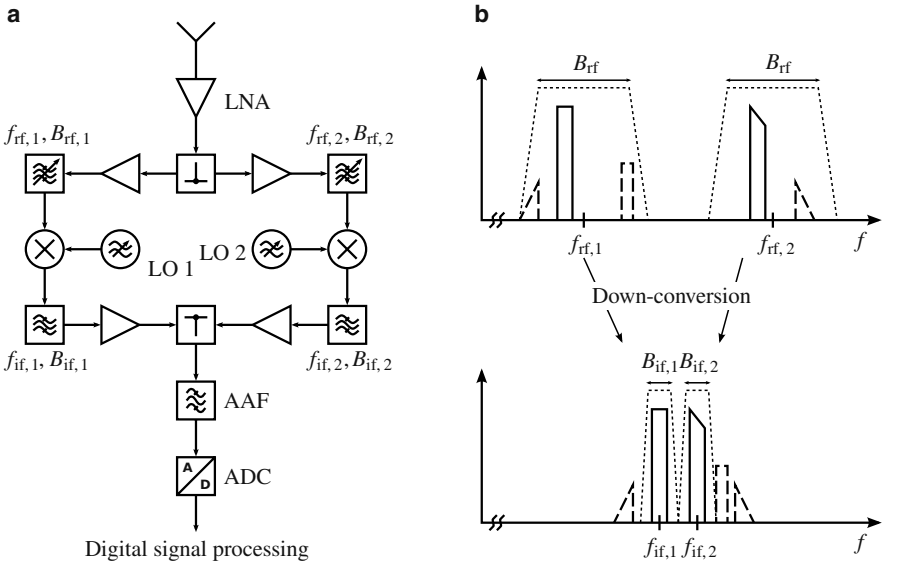



**Fig. 2** (**a**) Schematic view of an RF receiver front-end with spectrum aggregation capability. (**b**) Sketch of the RF signal spectrum and the IF spectrum after down-conversion. Undesired spectral parts are marked by *dashed lines*. The characteristics of the pre-selection filter and the IF filters are indicated by *dotted lines*

A receiving architecture capable of aggregating two arbitrary channels is depicted in Fig. 2a. After the LNA, the signal received is distributed to two signal branches by means of a power divider. Both branches are made up of equal components, including an amplifier to decouple the branches from each other, a tuneable bandpass to select the proper channel of bandwidth $B_{\mathrm{rf},1}$ with centre frequency $f_{\mathrm{rf},1}$ and bandwidth $B_{\mathrm{rf},2}$ with centre frequency $f_{\mathrm{rf},2}$, respectively, a down-converter, an IF filter of bandwidth $B_{\mathrm{if},1}$ with centre frequency $f_{\mathrm{if},1}$ and of bandwidth $B_{\mathrm{if},2}$ with centre frequency $f_{\mathrm{if},2}$, respectively, and another amplifier for decoupling. Further amplifying stages and variable attenuators can be included in the front-end. The idea behind the architecture shown in Fig. 2a is, basically, to divide the IF band into adjacent sub-channels, one per branch, of equal bandwidth $B_{\mathrm{if},1} = B_{\mathrm{if},2}$. Each pre-selected channel is thus down-converted to the respective IF sub-channel (cf. Fig. 2b), while afterwards the signals of all sub-channels are combined by a power combiner, followed by an AAF and an ADC. We notice that only one ADC is required.

The architectures portrayed in Fig. 1b and Fig. 2a essentially comprise similar components. While broadband amplifiers and mixers are commercially available, broadband or multiband antennas and LNAs as well as tuneable filters usually have to be individually tailored to the requirements of the system. They are therefore key components of a CR front-end.

## 3 Example Components: Design and Results

Having introduced the RF front-end architecture, we now want to look at design examples, addressing the key components identified of a CR system. First, we consider a dual-band dual-polarised antenna, covering the frequency ranges from 700 MHz to 862 MHz and from 2400 MHz to 2690 MHz. Second, a broad-band LNA design is proposed with a noise figure of at most 0.5 dB within its working frequency range. Last, the design of a tuneable bandpass is discussed.

### *3.1 Dual-Band Dual-Polarised Monopole Antenna*

Among the frequency region to be covered, the antenna design discussed here aims at a small size, allowing for the integration into a handset or a laptop device, and, in addition, at an implementation providing polarisation diversity. Preliminary simulation results are presented in this subsection to give an estimation on what can be achieved in terms of bandwidth, isolation, and gain.

The antenna concept is depicted in Fig. 3a. Two orthogonally-polarised dual-band antennas were chosen to reduce the number of outputs to two, as opposed to four outputs assuming one output for each frequency band and polarisation. The antenna elements are placed on the top layer of a PCB with a distance of

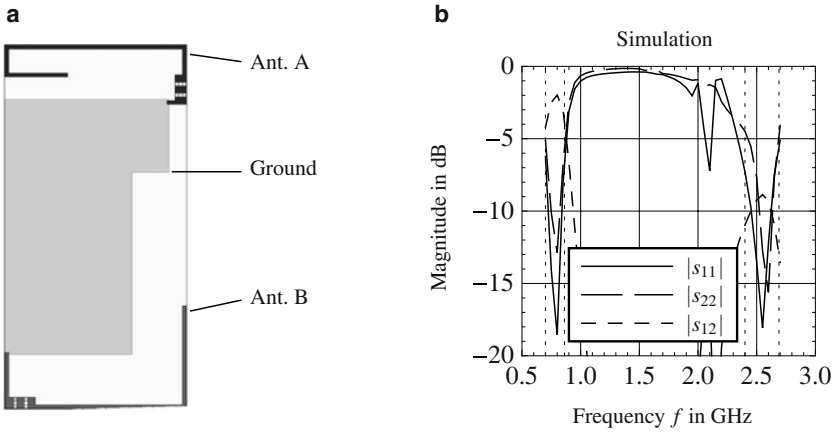



**Fig. 3** (**a**) Simulation model of the dual-band dual-polarised antenna. Antennas are located on the top layer, ground on the bottom layer. (**b**) Frequency response of S parameters obtained from simulation. *Dotted vertical lines* indicate band edges

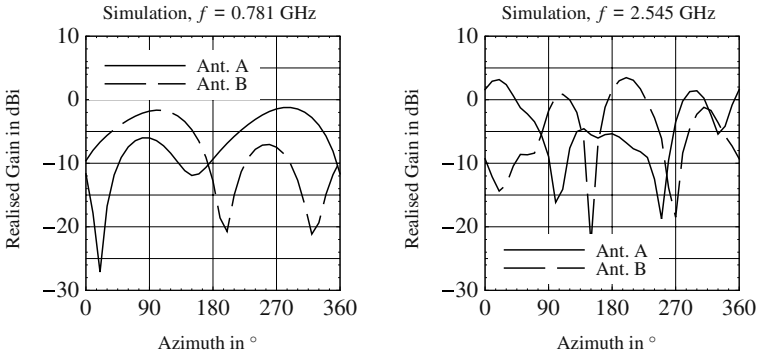



**Fig. 4** Realised gain over azimuth in the horizontal plane (elevation 0°) at 0.781 GHz and 2.545 GHz. The results were obtained from simulations

15 mm to the ground plane, which is located at the bottom layer. The substrate measures a thickness of 0.508 mm and has a relative permittivity of 3.55 (nominal values).[1] The size of the antenna was restricted to $100 \times 50\,\text{mm}^2$ (mobile phone format).

Figure 3b shows the frequency response of the input reflection magnitude of antenna A $|s_{11}|$ and antenna B $|s_{22}|$, respectively, as well as the transmission between both antenna ports $|s_{12}|$ obtained from simulation. We notice that both antennas display a matching of at least −6 dB within both frequency bands. The coupling $|s_{12}|$ within the upper frequency band does not exceed −9 dB, within the lower frequency band, however, a pronounced coupling occurs. The antenna configuration is, never-

[1] RO4003C from Rogers Corp.

theless, suited for diversity applications, even though a reduced diversity gain has to be expected in comparison with ideally decoupled polarisations.

Figure 4 portrays the realised antenna gain as a function of azimuth in the horizontal plane (elevation 0°), exemplary for one frequency inside the lower and the upper frequency band, respectively. In comparison with a single element with a minimum gain of −27 dBi at 0.781 GHz, at least −10 dBi is achieved with two antennas, at 2.545 GHz an improvement from −23 dBi to −12 dBi can be observed.

## *3.2 Broadband LNA*

This subsection deals with a broadband LNA, operating within the frequency range from 700 MHz to 3.0 GHz; TVWS (700 … 790 MHz), LTE–800 (792 … 862 MHz), and LTE–2600 (2500 … 2690 MHz) frequencies are therefore included. According to Friis' formula [6, p. 302], to minimise the noise figure of the front-end, the LNA directly follows the antenna, assuming it has a low noise figure.

Figure 5 shows a close-up of the implemented LNA mounted in metal case. It is based on a high-electron mobility transistor (p-HEMT) working within a frequency range from 50 MHz to 6.0 GHz.[1] As seen, the RF path and the DC part are separated from each other to avoid disturbances between both types of signals. To achieve a low noise figure over the working frequency range, a matching network implemented in microstrip technique is used. The LNA is built on a 0.508 mm thick substrate with a relative permittivity of 3.55 (nominal values).[2] The small design size of the actual circuit makes the LNA suitable for integration on an antenna.

Simulation and measurements results of the implemented LNA are illustrated in Fig. 6. Figure 6a shows the noise figure as a function of frequency $f$. We ob-

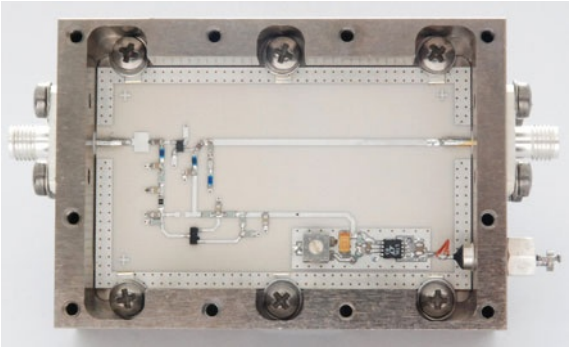


**Fig. 5** Close-up of the implemented LNA mounted into a case. RF connections are provided by SMA jacks. The PCB measures 63.8 mm × 37.8 mm

[1] ATF54143 from Agilent/Avago

[2] RO4003C from Rogers

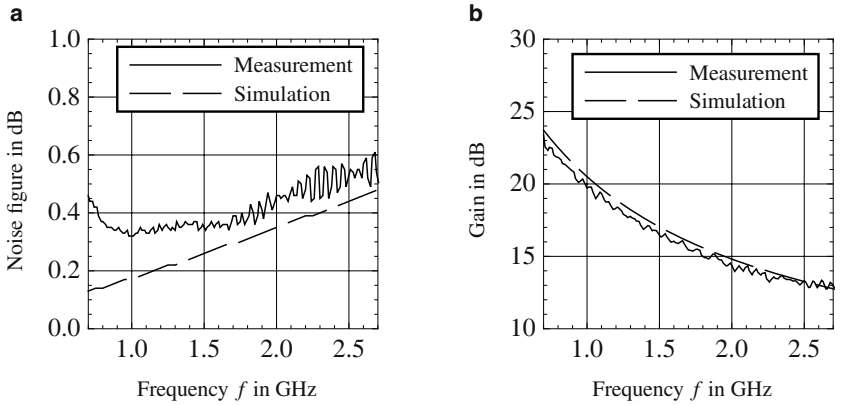



**Fig. 6** Simulation and measurement results of the broadband LNA: (**a**) noise figure and (**b**) gain over $f$

serve that a value of approximately 0.6 dB is not exceeded between 0.7 GHz and 2.69 GHz. Within this frequency range the gain amounts to at least 12 dB, as seen in Fig. 6b, where the frequency response of the gain is portrayed. The curve reflects the strongly pronounced frequency response of the transistor deployed, as the gain drops approximately 10 dB in between the lower and upper frequency.

## *3.3 Tuneable Pre-selection Filter*

This subsection focusses on the design of a tuneable bandpass filter. In terms of the centre frequency, a tuning range from 470 MHz to 960 MHz is aimed at, including TV band, LTE–800 band, and GSM band. The bandpass serves as pre-selection filter to suppress image frequencies. Assuming an IF of at least 200 MHz, the requirements on the bandwidth are relaxed, i.e., the focus during the design is on the tuneability of the centre frequency $f_{\mathrm{rf}}$.

Figure 7 shows a schematic of the filter circuit. The actual bandpass is built by two coupled resonators, each consisting of a varactor diode arranged in series to the fixed inductance $L_2 = L_4$. Depending on the reverse voltage $U_{\mathrm{r}}$ applied to the diodes, the resonant frequency of a resonator can be adjusted, as the junction capacitance varies with $U_{\mathrm{r}}$. Both resonators work at the same resonant frequency that is equal to $f_{\mathrm{rf}}$, only a single control voltage is therefore required. The coupling between the resonators is determined by the shunt inductance $L_3$.

The bandpass is followed by a lowpass based on a 3rd order type-II Chebychev filter [8]. The lowpass ensures passbands to be suppressed, possibly appearing due to parasitics, as those passbands may fall into the working frequency range of the antenna discussed in Sect. 3.1. The coupling of the bandpass to the lowpass is determined by $L_5$; to the source, in turn, by $L_1$.

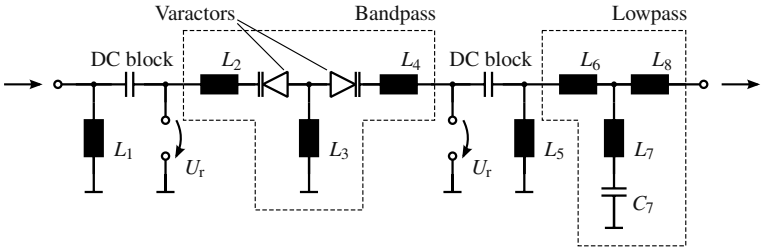



**Fig. 7** Schematic of the tuneable bandpass filter including two varactor diodes for tuning the centre frequency. Both diodes are biased with $U_r$; decoupling from RF is provided by DC blocks

The filter was implemented on a 0.508 mm thick substrate with a relative permittivity of 3.55 and dielectric loss factor of 0.0027 (nominal values).[1] The varactor diode chosen provides a junction capacitance tuneable from 12 pF at $U_r = 1$ V to 0.5 pF at $U_r = 28$ V.[2]

Figure 8 portrays the frequency response of the forward transmission magnitude $|s_{21}|$ obtained from measurement with $U_r$ serving as parameter. We notice that for a variation of $U_r$ from 0 V to 8 V the tuning range desired is already achieved, covering the frequency region from 430.6 MHz to 984.3 MHz. The centre frequency $f_{rf}$ grows linearly with $U_r$, since the varactor displays a hyperabrupt dop-

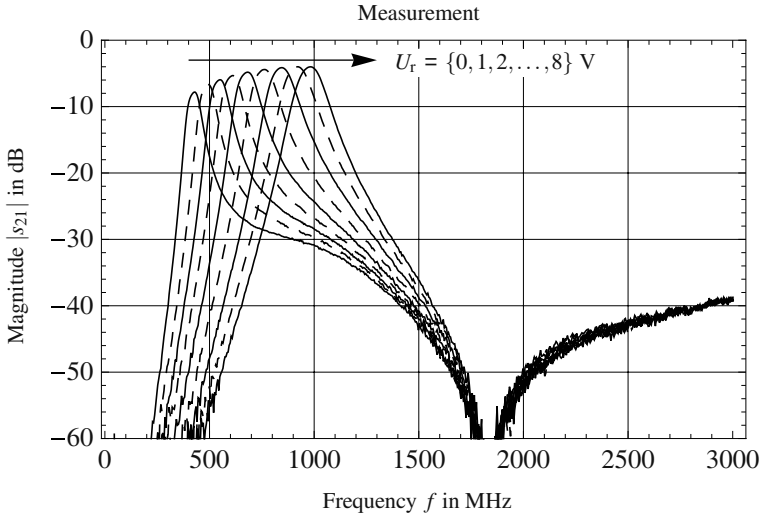



**Fig. 8** Measured frequency response of the forward transmission magnitude $|s_{21}|$ of the tuneable bandpass with parameter $U_r$

[1] RO4003C from Rogers Corp.

[2] BB833 from Infineon

ing profile. The variation of the insertion loss $1/|s_{21}|$ with $U_\mathrm{r}$ within the passband stems from the slightly pronounced dependence of the junction resistance on the reverse voltage: with increasing $U_\mathrm{r}$ the junction resistance decreases and so is the insertion loss. The relative bandwidth is approximately 14.7%, which is basically independent of $U_\mathrm{r}$. Yet a slight reduction with increasing $U_\mathrm{r}$ could be observed in the measurement due to the root caused by the lowpass stage.

## 4 Summary

The present work focussed on a reconfigurable RF receiver front-end providing spectrum aggregation capabilities. The architecture proposed allows for simultaneous use of channels being possibly spread over a wide frequency region, while only one ADC is required. The requirements on the IF filters, however, are strong for this approach. To exploit the whole input bandwidth offered by the ADC, the filters have to display a high selectivity. In terms of energy efficiency, on the other hand, minimising the number of ADCs leads to an reduced power consumption in total, provided that the insertion loss of the IF filters is comparatively small.

Three key components of a CR RF receiver front-end were identified: the antenna, the LNA, and the tuneable pre-selection filter. The design of these components were discussed by example in Sect. 3, where a dual-band dual-polarised monopole, a broadband LNA, and a tuneable bandpass were presented. The example implementations can not only be deployed in a CR RF receiver front-end as desired in SACRA and QoSMOS but possibly in other systems as well.

## Acknowledgements

This work was supported by the European Community's Seventh Framework Programme (FP7) under Grant 248454 (QoSMOS) and Grant 249060 (SACRA) and by the German Federal Ministry of Education and Research (BMBF) under Grant 01SF0708 (TEROPP) in collaboration with *Association des instituts Carnot*.

## References

1. Datta R et al. (2011) Flexible multicarrier PHY design for cognitive radio in white space. In Proceedings of the 6th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom), Osaka, Japan
2. European Community's 7th Framework Programm (2010) Quality of service and mobility driven cognitive radio systems (QoSMOS), www.ict-qosmos.eu. Grant 248454
3. European Community's 7th Framework Programm (2010), Spectrum and energy efficiency through multi-band cognitive radio (SACRA), www.ict-sacra.eu. Grant 249060

4. Jaschke A (2010) Entwicklung eines Mehrkanalempfänger Front-Ends für Cognitive Radio, Master's thesis, Friedrich-Alexander Universität Erlangen-Nürnberg, Germany. In german
5. Mitola J (2000) Cognitive Radio – An Integrated Agent Architecture for Software Defined Radio, Ph.D. Thesis, Royal Institute of Technology, Sweden
6. Nibler F (1998) Hochfrequenzschaltungstechnik, Expert-Verlag GmbH, 3rd edn.
7. Wyglinski AM, Nekovee M, Hou T (2010) Cognitive Radio Communications and Networks (Principles and Practice), Elsevier Verlag
8. Zverev AI (2005) Handbook of Filter Synthesis, John Wiley and Sons, Inc.

# Modelica – A Modelling Language for Heterogeneous Systems

**Christoph Clauß, Ulrich Donath, Olaf Enge-Rosenblatt, and Kristin Majetta**

**Abstract** The modelling language Modelica has been developed by a steadily growing community since 1996. Today, it is a well-established element of the design process in several technical application areas. Fraunhofer IIS/EAS has been involved in the language development from its very beginning. Research on dedicated fields like structural variability and control algorithms using Modelica as well as the development of complete packages (Modelica Standard Library, Complex and Statistics Library) are highlights of these Modelica activities. This article presents a compact overview about the variety of research results which have been achieved by the Dresden Division of Fraunhofer IIS during the last 15 years.

## 1 Introduction

The 8th International Modelica Conference which took place in March 2011 in Dresden with 316 attendees and 99 papers was organized by both the Modelica Association and Fraunhofer IIS/EAS. This event was a benchmark for the success of the Modelica activities at Fraunhofer IIS/EAS. 15 years ago Fraunhofer IIS/EAS was involved in the early language development of the object oriented model description language Modelica which was designed for multi-physical modelling based on hybrid DAEs using an equation-based description. Up to now Modelica has been used in many research projects, and Fraunhofer IIS/EAS has contributed to the development of model libraries. In this paper, highlights of Modelica related research at Fraunhofer IIS/EAS are presented. In the beginning this work was mainly focused on the development of the electronic packages of the Modelica Standard Library (MSL). Later investigations on structural variability of differential-algebraic systems (DAEs), the generation of control algorithms using Modelica as well as the

Christoph Clauß (✉)
Fraunhofer IIS/EAS, Zeunerstraße 38, 01069 Dresden, Germany
christoph.clauss@eas.iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_10, © Springer 2011

development of dedicated libraries such as Complex and Statistics Libraries became important supplemental topics.

## 2 Contributions to the Modelica Standard Library

The Modelica Standard Library (MSL) contains a huge amount of models sorted by their physical domain in different packages. These models can be connected to each other even if they are part of distinct packages. In the electrical domain, Fraunhofer IIS/EAS developed three packages: the Analog Library, the Digital Library, and the Spice3 Library.

### *2.1 Analog Library*

The first models of the `Modelica.Electrical.Analog` library have been developed since the very beginning of the MSL. Some of these basic textbook-like models are resistor, capacitor, inductor, transformer, lines, sources, ideal models, and simple semiconductor models. During the last years, the Analog Library has been strongly enlarged and improved [1]. More detailed models have been added such as thyristor, triac, operational amplifiers, switches with arc, AD/DA converters etc. A very important new model is a conditional heatport which allows the coupling of electrical and thermal networks. With these enlargements and improvements, the Analog Library became more important in the field of real electrical simulation.

### *2.2 Digital Library*

The `Modelica.Electrical.Digital` library contains digital electrical components based on the IEEE 1076 VHDL standard with 9-valued logic and conversion to 2-, 3-, 4-valued logic. The library consists of basic models like AND, OR, XOR, etc. Furthermore, flipflops, latches, gates, tristates, and memories are part of the Digital Library [2].

### *2.3 Spice3 Library*

The Spice3 Library [3] was developed between 2007 and 2010 mainly in the ITEA2 projects EUROSYSLIB and MODELISAR. This library contains most of the models of the de facto standard SPICE3 simulation tool developed by the University of California, Berkeley. The models of the `Modelica.Electrical.Spice3`

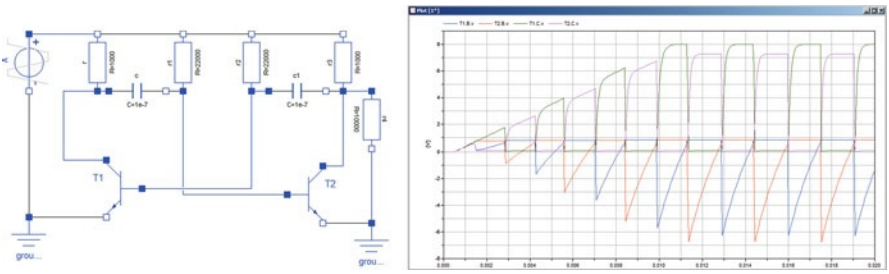



**Fig. 1** Schmitt trigger circuit with simulation result

library are thematically arranged in packages according to the grouping of the SPICE3 models. Additionally, MSL typical packages are added (User's Guide, Interfaces etc.). Besides typical electrical models like resistors or controlled sources, the Spice3 Library comprises semiconductor models (Mosfet level 1, bipolar transistor, diode, and resistor) which are the most important ones. These models are very complex and allow a much more detailed simulation with respect to more physical phenomena than the semiconductor models of the Analog Library. Since a huge amount of functions and data is needed to describe the models, a package *Internal* was introduced to store all these necessary data, records, and functions. Figure 1 shows a schematic of a Schmitt trigger that contains two bipolar transistors and further components of the Spice3 Library. The simulation result shows the typical oscillating behaviour.

## 3 Development of Dedicated Libraries

Besides the Modelica Standard Library a huge number of other, special Modelica libraries exists, e.g. a Belts Library, a Fuel Cell Library, etc. [4]. Fraunhofer IIS/EAS also developed such special libraries, e.g. a Complex Library and a Statistics Library.

### *3.1 The Complex Library*

The Complex Library is a free Modelica library for steady-state analysis of linear AC circuits with fixed frequency [5, 6]. This library contains special electrical components (resistors, inductors, capacitors, voltage sources, and current sources) using a particular interface definition. The behaviour of these components is described using the phasor domain (so-called time phasors). There are also components included from the electromechanical domain. In these cases, the electrical subsystem is modelled within the phasor domain while the mechanical subsystem is described using

time domain. The Complex Library is fully compatible with the Modelica Standard Library.

A quasi-stationary mode is usually not included within phasor domain method. However, under special assumptions, such an approach is allowed which yields enormous savings of simulation time. The quasi-stationary mode is understood as a sequence of steady states under the following condition: parameters (which would be constant at steady-state analysis) may vary extraordinary slowly compared to the system's dominant time constant. It is signalized by slow alterations of amplitudes and phases of the sinusoidal quantities.

### *3.2 The Statistics Library*

To investigate the influence of parameter variability on the systems behaviour, Monte Carlo methods are usually applied which need statistically varying parameters in repeated simulation runs. The simulation results are post-processed to measure the influence of parameter variations. Referring to the standard J 2748 prepared by the Electronic Design Automation Standards Committee of the Society of Automotive Engineers (SAE), a Statistics Library was developed which provides a set of distribution functions which can be assigned to Modelica model parameters. This allows the specification of statistical behaviour. The distribution functions are designed to ensure a statistical parameter variation for repeated simulation runs independently from the simulation tool [7]. The application of the Statistics Library is simple. E.g. the code `Modelica.Electrical.Analog.Basic.Resistor R3(R=uniform(100, 0.05));` specifies the resistance of R3 to be uniformly distributed between 95 Ohm and 105 Ohm (tolerance of 5% of the nominal value 100 Ohm). It is possible to generate both independent and correlated random values. All distribution functions are derived from the standard uniform (0, 1) distribution which is generated using a pseudo random number generation algorithm. By repeated choosing of the same seed number for the random number algorithm, it is possible to generate the same series of random values, e.g. for special investigations. Furthermore, the distribution functions can be applied to generate noise during transient simulation. A significant advantage of the Statistics Library approach is the documentation of statistical distributions in the model files.

## 4 Control Algorithms

Besides the physical model, e.g. for a machine, which describes the interactions of electrical, mechanical, or/and hydraulic components, the control algorithms are of particular interest to analyze the system's behaviour.

## 4.1 Graphical Design

Basically, open loop control systems are based on state machines which can be designed graphically by means of UML state diagrams [8]. Usually, they have the characteristics of both Mealy state machines and Moore state machines. This is achieved by the denotation of activities at transition arrows (Mealy machines) and in states (Moore machines). Besides the generation of output signals, the activities comprise assignments to variables, execution of simple algorithms, function calls, and generation of events. There are different kinds of triggers which can be used for state transitions: signal triggers, change triggers, and time triggers. An additional concept in UML state diagrams, similar to Harel statecharts, is realized by hierarchically nested states. The state hierarchy is realized by composite states which contain one or more regions, each with a set of states. Again, these states are either composite states or simple states. In the context of Modelica, expressions in the Modelica language are used for the denotation of activities and control operations, respectively.

## 4.2 Automatic Modelica Code Generation

The state diagram and the related type, variable, parameter, and signal declarations are translated automatically into a Modelica algorithm and corresponding declarations [9]. They are jointly handled in a Modelica model which is instantiated later on as control component of electrical, mechanical, or/and hydraulic machine components (Fig. 2). The state diagram translation into Modelica code is carried out by the Modelica Code Generator developed by Fraunhofer IIS/EAS. The Modelica Code Generator generates standard Modelica code from the UML state diagram.

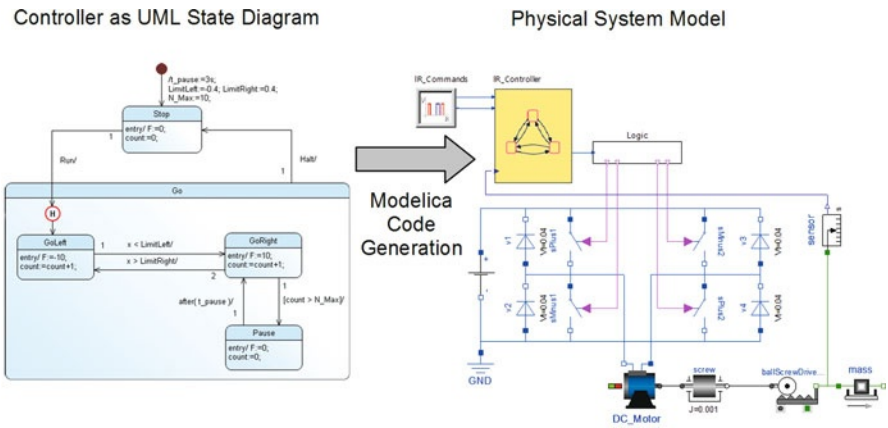



**Fig. 2** UML State Diagrams in SimulationX [9]

In addition, the Modelica Code Generator accomplishes Design Rule Checks. Besides Modelica code, the generator architecture is open to generate code for various targets such as formal verification tools [10] or production code for PLCs [11], or embedded controllers.

### *4.3 Test and Verification*

The main technique for the verification of control algorithms is the simulation. For checking the functionality of the whole system, the common view of internal variables of the controller and variables of the plant model are essential. Therefore, all variables of the state diagram may be accessed by the user. The causality of the system behaviour is reflected by the history of operator commands, controller variables, and plant variables. To increase the reliability of the simulation based verification, State Coverage Analysis and Transition Coverage Analysis are applied. For this purpose, state activation counters and transition activation counters are automatically inserted into the generated Modelica code and are incremented during the simulation run. The actual achieved counting results may continuously be back-annotated into the state diagram of the controller.

## 5 Structural Variability

Structural variability occurs in so-called hybrid systems. These are understood as dynamical systems consisting of components with both time-continuous and time-discrete behaviour. In order to model such systems correctly, the continuous and discrete behavioural phenomena have to be described separately. Furthermore, the interactions between them have to be considered. The dynamics of time-continuous behaviour is usually represented by DAEs. Time-discrete behaviour represents a loose succession of events and can, therefore, be described e.g. by Boolean equations, finite state machines, or statecharts. Sometimes, such a time-discrete event causes a change of the physical behaviour of the system and, hence, forces a change of the model equations (the DAEs). This is what we call a structural change [12].

### *5.1 Structure-Changing Phenomena*

Most systems which include structural changes are characterized either by existence of so-called unilateral constraints or by switching events which activate or deactivate parts of the system. Such a system either really changes its physical structure or the structure of the mathematical description is changed during operation. Examples may be found in different application areas, e.g.

- in mechanics: clutches, collision of masses, Coulomb friction, “maximum distance”-phenomena (see example string pendulum),
- in power electronics: diodes and thyristors (if considered as ideal switches).

There are different ways on which structural changes may arise. This includes substitution of one differential equation or a set of differential equations, replacement of behavioural or structural description of a component (e.g. a drastic variation of model order), addition or deletion of components, or of connections between them.

## *5.2 Hybrid Simulation with Modelica-Like Modelling*

Usually simulators for hybrid systems are able to handle continuous and discrete parts with appropriate numerical solvers. However, concepts and description means for models with structural variability are urgently needed. Such concepts were developed within the research project GENSIM which was a joint activity of four Fraunhofer institutes, including EAS. In an experimental simulator called Mosilab [13], an extension of the language Modelica by a concept for dealing with structural dynamics has been implemented. For this purpose, a description of statecharts was implemented both in graphical and textual way. Roughly spoken, every state of a statechart stands for a certain set of DAEs and every transition realises a change between different model structures. The basic algorithm consists of two phases, a discrete phase and a continuous phase (see Fig. 3). The main issue in the discrete phase is to update all state machines of the hybrid model and to establish the new set of differential-algebraic equations if necessary. During the discrete phase, no time step is made (zero time assumption). Between two successive discrete points in time, only analogue simulation is performed which is carried out by a numerical DAE

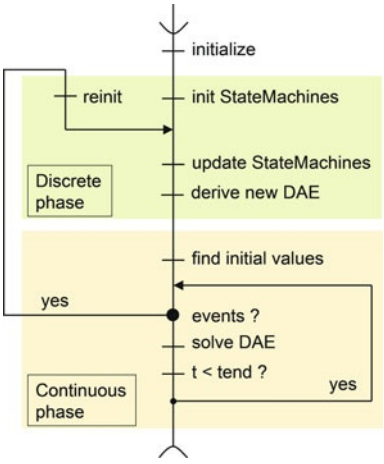



**Fig. 3** Basic algorithm of a simulator for hybrid systems

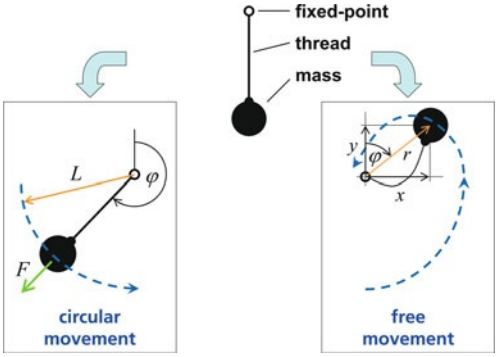



**Fig. 4** Two kinds of movements of string pendulum

solver. The next event may happen at a determined time instant (time-dependent event) or at an apriori unknown time because it has to be triggered by a zero crossing of an indicator function (variable-dependent event). The main challenge after every event is the determination of valid initial values for the new DAE structure.

### *5.3 Example of a String Pendulum*

The 2-D string pendulum (see Fig. 4) is a "maximum distance problem" (see also [12]) and well-suited as benchmark for structural variability. A point mass is attached to a fixed point by a non-elastic thread. The mass can perform a circular movement in case of a fully elongated thread (left part of Fig. 4) or a free movement in case of a non-elongated thread (right part of Fig. 4). Within the circular movement, one differential equation of second order is valid (and the distance condition of the thread). In the other case, two differential equations of second order are needed. Indicator functions are the cut force within the thread (circular movement) or the difference between thread length and mass' radius (free movement).

## 6 Summary and Outlook

Within this paper a compact overview of the Modelica activities at Fraunhofer IIS/EAS was given. Due to the development of several libraries and the research on structural variability as well as on control algorithms, today Fraunhofer IIS/EAS is a respected part of the Modelica community. Since Modelica has proven to be a sustainable modelling approach, Fraunhofer IIS/EAS will continue in contributing to both language and Standard Library within the Modelica Association. Furthermore, the language will be used intensively in research projects for modelling and the development of exploitable libraries will increase the commercial success.

## References


1. Majetta K, Clauß C, Franke M, Schneider P (2009) Improvement of the MSL Electrical Analog Library. In 7th International Modelica Conference, Como, Italy, Proc, pp 561–566
2. Clauß, C., Donath U, Schneider A, Weber E (2005) Standard Package Modelica.Electrical. Digital. In 4th International Modelica Conference, Hamburg, Germany, Proc, pp. 141–147
3. Majetta K, Clauß C, Böhme S, Schneider P (2011) MSL Electrical Spice3 – Status and Further Development. In 8th International Modelica Conference, Dresden, Germany, Proc
4. https://modelica.org/libraries
5. Enge O, Clauß C, Schneider P, Schwarz P, Vetter M, Schwunk S (2006) Quasi-stationary AC Analysis Using Phasor Description With Modelica. In 5th International Modelica Conference, Vienna, Austria, Proc, 2:579–588
6. Enge-Rosenblatt O, Clauß C, Schneider P, Vetter M, Schwunk S (2009) ComplexLib – a Modelica library for steady-state analysis of AC circuits within phasor domain. In 7th Int Modelica Conf, Como, Italy, Libraries in Proceedings
7. Haase J, Wolf S, Clauß C (2008) Monte Carlo Simulation with Modelica. In 6th International Modelica Conference, Bielefeld, Germany, Proc, pp 601–604
8. http://www.omg.org/spec/UML/2.2/Superstructure/PDF/ Accessed on: 10 May 2011
9. Donath U, Haufe J, Blochwitz T, Neidhold T (2008) A new approach for modeling and verification of discrete control components within a modelica environment. In 6th International Modelica Conference, Bielefeld, Germany, Proc, pp 269–276
10. Klotz T, Fordran E, Straube B, Haufe J (2009) Formal verification of UML-modeled machine controls. In 14th IEEE International Conference on Emerging Technologies and Factory Automation 2009, Palma de Mallorca
11. Seidel S, Donath U (2011) Error-free control programs by means of graphical program design, simulation-based verification and automatic code generation. In 8th International Modelica Conference 2011, Dresden, Germany
12. Enge-Rosenblatt O, Bastian J, Clauß C, Schwarz P (2007) Numerical Simulation of Continuous Systems with Structural Dynamics. In 6th EUROSIM Congress on Modelling and Simulation, Ljubljana, Slovenia, Proc on CDROM
13. Nytsch-Geusen C et al (2005) Mosilab: Development of a Modelica based generic simulation tool supporting model structural dynamics. In 4th International Modelica Conference, Hamburg, Germany, Proc

# Part II
# Information Systems

# Information Systems Driven by the Fraunhofer Institute for Integrated Circuits IIS

**Ernst Eberlein and Jörn Thielecke**

Communication and navigation have always been essential to the success of human societies. The term information society expresses the prominent relevance of information systems in general, and communications and navigation systems in particular for the society of today. Radio communications and radio navigation have a history of more than 100 years. Pioneers like Guglielmo Marconi and Christian Hülsmeyer, the inventor of radar, paved the way to modern information systems together with many others. For more than 25 years Fraunhofer IIS has contributed its share to the rapid development of information systems, concurrently training many young engineers in this seminal field.

Since 1985, the year when the institute was founded, communication technologies have been subject of Fraunhofer IIS activities. At the beginning, the focus was on analogue RF technologies. Hearing aids and weather sensors are examples that were addressed in the very beginning. In 1994 digital communications moved into focus. Due to a close cooperation with universities, Fraunhofer IIS was soon able to offer design services on all elements of the communication chain. Source coding, RF components, integrated circuit design, and VLSI technologies are counted among the core competences of Fraunhofer IIS. In combination with a customer requirement design philosophy, they form the basis for developing novel solutions. Examples are the satellite-based systems for mobile reception [1], which have been specified and developed with significant contributions from Fraunhofer IIS. Satellites are especially attractive for broadcasting applications, communication systems offering nationwide or even worldwide seamless coverage and navigation. The direct-to-home broadcasting applications are well known and offer many programs to the end user. In 1995 Fraunhofer IIS started to develop systems requiring only small antennas instead of satellite dishes. This allowed portable and mobile reception. Since 2001 systems like Sirius-XM Radio [1] have offered more than 150 audio channels available everywhere in the US.

Ernst Eberlein (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
ernst.eberlein@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_11, © Springer 2011

Especially for audio broadcasting to cars, the user expects seamless reception and a high quality of service (QoS). A satellite can offer nationwide or even European-wide coverage. The challenges are the characteristics of the land-mobile-satellite (LMS) channel. The propagation is characterized by fast variations of the signal strength and temporal blockages. To offer sufficient QoS, communication systems designed for full mobile reception typically use a high level of redundancy (e.g. powerful forward error correction schemes (FEC) combined with time-interleaving) or other diversity combining technologies (including MIMO concepts). The channel coding know-how combined with the source coding know-how are the enabling technologies offering attractive services to mobile users. The characteristic of the LMS channel and the theoretical capacity, taking into account different air-interface configurations, are detailed in the article by Marco Breiling et al. The paper demonstrates the advantages of the concept combining time interleaving with FEC schemes using low code rates. The concept is implemented in the latest standards such as ESDR (ETSI Standard for Satellite Digital Radio) and DVB-SH (DVB standard for satellite reception with handheld terminals) and is also considered for other future standards.

Time-slicing is a technology for optimization of the power consumption of battery powered terminals. For system optimization, the trade-offs between several performance criteria have to be taken into account. The article by Christian Forster et al. investigates the trade-off between the gain of statistical multiplex and delay under the constraints resulting from time slicing.

Broadcasting applications will more and more converge with other applications. In the framework of the European FP7 project SafeTRIP, the potential of satellite-based communication technology in the context of intelligent transport systems (ITS) is investigated. Bernhard Niemann et al. give an introduction to the project and describe the core technologies developed in the framework of this project.

Today, typically so-called bent-pipe satellites are common for commercial applications. The limited flexibility of existing on-board processing (OBP) or the long time required for space qualification are the main reasons. Bent-pipe satellites essentially amplify and forward the receive signal without any signal processing. For future networks and especially for terminals using small antennas, OBP becomes attractive. With an OBP a direct communication without hub station ("single hop") between terminals with small antennas becomes feasible or the required resources (satellite power, bandwidth, etc.) are significantly reduced and make the system commercially much more attractive. Fraunhofer IIS currently develops technologies for on-board processing. One project targets a flexible platform for various applications. The article by Alexander Hofmann describes the platform.

The history of Fraunhofer IIS contributions to navigation technologies is younger than the history of contributions to the communications community, but similarly versatile. For the last 15 years GPS and Galileo receivers have been designed as well as tracking systems for pedestrian navigation, sports monitoring or motion analysis in the context of health monitoring.

One particular focus is on the development of Global Navigation Satellite Systems (GNSS) receivers as part of Europe's effort to push its industry beyond the

state of the art and to create a wide variety of new jobs before the European navigation satellite system Galileo reaches its full operational capability. The essay on "Multi-Constellation and Multi-Frequency Automotive GNSS Receivers" by Cécile Mongrédien et al. explores the opportunities and challenges in the design and implementation of market-oriented receivers capable of handling various modern satellite navigation signals.

In open spaces, satellite-based navigation is the natural choice, especially for vehicle navigation. In the case of pedestrian navigation, however, it typically fails due to difficult signal propagation conditions indoors and in dense urban areas. Navigation based on wireless LAN explicitly exploits these propagation conditions and has proven to offer more robustness and room-accurate positioning. Furthermore, for a pedestrian, it is essential to always know in what direction she or he is facing. Outdoors, this can be accomplished by a magnetic compass. However, indoors or in dense urban areas, the disturbances of the earth magnetic field are often too severe. Wi-Fi signals from a set of directional antennas, which eventually can be integrated into mobile devices, can be exploited to sense the orientation of the antenna array, i.e. the device. The contribution by Jochen Seitz et al. "Wi-Fi Attitude and Position Tracking" examines this novel approach.

Besides navigating in a city, it is often of interest to humans to monitor their physical activities. In a society inclined to rather watch sports than to actively participate in them, devices can be of help which keep records of daily physical activities. Radio-based tracking systems can calculate the travelled distance, but they can hardly differentiate whether the distance was walked or cycled. Simple inertial sensors can be of use here. The article by Martin Rulsch et al. "Motion Sensing: From Single Sensors to Sensor Networks" addresses this topic. As the title suggests, the contribution goes beyond single sensors. It describes ways of monitoring complex human movements by applying a network of accelerometers integrated into sports clothing. It highlights the difficulties which have to be faced in this context.

There are inconceivably many applications of navigation technologies in the field of motion analysis. An exceptionally interesting application is the task to automatically track and analyze a football match or training in order to optimize the performance of the athletes' performance. In their paper "A Real-Time Tracking System for Football Match and Training Analysis", Thomas von der Grün et al. describe the architecture, signal processing and some implementation aspects of such a radio tracking system. Most interestingly, they discuss how to process low-level position and trajectory information in order to extract high-level event information like the occurrence of a double pass.

The article by Sebastian Lempert and Alexander Pflaum, concluding the chapter, sheds light on the difficult task of integrating real-time positioning technologies – as diverse as discussed above – into already existing enterprise infrastructures. It argues that in the future an internet of things has to provide mechanisms to ensure that diverse enabling technologies like real-time locating systems, radio identification technology and wireless sensor networks can coexist in parallel. The authors propose an integration and application platform for diverse enabling technologies and finally sketch a suitable software architecture.

The nine contributions in this chapter cannot describe all the activities of Fraunhofer IIS in the area of information systems, but the selection sheds light on the many facets tackled in the institute and the ambition to provide complete solutions when addressing theoretical, implementation, systems and application aspects.

## References

1. Michalski R (2002) An overview of the XM satellite radio system. International Communication Satellite Systems Conference AIAA-2002-1844, Montreal, Canada

# Choice of Physical Layer Parameters for Mobile Satellite Broadcast

**Marco Breiling, Albert Heuberger, Ernst Eberlein, Aharon Vargas, Daniel Arndt, and Alexander Ihlow**

**Abstract** This chapter assesses the choice of the physical layer parameters (code rate, modulation, time interleaving) that are best suited for satellite broadcast. The optimum choice depends on the system's envisaged spectral efficiency or available signal-to-noise ratio, and the target usage environment, like rural or sub-urban reception. The analysis is carried out by means of information theory and evaluation of satellite field measurements.

## 1 Introduction

Mobile broadcast systems have gained growing attractivity in the consumer world in the last decade. After pioneering standards like DAB, that were restricted to pure terrestrial transmission, the advantages of using satellites for mobile broadcast reception were soon discovered, as a large region can be covered with a single satellite. The terrestrial waveforms are yet inadequate for the land-mobile satellite channel, that is experienced in such reception scenarios. Due to the strong large-scale fading effect caused by trees, buildings etc. with a behaviour resembling an on-off channel, a long interleaver is mandatory to mitigate the fading. Moreover, a satellite-only network is not suitable for reception inside urban areas due to frequent and long obstruction of the direct line-of-sight path from the satellite to the mobile receiver.

Therefore, new waveforms were developed for hybrid, i.e. simultaneous reception from a satellite and a terrestrial complementary network with repeaters in larger cities. The pioneers for such systems were the commercial XMradio and the Sirius radio systems in the United States. The last couple of years saw quite a number of hybrid broadcasting standards evolve like S-DMB, IDB-S, CMMB, ETSI SDR (ESDR), DVB-SH, and – currently under specification – the hybrid component of DVB-NGH (Next Generation Handheld).

Marco Breiling (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
marco.breiling@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_12, © Springer 2011

The DVB-SH standard [1] and DVB-NGH, both propose the use of OFDM for the terrestrial component, and OFDM or preferably single-carrier (SC) transmission for the satellite (TDM for DVB-SH and SC-OFDM for NGH). Both allow the choice between several signal constellations for OFDM as well as SC, and they contain modern state of the art Forward Error Correction (FEC) schemes with selectable code rate like turbo codes (DVB-SH) and LDPC codes (NGH). Both offer long time interleaving on the physical layer.

Therefore, these DVB standards open the door for a host of configurations as concerns the modulation order, the code rate and the time interleaver. For the modulation order, the system operator has the choice between QPSK, 8-PSK and 16-APSK/16-QAM for the SC transmission modes, and QPSK or 16-QAM for the OFDM modes. The code rate can be chosen between $1/5$ and $2/3$. Therefore, spectral efficiencies $\Gamma$ between the extremes $2/5\,\mathrm{bit/s/Hz}$ and $8/3\,\mathrm{bit/s/Hz}$ can be achieved. Moreover, the channel time interleaver is configurable in a wide range, i.e. from short (several ten milliseconds) to long (more than 10 s), and with various distributions of the bits of each codeword over time.

For the spectral efficiencies $\Gamma$, there are overlap regions, where the same or similar values of $\Gamma$ can be achieved with different configurations. For instance, a spectral efficiency of $\Gamma = 1\,\mathrm{bit/s/Hz}$ can be achieved with QPSK and code rate $1/2$, 8-PSK with code rate $1/3$ or 16-APSK/16-QAM with code rate $1/4$. For an optimization of the system configuration, the system operator needs to know which of the different choices is best in terms of power efficiency or robustness. Moreover, the time interleaver length has to be traded-off between increasing robustness (when longer) and reduced receiver cost (when shorter).

The parameter space required by the operators and a set of meaningful configurations could be investigated in excessive field trials. However, satellite field trials are rather expensive. They occupy the full satellite beam for a considerable amount of time (weeks). Moreover, the satellite is usually launched only shortly before the planned service launch. Finally, the required parameter space should already be known at the specification of a standard (as is currently the case for DVB-NGH), not only when the satellite is available for field trials.

This chapter focuses on the analysis of possible satellite waveform configurations without performing field trials, i.e. without actually using a waveform for a real transmission over a satellite. In order to minimize the analysis effort, the analysis is split up into two nearly independent aspects: (1.) code rate and modulation order, and (2.) time interleaving. The first aspect can be dealt with by the help of information theory, while the second point is handled using results from field measurements.

## 2 System Model

In the sequel we will use the system model displayed in Fig. 1. $X$ are the transmitted symbols. They stem from the alphabet $\mathcal{X}$, which is for instance a 16-APSK constellation. The temporal order of these symbols is changed by the time inter-

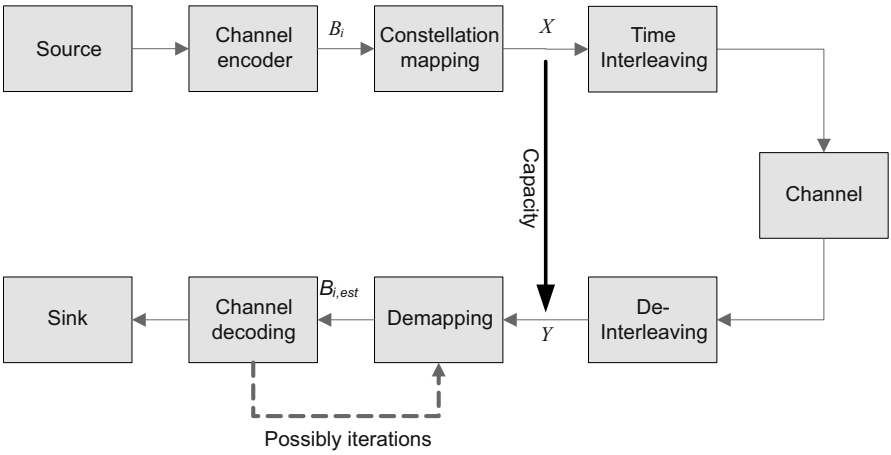



**Fig. 1** Generic model of the considered transmission chain

leaver. It spreads the symbols of one codeword over the given interleaver duration. After reception from the channel, the symbols are de-interleaved to the the received complex-valued symbols $Y$, whose energy is $E_s$. The one-sided power spectral density of the noise, that is added by the channel, is denoted by $N_0$.

Note that the block diagram applies directly for DVB-NGH, while it is slightly different for SH: the time (de-) interleaver and the (de-) mapping have to swap places. The results obtained in this chapter for the shown system model apply however in general also for SH.

## 3 Line-of-Sight Satellite Channel

Let us first analyze the choice of the code rate and modulation order for satellite reception. For stationary reception with a slightly directive antenna, and for signal bandwidths up to 5 MHz, the Line-Of-Sight (LOS) satellite channel can be characterized as a static channel with only a single relevant transmission path. There are only weak reflected paths, which moreover predominantly arrive from directions, where the directive antenna's gain is small (e.g. from the ground). Hence, the satellite LOS channel is a static flat (i.e. non-dispersive) Additive White Gaussian Noise (AWGN) channel.

From information theory, the maximum achievable Shannon capacity [2] over this channel for a given $E_s/N_0$ can be calculated for an optimum transmit signal, which is unconstrained of any signal constellation but has a continuous Gaussian amplitude distribution:

$$C_{\text{unconstr}} = \log_2\left(1 + \frac{E_s}{N_0}\right). \tag{1}$$

This and the following channel capacities assume perfect synchronization and channel estimation. The (energy/bandwidth) overhead for pilots/preambles and a guard interval is not taken into account here.

For a specific signal constellation $\mathcal{X}$, the Shannon capacity can be calculated as follows:

$$C_{\mathcal{X}} = \sum_{x \in \mathcal{X}} \int_{y \in \mathbb{C}} p_{X;Y}(x;y) \log_2 \left( \frac{p_{X;Y}(x;y)}{\Pr(X = x) \cdot p_Y(y)} \right) \mathrm{d}y \tag{2}$$

$p_{X;Y}(x;y)$ is the probability density function (pdf) of the transmitted symbol $X = x$ and the received symbol value $y$, while $\Pr(X = x)$ is the probability of the transmitted symbol $X$ taking the value $x$ (in general all values from the constellation X have identical probability), and $p_Y(y)$ is the pdf of the received symbol value $y$. The above capacity assumes optimum signal processing in the receiver. More precisely, this means that the signal has to be processed symbol-wise in the receiver, and the channel code has to work on these received symbols instead of (soft-)bits. Such a symbol processing is in fact carried out by schemes like (Turbo) Trellis Coded Modulation (T)TCM [3], but in most standards (including DVB-SH and -NGH), a Bit-Interleaved Coded Modulation (BICM) scheme was selected [4]. In this scheme, the channel code is binary, and the bits $B_i$ are then forwarded to the mapper, where each symbol is mapped from a vector of bits. In the receiver, the opposite is done, i.e. the demapper calculates a vector of soft-bits $B_{i,\mathrm{est}}$ from each received complex symbol, and after the demapper, the bits inside a vector are processed separately by the turbo or LDPC decoder. By decomposing a received complex symbol into $> 2$ soft-bits, there are necessarily statistical dependencies between these soft-bits. That means, the probability that two bits $B_i$, $B_j$, $i \neq j$, from the same symbol $X$ were a '1' in the transmitter is in general different from the product of the individual probabilities that each of the two bits was a '1' in the transmitter, which is reflected in the associated soft-bits $B_{i,\mathrm{est}}$ and $B_{j,\mathrm{est}}$. A feedforward BICM receiver, as is currently the state of the art, takes the demapper output (soft-)bit-wise and decodes these soft-bits in a binary turbo/LDPC decoder without paying attention to the statistical dependencies. However, the statistical dependencies between the received code bits carries information, and by ignoring this information, information theory predicts a loss of channel capacity.

In principle, such a loss can be avoided in a BICM scheme. For this purpose, the receiver has to carry out iterations between demapping and decoding. By this method, the statistical dependencies are implicitly taken into account, as the demapper is revisited several times during the processing of the received signal. Hence, such an iterative BICM receiver can be considered similar to optimum symbol-wise processing. There exists a third scheme, that can reach the signal constellation's capacity: Multi-Level Codes (MLC) with Multi-Stage Decoding (MSD) in the receiver [5]. MSD also involves iterations between demapper and decoder, and we will not further investigate this scheme as it is different from the BICM scheme selected for, e.g., DVB-SH and -NGH.

For the considered BICM scheme with a receiver processing the symbols Bit-Wise and Non-Iterative (BWNI), the Shannon capacity can be calculated as follows:

$$C_{X,\mathrm{BWNI}} = \sum_{i=0}^{\log_2 \|X\|-1} \sum_{b_i \in \{0;1\}} \int_{y \in \mathbb{C}} p_{i;B_i;Y}(b_i; y) \cdot \log_2 \left( \frac{p_{i;B_i;Y}(b_i; y)}{\Pr(B_i = b_i) \cdot p_Y(y)} \right) \mathrm{d}y \ . \tag{3}$$

In this equation, $||\mathcal{X}||$ is the cardinality of the set of constellation points (e.g. 8 for 8-PSK), and $p_{i;Bi;Y}(b_i; y)$ is the probability density function that the $i$-th bit of the bit vector mapped into the signal constellation $\mathcal{X}$ is $b_i$, while all other bits of this bit vector can be arbitrary, and that the received complex symbol is $y$. This pdf can be calculated as follows:

$$p_{i;B_i;Y}(b_i; y) = \sum_{x \in \mathcal{X}_{i;b_i}} p_{X;Y}(x; y) \tag{4}$$

with $\mathcal{X}_{i;bi}$ being the subset of $\mathcal{X}$, where the associated bit vectors have the value $b_i$ in the $i$-th bit. The difference between the capacities $C_{\mathcal{X}}$ and $C_{\mathcal{X},\mathrm{BWNI}}$ is caused from ignoring the statistical dependencies between the bits.

Figure 2 displays the Shannon capacity curves for the optimum unconstrained transmitted signal and for the three different signal constellations of the DVB-SH TDM mode for optimum (symbol-wise) processing of the received signal. Moreover, the maximum achievable capacities for a BWNI receiver are displayed. Note that for QPSK, there is no loss for a BWNI receiver, as the two bits from each symbol are in different components (I and Q) of the complex symbol and hence there are no statistical dependencies.

Figure 3 shows the loss of the different signal constellations with respect to an unconstrained transmit signal. If a receiver could use optimum symbol-wise processing, we see that the loss from 16-APSK to the theoretically best signal constellation is small, not exceeding 0.4 dB for spectral efficiencies $\Gamma < 2$ bit/s/Hz. We find that the loss of the BWNI power efficiency compared to that for optimum symbol-wise processing of the same signal constellation (e.g. 16-APSK opt. vs. 16-APSK BWNI) depends on the desired spectral efficiency. For the values relevant within DVB-SH/NGH ($\leq 8/3$bit/s/Hz), the loss is between 0.9 and 0.1 dB for 16-APSK (note that for NGH, 16-QAM is used instead of 16-APSK). For dimensioning a system for the first-generation chipsets, the BWNI capacity should be taken as the benchmark. It appears therefore for a pure LOS channel, that a QPSK signal constellation is optimum for spectral efficiencies up to 1.5 bit/s/Hz, while a 16-APSK/QAM constellation is better for higher spectral efficiencies. 8-PSK is best only for a very small area around 1.5 bit/s/Hz, such that it is possibly not needed at all. Therefore, in the DVB-NGH standard, only QPSK and 16-QAM were included. However, for second generation chipsets, more elaborated algorithms applying iterations between the demapper and the decoder might be employed and the full capacity of the signal constellation may be exploited. For this future scenario, use of the 16-APSK (or 16-QAM) constellation proposes itself over the full range of

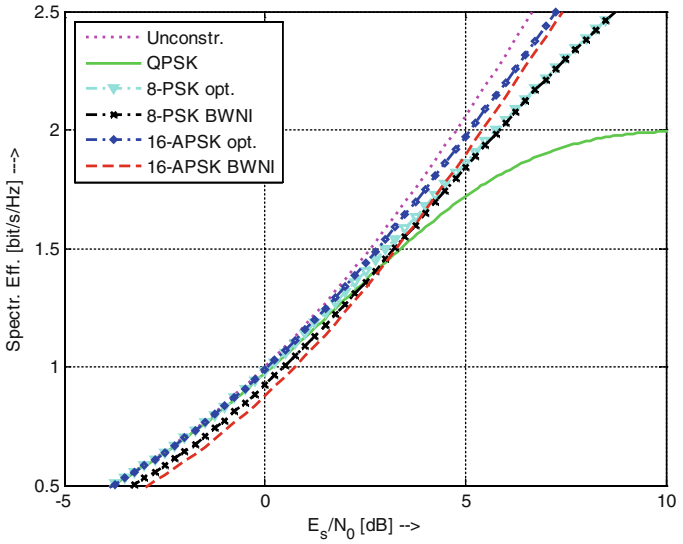



**Fig. 2** Capacity curves for the SC transmission mode

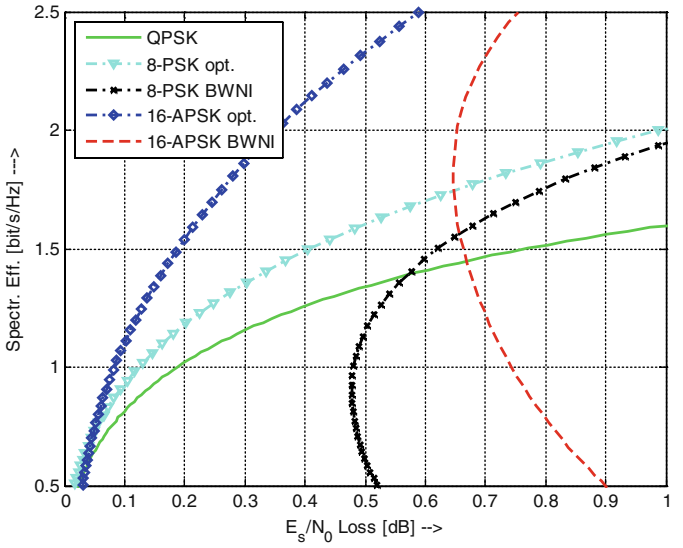



**Fig. 3** Loss of the signal constellation with respect to the optimum unconstrained transmit signal

spectral efficiencies achievable by this constellation with the DVB-SH/NGH code rates, i.e. for $4/5$ bit/s/Hz and above.

In practice, the shown power efficiencies cannot quite be met. In simulation, turbo and LDPC codes typically operate at $E_s/N_0$-values approx. 1 dB above the theoretical thresholds. This difference is due to several causes:

- the employed turbo code or LDPC code, resp., has an information word length of several thousand bits, while information theory always assumes infinitely long information words
- turbo codes and LDPC codes are only nearly-random codes, such that they have a different (sub-optimum) weight spectrum than the random codes used in the derivation of the channel capacity
- the iterative turbo/LDPC decoder is a sub-optimum approximation of the optimum non-iterative Maximum Likelihood Sequence Estimation (MLSE) decoder

All these effects add up to the difference of approx. 1 dB to information theory. Additionally an implementation loss has to be taken into account, for instance caused by the quantization inside the decoder. It should be noted that the 16-APSK/QAM constellation has a higher Peak to Average Power Ratio (PAPR) than QPSK and 8-PSK constellations due to the presence of two different amplitudes in the signal constellations. Therefore, the transmitter has to provide a larger input and output power backoff for 16-APSK/QAM than for the other two constellations, which reduces the received $E_s/N_0$. Finally, observe that for higher code rates like $2/3$ of the DVB-SH turbo code, an error floor is observed for BERs below $10^{-5}$.

Figure 4 shows the theoretical Shannon capacities of the LOS channel for the two signal constellations available for the OFDM mode. It can be seen that similar conclusions can be drawn as for the SC mode, namely that for first-generation BWNI receivers, there is a break-even point around $\Gamma = 1.5\,\mathrm{bit/s/Hz}$, below which QPSK is favourable while above, a 16-QAM signal constellation should be used in a LOS channel. Note that the 16-APSK signal constellation defined in DVB-SH has for higher spectral efficiencies a slightly lower channel capacity than a 16-QAM con-

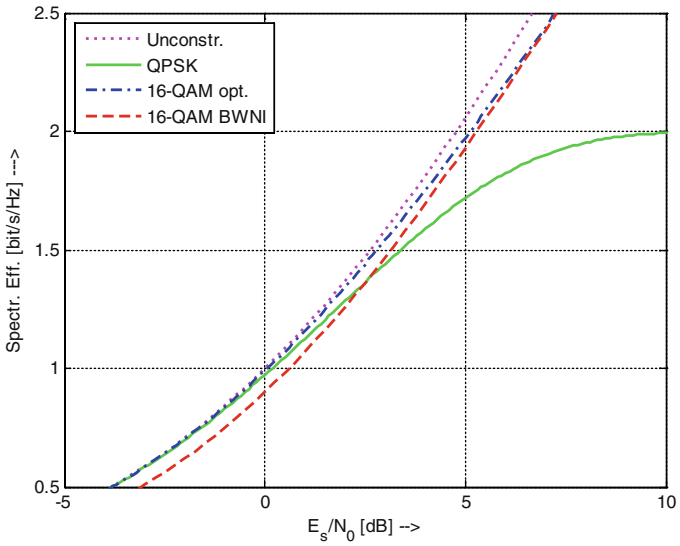



**Fig. 4** Capacity curves for the OFDM transmission mode

stellation. This is due to the fact that the Euclidian distance between the 12 points of the APSK's outer ring (i.e. $6/\sqrt{7} \cdot \sin(\pi/12) \approx 0.587$) is slightly smaller than the smallest Euclidian distance inside a 16-QAM constellation (i.e. $2/\sqrt{10} \approx 0.632$). For instance, for a spectral efficiency of 3 bit/s/Hz, the loss is approx. 0.2 dB.

In the sequel we will assume that the single-carrier (SC) mode, that was specifically designed for this task, is employed for the satellite component. However, for DVB-SH, the conclusions can be readily adapted to the OFDM mode as the 16-APSK constellation has very similar properties to the 16-QAM constellation. In DVB-NGH, both OFDM and SC use 16-QAM, such that the performance figures are identical as long as we neglect the effects of increased Peak-to-Average Power Ratio (PAPR) of an OFDM signal.

## 4 Perfectly Interleaved Satellite Channel

Next we will consider a channel that is more relevant for the dimensioning of a broadcasting system to mobile receivers: the Land-Mobile Satellite (LMS) channel. In fact, various such channels exist for the different environments like suburban, rural etc., and the channel capacity depends very much on the individual environment and the parameters chosen for modelling the specific channel. For an information theoretical analysis, it is useful to define a more abstract LMS channel model: the perfectly-interleaved On-Off channel. This model comprises two channel states. In the "On" state, the receiver experiences LOS reception conditions with a given $E_{\mathrm{s,LOS}}/N_0$, while the "Off" state corresponds to complete signal blockage, where no signal power is received at all. Any transition between these states is immediate without any intermediate channel states. A signal blockage hence erases the affected symbols. As usual in information theory, it is assumed that the codeword has infinite length and hence the interleaver has infinite duration. "Perfectly interleaved" means in this context that each transmitted symbol (PSK or QAM) is affected by this selective channel randomly and independently from the other symbols. A given erasure rate $R_{\mathrm{er}}$ means that this percentage of all transmitted symbols is erased by the channel, while the rate $1 - R_{\mathrm{er}}$ is received under LOS conditions. This channel is commonly refered to as an AWGN Erasure Channel (AWGN-EC) in the literature with given $E_{\mathrm{s,LOS}}/N_0$ and $R_{\mathrm{er}}$ [6].

The capacity of the AWGN-EC can be derived very simply from the capacity $C_{\mathrm{AWGN}}$ of the AWGN channel for the same $E_{\mathrm{s}}/N_0$: In the "On"-state (with relative frequency $1 - R_{\mathrm{er}}$), the capacity is the same as $C_{\mathrm{AWGN}}$, while it is zero for the "Off"-state. Hence, the overall capacity is

$$C_{\mathrm{AWGN\text{-}EC}}(E_{\mathrm{s}}/N_0) = C_{\mathrm{AWGN}}(E_{\mathrm{s}}/N_0) \cdot (1 - R_{\mathrm{er}}) \; . \tag{5}$$

Let us for instance assume that one target channel, where the system must be able to operate, has an erasure rate $R_{\mathrm{er}} = 1/3$. Therefore, the capacity curves are scaled down by a factor $2/3$ with respect to the AWGN channel capacity curves. For instance, if a spectral efficiency of 1 bit/s/Hz is targeted for this channel,

a $\Gamma = 1.5$ bit/s/Hz must be headed for in the LOS channel capacity of Fig. 2, such that an $E_{s,LOS}/N_0$ of approx. 3.3 dB must be available (plus implementation losses). We find for this AWGN-EC channel with $R_{er} = 1/3$ that we have reached the region, where QPSK is inferior to 16-APSK/QAM (and 8-PSK) with a BWNI receiver.

Figure 5 displays the acceptable $R_{er}$ for varying $E_{s,LOS}/N_0$ for the QPSK, 16-APSK/QAM (with a BWNI receiver) and unconstrained signal constellations and for several spectral efficiencies $\Gamma$. The curves are gained as follows: for a given $E_{s,LOS}/N_0$ and a given signal constellation, the channel capacity of the LOS channel $C_{AWGN}$ is calculated. Then the max. tolerable $R_{er}$ is determined by

$$R_{er} = 1 - \Gamma / C_{AWGN}(E_s/N_0) \tag{6}$$

such that after the scaling down of the capacity described above, $C_{AWGN\text{-}EC}$ becomes identical to the target $\Gamma$. For instance, $\Gamma = 1$ bit/s/Hz can be achieved by QPSK with code rate $1/2$ or 16-APSK/QAM with code rate $1/4$. The curves show that the QPSK curve runs very close to the unconstrained curve for low $E_{s,LOS}/N_0 < 3$ dB, but for higher $E_{s,LOS}/N_0$ values, only the 16-APSK/QAM BWNI curve can almost follow the curve of the optimum unconstrained signal constellation. The cross-over point at 3 dB is the $E_{s,LOS}/N_0$ value, where the curves cross in Fig. 2. That means, for higher $E_{s,LOS}/N_0$ the gap between the unconstrained curve and the QPSK curve opens up, while it closes down for the 16-APSK/QAM BWNI curve. The reason for this effect is the significantly increased constellation diversity that is present inside a 16-APSK/QAM constellation. Exemplarily for $E_{s,LOS}/N_0 = 5$ dB Fig. 2 shows that a 16-APSK/QAM can collect much more instantaneous capacity from the channel, when it is in the "On"-state than a QPSK constellation. The perfect interleaving averages out the time-variant channel, capacity collected in the "On"-state fills the capacity holes in the "Off"-state, such that this extra capacity from the LOS conditions increases the average capacity of the AWGN-EC and the robustness against blockages. Viewed from the channel coding perspective, for 16-APSK/QAM the codewords comprise twice the number of code bits than for QPSK, such that erasing a certain percentage of the code bits by the channel is better tolerable and enough bits remain to ensure the decoding.

Besides the information-theoretical curves, Fig. 5 also shows simulation results for DVB-SH. The figure shows for each examined $\Gamma$ the simulated acceptable $R_{er}$ for varying $E_s/N_0$ both for QPSK (asterisks) and 16-APSK (circles). Observe that the the simulation results are always on the right-hand side of the corresponding theoretical curve. We find that for low erasure rates, the difference between theory and the simulated performance is approx. 1 dB, which is quite in line with the loss that is usually observed for the static flat LOS channel, too. Only for higher erasure rates, the difference becomes more pronounced.

Note that in this and the following simulations, perfect synchronization and channel estimation is assumed, and the energy of the pilots/preambles and the guard interval is not taken into account. On the other hand, two DVB-SH-specific effects are taken into account in the simulations:

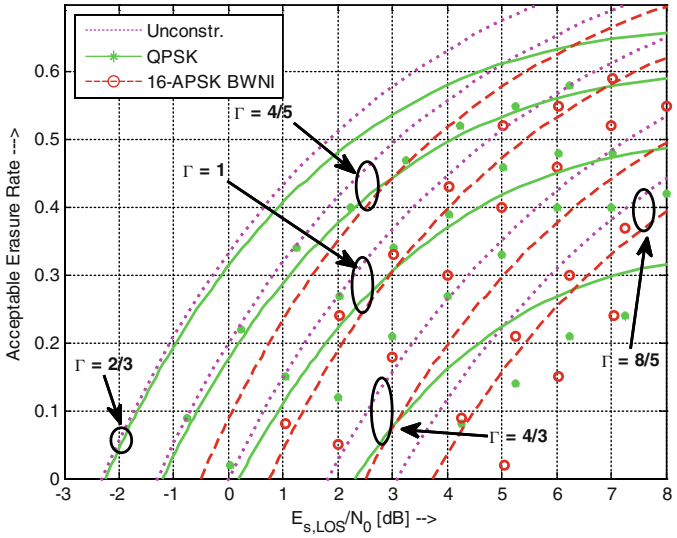



**Fig. 5** Acceptable erasure rates (theoretical and simulated) for the AWGN-EC channel for QPSK and 16-APSK BWNI transmission. *Lines* are information-theoretic curves, *asterisks* and *small circles* are simulation results for QPSK and 16-APSK BWNI, respectively. They are approx. 1 dB right for the theoretical curves. The *big circles* embrace a pair or triple of curves (unconstr., QPSK and 16-APSK) for spectral efficiencies $\Gamma = 2/3, 4/5, 1, 4/3$, and $8/5$

- the rate adaptation procedure as described in the DVB-SH standard is included, i.e. 2 out of the 128 bits of each Interleaver Unit (IU) are pruned away
- the structure of the DVB-SH channel interleaver with the temporal distribution of the IUs into at most 48 clusters (one for each delay line of the convolutional interleaver) is also taken into account
- In the presence of large-scale fading, which is the dominant effect in LMS channels, the interleaver is therefore able to disperse the codeword only over (at most) 48 different time instants, which may individually suffer from independent large-scale fading. Note that in practice, the correlation between the fading in these time instants depends on the configured interleaver profile and the velocity of the terminal. It is therefore reasonable to model an LMS channel by a perfectly interleaved On-Off-channel with perfect block-wise-interleaving, i.e. not the symbols fade individually, but complete blocks of IUs (i.e. 128 code bits minus two pruned bits), which are transmitted around the same time instant. For instance for code rate $1/4$ with a codeword length of 49 152, each such block contains 384 IUs/48 clusters = 8 IUs/cluster.

Moreover, the simulations do of course take into account the aforementioned properties and losses of the DVB-SH turbo codec being a real-world code structure and decoder. In the turbo decoding process, 8 iterations were used, and a fixed-

point decoder was employed that loses 0.1 to 0.2 dB with respect to a floating point decoder.

Examining the break-even points, where the theoretical curve for QPSK crosses that of 16-APSK BWNI, it appears that a good rule of thumb for the corresponding $R_{\text{er}}$ is $R_{\text{er}} = 1 - 2R_{\text{code}}/1.44$, where $R_{\text{code}}$ is the turbo code rate of the QPSK scheme, as the QPSK and the 16-APSK BWNI curves cross at a spectral efficiency of 1.44 bit/s/Hz in Fig. 3. The simulation results show additionally that only approx. 85% of this theoretical value $R_{\text{er}}$ is acceptable.

Extrapolation of the simulation results to the performance of an implemented system has to take into account the implementation loss. It is higher for 16-APSK/QAM than for QPSK mainly due to the presence of non-linearities at the transmitter side (e.g. satellite Travelling Wave Tube Amplifiers – TWTA) and possibly also in the receiver's tuner, as 16-APSK/QAM needs higher dynamics for the amplitude and has a higher PAPR. A second impact on the implementation loss is the demapping algorithm, that can be implemented optimally for QPSK but only sub-optimally for 16-APSK/QAM in order to save hardware complexity. Note that a 16-QAM demapper has a significantly lower complexity than one for 16-APSK, as the I and the Q-component can be processed completely independently from each other for the former. The non-linear distortion introduced by the satellite power amplifier can be reduced, however at the price of a larger amplifier back-off, such that the satellite's Effective Isotropic Radiated Power (EIRP) is reduced. This would on the one hand reduce the implementation loss, but on the other hand lower the link margin. A trade-off between both effects can be made by evaluating the total degradation [7]. Apart from these effects, phase noise should for soft-decoding not increase the implementation loss markedly for 16-APSK/QAM compared to QPSK, as it can be modelled as (non-Gaussian) noise and is hence already accounted for in the capacity calculations and the simulations.

## 5 Constrained Time-Interleaving

Having analysed the favourable pairing of code rate and modulation order for infinite time interleaving, let us now proceed to LMS channels with limited time interleaver duration. Models exist for the simulation of such channels [8], but in this chapter, we have the chance to exploit the results of field measurements including all real-world effects that cannot be represented in a simple channel model. The measurement campaign was conducted in October 2009 along the east coast of the United States of America. Over a traveled distance of approximately 4300 km, amongst others, the power levels of a high-power S-band satellite (ICO-G1) were recorded with a sample rate of 2.16 kHz. The elevation angle of the satellite varied between 29 and 46 during the measurement campaign. More details about the measurement campaign and the analysis can be found in [9]. The measurement route covered highways, rural, suburban and urban areas as well as areas with trees.

- Urban environments cover typical city centers with buildings with at least three floors.
- Suburban environments cover residential areas of cities and greater villages. The building height is limited to three floors, also foliage on both sides of the roads is usual.
- Commercial areas include business parks and industrial parks with wide roads (often four lanes) in the outskirts of a city. The buildings aside the roads are often flat and the foliage is light.
- Highways, freeway and interstates are characterized by wide highspeed roads with at least two lanes. Vegetation aside the road is usual.
- Intermediate tree shadowed rural areas (ITS) include two-lane roads in the countryside with a mixture of line-of-sight conditions and sections with groups of trees beside the road.
- Forested rural areas include two-lane roads in the countryside with forests aside the road.

Especially the urban and suburban environments resemble very much the scenario that we considered by the AWGN-Erasure Channel in the previous section: the buildings in these environments block the signal very strongly (i.e. erasure), while in the remaining time, the receiver experiences a Line-Of-Sight channel.

For the analysis of the recorded time series, the instantaneous SNR C/N for each recorded sample is calculated, where each sample consists of two values: the useful signal power C and the in-band noise power N (in fact, the noise power spectral density was measured not measured in-band, but in a vicinate unoccupied band).

Then the effect of time interleaving at the physical layer was emulated by averaging the instantaneous SNR over a sliding time window of a given duration. This window represents a uniform convolutional interleaver of the same duration. A successful trasmission is assumed, whenever the average SNR meets or exceeds a target SNR. This approach is independent of the choice of code rate and modulation order. Besides the time diversity gain that is discussed in this chapter, [9] also investigates the gains from antenna diversity.

Note that the results might differ strongly for time interleaving at a higher layer, to which the SNR method exhibited above cannot be applied. For such a comparison, the instantaneous capacity has to be calculated and averaged over the interleaver length instead. The less an environment resembles an on-off-channel, the larger average capacities disagree for time interleaving on the physical and higher layers, because the physical layer does not output soft-information about the code symbols of the higher layer channel code.

Signal power fluctuations are caused by wave propagation effects. Thus, they are fully dependent on the location of the receiver and not of the time instant of reception. Therefore, the time interleaver length is given as a spatial distance in the sequel rather than a time interval to enable an analysis without considering the user's moving velocity.

Figure 6 shows the required margin for various interleaver lengths in the different environments. The margin is defined as the difference between the available $E_s/N_0$

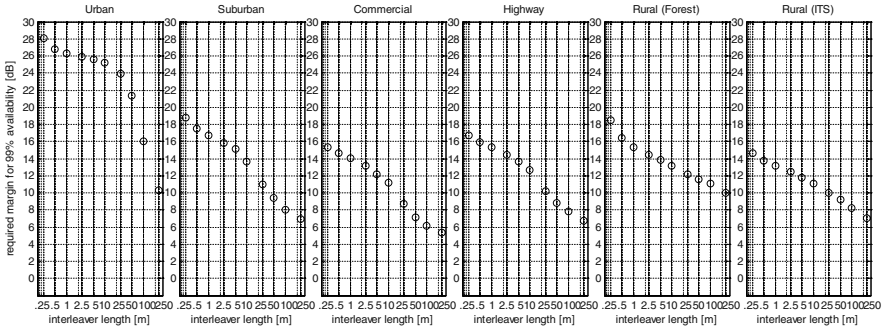



**Fig. 6** Required margin for a signal availability of 99% for various interleaver lengths

during the LOS reception instants of the time-variant channel, and the minimum $E_s/N_0$ value that is required for successful reception under constant LOS conditions. The figure was calculated assuming that the average $E_s/N_0$ has to exceed the minimum required $E_s/N_0$ for at least 99% of the time. That means, providing the required margin ensures that the signal can be decoded successfully 99% of the time. This approach is independent of the target spectral efficiency of the system. On the other hand, when the alternative approach is chosen, where capacities are calculated and averaged instead of SNRs, the required margins depend somwhat on the target spectral efficiency, i.e. code rate and modulation order.

A typical margin for a satellite broadcast system is approximately 10dB. It can be seen in the figure that for such a margin, 99% signal availability can hardly be achieved in the urban environment. Therefore, in urban areas supplementary terrestrial repeaters might be used to meet the quality-of-service requirements. Also the Forest environment could fail to meet the 99% availability requirement for 10 dB margin. In other environments, lower margins are sufficient.

Table 1 shows for each environment the required spatial interleaver length for 10 dB margin and 99% availability, the minimum typical vehicle speed and the resulting required temporal interleaver duration.

While long interleavers are impractial for bi-directional communication systems, they offer a great potential for broadcasting systems where time delays of several seconds are feasible.

**Table 1** Generic model of the considered transmission chain

| Environment | Req. interleaver length [m] | Min. vehicle speed [km/h] | Req. interleaver duration [s] |
|---|---|---|---|
| *Urban* | $> 250$ | 30 | $> 30$ |
| *Suburban* | 50 | 30 | 6 |
| *Commercial* | 25 | 30 | 3 |
| *Highway* | 25 | 100 | 0.9 |
| *Forest* | 250 | 50 | 18 |
| *ITS* | 25 | 50 | 1.8 |

It is found from Table 1 that the urban and the forest environments require unrealistically long time interleaving. In the urban case, a high service availability can be achieved using terrestrial repeaters. In the forest environment, the consumer will have to live with signal dropouts from time to time. The most critical environment among the others is the suburban one. The interleaver should be at least 6s for this environment. Some time margin should be included to cope also with the implementation loss of a real system and for cases, when the vehicle speed sometimes drops below the values in the table. Therefore, as a rule of thumb, 10 s time interleaving seems to be a reasonable choice. This value has also been verified in field trials for DVB-SH.

## 6 Conclusion

We have studied the selection of the optimum pair of code rate and modulation order separately from the optimum interleaver duration. In all studied cases we have found that the trade-off between using a lower modulation order (QPSK) or a higher one (16-APSK/QAM) is a question of the target system properties, mainly the target spectral efficiency and required tolerable blockage rate. It was found that 8-PSK modulation can be omitted in a standard without compromising the system's performance significantly.

For dimensioning of the time interleaver, we provided results from field measurements for various environments. For a link margin of 10 dB, we found that an adequate interleaver length is approx. 10 s. Most environments can be covered appropriately by such a system, except for the urban and the sheer forest environments.

The Shannon capacity was used for the analyses, and it seems well suited to predict the performance of the simulations, and hence this theoretic measure can probably also be used for a performance prediction of a fully implemented system in real-world environments.

## References

1. ETSI EN 302 583 V1.1.0 (2008-01) Framing structure, channel coding and modulation for Satellite Services to Handheld devices (SH) below 3 GHz
2. Cover T, Joy T (2006) Elements of Information Theory, Wiley, ISBN 978-0-471-24195-9
3. Robertson P, Wörz T (1998) Bandwidth-Efficient Turbo Trellis-Coded Modulation Using Punctured Component Codes. IEEE Journal on Selected Areas in Communications 16(2):206–218
4. Caire G, Taricco G, Biglieri E (1998) Bit-interleaved coded modulation, IEEE Transactions on Information Theory, 44(3):927–946
5. Wachsmann U, Fischer R, Huber J (1999) Multilevel codes: theoretical concepts and practical design rules. IEEE Transactions on Information Theory, 45(5):1361–1391
6. Sahai A, Xu Q (2004) The Anytime Reliability of the AWGN+erasure channel with Feedback

7. Thompson S, Proakis J, Zeidler J (2005) The Effectiveness of Signal Clipping for PAPR and Total Degradation Reduction in OFDM Systems. IEEE Conference on Global Communications (GlobeCom), pp 2807–2811
8. Pérez Fontán F, Vázquez-Castro M, Enjamio Cabado C, Pita García J, Kubista E (2001) Statistical Modeling of the LMS Channel. IEEE Transactions on Vehicular Technology, 50(6):1549–1567
9. Arndt D, Ihlow A, Heuberger A, Eberlein E (2011) Measurement-Based Evaluation of Antenna- and Time Diversity for Mobile Satellite Systems. IEEE Proceedings of the 11th International Conference on Telecommunications, Graz, Austria

# Enhancements in DVB-H and DVB-SH Based Mobile-TV Multiplexing

**Christian Forster and Nikolaus Färber**

**Abstract** Energy efficient broadcasting of mobile-TV content through DVB-H and DVB-SH suffers from signaling limitations introduced by the DVB-H and DVB-SH multiplexing scheme resulting in an increased system delay. This paper presents a novel scheduling and multiplexing approach to overcome these limitations. The proposed algorithm allows to trade time slicing efficiency for lower system delay and more flexible data rate assignments possibly resulting in better video quality. The algorithm is evaluated through extensive cross layer simulations for different parameter sets.

## 1 Introduction

The idea of mobile television is almost as old as color-TV itself (e.g. [13]). But it was not until the introduction of dedicated broadcast networks, especially designed for mobile reception, that a significant population of mobile-TV receivers appeared to be within grasp. One of the driving technologies behind mobile-TV is "Digital Video Broadcasting for Handheld Terminals" (DVB-H) [4], which is considered to have a high market potential by the European Union [2, 10]. Although political and strategic decisions of the market players have seriously slowed down the introduction of the technology, DVB-H and corresponding successors such as "Digital Video Broadcasting for Satellite Services to Handheld Devices" (DVB-SH) [7] are available to ensure that broadcast data is distributed to portable devices economically and on a large scale.

As most recent examples [14] and [11] have demonstrated, that standard telecommunication networks (GSM, UMTS) have already been stretched to their limits due

Christian Forster (✉)
Chair of Information Technologies, University of Erlangen-Nürnberg, Am Wolfsmantel 33, 91058 Erlangen, Germany
forster@like.eei.uni-erlangen.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_13, © Springer 2011

to the growing popularity of smartphones [15], it is safe to assume that in the near future we will need those specialized distribution grids even more than before.

DVB-H and DVB-SH are meant to be received on devices that suffer from a limited energy supply – namely a battery –, therefore they both support a special time slicing method [8] to reduce the power consumption of the receiver. This enables the receiver – based on its parameter setting – to reduce its front-end power consumption in the range of 90%. These high values however, lead to a certain inflexibility that can be observed as additional system latency when variable data rate service are transmitted. Nevertheless a common reason for using variable service rates is the possibility to significantly improve the video- and audio-quality of the broadcast services through statistical multiplexing and joint encoding.

This optimization is especially important for DVB-H and DVB-SH, because both transmission methods broadcast a large number of mobile-TV services packed together in one bit stream (systems multiplex). Compared to other methods such as "Digital Multimedia Broadcasting" (DMB) [6], this significantly increases the optimization potential [1]. For video services a good overview of the algorithmic principles can be found in [12].

However due to the design of DVB-H and DVB-SH an additional system delay is introduced when variable rate services are being multiplexed together. Especially for time-critical services[1] this might be undesirable. This paper provides a new approach for multiplexing variable bit rate services and limiting at the same time additional latency. To achieve this goal, a backward compatible modification of the multiplexing method is proposed. The advantages of this method come from its ability to trade energy efficiency of the time slicing method for the reduction of additional system latency and data rate flexibility. Because of its primary application area in improved video coding, we will also examine the method presented here in a cross-layer approach with regard to the extent, to which the video quality of a mobile-TV broadcast can be improved by using the additional data rate flexiblity.

Furthermore, this paper is structured as follows: In Sect. 2 we briefly present the time slicing method of DVB-H and DVB-SH. Also, we will elucidate the limitations of the time slicing method that is usually used. Based on that, in Sect. 3 we introduce the proposed modification of the time slicing method. Finally, we discuss the results of detailed simulations (Sect. 4) in Sect. 5.

## 2 Time Sliced Mobile-TV and Its Limitations

DVB-H and DVB-SH support a time slicing method to broadcast media streams energy-efficiently. This method groups all datagrams of each service into separate bursts, which are then transmitted at a higher rate. As each burst only contains the data of a single service (see Fig. 1), receivers may deactivate their front-end while data for other services is transmitted, yielding energy savings of more than 90% [3].

[1] For example, live broadcast of sports events or time-critical test- and control-information.

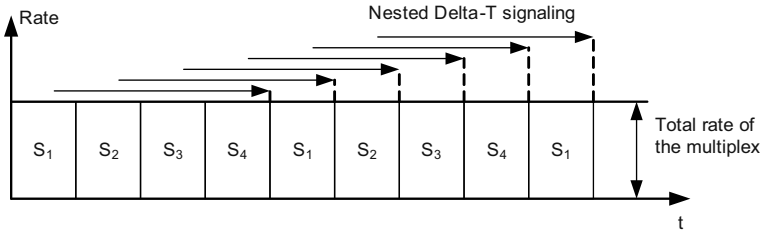



**Fig. 1** DVB-H multiplex with four services

As the sequence of individual bursts neither follows a specific pattern, nor does it have to be cyclic, this time slicing method leads to the problem of when to turn on a receiver's front-end, so that all bursts belonging to one service can be successfully received. For that, DVB-H and DVB-SH use a Delta-T signaling method: Each burst contains the relative waiting time until the next burst belonging to the same service starts. Using this information, receivers can calculate the exact front-end activation time.

This method works for constant rate services without further side effects. But if services are supposed to be transmitted with variable bit rate, it must be taken into account that the nested structure of the Delta-T signaling requires ahead of time scheduling of future burst sizes. For the well-known and often used round-robin burst scheduling pattern (Fig. 1), we can therefore assume that the size of every burst has to be known at least one cycle in advance.

To solve this problem, variable rate streams are usually delayed by one cycle, so that the data that has to be transmitted is already available at the point of the Delta-T signaling time, making it possible to directly determine the individual burst sizes. This, however, leads to an additional increase in system latency.

## 3 Dynamic Modifications of the Multiplex Structure

The problems mentioned in the previous section regarding the transmission of variable rate services can be minimized through the method that is proposed in this section. We first present the basic concept with the help of an example.

### *3.1 Concept*

The basic problem in broadcasting variable rate services via DVB-H and DVB-SH is based in the fact that the required data rates have to be known in advance. The core of the proposed solution is based on the definition of a so-called *flexible multiplex*. This refers to a multiplex bit stream that can change the data rate of individual DVB-H

or DVB-SH services in place and within certain limits, although the corresponding Delta-T values have already been signaled to the receivers. Maximum efficiency can be achieved with a two-fold approach:

1. First, the burst sizes (durations) are predicted as well as possible through a conventional scheduling-algorithm. Those values are then signaled by the means of the Delta-T values. Although the prediction can be achieved with any scheduling algorithm, an AR(1)-based scheduler was selected for the simulation (Sects. 4 and 5).
2. In a second step, the bit stream will be adjusted to match the final data rates of the services by coping with unexpected peaks. This step will be performed by modifying the multiplex, yielding a flexible multiplex bit stream.

These modifications are based on extending bursts beyond their original end, so that they use parts of the transmission time that has originally been reserved for the subsequent burst. In doing so, all the bursts that follow an extended burst are delayed by the period of time, by which the first burst was extended. This continues an entire cycle until a fresh burst is scheduled by the service that originally created the extended burst. This fresh burst is then shortened to the extent needed to compensate the shifting. Through this method, transmission capacity that will be available in the future can already be used in advance. In addition the scheme makes sure that the long term average data rates of all participating services are not altered by the proposed method.

Let's shed some light on this method using a simple example: At first a sequence of bursts originating from the services $S_1$, $S_2$, $S_3$ and $S_4$ is assumed (see Fig. 1). If we now extend the first burst of the service $S_1$, the bursts of $S_2$, $S_3$ and $S_4$ will shift accordingly. To compensate the offset the second burst of $S_1$ will finally be shortened, and that compensates for the resulting shift. Figure 2 shows the resulting scenario.

If, on the other hand, different services extend several bursts, the total result is determined through superposition of individual processes. In addition a burst can also be shortened if it needs a lower transmission capacity than the one allocated to it. However, this is only possible if the burst has already been delayed in advance through the flexible multiplex modification.

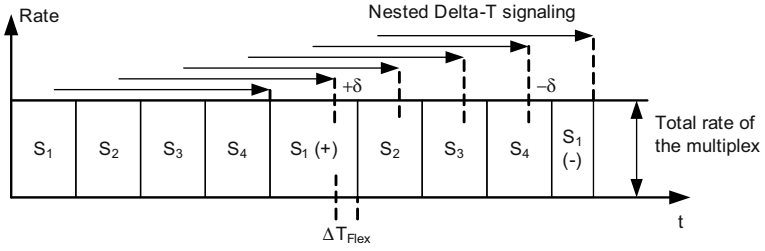



**Fig. 2** Example of a shift in a flexible DVB-H multiplex.

On the receiver side the flexible multiplex modification has the following impact: As the size and the position of the bursts is altered after the Delta-T signaling has been sent the receivers will activate their front-ends at the time that was initially signaled. Therefore, they first receive the additional data of the preceding burst until the transmission of the requested burst begins. Thus receivers will need a minor amount of additional power as their front-end will stay enabled a little bit longer. Nevertheless the ability to decode the data is sustained. As stronger modifications will result in longer receiver on-times, the proposed method allows a network operator to find a trade-off between these key figures.

The flexible multiplex method presented here is topically related to the context of the concept of the "bit-savings bank": In a bit-savings bank, the unused transmission bits are "saved" and used for cushioning of future peak data rates. However, in case of flexible multiplex modification this option cannot be used, because the Delta-T time data has already signaled and determined the earliest limits for the start of every burst. The method therefore in the "bit-savings bank" jargon should rather be referred to as a "bit-overdraft credit", because those additional bits that are added to a burst are exclusively compensated for by future bursts.

## *3.2 Rule Based Algorithm*

When executing the multiplex modification described in Sect. 3.1, narrow limits have to be taken into account, because otherwise signals cannot be received or played properly on receiving terminals: Let $\delta[b]$ be the change in duration of the bursts $b$. Then one would have to make sure that

$$\Delta T_{\text{Flex}}[b] = \sum_{j=1}^{b-1} \delta[j] \geq 0 \tag{1}$$

applies to every $b$. $\Delta T_{\text{Flex}}[b]$ is referring to the delay in start of the burst $b$ compared to the point of time signaled via the Delta-T method. Violating this condition, will shift some burst before the signaled starting point and thus receivers are not yet prepared for the beginning of data transmission.

In addition to this constraint it is equally important that the delay of bursts does not become too long, because receivers can otherwise experience buffer underflows that can interrupt the playback of the service. Owing to the special structure of DVB-H and DVB-SH burst signaling information, it is sufficient to make sure that the *start* of a burst does not exceed a certain maximum delay $\Delta T_{\text{Flex, max}}[b]$:[1]

$$\Delta T_{\text{Flex}}[b] \leq \Delta T_{\text{Flex, max}}[b] \forall b \tag{2}$$

[1] This condition results from the weakly protected end of burst signaling, in addition to the fact that error-correction data (MPE-FEC) are always transmitted at the end of a burst. We therefore must proceed from the assumption that receivers (under unfavorable receiving conditions) for every burst expect at first the maximum burst duration before they start to process the received bursts.

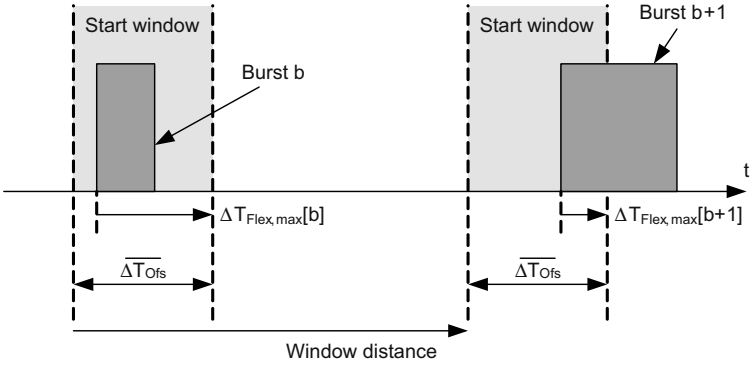



**Fig. 3** Example for a recurring start window associated to a service. Bursts must start within the window region to avoid media buffer under- and overrun

The extent of the $\Delta T_{\text{Flex, max}}[b]$ value for every burst $b$ results from its pre-scheduled positioning by the multiplex-scheduler and the requirements of the application protocol. For media data-streams this means that a recurring window has to be assigned to every service, during which a burst has to start, so that the buffer does not experience any underflow or overflow conditions (Fig. 3). With respect to the figure the size of this start window will be specified by $\overline{\Delta T_{\text{Ofs}}}$. The $\Delta T_{\text{Flex, max}}[b]$ value can then be determined from the distance of the scheduled burst $b$ to the right margin of the window.

An algorithm for determining the flexible multiplex modification can be derived from the mentioned constraints. It will be presented as a rule base algorithm that will be executed burst by burst:

1. If the burst $b$ is extended by the duration $\delta[b]$, the follow-up burst of the same service has to be scheduled using a value big enough to compensate for the delay.
2. Equation 1 must not be violated. The lower limit for $\delta[b]$ that results from this has to be observed.
3. Equation 2 must also not be violated at any point of time. This applies especially to future bursts. This means that a burst extension $\delta[b]$ not only becomes limited by $\Delta T_{\text{Flex, max}}[b]$, but also to a greater extent by the maximum value of all bursts that are shifted because of the extension.
4. If a burst is shortened ($\delta[b] < 0$), this reduction is proportionally distributed to all the preceding bursts that have executed an extension, which was not yet compensated for. This prevents overcompensations for executed burst extensions.

## 4 Simulation

As the main scope of application of the proposed method is within the field of low latency mobile-TV broadcasting, detailed simulations have been conducted focusing on achievable video coding performance. Table 1 lists the properties of the used video streams ("Capability B" from [5]).

A regular Round-Robin method was selected as the burst-scheduling-algorithm. The corresponding burst values are determined for every burst based on a modified version of the Lagrange Rate-Distortion (R-D) optimization. The modified optimization differs from the plain Lagrange R-D optimization ( [12]) in the fact that it performs burst by burst optimization avoiding additional latency. The outcome of the optimization is used two times: First it is used for signaling of future burst size of the service (Delta-T signaling), and second, it is also the input to the flexible multiplex modification algorithm for the current burst. Where applicable, Sect. 5 will also show the isolated effect of the modified R-D optimization disregarding the benefit from the flexible multiplex modification.

The video material used for the simulation was taken from a test recording of the DVB-H pilot-broadcasting-project in Erlangen [9]. This is a recording of twelve re-encoded TV network programs broadcast over a period of 24 hours, which provides for a well representative mix.

**Table 1** Video format for encoding the test sequences

| Property | Value |
|---|---|
| Video format | H.264/AVC, Baseline, QVGA, 12.5 fps |
| Average bit rate | 256 kbps |
| Duration of the GOP | 2 s |
| Average cycle duration | 2 s |

## 5 Simulation Results

The following presents the results of the simulation and their interpretation. It deals with the three areas of image quality, latency and energy efficiency.

### *5.1 Image Quality*

Simulation results show that the method presented here significantly improves the image quality of the encoded video streams. Figure 4 shows the improvement of the average Y-PSNR over the selected burst start window size $\overline{\Delta T_{\mathrm{Ofs}}}$. The improvement is measured in comparison to constant rate encoding of the same sequence. As a first

reference the chart shows the performance of the modified R-D optimization without the proposed flexible multiplex modifications. It is noticeable that only a small Y-PSNR gain (approx. 0.5 dB) can be observed. This is due to the fact that future data rates must be predicted to a certain degree by the modified R-D optimization. However by also allowing flexible burst modifications the image quality raises further as prediction mismatches will be compensated at the time when the burst is transmitted. As a second reference Fig. 4 also contains the possible image improve-

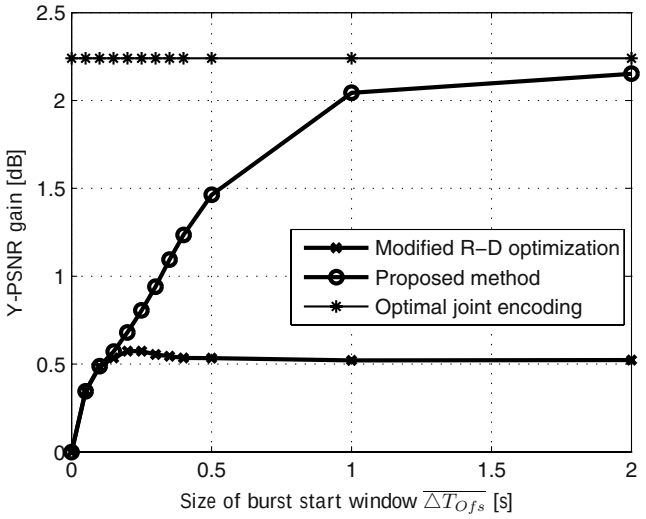



**Fig. 4** Achieved Y-PSNR gain for different burst start windows $\overline{\Delta T_{\text{Ofs}}}$

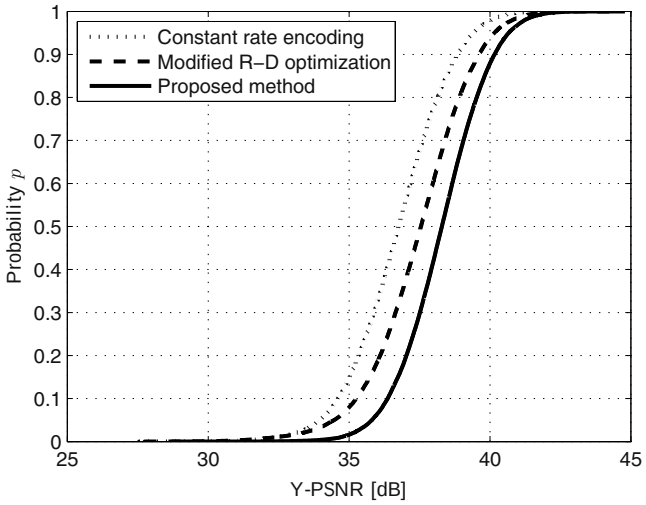



**Fig. 5** Cumulative probability density function of the average Y-PSNR

ment that will be achieved when all constraints are neglected (optimal encoding, graph marked with asterisk symbols). It becomes apparent that the quality improvement of the proposed method converges to that optimum quality by extending the burst start windows.

By evaluating the simulation data in more detail it becomes apparent that the proposed method can make use of its advantages especially in the range of below-average Y-PSNR values. As an example Fig. 5 shows the cumulative distribution density function of the achieved Y-PSNR values for the different multiplex variants. Within the range of $p = 0.1$ average improvement compared to a constant rate distribution amounts to 1.76 dB, whereas for $p = 0.9$ this "only" comes to 1.32 dB.

## *5.2 Latency Saving*

The achieved latency reduction heavily depends on the basic system latency introduced by the other processing stages of the transmission chain. This relationship is demonstrated in Fig. 6. Two effects determine the achievable latency reduction:

- The basic system latency $T_{\text{Base}}$. The value is determined by external effects like the time required for video en- and decoding and the basic transmission delay.[1]
- The size of the burst start windows $\overline{\Delta T_{\text{Ofs}}}$. The bigger this window, the longer the required buffers. Thus decreasing the effective latency saving.

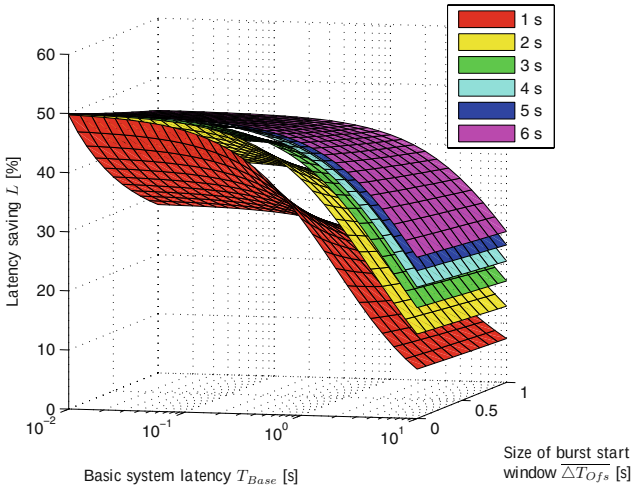



**Fig. 6** Saving on latency for various average cycle durations (shown in *different colors*)

[1] Especially the single frequency network synchronization time can make a significant contribution to this delay.

The percentage of latency reduction was calculated by the following formula:

$$L = 1 - \frac{T_{\text{Flex}}}{T_{\text{Lat}}} \cdot 100\% \tag{3}$$

$T_{\text{Flex}}$ refers to the system latency that can be achieved through the proposed method, whereas $T_{\text{Lat}}$ refers to the latency of a legacy system. For instance a transmission delay $T_{\text{Base}}$ of 500 ms, a cycle time of 2 s and an optimization interval $\overline{\Delta T_{\text{Ofs}}}$ of also 500 ms, will result in an end-to-end delay $T_{\text{Lat}}$ of 10.5 s for a legacy transmission. In the optimized case, on the other hand, only 6.5 s ($T_{\text{Flex}}$) are needed. This corresponds to a latency reduction of 38%.

## 5.3 *Efficiency of the Time Slicing Method*

The simulation of different parameter combinations showed the effects of the proposed modification on the efficiency of the time slicing method. Section 3.1 pointed out that the flexible multiplex modification results in longer on-times at the receiver resulting in a reduced time slicing efficiency. For example, time slicing efficiency drops by approx. 4.2% if a start window $\overline{\Delta T_{\text{Ofs}}}$ of 500 ms is selected. On the other hand, this loss of time slicing efficiency can be matched to the gain in Y-PSNR improvement (1.63 dB). Figure 7 shows all resulting combinations in a single graph. From the chart we notice that the curve can be divided in two parts: The significant rise at the beginning of the curve is due to the modified R-D optimization (including the rate prediction). As this method already achieves a certain optimization gain

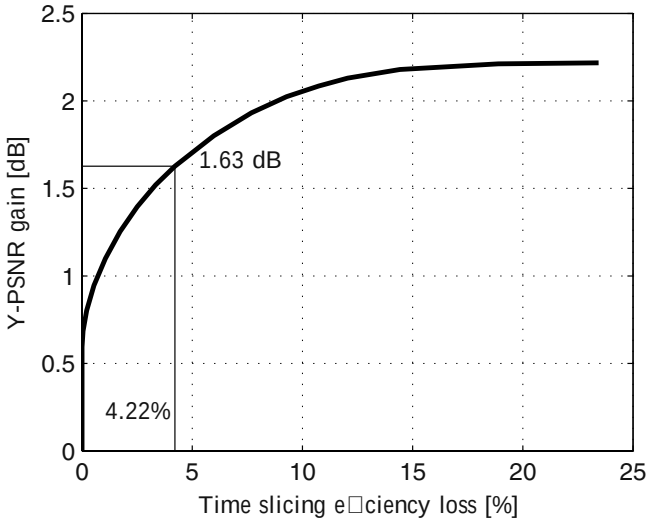



**Fig. 7** Y-PSNR gain over the loss in time slicing efficiency

even without the flexible modification scheme, the curve rises almost vertically in the diagram. To achieve further optimization gain the burst sizes must be modified thus the time slicing efficiency starts to decrease. This leads to the second part of the curve.

## 6 Conclusion and Summary

In this paper the objective was to exhibit how to optimize the transmission and joint encoding of several variable rate mobile-TV streams via DVB-H or DVB-SH. The main optimization constraints were low latency transmission, high time slicing efficiency, and equal treatment of all data streams.

The proposed cross layer method allows for choosing a trade-off point defined by image quality gain, latency savings, and time slicing efficiency. Extensive simulation were used to quantify the presented relations. The simulation data was extracted from a typical mobile-TV bouquet. By selecting a typical start window size of 500 ms an average image improvement of 1.63 dB Y-PSNR can be achieved by sacrificing 4.22% of the time slicing efficiency. Compared to a legacy transmission system the end-to-end delay is also reduced by 38% (based on the configuration of the broadcasting network).

## Acknowledgements

This contribution is based on the project “DVB-H Pilot Transmitter Erlangen”. This project has particularly been supported by the Fraunhofer Institute IIS, not only financially, but also in terms of organization and staff. We, the authors, would hereby like to express our gratitude for this support. We also would like to thank our colleagues from the Chair of Information Technologies, who have always been there for us with help and advice.

## References

1. Chandra K (2003) Statistical Multiplexing. In Proakis JG (ed) Wiley Encyclopedia of Telecommunications, vol 5, pp. 2420–2432, Wiley, Hoboken, NJ. www.mrw.interscience.wiley.com/eot
2. Europäische Union (2007) An EU Strategy for Mobile TV – Frequently Asked Questions. Brüssel http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/07/298
3. European Telecommunications Standards Institute (ETSI) (2004) EN-301-192 – V1.4.1 – Digital Video Broadcasting (DVB); DVB Specification for Data Broadcasting
4. European Telecommunications Standards Institute (ETSI) (2004) EN-302-304 – V1.1.1 – Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals (DVB-H)

5. European Telecommunications Standards Institute (ETSI) (2006) TS-102-005 – V1.2.1 – Digital Video Broadcasting (DVB); Specification for the Use of Video and Audio Coding in DVB Services Delivered Directly Over IP Protocols
6. European Telecommunications Standards Institute (ETSI) (2006) EN-300-401 – V1.4.1 – Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers
7. European Telecommunications Standards Institute (ETSI) (2007) Digital Video Broadcasting (DVB); System Specifications for Satellite services to Handheld devices (SH) below 3 GHz
8. Faria G, Henriksson JA, Stare E, Talmola P (2006) DVB-H: Digital Broadcast Services to Handheld Devices. Proceedings of the IEEE 94(1):194–209. doi:10.1109/JPROC.2005.861011
9. Forster C DVB-H Versuchssender Erlangen. http://www.like.e-technik.uni-erlangen.de/dvb-h
10. Gozalvez J (2008) The European Union Backs the DVB-H Standard. IEEE Vehicular Technology Magazine 3(2):3. doi:10.1109/MVT.2008.923967
11. Jurran N (2009) Handy-TV: Telekom beschränkt Bundesliga-Übertragung auf Konferenzschaltung. http://heise.de/-751305
12. Ortega A, Ramchandran K (1998) Rate-Distortion Methods for Image and Video Compression. IEEE Signal Processing Magazine, pp 23–50. doi:10.1109/79.733495
13. Sony (1964) Werbung "This is the smallest TV set in the world". LIFE 57(12):R2
14. Windeck C (2009) Holpriger Handy-TV-Start in die Bundesliga. http://heise.de/-750491
15. Wortham J (2009) Customers Angered as iPhones Overload AT&T. The New York Times (New York Edition), B1

# SafeTRIP – Interactive Satellite Services for Automotive Applications and Road Safety

**Bernhard Niemann, Thomas Heyn, Aharon Vargas, Sabino Titomanlio, Guillermo Grau, and Ashweeni Beeharee**

**Abstract** The acronym SafeTRIP stands for "Satellite Application for Emergency handling, Traffic alerts, Road safety and Incident Prevention". Being a European (FP7) project, it brings together 20 partners from seven countries having the common goal of improving road safety, mobility, and environmental protection for all kinds of passenger vehicles. The general objective of the SafeTRIP project is to improve the use of road transport infrastructures and to improve the alert chain (information/prevention/intervention) in case of incidents by offering an integrated system from data collection to service provision. SafeTRIP directly contributes to the achievement of the European Commission's objectives regarding road transport safety and road mortality reduction.

The SafeTRIP system basically is an Intelligent Transport System (ITS) aiming at the combination of digital broadcasting technology with satellite and complementary communication channels to enable novel bi-directional service offerings to the automotive user with pan-European coverage. The satellite will provide seamless coverage for mobile reception whereas the terrestrial components are mainly used to increase the high system capacity in areas with many users or it will increase the quality of service (QoS) in areas with limited satellite coverage. The SafeTRIP project focuses on the test of the core technologies for different applications and usage scenarios. An open architecture is created allowing the development of innovative applications by leveraging a set of enabling services. The demonstrator will support communication via a real satellite, corresponding ground stations or other 3G/4G infrastructure, an on-board unit and end-to-end applications to evaluate the benefits of the architecture and the applications during a trial phase.

Bernhard Niemann (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
bernhard.niemann@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_14, © Springer 2011

## 1 Intelligent Transport Systems (ITS)

Intelligent Transport Systems make use of information and communication technology in order to improve one or several aspects of transportation systems. There are many different flavors of ITSs depending on the traffic system for which they are employed and the objective they try to achieve. In the context of this article we concentrate on ITSs for road traffic.

With a growing amount of traffic, on-going urbanization and the increasing popularity of electro mobility, Intelligent Transport Systems are a crucial factor in making road traffic reliable, safe, and sustainable. These systems usually comprise a server infrastructure in the service centers, a communication sub-system and an on-board unit with one or several modems and general purpose computer resources. Typical communication scenarios for road traffic ITSs are:

- Vehicles sending messages to the road-side infrastructure or a centralized service center.
- Vehicles receiving messages from the road-side infrastructure or a centralized service center.
- Vehicles directly exchanging messages between each other.
- Connection-oriented, bi-directional, real-time communication between vehicles and a centralized service center.

The first three scenarios are employed to implement data-centric applications like vehicle tracking, traffic measurement using floating-car data, traffic alerts, etc. The last scenario is of great importance for emergency notification systems like the European eCall [1].

ITS systems are subject of several research projects. Examples are the SPITS (Strategic Platform for Intelligent Traffic Systems) project [2] and the SISTER (Satcoms In Support of Transport on European Roads) project [3]. Both projects are typical examples of ITS research involving many partners with different areas of expertise.

The SPITS project envisions a platform which is composed of on-board units in the vehicle, roadside units for the road infrastructure and back offices for the application management. Communication relies on existing 3G infrastructure and additional 802.11p-based Car2X communication, which is the base-line for many current ITS projects.

The SISTER project was concerned with researching the potential benefits of satellite communications for ITS systems. A proof-of-concept demonstrator has been developed showcasing a combination of several existing satellite communication technologies.

First commercial deployments are based on terrestrial communication systems and do not benefit from the seamless coverage of satellite-based systems or the high capacity for data casting of broadcasting systems. Efficient communication technologies, including bi-directional (real-time) communication, data casting (unicast, multicast and broadcast), message-oriented information collection from many users

and direct car-to-car links (e.g. based on IEEE 802.11p) are essential core technologies of ITS. Broadcasting technology is not mandatory for ITS, but many ITS applications will benefit from broadcast-like network topologies. Mobile entertainment is a standard application in most cars anyway. Combining systems originally optimized for multimedia applications with ITS may provide a cost efficient implementation. To fully benefit from ITS systems, many cars should be equipped with such a system. Offering an attractive package of applications will also help to increase the acceptance rate of ITS.

## 2 Broadcasting Systems

A variety of digital broadcasting systems for mobile reception without a return channel is currently in use all over the world. They can be categorized regarding their distribution system (e.g. terrestrial, cable, satellite) and regarding the content being distributed (e.g. audio, video, data services).

Examples for terrestrial broadcasting systems include DAB [4] and DVB-T [5] with the first being focused on audio content with some additional data services while the latter is used for mainly fixed television.

Especially for vehicular reception, the combination of a terrestrial network with one or several satellites is an interesting concept. These so-called hybrid broadcast systems employ a Satellite Component (SC) in combination with a Complementary Ground Component (CGC) to achieve the highest possible territorial coverage at moderate expenses for the terrestrial network. An example of a hybrid broadcast system is SiriusXM Satellite Radio in the US.

## 3 SafeTRIP System Concept

SafeTRIP benefits from the assignment of the S-band for MSS applications and a new satellite W2A that has been successfully launched in 2009. Opening new perspectives for European telecommunications, the S-band is well suited for terminals combining terrestrial 3G technologies with satellite-based applications, allowing small and cheap terminal antennas. Satellite technology will be used for positioning, broadcasting, messaging communication and connection-oriented bidirectional communication services (e.g. voice). In contrast to conventional terrestrial communication technologies like GSM or 3G cellular networks, satellite technology for communication has the advantage of large territorial coverage without the necessity to deploy a ground network. State of the art satellite communication systems for mobile users provide already high service availability for mobile applications. Hence, in principle, a satellite-only system is feasible. Combining the satellite system with ground networks gives more flexibility for the trade-off between system capacity and system cost. Using, for example, ground components

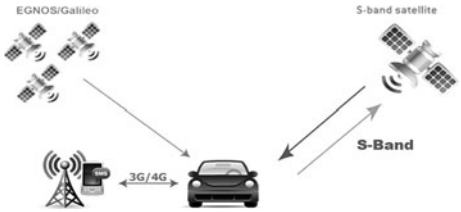



**Fig. 1** SafeTRIP system concept

for densely populated areas, significantly increases the system capacity at moderate cost. Targeting urban areas from the satellite would require high link margins, resulting in high satellite costs. Covering urban areas with terrestrial repeaters/collectors allows the selection of air-interface parameters providing link margins sufficient for rural and suburban areas, resulting in a much higher capacity of the satellite signals.

Accordingly, the SafeTRIP system will integrate several communication systems. To avoid the need for the user and also the application developer to deal with a variety of different communication channels, each with its own characteristics, the SafeTRIP concept includes a smart middleware to make the selection of the communication channel transparent to the applications. The middleware is split in a server-side implementation (Service Enabling Platform, SEP) and an on-board unit (OBU) (client)-side implementation (Service Enabling Layer, SEL). The most important functionality of the middleware is to make the applications and services in the OBU running independently from the physical communication channel.

The SafeTRIP system basically aims at combining hybrid broadcasting technology for mobile users with novel satellite-based messaging-oriented M2M (machine-to-machine) communication technologies and complementary state of the art communication system offerings to the automotive user. Also, it provides pan-European coverage and will be suitable for various applications. Within the project, the core technologies will be integrated to a demonstrator system serving as basis for application development. Main focus of the project is the development of applications benefiting from this powerful and flexible communication infrastructure. This includes broadcasting applications like map updates, traffic information, weather forecast and news, unicast information embedded in the forward link multiplex, collecting of message-oriented status information (e.g. traffic status, alerts) and individual two-way communication, including emergency communication.

## *3.1 SafeTRIP System Architecture*

The SafeTRIP system architecture includes a satellite-based broadcasting system with extensions for two-way communication and complementary ground components for areas with limited satellite coverage (Fig. 2).

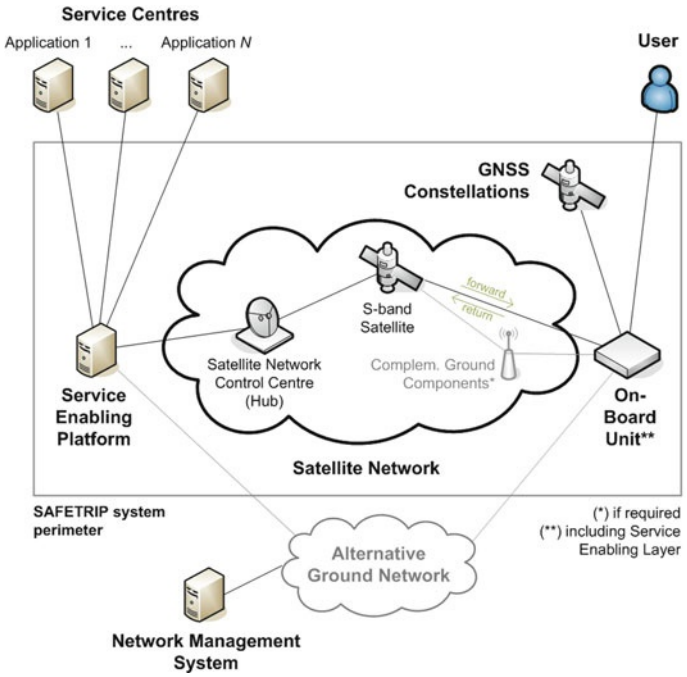



**Fig. 2** SafeTRIP system overview

The SafeTRIP infrastructure is composed of central items (Hub, Network Control Centre) and distributed items (service centers). The SEP provides a well-defined interface for the applications to the different communication channels of the SafeTRIP architecture. On the user side, the onboard items are aggregated into an on-board unit (OBU). A middleware, software running both onboard and centrally, provides an abstraction layer to the services, making the SafeTRIP architecture:

- Capable of integrating alternative communication technologies such as 3G.
- Open to the integration of third party services.

In the SafeTRIP reference architecture, low-price on-board units installed in vehicles provide personalised services, e.g. emergency calls and messages, traffic alerts, incident/accident warning etc.

## *3.2 Satellite Segment and Related Ground Network*

For satellite communication, the S-band spectrum in the range from 2170–2200 MHz for the downlink (towards user terminal) and 1980–2010 MHz for the uplink (from user terminal) will be used. The same band is used for the satellite link and the related terrestrial components. This part of the S-band has exclusively

been reserved for the use of Mobile Satellite Services (MSS) by the European Commission according to the decision 626/2008/EC. For the demonstration phase of the SafeTRIP project, the Eutelsat W2A satellite will be used.

### *3.3 On-Board Unit*

The on-board unit (OBU) is divided in three main parts: the Communication modules, the Middleware (Service Enabling Layer SEL) and the application parts (Fig. 3). Main focus of SafeTRIP is the communication module for S-Band ("S-Band Terminal" in the block diagram). To support different communication modules, a hardware abstraction layer (HAL) is inserted.

The S-Band terminal performs the forward and return link processing comprising the messaging and connection-oriented return link, respectively. For reception of the S-band satellite signal, antennas with small form factors compliant to requirements from the car industry are sufficient.

The communication protocols are based on Internet Protocol (IP). The SEL processes the communication requests from the application layer performing the authentication, authorization, accounting tasks and necessary adaptation to the communication module. Besides, the SEL receives and processes the positioning information from the GNSS terminal (GPS or Galileo) as well as the alternative ground network interface (e.g. 3G technology). Finally, the application layer presents the IP information to the user and transforms the user requests into IP requests to the SEL. Besides the direct interaction with the user using an HMI (human machine

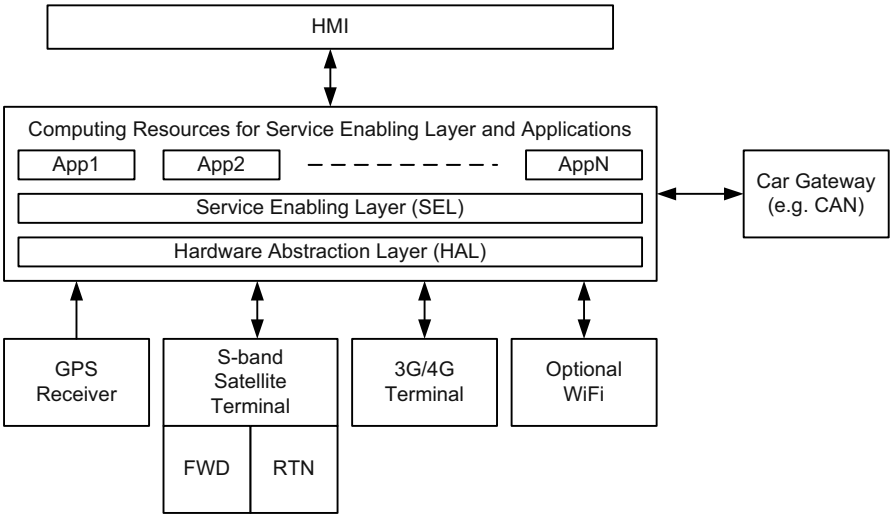



**Fig. 3** On-Board Unit (OBU) architecture

interface), the OBU includes also interfaces to the car infrastructure. For the OBU prototype, the CAN bus has been selected.

## *3.4 Service Enabling Platform*

The part of the middleware provided by the network operator is called "SEP" (*Service Enabling Platform*) (SEP) and provides a set of features independent from the communication channel which are offered to the service providers and application developers. The OBU part is called "SEL" (Service *Enabling Layer*). The SEP connects to the *Existing Service Providers* (ESP) and the *Value Added Service Providers* (VASP), providing a virtual communication channel over the S-band satellite or over *Alternative Ground Networks* (AGNs) (Fig. 4). AGNs are typically GPRS, UMTS, HSDPA or Wi-Fi.

In the OBU, the SEL works as a *mediator* between the Client Applications and the SafeTRIP network. It exposes a set of standardized interfaces of the enabling services to the applications running on the OBU. The SEL communicates with the SafeTRIP network via S-band or AGNs. Among the main enabling services are the following:

- *Non-Real-Time Messaging* provides a specific messaging protocol and can be adopted by all applications which need to exchange non-real-time messages between the service providers and the SafeTRIP end users.
- *Authentication, Authorization and Accounting* (AAA) provide *single-sign-on*, regardless to the used channel, *import*, *normalize* and *manage user profiles* to control access to services, *register raw packet access* for traffic-based billing.
- *Vertical Handover* (VHO): virtualizes the communication channel, switching from one network to another (i.e. from satellite to AGN) without impact on the running applications.

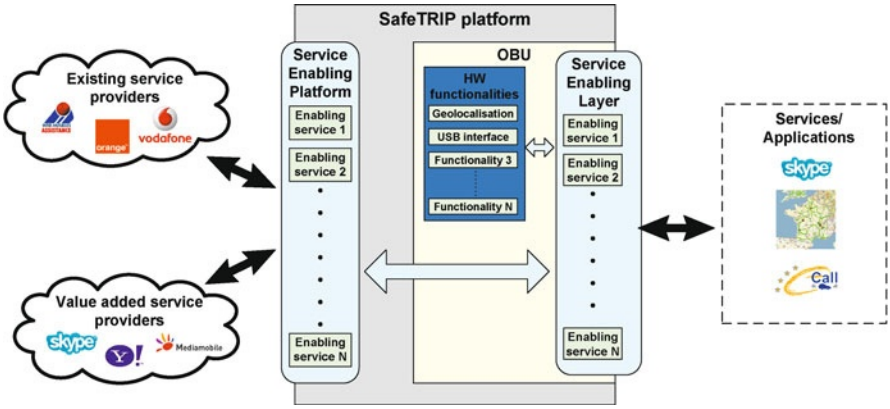



**Fig. 4** The role of Service Enabling Platform in SafeTRIP

- *Broadcasting and Data casting* broadcasts Electronic Service Guide of the services/programs via satellite channel. It provides also reliable file delivery with multicast protocols over the satellite channel, possibly with return link via AGN or S-band.

### *3.5 SafeTRIP Services*

The SafeTRIP communication concept follows the "always connected" principle in the sense that the user does not need to bother about the currently available type of connectivity. The Service Enabling Middleware seamlessly selects the most suitable connection. Novel integrated services can be implemented by the combination of ubiquitous connectivity with positioning and car sensor data. Professional services for emergency vehicles benefit from network redundancy (due to the combination of satellite and terrestrial communication) and the possibility to select the most appropriate communication scheme from a variety of possibilities.

The service provider can take advantage from pan-European coverage due to the S-band satellite while having the possibility to also leverage existing mobile network infrastructure. Novel service opportunities are offered:

- Satellite broadcasting can efficiently and seamlessly distribute content to a virtually unlimited number of users.
- Messaging directly from the OBU to the satellite
- Bi-directional communication via satellite for emergency services
- Immediate access to service by a large user base, supporting a strong business case for service providers
- Access to information about the vehicle – including its occupants or cargo
- Monitoring of driving behavior (including detection of alertness) from the OBU
- Aggregated sensor data can be used from a platoon of vehicles to infer traffic, road and weather conditions
- Personalization of traffic information as a large amount of broadcasted information can be processed by OBU and customized to fit the driver's needs
- Access to large data streams – such as infotainment through the OBU by passengers

These customer-oriented applications dedicated to car drivers & passengers, bus or coach drivers & passengers and road operators will be tested with a SafeTRIP demonstrator in the field.

## 4 Satellite Communication Technology in SafeTRIP

In SafeTRIP three communication schemes are combined. Table 1 gives an overview:

**Table 1** SafeTRIP system concept

| Communication scheme | Air interface forward link | Air interface return link | Connection type |
|---|---|---|---|
| Broadcast | DVB-SH | Not applicable | Continuous transmission, one-to-many |
| Messaging | DVB-SH-LL | E-SSA | Asynchronous burst, one-to-one |
| Connection oriented bi-directional | DVB-SH-LL | QS-CDMA | Continuous transmission, one-to-one |

The system architecture of the satellite communication system has mainly been defined in projects funded by the European Space Agency (ESA), namely the MiReSys [6] and DENISE [7] projects and is now subject to standardization within ETSI.

The fundamental idea is to combine a shared forward link using DVB-SH [8, 9] extended with a low-latency feature (DVB-SH-LL [10]) with two types of return links. The first type based on the Enhanced Spread Spectrum Aloha (E-SSA) standard [11] is optimized for asynchronous transmission of messages from millions of cars directly to the satellite using low-power transmitters in the OBU. The second type based on Quasi-Synchronous Code-Division Multiple-Access (QS-CDMA) [12] is intended for connection-oriented bi-directional communication.

This system architecture is particularly attractive because it offers the possibility to dynamically partition the bandwidth between regular latency transmission for broadcasting and low-latency transmission for signalling and bi-directional communication on an on-demand basis.

DVB-SH, being a hybrid (satellite and terrestrial) broadcasting system, provides the possibility to deploy a network of terrestrial repeaters to complement satellite reception in dense urban areas or other regions with limited line-of-sight to the satellite. This concept can be leveraged for the satellite return link as well by extending the concept of repeaters to collectors thus allowing return link transmission to either the satellite or the terrestrial collectors. This concept gives additional flexibility in network planning.

## *4.1 SafeTRIP Forward Link*

In SafeTRIP, the forward link is used for two main purposes, namely streaming of multi-media content and data casting. Most of the multi-media applications are broadcast-like applications or address several users (multicast).

Multi-media content is characterized by a huge amount of information, as for example, TV shows, films or sporting events, which require a continuous transmission scheme. For such information, latency does not play an important role, where a reception delay of several seconds is acceptable by a user, but interruptions in the

received stream are annoying. If latency is not very critical, time interleaving and other technologies for the mitigation of temporal signal outages offer a high gain and are enabling technologies for systems with many temporal short blockages. Especially for satellite-based systems, time interleaving is a very powerful technology to overcome deep fades caused by tunnels or high buildings leading to interruptions of the line-of-sight reception (LOS) for a couple of seconds. Accordingly, long time interleaving (or time diversity) is common for satellite-based systems targeting broadcast to mobile users.

In contrast to multi-media content, interactive applications, voice communications and time-critical control information demands a low-latency transmission. Typically, the average bitrate is low, allowing a transmission of the data as short bursts within a high bitrate multiplex. Obviously, long time interleaving cannot be applied in this case. Therefore, SafeTRIP requires a hybrid broadcasting scheme supporting the parallel transmission of non-real-time and low-latency information.

The DVB-SH standard [8, 9] proposes a transmission scheme based on satellite communications and mobile receivers, which matches exactly with the requirements of SafeTRIP for the broadcasting part. Figure 5 shows a simplified model of the physical layer of the DVB-SH system. At first, the mode and stream adaptation block adds a CRC to each MPEG packet to provide error detection capability. Besides, a header is added to signal the input stream features and to support other stream formats. Additional error protection is then provided by a turbo encoder as defined by 3GPP2 [13]. This is achieved by adding redundancy to the information stream. The output code words are then interleaved in order to spread the information over time. A configurable convolutional interleaver is defined, which allows a wide range of delays. The bits are then mapped onto channel symbols, and pilots (a known sequence of symbols) are inserted in the information stream. The receiver uses these pilots to estimate the channel conditions. Finally, a modulation step is required prior to the transmission of the information.

The interleaver scheme proposed by the DVB-SH standard offers a configurable delay, ranging from few hundreds of milliseconds to few dozens of seconds. However, only one static interleaver profile can be employed in a given transmission, i.e. either a low-latency or a non-real-time transmission can be supported. This is in conflict with the SafeTRIP requirement of simultaneous transmission of low-latency and non-real-time information, where two different interleaver profiles are

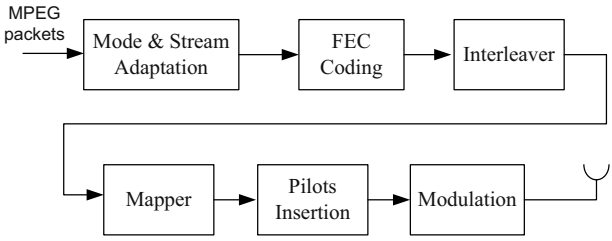



**Fig. 5** Block diagram of the DVB-SH broadcasting system

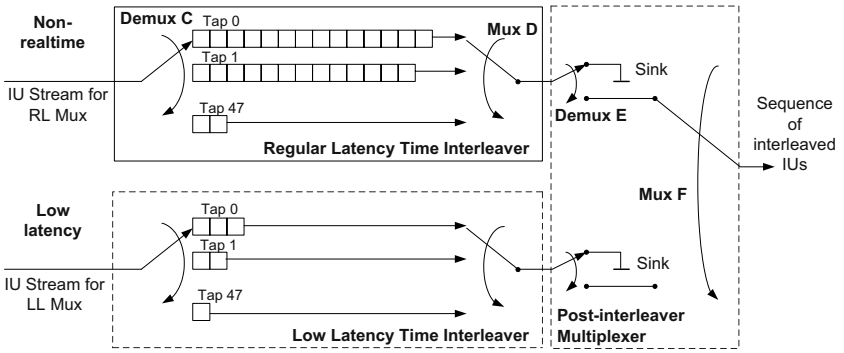



**Fig. 6** Extension of the DVB-SH standard to support simultaneous non-real-time and low-latency streams

desirable. Therefore, the interleaver concept of DVB-SH is extended in order to fulfil the SafeTRIP requirements, by adding an option for low-latency transmission simultaneously with non-real-time transmission.

Figure 6 shows the low-latency extension to the DVB-SH standard (DVB-SH-LL) [10]. The information is split into two streams, non-real-time and low-latency, depending on the delay requirements. The upper path corresponds to a regular DVB-SH transmission, where convolutional interleaving is employed to spread the information using 48 taps with different delays. Besides, dummy information is inserted in this interleaver, which is first interleaved by *Demux C* and later discarded by *Demux E* to create some gaps inside the non-real-time stream. These gaps are used to insert the low-latency information. The low-latency stream is generated using the bottom path of Fig. 6, where a very short time interleaving profile is used to assure low-latency transmission. Obviously, this interleaving profile must have some properties in order to fit into the non-real-time stream, which is achieved by applying a modulus operation to the tap lengths of the regular interleaver. At the output, both streams are time multiplexed to create the transmitted sequence. The main advantage of the low-latency extension is the backward compatibility with legacy DVB-SH receivers. A more detailed description of the molded interleaver can be found in [14].

## *4.2 SafeTRIP Message-Oriented Return Link*

For the return link, i.e. communication from user to provider, the E-SSA standard [11] is proposed, in combination with the return link encapsulation (RLE) standard [15]. As the E-SSA standard supports only a fixed frame length, an encapsulation is required to break down the information packets into suitable sizes for the transmission over the E-SSA communication channel.

Figure 7 shows the flow of the data information through the RLE module and the E-SSA modulator. The format of the input information is assumed to be IP data of

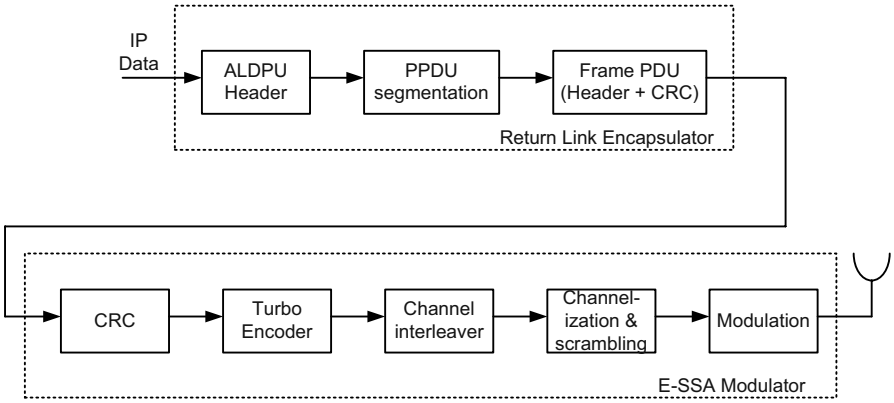



**Fig. 7** Block diagram of the E-SSA return link modulator

variable length. The RLE block slices the input data into packets of a fixed length. First, a header is prepended to the IP data containing an Automatic Repeat reQuest (ARQ) mechanism, together with information about the encapsulated IP data, like the protocol type or the priority of the packet. The resulting packet is segmented in a sequence of payload-adapted protocol data units (PPDUs), where the length is chosen according to the E-SSA standard requirements. Some extra information is prepended and appended to the PPDUs, as for example an error detection mechanism in order to discard wrong PPDUs at the receiver. The output protocol data unit (PDU) frames are then processed by the E-SSA modulator. An error detection control (CRC) and an error correction control (Turbo code) are applied to overcome errors introduced by the channel. An interleaver is applied to each code word to avoid dependencies between the bits involved in each channel symbol. A scrambling and spreading process is applied to the output of the interleaver prior to the modulation step.

## *4.3 Connection-Oriented Communication*

An asynchronous communication based on short messages is not enough to fulfill the requirements of some services supported by the SafeTRIP project. The asynchronous media access scheme and the related multiplexing introduces mainly a jitter for the round trip delay. This jitter combined with the latency introduced by the high distance earth to satellite exceeds the delay/latency requirements of some applications. For example, a voice call between a snow plough driver and the operator at the motorway operations center requires a bidirectional, continuous transmission with high data rate, which in principle is not supported by the E-SSA modulator. In other services, unicast or multicast video streaming from a patrol vehicle also demands connection-oriented communications layers for the return (vehicle-

to-infrastructure) link, as a complement to the infrastructure-to-vehicle broadcast services described above.

The requirements for these kinds of services are relatively demanding. Jitter, or delay variation, is important for all multimedia broadcast services, but voice calls in particular also require very low end-to-end latency, above which the perceived quality of the call quickly decreases. Considering the high round-trip delay inherent to satellite communications, special equipment is required to ensure this high quality of service.

These requirements are even more challenging for a mobile satellite system with a low-cost omnidirectional antenna with hemispheric pattern, considering the reduced available transmission power for the return link and the susceptibility of such low-directivity antennas to multipath and interference. In addition to this, spectral efficiency needs to be maximized in order to provide high capacity and an efficient use of satellite resources.

All these requirements make the design and implementation of a suitable waveform a crucial task. In order to accomplish this task, the SafeTRIP project builds upon and demonstrates the key results of the ESA DENISE studies [7], which include the design of a Quasi-Synchronous Code-Division Multiple-Access waveform (QS-CDMA). Spread-spectrum techniques are inherently robust against interference and provide good capacity and spectral efficiency, especially when using synchronous or quasi-synchronous variants of these techniques. Adequate power control techniques also provide high power efficiency.

Figure 8 shows an overview of the QS-CDMA return-link modulator. The stream of user and signalling data is passed to the physical layer, which generates packets prepended with some header information. Packets then are passed through the channel coding block and the output bits are mapped to channel symbols in the I- and Q-branches. Scrambling is then applied to improve the shape of the signal spectrum. Afterwards, a unique word is added at the beginning of each physical-layer data frame. The final stages of spreading, pulse shaping, and (QPSK) modulation generate the spread-spectrum signal prior to transmission.

The connection-oriented communication technology is currently standardized within ETSI [12].

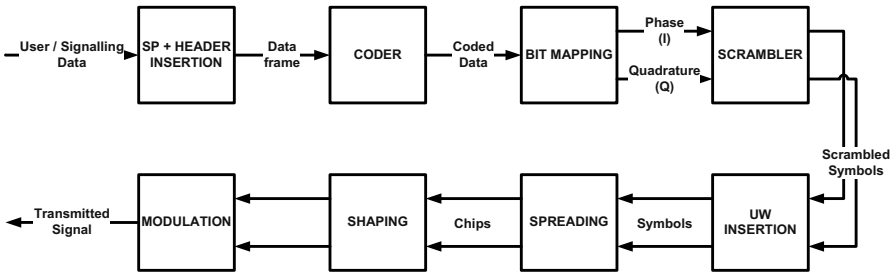



**Fig. 8** Block diagram of the QS-CDMA return link modulator

## 5 Conclusions

Integrating satellite-based systems with other communication technologies (e.g., 3G/4G) into a single communication platform offers the possibility for novel service concepts while making efficient use of scarce spectrum resources. The satellite provides pan-European coverage thus giving the user ubiquitous connectivity independent of ground infrastructure. The system combines efficient data distribution via broadcasting with very low overhead messaging and full bi-directional communication. It allows efficient implementation of a variety of different applications like multimedia entertainment, traffic management, and emergency communication. Using a common infrastructure for these types of services allows a flexible radio resource management thus ensuring an efficient use of the available spectrum. The platform developed within the project offers open interfaces for application developers. The SafeTRIP project covers the development of a demonstrator and the evaluation of the system performance. For field trials the existing W2A satellite will be used.

## Acknowledgements

SafeTRIP is a project co-funded by the European Commission in the 7$^{th}$ framework program.

## References

1. www.ec.europa.eu/ecall
2. SPITS project, Strategic Platform for Intelligent Transport Systems. www.spits-project.com
3. SISTER project, Satellite communications for road transport telematics applications. http://www.cvisproject.org/en/links/sister.htm
4. ETSI EN 300 401 V1.4.1, Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers
5. ETSI EN 300 744 V1.6.1, European Standard (Telecommunications series) Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television
6. MiReSyS project, S-Band Mission Requirements and Preliminary System Specifications Definition
7. DENISE project, Demonstrator Emergency aNd Interactive S-band sErvices, http://telecom.esa.int/telecom/www/object/index.cfm?fobjectid=30191
8. ETSI TS 102 585 V1.1.2, Technical Specification Digital Video Broadcasting (DVB); System Specifications for Satellite services to Handheld devices (SH) below 3 GHz
9. ETSI EN 302 583 V1.1.2 European Standard (Telecommunications series) Digital Video Broadcasting (DVB); Framing Structure, channel coding and modulation for Satellite Services to Handheld devices (SH) below 3 GHz
10. Digital Video Broadcasting (DVB) Bluebook A111, Framing Structure, channel coding and modulation for Satellite Services to Handheld devices (SH) below 3 GHz, May 2011

11. Draft ETSI TS 102 721-3, Satellite Earth Stations and Systems; Air Interface for S-band Mobile Interactive Multimedia (S-MIM); Part 3: Physical Layer Specification, Return Link Asynchronous Access
12. Draft ETSI TS 102 721-4, Technical Specification, Satellite Earth Stations and Systems; Air Interface for S-band Mobile Interactive Multimedia (S-MIM); Part 4: Physical Layer Specification, Return Link Synchronous Access
13. Third Generation Partnership Project 2 (3GPP2) (2004) Physical layer standard for CDMA2000 spread spectrum systems
14. Vargas A, Breiling M, Gerstacker W, Stadali H, Eberlein E, Heuberger A (2010) Adding Different Levels of QoS to the DVB–SH Standard. In Proceedings of the Advanced Satellite Multimedia Systems (ASMS) Conference, Sardegna, Italy
15. Draft ETSI TS 102 721-5 V0.1.2; Satellite Earth Stations and Systems; Air Interface for S-band Mobile Interactive Multimedia (S-MIM); Part 5: Protocol Specifications, Link Layer

# An On-Board Processor for In-Orbit Verification Based on a Multi-FPGA Platform

**Alexander Hofmann, Robért Glein, Bernd Kollmannthaler, and Rainer Wansch**

**Abstract** The increasing demand for higher data rates, smaller antenna apertures, or less power at the uplink for mobile devices requires air-interface and application specific processing, especially for telecommunication satellites. Only on satellites for dedicated applications or with a short limited lifetime, on-board processing is partly used, but processing on-board improves the system performance or increases the system capacity in several cases. Today's on-board processing for satellite communication is mostly based on ASIC (Application Specific Integrated Circuit) chips, which have their main drawback in the limited flexibility. In order to demonstrate and validate the flexibility of an FPGA-based on-board processor (OBP) for space applications the Fraunhofer IIS is involved in a so-called in-orbit verification (IOV) payload on the Heinrich-Hertz-Satellite. During the development of an on-board processor for space applications, the main challenge is to ensure a typical life time of 15 years for the hard-, firm- and software under the given environmental conditions. Alternatively, an FPGA platform can be reconfigured for novel communication protocols. In order to investigate new standards for telecommunication satellite systems, the Fraunhofer IIS is developing an OBP platform, based on four FPGAs. The OBP will be embedded into the H2Sat satellite, which will be launched in 2016 and will be located on a geostationary earth orbit (GEO). To the best of our knowledge, this is the first completely reconfigurable platform based on four leading-edge radiation-hardened FPGAs in the geosynchronous orbit for telecommunication satellites.

Alexander Hofmann (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
alexander.hofmann@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_15, © Springer 2011

# 1 Introduction

The H2Sat mission aims to explore and test new communication technologies in space at a technical and scientific level in order to determine how broadband communications, for example, can result in high data rates for a mobile end user [6, 7].

The satellite itself is based on the "SmallGEO" (SGEO) platform of OHB System AG which was developed especially for communication applications. This version offers a maximum of 400 kg and 3.6 kW. The payload includes two parts. One part of the satellite will be equipped with a commercial payload. The other part is a scientific payload, which is placed to perform IOV of different kinds of new payload subsystems developed in Germany. This is to extend Germany's capabilities in the area of satellite and telecommunication payloads.

Today most of the common GEO satellites for telecommunication systems are based on traditional bent pipes [4], which consist of filters, mixers and traveling wave tube amplifiers (TWTA). This "amplify and forward" configuration converts the signal to the downlink frequency and only compensates the effect of total signal attenuation. In most scenarios, meaning traditional unidirectional broadcasting of a signal to a wide range of customers (e.g. satellite television), a GEO bent pipe configuration in the satellite is the best choice. In these cases an uplink station with enough power and a large antenna aperture is used, so the power of input signals at the satellite is not an issue.

In telecommunication scenarios with more than one nomadic or mobile user an OBP helps to improve the system performance (regenerative payload) or increases the system capacity (e.g. on-board routing). In these scenarios the uplink power and antenna accuracy is strictly limited which causes a weak signal on the satellite.

The system performance can be increased in two different ways. The first increase for frequency division multiple access (FDMA) systems can be achieved by an accurate selection of the correct input band, which is normally done by many analog filter banks. On such a reconfigurable platform every kind of filter can be implemented in the digital domain. In combination with a flexible automated gain control (AGC) or level control the variable filters offer the ability to gain system performance. The second opportunity to increase system performance is to decode and reencode the data stream completely. This will cause a better performance on noise immunity and reaches a lower bit error rate at the receiver on the ground. The reencoding receiver contains the complete demodulation and decoding (e.g. channel coding) of the data stream to correct the bit errors, followed by new channel encoding and modulation. As a result of this operation a fully reconstructed data stream can be transmitted back down to earth without any impact of the uplink. So the physical effects of the uplink (e.g. noise) are separated from the downlink signal, which is important for a power limited uplink. The OBP additionally offers the ability to do a so-called switching on board of the satellite (on-board routing) for a direct connection between two users (single-hop connection). On satellite systems without an OBP the routing or switching is only done on the ground at a so-called hub station, which results in a multi-hop connection. A direct terminal-

to-terminal communication (single-hop) improves not only the delay of the connection, it improves the system capacity, because it requires only 50% of the spectrum resources, if the system offers the same spectrum efficiency per link as a hub based system.

As a result an OBP is relevant for the following cases or scenarios:

- low signal-to-noise ratio (SNR) at the uplink (SNR of uplink is in the same order of magnitude as downlink)
- direct routing on board
- remultiplexing in case of more beams or carriers
- interference of the uplink.

Nowadays some of the telecommunication satellites use application specific integrated circuits as an OBP on the satellite. A huge disadvantage of this solution is the fact that telecommunication standards change frequently and that the ASIC is only designed for one technology or standard. In addition to that, typical lifetimes of communication satellites are up to 15 years due to the limited fuel, which is necessary for the correction of the orbit position and the lifetime of the solar panels. Including the development time of the chip and the satellite, the technology has to cover a long period of operation without the ability of updates. In order to be prepared for the telecommunication standards of tomorrow and to gain additional improvement of the SNR, a completely reconfigurable platform is essential.

To validate a new on-board processing architecture, to demonstrate its flexibility and to avoid the disadvantages of common OBPs, Fraunhofer IIS is developing a new kind of reconfigurable OBP based on four leading-edge rad-tolerant FPGAs (Xilinx Virtex-5). With this platform, it is possible to develop and test new data-protocols and standards for satellite telecommunication directly on a satellite link.

Common FPGAs provide the level of flexibility which is needed for a completely reconfigurable platform because of their programmable architecture. Accordingly, an FPGA platform is a good base for an OBP. Furthermore, the Virtex-5 also provides a lot of digital signal processing (DSP) performance.

Current FPGAs for space applications also provide a high level of hardness against all the radiation effects in space. Hence, an FPGA-based platform fulfills the three major needs

- completely reconfigurable
- high amount of processing power (especially for DSP)
- radiation hardened for space applications

to be the perfect choice for an OBP. Having access to such a variable platform Fraunhofer is in the position to demonstrate and test system configurations based on OBP. This is considered as an enabler for commercial applications in satellite telecommunication.

## 2 Environmental Conditions

The enormously harsh environment of space with all its radiation effects like gamma rays, protons and heavy ions are one of the main challenges to deal with. In order to classify these kinds of effects, two of the basic parameters are discussed below [3],

- total ionizing dose (TID), and
- single event effects (SEE).

TID is the measured energy, deposited in a medium, e.g. silicon, caused by the gamma rays and has the unit rad or gray [Gy]. 100 rad are equal to 1 Gy, which is defined as shown in the following equation.

$$1\,\mathrm{Gy} = 1\,\frac{\mathrm{J}}{\mathrm{kg}} = 1\,\mathrm{m}^2 \cdot \mathrm{s}^{-2} \tag{1}$$

It has long term impact, limiting the lifetime of the electronic components. Furthermore, high-energy particles, e.g. heavy ions, can cause single event effects, which could lead to a change of state in the digital components. The SEE immunity depends on the maximum linear energy transfer (LET) threshold of the electronic parts, measured in [MeV $\cdot$ cm$^2$/mg]. In hardware, a single event effect caused by a high-energy particle leads to a transfer of the energy of this particle into the transistors of the device.

Two of the most important single event effects are single event upset (SEU) and single event latch-up (SEL). An upset is a non-destructive short change of state (soft error) and could be detected by parity checksums. On the contrary, a latch-up could result in a long term change of state (hard error), which could harm an electronic device forever, if it is unnoticed.

In order to prevent radiation effects like TID, shielding in the form of an aluminum box with a 5 mm thickness around the OBP is planned. For this satellite in the GEO the total ionizing dose could be estimated as 10 krad/year for the OBP in the box. So the electronic parts have to cover about 150 krad in a lifetime of 15 years. In order to reduce the probability of a latch-up, the electronics should be able to handle a linear energy transfer of minimum 50 MeV $\cdot$ cm$^2$/mg.

SEL detection can be done by current control, because normally the current consumption increases in case of a latch-up. A complete reset of the affected electronic parts could then prevent any subsequent errors. SEU are not that crucial, because of a non-steady change of state without any follow-up damage of the electronic part. For instance, the probability of a single event upset in the FPGA predicted by Xilinx [10] is specified with $3.8 \times 10^{-10}$ Upsets/bit/day. This effect can not be avoided by an extra shielding and occurs randomly depending on the sun activity. In order to detect such an effect and prevent any errors, a cyclic redundancy check (CRC) is required, especially for the FPGA configuration itself.

# 3 System Overview

## *3.1 Hardware*

The OBP can be separated into four functional units, i.e. analog front ends, a platform for digital signal processing (DSP), the periphery (e.g. memory and clock) and a unit which handles the interfaces as shown in Fig. 1.

The OBP will support only 3 inputs to reduce the total weight. The first input is the received uplink signal from an anchor station in Weilheim, Germany. In the current state of development, the other two inputs are the south-east and north-west beam located over Germany, as shown in Fig. 2. These three inputs at an intermediate frequency of 1.65 GHz could be switched to either three redundant RF receiver paths with a down mixing stage or to one direct sampling path. For experiments typically 1 or 2 inputs are used. The architecture accordingly supports 4 for 3 redundancy for the RF input part. Within the DSP unit the signal can be processed in many different ways like filtering and reencoding. Afterwards the data stream is up-converted via one of three redundant analog transmitter front ends before it can be connected up to two different downlink beams. So the transmitter side offers a 3 for 2 redundancy. All details about the architecture of the OBP are shown in Fig. 1.

A typical use case for the OBP is the single-hop communication between two terminals in different beams as it is shown in Fig. 2. For testing it is sufficient to process data from two spots only. In order to simplify the hardware and reduce weight of the demonstrator only 2 out of 4 spots are connected to the OBP.

## *3.2 Firmware*

The harsh environment introduces also challenges for the firmware, which requires a redundant design, too. This will be provided by the Xilinx Triple Module Redundancy (XTMR) Tool. This technology was developed to address the special needs of FPGAs in high-radiation environments. It is originally designed for space applications and is proven through numerous mission-critical projects [8]. With its internal triple redundancy it provides full SEU and SEL immunity for any high reliability Virtex FPGA design. To verify the correctness of the design file on the satellite, a classical CRC is done by the FPGAs. Also, the configuration handling of the FPGAs is redundant, too. The Virtex devices can be programmed either by JTAG or by a serial interface via slave serial. A verification of the FW before the upload is done with an identical OBP on the ground. In order to get a comparable result, the OBP on the ground is tested in a thermal vacuum chamber to simulate the exact environmental conditions of the OBP in space. The firmware itself is uploaded to the OBP on the satellite via a telemetry and telecommand (TM/TC) link with an AES (Advanced Encryption Standard) encryption.

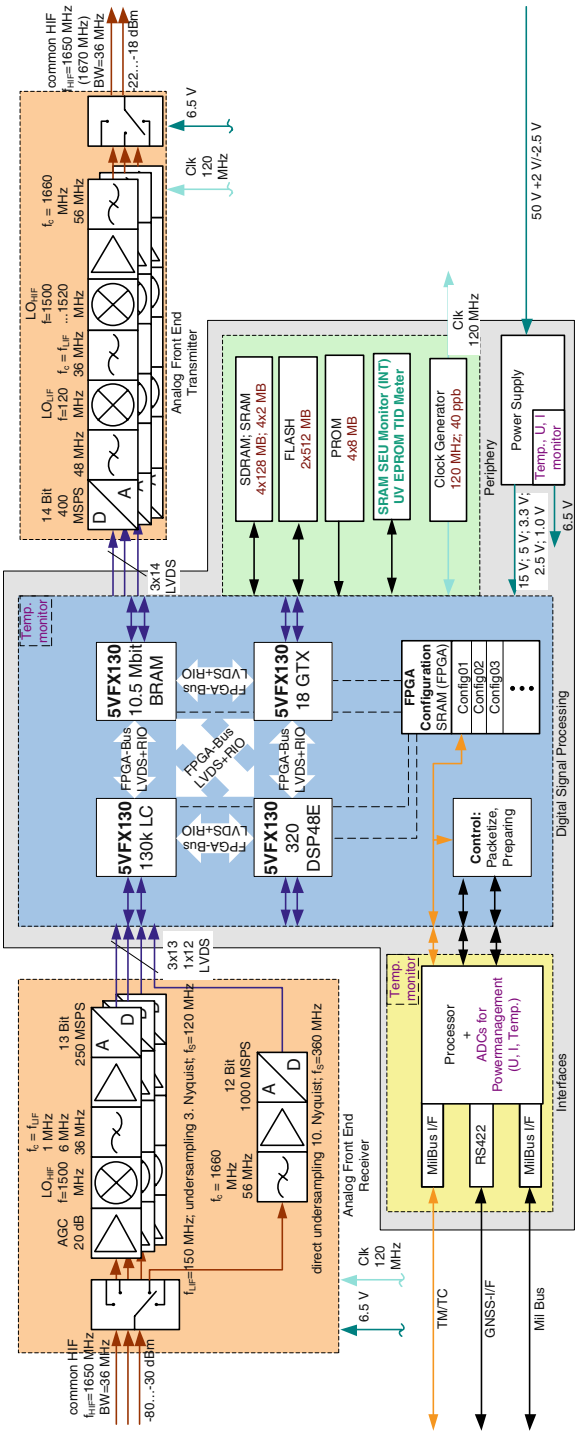



**Fig. 1** On-board processor overview

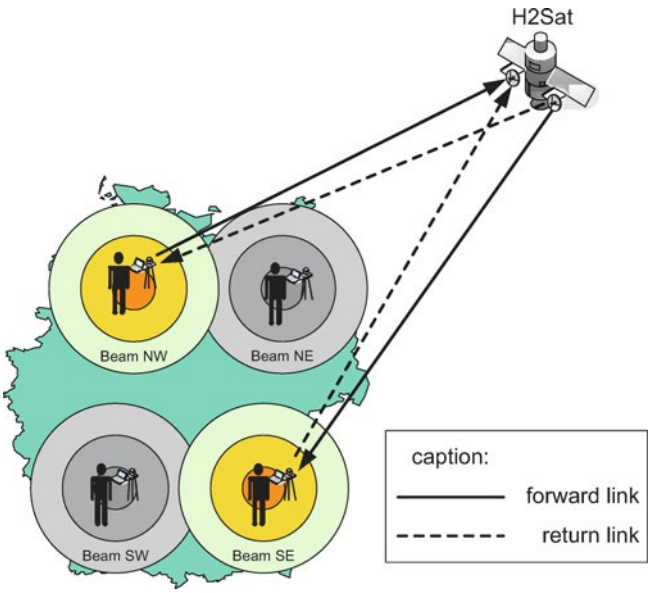



**Fig. 2** Schematical diagram of the four satellite beams with processed links located over Germany

## 4 First Testcase: DVB-S2

The OBP will support several functionalities. As a reference test case the DVB-S2 standard has been selected to demonstrate and test the performance gain provided by the OBP. A wide range of different modulation and coding schemes could be chosen as shown in Table 1. This will cause a variation of the required $E_S/N_0$ from $-2.4$ dB (QPSK, $R = 1/4$) up to 16.1 dB (32APSK, $R = 9/10$). As a result, the spectrum efficiency will vary from 0.49 bit/s/Hz to 4.45 bit/s/Hz. The combination of the lowest modulation and coding scheme (QPSK, $R = 1/4$) allows a very low SNR. In case of a classical amplify-forward architecture on the satellite, the received noise power will be retransmitted even though more than 50% of the downlink power is noise. So power robbing is the result. On the other side, the highest modulation and coding scheme (32APSK, $R = 9/10$) sets very high requirements in the overall SNR. Within the DVB-S2 standard [2], the ability of changing the modulation scheme is defined, depending on the environmental condition. Changing the modulation scheme could be done by the new partial reconfiguration flow [11], which is provided by the new Xilinx PlanAhead software [9]. As a result of the two higher modulation schemes (e.g. 16 APSK and 32 APSK) in DVB-S2, the overall performance of passband and group-delay ripple is very important. In order to achieve a good performance for the ripple related to passband and group-delay of the transmitter path a digital predistortion may be used in the FPGAs.

Figure 3 shows the symbolic architecture of a DVB-S2 transmitter, as it is defined in the ETSI (European Telecommunications Standard Institute) standard [2]. The

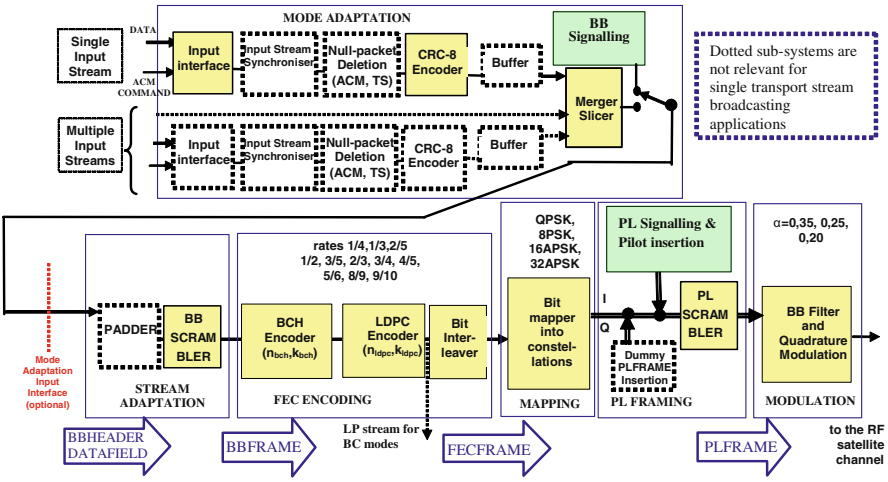



**Fig. 3** Functional block diagram of the DVB-S2 transmitter [2]

input stream is defined as an MPEG-TS (Moving Picture Experts Group – Transport Stream) single or multiple stream related to the OSI (Open Systems Interconnection) layer 2. In the physical layer (OSI layer 1) after merging, a scrambler is used to grant alternate pseudo-random input sequences in order to prepare the data for the forward error correction (FEC). The FEC is done in DVB-S2 by an outer and inner encoder. For outer encoding a Bose-Chaudhuri-Hocquenghem-Code (BCH-Code) is used. The inner coding is realized by an low-density parity-check (LDPC) encoder. All available code-rates are defined in Table 1.

The most complex algorithm to be managed within DVB-S2 is the LDPC channel decoding in the receiver [5]. An LDPC decoder, which is the bottle neck of the system in terms of the data rate, is based on a high number of matrix permutations. In order to handle this kind of permutations a complex structure of memory reordering is necessary. Memory reordering is done by the internal block RAM (BRAM) of the FPGA, which is strictly limited. So the channel decoder in the receiver needs a high amount of valuable FPGA resources.

Compared to a classical telecommunication satellite in the GEO with bent pipe technology, the gain in signal-to-noise ratio could be estimated to maximum 3 dB for uplink and downlink respectively [1], which may help to either reduce the antenna size or the power of the transmitter or increase the data rate.

**Table 1** Overview of the parameters of DVB-S2

| Parameters of DVB-S2 | |
|---|---|
| Coderates | 1/4, 1/3, 2/5, 1/2, 3/5, 2/3, 3/4, 4/5, 5/6, 8/9, 9/10 |
| Modulation | QPSK, 8PSK, 16APSK, 32APSK |

## 5 Main Applications

For the H2Sat mission several applications and experiments are defined, which will be analyzed in order to develop adequate ground terminals. One significant advantage of a reconfigurable processor on a satellite is to command an evaluation platform for new developments. Firmware upload will be possible either over the TM/TC link or through the uplink via the OBP. The operational states – like configuration, bandwidth, temperature and power consumption – can be monitored over the TM/TC link. Prompt tests can significantly reduce time to market for inventions e.g. referring to new modulation processes or waveforms.

The terminals shall be suitable for nomadic and mobile users. As opposed to mobile users, nomadic users don't change their position during the application. For nomadic terminals, antennas with reduced aperture sizes with a diameter of 22 cm

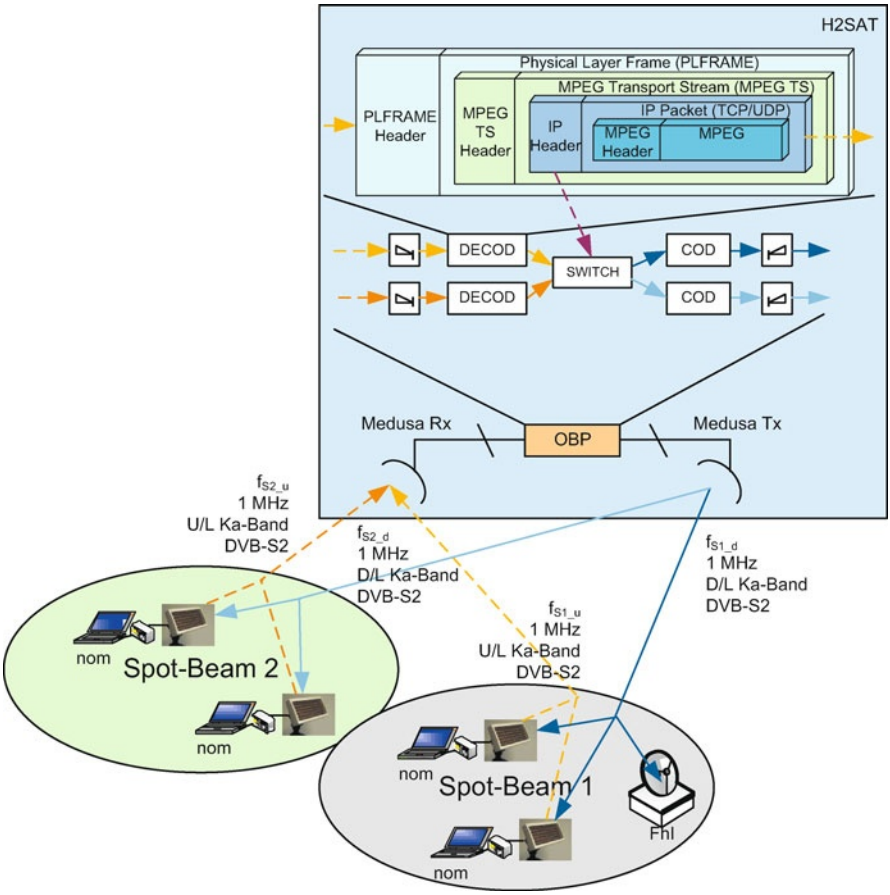



**Fig. 4** Functional diagram for the on-board switching use case

for receiving, 14 cm for transmitting and with wide radiation patterns can be applied. These antennas can be set up and run fast and simple, for example mounted in or on vehicles. Based on these antennas, the link budgets were calculated to estimate the margins of the communication channel. In order to communicate with terminals, the OBPs have to process the information. The terminals also need the capability of a backward channel, that will be implemented and tested. The first step is to realize an asymmetrical communication with higher data rates in forward transmission. Based on this, the possibility of higher data rates for the backward channel will be analyzed.

Other experiments concern single-hop communication. The OBP shall be utilized as a satellite hub station that provides direct switching between terminals. Most applications for this concept will be useful for action forces that do not have a centralized infrastructure in their operation area. Because of their huge vehicles, bigger antennas are applicable. This experiment will be extended to use the OBP for data packets (IP) switching, which also can be used for VoIP (Voice over IP). For switching based on the IP the demodulation has to take place in the network layer as shown in Fig. 4. An appropriate periodically updated routing table is applied in orbit to send the data via the considered beam. Most of today's switching between users is done in the ground station, even for satellite communication with the disadvantage of higher switching delays for mobile to mobile communication links.

Another experiment evaluates the possibility of capturing a huge amount of telemetry stations by narrowband communication. In case of messaging applications the OPB may receive data from millions of terminals using access schemes optimized for short messages. This data has to be sampled and stored in the OBP before sending it back to the earth station. In this way, the communication will be evaluated referring to the restoration of the information.

## 6 Conclusion

The proposed on-board-processor platform establishes a scientific vehicle to perform various in-orbit tests. It is set up as a platform for conducting communication experiments, which rely on the regeneration of the received data. Therefore, reconfiguration is a main feature of this processor, especially for implementing standards of the future. This also opens the path for reconfigurable and reliable processor platforms to be used in terrestrial and avionic communication systems. It is shown that current FPGA technology fits well for an OBP, because of its potential processing power and hardness for the radiation effects in space. In addition to that, the programmable architecture allows a complete reconfiguration of the hardware, which makes it extremely flexible for novel technologies in satellite communication. In order to enable a high level of safety, an additional redundancy for most of the components is chosen.

## Acknowledgements

This paper is a result of the ideas and work of many people in the RF and Microwave Department at Fraunhofer IIS. The authors would like to thank Ernst Eberlein, Elfriede Fritschle, Niels Hadaschik, Markus Polster and Mario Schühler for supporting and reviewing this paper.

## References

1. Dodel H, Eberle S (2007) Satellite communication. Springer
2. European Telecommunications Standards Institute (2009) EN 302 307 – V1.2.1 – Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications (DVB-S2)
3. Koons HC, Mazur JE, Selesnick RS, Blake JB, Fennell JF, Roeder JL, Anderson PC (1999) The Impact of the Space Environment on Space Systems
4. Maral G, Bousquet M (1998) Satellite Communications Systems – Systems, Techniques and Technology 3. Edition. Wiley
5. Massaro F, Elsner J, Vedova FD, Jondral FK Components for Software Radio Wideband Receiver: A space segment survey
6. OHB System AG: OHB System ENG – Heinrich Hertz. https://www.ohb-system.de/
7. Voigt S (2010) The German Heinrich Hertz Satellite Mission. IEEE
8. Xilinx (2009) Aerospace and Defense, Xilinx TMRTool Product Brief
9. Xilinx (2009) PlanAhead User Guide
10. Xilinx (2010) Radiation-Hardened, Space-Grade Virtex-5QV Device Overview
11. Xilinx (2011) Partial Reconfiguration User Guide

# Opportunities and Challenges for Multi-Constellation, Multi-Frequency Automotive GNSS Receivers

**Cécile Mongrédien, Alexander Rügamer, Matthias Overbeck, Günter Rohmer, Philipp Berglez, and Elmar Wasle**

**Abstract** In this paper, the implementation of a multi-constellation, multi-frequency automotive GNSS receiver is discussed. The main objective of this paper is threefold. First, to identify, in the context of automotive applications, the most promising GNSS signal combination and analyze its benefits and limitations. Second, to propose a receiver architecture that offers sufficient robustness and flexibility to maintain high-accuracy and high-availability navigation capabilities in challenging automotive signal environments as well as to accommodate the particulars of the legacy, new and modernized signals. Third, to optimize the receiver's implementation so that it meets the automotive requirements in terms of size, cost and power consumption. To this end, several front-end architectures are compared and some key aspects of the baseband hardware implementation discussed. Additionally, robust acquisition and tracking algorithms that respectively account for the availability of a second frequency and for the introduction of advanced modulations are presented. Finally, some insights are provided regarding optimization of the PVT performance in terms of multipath mitigation and ionospheric corrections.

## 1 Introduction

Global navigation satellite system (GNSS) receivers greatly benefit from the modernization of existing GNSS constellations such as GPS and GLONASS as well as from the launch of new ones such as Galileo and COMPASS. First, the combining of these constellations can significantly improve the navigation solution availability in urban canyons and heavily-shadowed areas. Second, increased satellite availability translates into higher measurement redundancy and improved reliability. Additionally, the excellent inherent noise and multipath mitigation capacity of the new and

Cécile Mongrédien (✉)
Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany
cecile.mongredien@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_16, © Springer 2011

modernized wide-band signals, combined with the ionospheric mitigation capacity brought by frequency diversity, notably improves accuracy in both measurement and position domains.

GNSS receiver performance, however, remains strongly affected by the user environment. In particular, developing a receiver for automotive applications implies that a continuous and reliable navigation solution has to be delivered in rapidly changing environments where signal reception can be degraded by the presence of buildings, trees or tunnels and suffer from heavy attenuation, high multipath or short loss-of-lock.

Following an overview of the current and upcoming global satellite navigation systems and their signals, the trade-offs in terms of hardware, software, and algorithm complexity for a multi-constellation, multi-frequency automotive-grade GNSS receiver will be discussed. Having identified a promising GNSS signal combination and analyzed its benefits and limitations in the context of automotive application, this paper will present the implementation and optimization of the different receiver blocks.

## 2 GNSS Signals

Currently there is only one fully deployed global navigation satellite system available, the American NAVSTAR-GPS. Nevertheless the Russian GLONASS-system has almost reached its full operational capability again. Both systems are continuously evolving to enhance the signals and services they will provide in the coming years. Aside, two other GNSS are being developed – the European Galileo and the Chinese COMPASS systems. Therefore, in a few years, at least four independent but interoperable GNSS will be available and will better support an increasing range of applications.

GPS and Galileo are both using a code division multiple access (CDMA) multiplexing scheme. Additionally, they are sharing two frequency bands, making dual-frequency GPS/Galileo reception and processing very attractive. In contrast, the first GLONASS satellites (*Glonass* and *Glonass-M*) are using frequency division multiple access (FDMA). Since all these GLONASS satellites use the same pseudo random noise (PRN) code to spread their navigation message and to provide the required correlation gain, they are distinguished through 12 individual carrier frequencies. In addition to relying on an FDMA scheme, GLONASS uses frequency bands that are offset with respect to those used by GPS and Galileo, currently making GLONASS receiver integration a challenging task. But the new *Glonass-K1* and *-K2* satellite generations, whose deployment started early this year, will also broadcast CDMA signals on frequency bands partly shared with GPS and Galileo. This will significantly increase the GPS/Galileo/GLONASS interoperability and, therefore, make GLONASS reception a much more attractive option. The Chinese COMPASS system is expected to broadcast signals that are closely related to those offered by the Galileo system (e.g. in terms of center frequency, bandwidth or modulation).

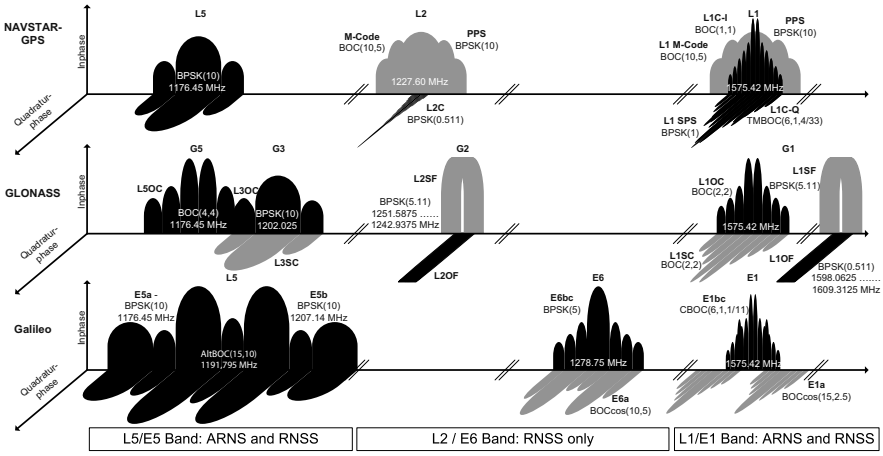



**Fig. 1** Current, renewed and upcoming GPS, GLONASS, and Galileo signals

However the lack of a public released interface control document (ICD) makes the current use of COMPASS unattractive for any purposes other than research.

Figure 1 shows the L-band spectrum of the current and planned GNSS signals with the notation of their modulation names and carrier frequencies. The grey signals are classified (e.g. signals for military purpose only), the black ones are open signals. All current and upcoming GNSS signals are within the protected Radio Navigation Satellite Services (RNSS) band but only the L1/E1 and L5/E5 bands are within the even better protected spectrum allocated to Aeronautical Radio Navigation Services (ARNS). The other two GNSS bands, E6 and L2, only protected through the RNSS, suffer from radar, military transmissions and other potentially strong interferers.

Thus, the combination of both L5/E5 (including the GPS L5, Galileo E5a and E5b) and L1/E1 (featuring the GPS L1 C/A, L1C and Galileo E1-B/C) are deemed the most appropriate for an advanced multi-constellation, multi-frequency automotive GNSS receiver. For fast acquisition the relatively narrow-band L1/E1 signals (GPS C/A and Galileo E1-B/C with BPSK(1) and BOC(1, 1) modulation, respectively) are typically used. The estimated Doppler and code delay can then be used for high performance tracking with the wide-band L1/E1 MBOC(6, 1, 1/11), L5/E5a BPSK(10) or E5b BPSK(10) signals.

## 3 Opportunities and Challenges in the Automotive Area

### *3.1 Opportunities*

The performance of a navigation system is usually quantified using the following measures: availability, accuracy, reliability and integrity.

Availability refers to the number of satellite ranging signals available to the user. Accuracy quantifies how close the navigation solution provided by the GNSS receiver is to the true user's position and velocity and mainly depends on two quantities: the user equivalent range error (UERE), which maps all GNSS errors into a single generic receiver-satellite range error, and the geometric dilution of precision (GDOP), which measures the geometrical strength of the solution. GNSS reliability often refers to a receiver's ability to to detect a blunder through statistical testing and to estimate the effects of undetected blunders on the estimated navigation solution. Integrity is theoretically defined as the receiver's ability to provide users with timely warnings when system failures make the estimated navigation solution unusable. In practice, integrity and reliability are often combined in the so-called receiver autonomous integrity monitoring (RAIM) concept.

Using the aforementioned definitions, it is possible to examine the benefits of multi-constellation and multi-frequency navigation.

Upon completion of the Galileo full operational capability (FOC), the combining of the GPS and Galileo constellation will approximately double the number of visible satellites. This, in turns, will significantly improve the navigation solution availability in urban canyons and heavily-shadowed areas. The excellent inherent noise and multipath mitigation capabilities brought by the new and modernized wide-band signals, combined with the ionospheric mitigation capacities derived from frequency diversity will notably improve accuracy in the measurement domain. Additionally, the superior geometrical strength brought by multi-constellation satellite coverage will further enhance accuracy in the position domain. The increased measurement redundancy and accuracy combined with an improved geometry will make the receiver significantly more reliable. Multi-frequency reception will offer an additional protection against jamming and interference since, if one frequency band is malfunctioning, the receiver will still be able to provide a navigation solution relying on another frequency band.

## *3.2 Challenges*

Designing a GNSS receiver for automotive applications is challenging because it implies that the receiver needs to withstand a wide range of signal degradations. In addition to the accelerations and vibrations inherent to driving, the receiver needs to cope with frequent multipath and blockages. Multipath often occur when driving along buildings and traffic signs. Partial and complete blockages typically occur when driving under a tunnel or in a dense forest.

In addition to facing a challenging signal environment, the receiver's modules must also be designed to cope with new signal specifications, including but not limited to, frequency, modulation and navigation message. In the position domain, this implies that the receiver must implement new algorithms to optimize ionospheric corrections and measurement weighting. Regarding baseband processing, it means that innovative tracking algorithms have to be developed to fully benefit from the

improvements brought by the new and modernized signals. One particular example of such innovation is the introduction of new signal modulations that outperform the legacy BPSK(1) in terms of ranging accuracy and frequency compatibility. Specifically, the MBOC modulation, proposed as the common GPS/Galileo baseline in the L1/E1 frequency, is defined in the frequency domain as the sum of 10/11 of the normalized BOC(1, 1) spectrum and 1/11 of the normalized BOC(6, 1) spectrum. As mentioned in Sect. 2 this modulation supports fast acquisition and high-accuracy tracking but, as will be discussed in Sect. 4.5, it also requires specific algorithms to ensure reliable and accurate code tracking.

## 4 GNSS Receiver Blocks

When designing a multi-constellation, multi-frequency receiver some trade-offs are required to meet both accuracy and cost requirements of the automotive sector. To maintain cost in an acceptable range, the receiver complexity, size, and power consumption should be kept as low as possible. However, to deliver a high-performance navigation solution, computationally expensive reception of multi- and wide-band signals is unavoidable.

### *4.1 RF Front-End Hardware*

Traditional GNSS front-ends but also current mass-market GNSS receivers typically feature a low-IF architecture with RF-bandwidth of approx. 2 to 4 MHz and a low resolution analog-to-digital converter (ADC) of 1 to 3 bit. This is sufficient for the legacy GPS L1 C/A or the narrow-band Galileo E1 BOC(1, 1) signals but not for most of the new GNSS signals, especially if their full potential in terms of accuracy and multipath resistance is to be reached. For these, considerably larger bandwidths are necessary (e.g. at least 14 MHz and 20 MHz for the GPS/Galileo L1/E1 MBOC(6, 1, 1/11) and L5/E5a BPSK(10) signals, respectively) which leads to higher sampling rate requirements. Wider RF-filters also make the front-end more susceptible to jamming and unintentional interferers. Therefore, a higher ADC resolution is required to offer a higher dynamic range and to enable digital mitigation algorithms.

The straight forward approach is to widen the bandwidth and to use higher sampling rates for each desired GNSS-signal while keeping the original low-IF architecture approach where the intermediate frequency (IF) is within the range of the RF-bandwidth. These solutions can already be found as integrated circuits and can easily be tuned to the required GNSS signal band.

However, for the wide-band BOC signals such as the Galileo E5 AltBOC(15,10), a zero-IF architecture can be very advantageous since the inherent zero-IF problems, namely DC-offset and flicker noise, are not so relevant to the DC-free BOC signals.

If needed a Hilbert transformation can be implemented in the digital signal processing to select the lower or upper band of the complex signal received. Moreover, quasi zero-IF architectures are also a popular solution to enable simultaneous reception of the L1/E1 GPS/Galileo signals and the GLONASS G1 FDMA signals by placing the local oscillator between both signal bands.

The continuously improving ADC performance makes RF- or sub-sampling front-ends feasible even for the L-band GNSS signals [7]. However, this type of architecture still suffers from several limitations such as higher power consumption in the front-end and the following digital baseband signal processing, stringent ADC sampling jitter requirements, potential instability due to the high amplification needed on one frequency range, or susceptibility to interference. Therefore, the sub-sampling architecture is currently not the best choice for an integrated multi-band GNSS front-end receiver. However, since it closely matches the software defined radio (SDR) philosophy, it is expected to gain importance in the coming years.

All the architectures mentioned so far are generic in a way that they do not exploit any properties of the GNSS signals CDMA structure. This makes their implementation straight forward but is not necessarily efficient. A new front-end architecture based on intentional overlay of the GNSS CDMA signals with an appropriate path control for simultaneous multi-band reception has been proposed in [8] and [9]. It has been demonstrated that such an overlay architecture exploiting the GNSS signal characteristics can be more efficient than the aforementioned generic approaches in terms of cost, size and power consumption.

## *4.2 GNSS Baseband Hardware*

The GNSS baseband hardware, typically realized in an FPGA or ASIC, consists of two main parts: the digital signal processing blocks including the acquisition, tracking or management units, and an embedded processor or external DSP where software tracking loops are closed, measurements generated and final position, velocity, and time (PVT) information computed.

The sampling frequency is fixed by the Nyquist criterion and the front-end bandwidth. In order to process the digital data stream coming from the front-end in real-time, the tracking hardware must at least use the same clocking frequency. As the power consumption in CMOS technology increases proportionally to the clocking frequency, a trade-off between optimal power consumption and received bandwidth needs to be made. Nevertheless, the processing of new wide-band signals such as GPS L5 or Galileo E5a/b is needed to reach the desired level of code tracking accuracy. This, in turns, justifies the use of sampling frequencies 10 times higher than that required by the legacy GPS L1 C/A signal.

The benefit of a higher availability in multi-constellation, multi-frequency receivers comes at the expense of a sharp increase in the number of channels that needs to be implemented. For each given GNSS signal, the tracking channel consists of a complex correlator with up to five output taps combined with a code generator.

**Table 1** Code length of the different GPS/Galileo signals to be tracked

| Signal | | L1 | E1-B | E1-C | L5-I | L5-Q | E5a-I | E5a-Q | E5b-I | E5b-Q |
|---|---|---|---|---|---|---|---|---|---|---|
| Length | Prim. | 1023 | 4092 | 4092 | 10230 | 10230 | 10230 | 10230 | 10230 | 10230 |
| | Sec. | N/A | N/A | 25 | N/A | N/A | 20 | 100 | 4 | 100 |
| Shift reg. | | Yes | No | No | Yes | Yes | Yes | Yes | Yes | Yes |

For a legacy GPS L1 C/A receiver seeing a maximum of 12 satellites, 12 tracking channels were sufficient. However, accounting for the presence of a pilot channel on all modernized signals, the envisioned GPS/Galileo L1/E1 L5/E5a/b receiver would require approximately one hundred tracking channels. This dramatically increases the digital hardware complexity, size and power consumption.

While correlating the incoming signal with the desired local code replica, it is of the utmost importance to ensure timely delivery of the PRN code chips to the tracking channel. These PRN sequences, which are unique for each GNSS signal component, can either be stored in memory or generated in real-time using a linear feedback shift register (LFSR) with individual feedback tap settings.

Attaching a memory cell to each tracking channel provides the greatest flexibility for legacy, modernized and future GNSS signals. Indeed it guarantees an universal implementation for CDMA signals since new PRN sequences can be easily uploaded or updated. However, memory cells are burdensome in terms of cost and size, especially in an ASIC design. An alternative is to use external shared memory combined with "on-demand" uploading of the PRN sequence via the system bus. But since the bus system is typically also used by other components, it is easy to understand that this approach can rapidly reach its limits when the number of tracking channels increases. Despite the flexibility loss, it is therefore recommended to use shift registers whenever possible (see Table 1).

## *4.3 Acquisition*

GNSS acquisition is usually performed as a 2-D search in time and frequency where the unknown time (PRN code delay) corresponds to the receiver-satellite range and the unknown frequency (Doppler frequency) corresponds to the receiver-satellite relative motion. To detect the signal, the incoming signal first has to be correlated with all possible local code and carrier replicas. Once the maximum correlation value is found, the signal's presence is decided using statistical testing. The very large number of correlation performed makes acquisition the most computationally expensive stage in the receiver chain.

Dual-frequency receivers can achieve significant savings in hardware computational requirements, using a sequential acquisition approach. The main idea is to use time and frequency information exchange between the E1/L1 and E5/L5 acquisition blocks to reduce of the uncertainty region in the latter block. The underlying

assumption behind this algorithm is that the E1/L1 and E5/L5 signals are perfectly coherent or, in other words, that their chip transitions are perfectly aligned and their Doppler frequencies are directly related through the theoretical scale factor given by the ratio of their carrrier frequencies. Unfortunately, this assumption does not fully hold true since signal's propagation through the ionosphere results in additional ranging delays and Doppler shifts and frequency-dependent hardware paths in the satellite's transmission chain and receiver front-end create additional group delays. Both effects degrade signal coherency and can therefore limit the benefits of this sequential algorithm.

## *4.4 Robust Carrier Tracking*

In order to maintain robust tracking regardless of the car dynamics and signal obstructions, the quality of the carrier tracking needs to be permanently monitored. This can be done, as suggested in [5], using phase and frequency lock indicators (PLI and FLI). Based on their return values, it is possible to decide on-the-fly how to optimally close the carrier loop:

- Frequency lock loops (FLLs) are robust against large frequency errors and high signal dynamics. However they produce noisy estimates that still exhibit residual phase error. FLLs are therefore best suited to ensure fast convergence after acquisition.
- Phase lock loops (PLLs) produce accurate estimates with no residual phase error. However, in presence of high signal dynamics, they are likely to loose lock. PLLs should therefore only be used once tracking has reached its steady-state.
- FLL-assisted-PLLs offer a good trade-off between accuracy and robustness. They are therefore well suited either as an intermediate step during the convergence phase or as a fall-back step to avoid loss-of-lock in presence of high signal dynamics.

To increase the carrier loop resistance to dynamics, the FLL and PLL should be implemented as second and third order loops with relatively large bandwidth and fast update rate.

## *4.5 Unambiguous Galileo E1-B/C Code Tracking*

One of the main challenges when designing a code tracking algorithm for the Galileo E1-B/C signal is to avoid converging to one of the auto-correlation side peaks (shown in Fig. 2 as VE and VL). To this end, the two-step tracking algorithm proposed by [2] can be implemented. The main idea of this algorithm is to use two different discriminators, shown in Fig. 3, to both benefit from the narrowness of the

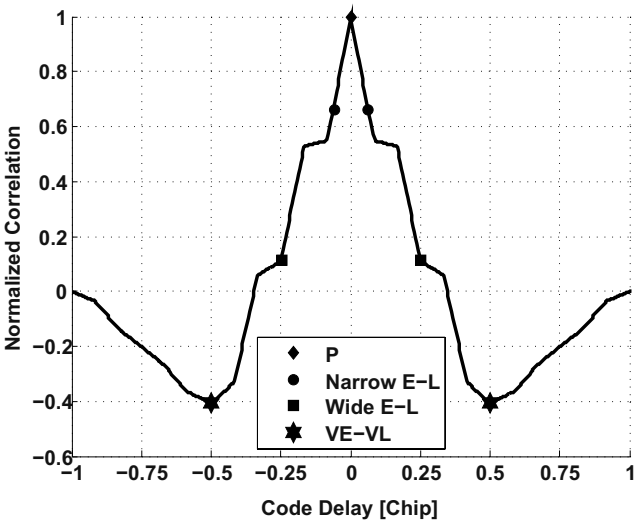



**Fig. 2** Galileo CBOC(6, 1, 1/11) autocorrelation function with correlator taps

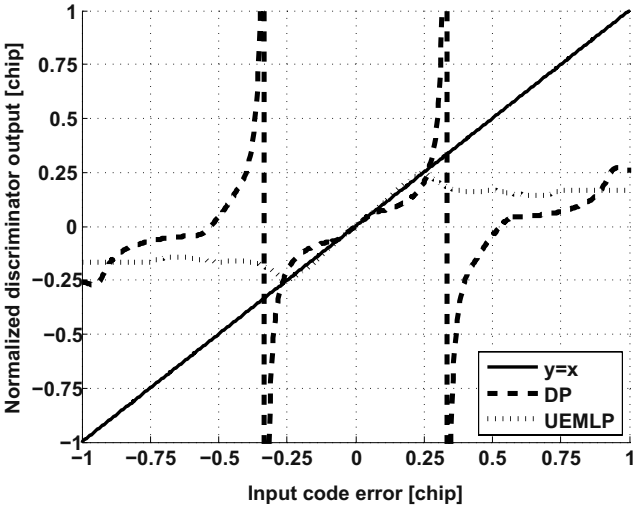



**Fig. 3** Galileo E1-B/C coarse (UEMLP) and fine (DP) discriminators

CBOC(6, 1, 1/11) autocorrelation main peak and minimize the risk of getting locked on one of the side peaks.

The coarse tracking step makes use of an unambiguous discriminator hereafter referred to as the unambiguous early-minus-late-power (UEMLP) discriminator. This discriminator is an extension of the well known early-minus-late-power (EMLP) discriminator in that it uses the information carried by the very early (VE) and very late (VL) correlator outputs to provide an unbiased response for input code

errors below 0.25 chips. As highlighted in [2], this discriminator, even though unambiguous, offers limited noise and multipath mitigation capabilities. Therefore, it is only used directly after acquisition to ensure convergence to the correct peak. Once proper convergence is ensured, fine tracking can be used.

The goal during the fine tracking step is to take advantage of the narrowness of the CBOC(6, 1, 1/11) main autocorrelation main peak. To this end, a dot-product (DP) discriminator is used. This discriminator exhibits the two false-lock point issue inherent to BOC(1, 1) modulation. As explained in [3], the zero-crossings observed at approximately $\pm 0.5$ chips provide stable lock points which lead to a ranging bias of 150 m. Even though such ranging bias can be easily detected in the position domain using a Fault-Detection and Exclusion (FDE) algorithm, the bump-and-jump (BJ) algorithm described in [3] is implemented to detect such occurrences at the tracking level and provide shorter time-to-alert.

## *4.6 Multipath Mitigation*

The mathematical fundament of the satellite navigation signal design is the orthogonality, which is exploited within the correlation function. The incoming mix of signals $S(t)$ is correlated with the locally generated signal $\bar{s}_i(t)$ of one particular satellite $i$ by

$$\int \bar{s}_i(t+\tau)S(t)\mathrm{d}t = \int \bar{s}_i(t+\tau)\left[\sum_j s_j(t) + n(t)\right]\mathrm{d}t = \int \bar{s}_i(t+\tau)s_i(t)\mathrm{d}t \ . \tag{1}$$

The correlation of all signals $s_j(t)$, including here both interferer and desired navigation signals [10], with $\bar{s}_i(t)$ erases all but the desired signal broadcast by satellite $i$ whereby only the autocorrelation of $\bar{s}_i(t)$ becomes maximum on $\tau_{\mathrm{PR}}$, which finally indicates the pseudorange between satellite and receiver. The prefix *pseudo* accounts for the unknown time-offset of the receiver to the atomic clocks of the satellites. Any signal in the same frequency band with very high power or which is not orthogonal in the sense of the mathematical formulation deteriorates the correlation peak measurement and the pseudorange determination. These two deficiencies drive all strategies for multipath and interference mitigation.

Multipath is the effect that a signal does not propagate along direct line-of-sight between satellite and receiver but it is reflected, e.g., by buildings. Reflections are a desired effect for communication in order to provide service even where no line-of-sight between transmitter and receiver is available. However for satellite navigation, where position determination is based on the measured distance between satellite and receiver, this effect reduces the positioning accuracy.

Multipath is still the major error source and it affects all GNSS signals even the upcoming modernized ones. The increasing number of civil signals allows to set-up new multi-signal mitigation techniques [4].

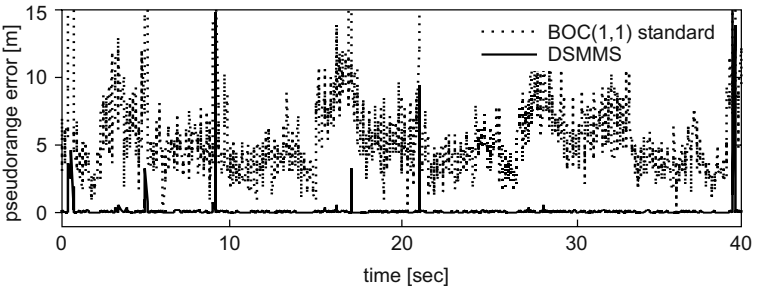



**Fig. 4** Pseudorange error driven by multipath with and without mitigation action

Two signals of one and the same satellite, modulated on the same carrier frequency propagate the same signal path. The difference in their modulation will result in a different multipath error within the tracking loops. The observation equation of the pseudoranges reads in a simplified form

$$R_i = \rho + c\Delta\delta + \Delta_{\text{iono}} + \varepsilon_{\text{MP}i} \ , \tag{2}$$

where $R_i$ denotes the pseudorange measurement of the signal $i$ [1]. The true geometric distance between receiver and satellite is represented by $\rho$ and includes the three dimensional coordinates of the unknown receiver position. The parameter $c\Delta\delta$ denotes the unknown receiver clock error, $\Delta_{\text{iono}}$ the ionospheric effect discussed in the next subsection, and $\epsilon_{\text{MP}i}$ is the range error of the signal $i$ induced by the satellite multipath channel. The multipath range error is a nonlinear function of the channel impulse response, the discriminator type, and signal type. Subtracting two pseudorange measurements to one satellite results in the difference of the corresponding multipath range errors

$$R_i - R_j = \varepsilon_{\text{MP}i} - \varepsilon_{\text{MP}j} \ . \tag{3}$$

Based on this equation, it is possible to eliminate the multipath channel using an optimization function. Figure 4 shows the result of a simulated urban automotive scenario for a high-elevation satellite [6]. Two signals modulated by BOC(1, 1) and BPSK(10) are used jointly to eliminate the multipath error and to compute the dual-signal multipath mitigation solution (DSMMS).

Research has shown that the estimation and mitigation of multipath exploiting multi-signal measurements is feasible but a processing intensive operation. Multipath estimation will allow in future deriving and mapping information about the surrounding of the receiver, thus, the buildings, building structure, or even the characteristic of building faces.

## 4.7 Ionospheric Correction

The major effect of the ionosphere onto satellite navigation is the signal delay. GNSS deal with it by transmitting model parameters, which estimate the ionospheric activity. These models, however, are of limited spatial and temporal resolution, therefore major errors on the pseudorange measurements remain during high solar activity.

GPS uses the Klobuchar model, which adopts a cosine function to describe the characteristics of ionospheric diurnal variation. The European navigation system Galileo implements the Nequick model, which applies the Epstein principles. Simulations have shown that the Nequick algorithm outperforms the Klobuchar model.

The ionosphere is a dispersive medium, thus, the influence onto the electromagnetic waves is frequency dependent. This difference is taken advantage of in order to eliminate the ionospheric effect. A linear combination of two observation equations eliminates the ionospheric term $\Delta_{\text{iono}}$. In case of measurements on the carrier frequencies $f_1$, $f_2$ the ionosphere-free linear combination reads in a simplified form

$$\left[R_1 - \frac{f_2^2}{f_1^2} R_2\right] \frac{f_1^2}{f_1^2 - f_2^2} = \rho + c\Delta\delta \; . \tag{4}$$

Note, that for simplification the ionospheric correction and the multipath mitigation have been considered separately in the algorithms shown. Considering multi-signal measurements on multiple frequencies an appropriate combination mitigates ionosphere and multipath influences. An ionosphere-free linear combination has been used to compute the position solution shown in Fig. 5.

Instead of eliminating the ionospheric effect, multi-frequency measurements are used to estimate the number of electron particles. This opens the door for numerous fields of scientific research.

## 4.8 Position, Velocity, Time

The tracking loops output pseudoranges, phase measurements, Doppler information, signal to noise information, a time indicator and a navigation bit train. The positioning software is responsible for navigation message recovery, the computation of the navigation solution, its statistics, receiver autonomous integrity monitoring, and for feedback information to the tracking loops.

The influence of multipath onto the phase measurements is lower compared to the influence onto the pseudoranges due to the measurement principle. Therefore a filter algorithm combines the two measurements over several epochs to minimize the multipath effect. Within the automotive application short initialization times and low rates of cycle slips in the phase measurements are the main challenge to achieve high position accuracy. Figure 5 shows the single point solution applying dual-frequency correction methods and pseudorange smoothing.

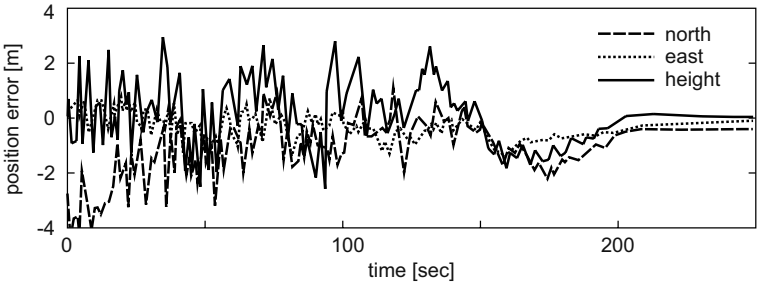



**Fig. 5** Position solution using dual-frequency ionospheric correction and pseudorange smoothing

Redundant measurements on signals on multiple frequencies are used to monitor the integrity of the position solution and provide reliable results.

## 5 Conclusion

The design of a robust multi-constellation multi-frequency automotive GNSS receiver was presented. In particular, several architectures and algorithms designed to meet the automotive accuracy and reliability requirements were introduced and discussed for each receiver block.

The trade offs in terms of hardware complexity were analyzed for both front-end and baseband. To reduce the acquisition hardware requirements, a sequential dual-frequency acquisition scheme was suggested. To achieve robust and accurate tracking, two algorithms were recommended: a reliable carrier tracking state logic and a two-step Galileo E1-B/C code tracking. An innovative dual-signal algorithm was introduced for multipath mitigation. Several multi-constellation ionospheric correction approaches were compared. Finally, the accuracy of a dual-frequency PVT algorithm using carrier-smoothing simulation was presented.

Following this performance analysis, multi-constellation, multi-frequency GNSS receiver is confirmed as a key technology to support new applications such as advanced driver assistant systems (ADAS).

## References

1. Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2008) GNSS – Global Navigation Satellite Systems – GPS, GLONASS, Galileo & more. Springer, Wien New York
2. Jovanovic A, Mongredien C, Botteron C, Tawk Y, Rohmer G, Farine P (2010) Requirements and Analysis of The Robust E1 Galileo Tracking Algorithm in the Scope of the GAMMA-A Project. In Proceedings of the International Technical Meeting of the Institute of Navigation, ION ITM 2010, San Diego, California

3. Julien O (2005) Design of Galileo L1F Receiver Tracking Loops. Ph.D. thesis, UCGE Report No. 20227, The University of Calgary, Department of Geomatics Engineering
4. Lehner A, Steingass A (2007) Verfahren zum Reduzieren von durch Mehrwegeempfang verursachten Störungen in Satellitennavigationsempfängern. DE 10 2006 001 794 B1
5. Mongredien C, Overbeck M, Rohmer G (2010) Development and Integration of a Robust Signal Tracking Module for the Triple-Frequency Dual-Constellation GAMMA-A Receiver. In Proceedings of the 23th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2010, Portland, Oregon
6. Pedross A, Wasle E (2010) Multipath and Interference Cancellation, Galileo Receiver for Mass Market Applications in the Automotive Area. Project internal document
7. Psiaki M, Powell S, Jung H, Kintner P (2005) Design and practical implementation of multi-frequency RF front ends using direct RF sampling. Microwave Theory and Techniques, IEEE Transactions on 53(10):3082–3089. doi:10.1109/TMTT.2005.855127
8. Ruegamer A, Mongredien C, Urquijo S, Rohmer G (2011) Optimal Path-Control for Dual-Frequency Overlay GNSS Receivers. In Proceedings of the 1st International Conference on Localization and GNSS – ICL-GNSS 2011, Tampere, Finland
9. Ruegamer A, Urquijo S, Rohmer G (2010) Multi-band GNSS Front-end Architecture Suitable for Integrated Circuits. In Proceedings of the 2010 International Technical Meeting of The Institute of Navigation – ION ITM 2010, San Diego, CA
10. Wasle E, Seybold J, Hofmann-Wellenhof B (2011) GNSS signal generation for interference analysis. In International Symposion on Certification of GNSS Systems and Services (CERGAL), Oberpfaffenhofen, Germany

# Wi-Fi Attitude and Position Tracking

**Jochen Seitz, Thorsten Vaupel, Stephan Haimerl, Steffen Meyer, Javier Gutiérrez Boronat, Günter Rohmer, and Jörn Thielecke**

**Abstract** An approach for pedestrian navigation in indoor environments is presented. It addresses mobile platforms with low processing power and low-cost sensors. Outdoors the horizontal attitude of a device can be easily detected using electronic compasses. Indoors magnetic disturbances lead to unreliable compass outputs. In this paper a novel approach for attitude and position tracking is introduced. Four horizontally arranged directional antennas are used to collect the Wi-Fi signal strengths of transmitters (access points) in range. For attitude estimation an extended Kalman filter is used, and for position tracking Wi-Fi fingerprinting. With this approach the attitude of a mobile device can be estimated and the position can be tracked in indoor environments like e.g. museums. This enables the use of electronic guides that offer additional information by means of augmented reality on exhibits in visual range. Possible accuracies are evaluated in simulations. A test with measurements collected in a museum demonstrates the functionality of the approach.

## 1 Introduction

Modern smart phones are equipped with a variety of sensors. For positioning, satellite receivers, GSM (Global System for Mobile Communications) and wireless LAN (Local Area Network) modules can be used. Based on them, new and cheap approaches to pedestrian navigation can be provided. This enables new types of location based services for pedestrians ranging from calls for taxis, finding points of interests to city and museum guides.

Commonly the first choice for navigation is the Global Positioning System (GPS). However, the lack of precision and availability of GPS in urban and indoor

Jochen Seitz (✉)
Chair of Information Technology, University of Erlangen-Nürnberg, Am Wolfsmantel 33, 91058 Erlangen, Germany
seitz@like.eei.uni-erlangen.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_17, 

environments is a prevalent problem. With the popular use of assisted GPS (A-GPS) in smart phones the startup time to the first GPS position fix and power consumption can be reduced. But, if signals are too week for detection, positioning fails.

As an alternative or complementary solution for indoor environments Bahl et al. [2] suggested a positioning approach based on the received signal strength (RSS) in Wi-Fi™ [18] networks. Nowadays, because of an increasing number of public and private access points, Wi-Fi positioning becomes more and more attractive for pedestrian navigation [10] and is already integrated into many smart phones.

One remaining challenge in tracking pedestrians is estimating the heading of a person. Pedestrians move very slow and can turn anytime without changing their position. So, the speed vector of a pedestrian calculated from consecutive positions has a very low accuracy. The positioning accuracy can be improved by combining Wi-Fi positioning with dead reckoning, using low cost sensors as proposed in [12, 13]. For pedestrians, dead reckoning can be improved by step detection, as analyzed in [6]. But estimating the attitude is still challenging.

In this paper we present a different approach for attitude and position tracking for pedestrians using only the received signal strength (RSS) measurements of a Wi-Fi device. Instead of additional sensors a special antenna setup with four directional antennas is used. In Sect. 2 the most common techniques and the Fraunhofer IIS awiloc® [1] system for Wi-Fi positioning are discussed. In Sect. 3 the components and equations of the novel approach are presented. In Sect. 4 simulation results and in Sect. 5 experimental results with measurements collected at the "Museum Industriekultur" in Nuremberg are used to discuss the general performance, drawbacks and possibilities of the new approach.

## 2 Wi-Fi Positioning and Related Work

Wi-Fi positioning methods can be divided into two groups. The first group needs a database with the positions and the signal strengths of known Wi-Fi access points, see e.g. in [9] and [11], and the second group needs a database of so called fingerprints, e.g. in [2–5, 14] and [19]. Our approach for citywide localization belongs to the second group. Reasons for choosing this approach are that fingerprinting is reported to achieve higher precision than access point based methods and generally positions of access points are not public.

A fingerprinting database is created by previously gathered RSS measurements. These are then referenced with the coordinates of the positions where they have been observed. Thus, one fingerprint contains a geo-referenced position, RSS values and the corresponding identifiers of the received access points. For positioning, fingerprinting is done by correlating current RSS measurements with the entries of the fingerprints in the database. Then, after selecting the best matching fingerprints, the user position can for example be calculated by a mean of the fingerprint positions weighted by their correlation results. More details on fingerprinting can be found in [2].

Each environment has a characteristical signal propagation. The RSS at a specific position depends on the path loss, shadowing by objects and multipath propagation. The higher the density of shadowing objects, the higher is the accuracy of Wi-Fi positioning, as different fingerprints are less similar in signal space. Therefore, indoors Wi-Fi positioning works very well because of the building structure and furniture. Outdoors, especially on large squares, the database correlation results in ambiguities.

To get meaningful Wi-Fi positioning results, in practice at least three access points must be observed. An advantage of Wi-Fi positioning in urban environments is that the infrastructure is already set up. Existing private and public access points can be used. But on the other hand, positioning suffers from unobserved changes over time and the number of available access points varies from one place to another. An analysis of database changes can be found in [10].

As reported in [10] and [15], several methods are used to collect the measurements to build up the fingerprinting database. As a testbed for positioning, metropolitan areas of several major cities in Germany (including Berlin, Hamburg, Nuremberg and Munich) are used by the Fraunhofer IIS. In Fig. 1 a part of the database covering the city center of Nuremberg is presented. There, on average a fingerprint contains 21 access points, if there is coverage at all.

**Fig. 1** Example extracted from the Fraunhofer IIS awiloc® fingerprinting database in Nuremberg, visualized on an openstreetmaps.org map. *Dots* indicate fingerprint positions and the amount of detected access points at each position, as depicted in the scale-up

## 3 Approach to Simultaneous Attitude and Position Tracking

Wi-Fi positioning can be well used for localization in urban areas, because the density of receivable access points is high enough there. Especially indoors, Wi-Fi positioning offers reliable localization results, but a cheap and reliable attitude estimation system for indoor environments is missing. Ferromagnetic materials in building structures cause large magnetic disturbances that lead to unreliable compass headings. Inertial Navigation Systems (INS) based on micro electromechanical systems (MEMS) suffer from large drift errors with increasing time. This problem can be partially solved by sensor data fusion [7], which is also a research topic at the Fraunhofer IIS.

In the following we present a novel approach for indoor attitude and position estimation. No additional sensors are used, only a special antenna setup. In Fig. 2 the iterative process is presented. RSS measurements are collected simultaneously with a setup of four directional antennas horizontally headed in orthogonal directions. Details on the antenna setup are presented in Sect. 3.1. The RSS measurements $\mathbf{r}_t(x, y, \varphi)$ at time $t$ depend on the position $[x, y]$ and the attitude angle $\varphi$ of the setup. Using these measurements, at first the attitude of the setup is estimated, following the procedure descibed in Sect. 3.2. Secondly, the corresponding RSS values for a standard omni-directional Wi-Fi antenna are calculated, according to Sect. 3.3. Then, the position is estimated using fingerprinting, as introduced before in Sect. 2.

### *3.1 Setup with Directional Antennas*

Similar to [16] the RSS value $P_{Rx}(d, \varphi)$ in dB, received in a distance $d$ to a transmitter (access point) and with an attitude $\varphi$ is modeled as follows:

$$P_{Rx}(d, \varphi) = P_{Tx}(d_0) - 10\eta \log\left(\frac{d}{d_0}\right) + G_{Rx}(\varphi) - \sum_{k=1}^{K} n_k a_k \qquad (1)$$

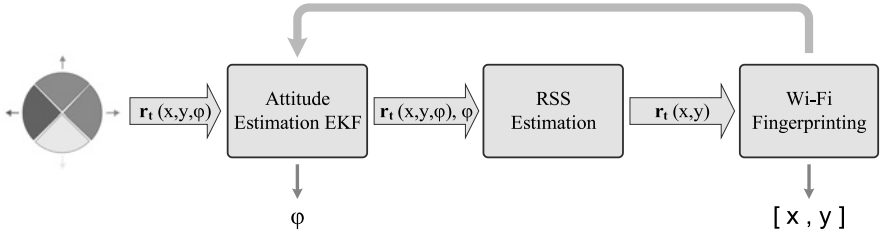



**Fig. 2** Iterative estimation process using RSS measurements of an antenna setup with four directional antennas

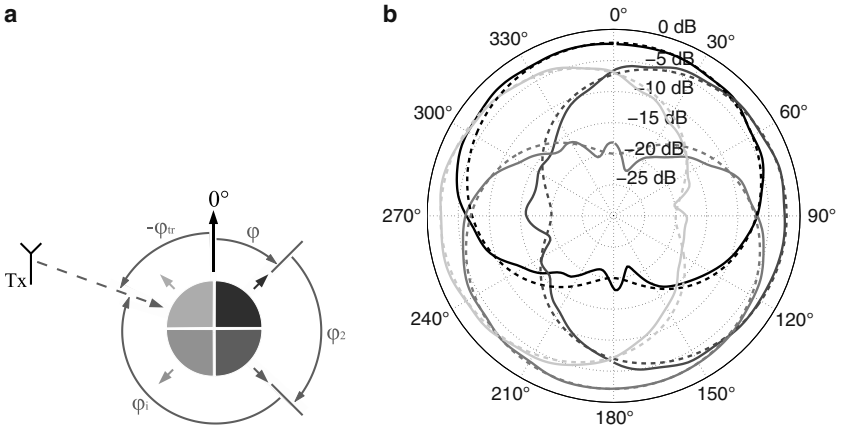



**Fig. 3** (**a**) Sketch of the antenna setup and the used angles for the directional antenna number $n = 2$, rotated with $\varphi_2 = 90°$ clockwise relative to the attitude of the antenna setup $\varphi$. (**b**) Polar diagram with measured (*solid lines*) and simulated antenna gains (*dashed lines*) for the proposed antenna setup with $\varphi_n \in [0° \; 90° \; 180° \; 270°]$

We get one RSS value per directional antenna. $P_{Tx}(d_0)$ is the reference signal strength measured in a distance $d_0$ to the transceiver. It is followed by the path loss, whereas $\eta$ is the path loss exponent, $G_{Rx}(\varphi)$ is the antenna gain. The last subtrahend represents $k$ different shadowing objects on the path of the signal. The objects are grouped to $n_k$ objects with equal attenuation $a_k$. Multipath propagation is not considered.

In Fig. 3a the used angles are presented. The angle of incidence $\varphi_i$ for one antenna is the angle between the positions of transceiver and receiver $\varphi_{\text{tr}}$ relative to a reference direction minus the attitude of the antenna setup $\varphi$ and the rotation angle $\varphi_n$ of the directional antenna $n$ of the setup that received the signal:

$$\varphi_i = \varphi_{\text{tr}} - \varphi - \varphi_n \tag{2}$$

As an example for an antenna setup, in Fig. 3b the polar diagram of a special, compact antenna setup is presented together with the results of an approximation of the antenna gains. It has been designed by the Fraunhofer Institute for Integrated Circuits IIS and is based on four dipole antennas and phase shifters. To approximate the directive gains $G_{Rx}(\varphi)$ of the setup a polar equation is used:

$$G_{Rx,n}(\varphi) = 10 \log (A + B \cos(\varphi_{\text{tr}} - \varphi - \varphi_n)) \tag{3}$$

$A$ is the isotropic part and $B$ the dipole part of the antenna, e.g. with $A = 1$ and $B = 0$ we get an omni-directional antenna, with $A = 0$ and $B = 1$ a dipole. For the presented antenna setup $A = 0.44$ and $B = 0.34$ are the best fit for our antennas.

### 3.2 Attitude Estimation Using an Extended Kalman Filter

To estimate the attitude $\varphi$ of the antenna setup differences of the RSS values, collected by the directional antennas, are used in an extended Kalman filter (EKF). An introduction to the EKF can be found in [17]. The main advantage of using the differences instead of the absolute values in (1) is that the reference signal strength, path loss and attenuation by objects on the direct path can be ignored:

$$P_{Rx,1}(d,\varphi) - P_{Rx,2}(d,\varphi) = G_{Rx,1}(\varphi) - G_{Rx,2}(\varphi) \tag{4}$$

Another positive effect is that the sensitivity increases by using the differences. The measurable range increases from 17 dB in Fig. 3b to 34 dB. This is important as commercial Wi-Fi cards output RSS values with a sample interval of 1 dB or multiples of that.

In the time update for the presented attitude estimation the a priori estimate of the attitude $\hat{\varphi}_k^-$ is calculated. It is assumed that there is no change between two time steps. The a priori estimate of the covariance is then simply the a posteriori covariance from the last step plus the process noise covariance:

$$\hat{\varphi}_k^- = \hat{\varphi}_{k-1} \tag{5}$$

$$P_k^- = P_{k-1} + Q \tag{6}$$

The random variable $w_k$ represents the process noise assumed with a normal probability distribution $p(w) \propto N(0, Q)$. The Kalman gain $\mathbf{K}_k$ is calculated according to [17]:

$$\mathbf{K}_k = P_k^- \mathbf{H}_k^H \left(\mathbf{H}_k P_k^- \mathbf{H}_k^H + \mathbf{R}_k\right)^{-1}, \quad \text{with} \quad \mathbf{H}_k = \left.\frac{\partial \mathbf{h}_k(\varphi)}{\partial \varphi}\right|_{\hat{\varphi}_k^-} \tag{7}$$

Whereas $\mathbf{R}_k$ is the measurement noise covariance matrix and $\mathbf{H}_k$ relates the state to the measurements. Using RSS differences the correlations between each pair of differences need to be considered in $\mathbf{R}_k$. Because $\mathbf{h}_k$ is a non-linear function, it is linearized by calculating the Jacobian matrix of the derivative of $\mathbf{h}_k$ at $\hat{\varphi}_k^-$. The non-linear measurement equation $\mathbf{h}_k(\varphi)$ consists of the six possible RSS differences. Vector $\mathbf{v}_k$ represents the measurement noise assumed with $p(\mathbf{v}) \propto N(0, R)$:

$$\mathbf{h}_k(\varphi) = \begin{bmatrix} G_{Rx,1}(\varphi) - G_{Rx,2}(\varphi) \\ G_{Rx,1}(\varphi) - G_{Rx,3}(\varphi) \\ G_{Rx,1}(\varphi) - G_{Rx,4}(\varphi) \\ G_{Rx,2}(\varphi) - G_{Rx,3}(\varphi) \\ G_{Rx,2}(\varphi) - G_{Rx,4}(\varphi) \\ G_{Rx,3}(\varphi) - G_{Rx,4}(\varphi) \end{bmatrix} + \mathbf{v}_k \tag{8}$$

Finally, in the measurement update the a posteriori state $\hat{\varphi}_k$ and covariance $P_k$ are calculated using the differences of measured RSS in the measurement vector $\mathbf{z}_k$:

$$\hat{\varphi}_k = \hat{\varphi}_k^- + \mathbf{K}_k \left(\mathbf{z}_k - \mathbf{h}(\hat{\varphi}_k^-)\right) \tag{9}$$

$$P_k = P_k^- - \mathbf{K}_k \mathbf{H}_k P_k^- \tag{10}$$

In the presented EKF, one iteration step is performed for each detected access point with each measured $\mathbf{r}_t$. So, more than one iteration step is executed per time step. In practice this showed a better performance than using all detected access points in one step. A deeper analysis follows in future work.

### *3.3 Omni-directional RSS Estimation*

Usually, the fingerprinting database is collected with omni-directional Wi-Fi antennas. To be able to use the comprehensive fingerprinting database of the Fraunhofer IIS, we calculate an equivalent $\mathbf{r}_t$ vector for a virtual omni-directional antenna. $r_{t,i}$ for access point $i$ is calculated by the mean of the results for all four directional antennas using the estimated attitude $\varphi$:

$$r_{t,i}(x, y) = \frac{1}{4} \cdot \sum_{n=1}^{4} \left(r_{t,i}(x, y, \varphi) - 10 \log\left(A + B \cos(\varphi_{tr,i} - \varphi - \varphi_n)\right)\right) \tag{11}$$

Finally, the position $[x, y]$ can be calculated using fingerprinting, as in Sect. 2.

## 4 Simulation Results

In the simulations presented, the performance of the novel approach is evaluated and the positioning errors are analyzed when measured RSS values vary from the database entries. We used a regular grid with a spacing of 1 m as our database. In Fig. 4 the size of the grid and the positions of four access points are depicted. Four more access points have been placed in a distance of 10 m away of the edges of the simulated room. To build up the fingerprinting database for each grid position and each access point we calculated $P_{Rx}(d)$ according to (1), but for an omni-directional antenna ($G_{Rx}(\varphi) = 0$).

One simulation path, as depicted in Fig. 4, is divided into 45 RSS measurements $\mathbf{r}_t(x, y, \varphi)$. It starts at position $[5, 5]$ going up to $[5, 25]$ following the gray arrows. In the corners the attitude has been rotated in 90° steps to simulate a person in a museum looking at different objects. Finally, the path ends at the starting position.

To simulate RSS measurements database entries are perturbated in two ways. At first, RSS variations for each access point $i$ are created by adding random noise to database entries:

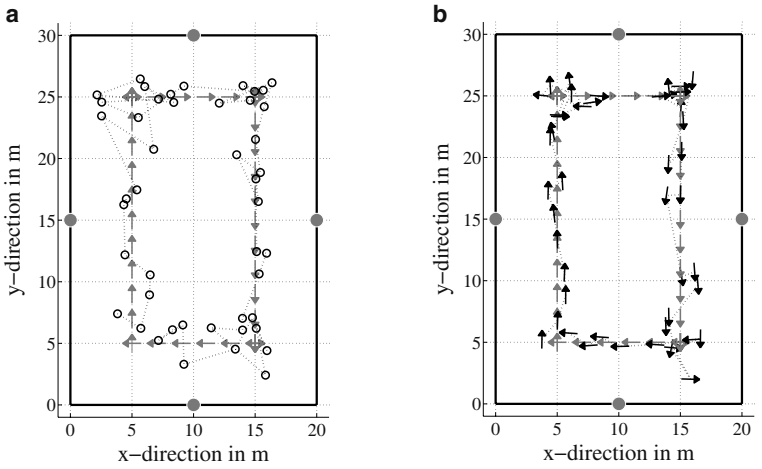



**Fig. 4** Layout of the simulated area with access points (*gray dots*), reference path with attitudes (*gray arrows*) and (**a**) results from standard Wi-Fi positioning (*circles*), and (**b**) results using the new attitude and position tracking (*black arrows*), with $\sigma_\mathrm{o} = 1\,\mathrm{dB}$ and $\sigma_\mathrm{d} = 1\,\mathrm{dB}$

$$r_{t,i}(x,y) = P_{Rx,i}(d) + v_\mathrm{o}\ , \quad \text{with} \quad p(v_\mathrm{o}) \propto N(0,\sigma_\mathrm{o})\ . \tag{12}$$

These RSS variations $v_\mathrm{o}$ affect all antennas in the same way. Therefore, we call them omni-directional RSS variations. In reality this can represent environmental changes. The RSS variations are local, but local correlations between measurements have not been accounted for.

Secondly, we simulated directional RSS variations by adding random noise from four directions $\varphi_\mathrm{d}$, the four antenna directions for simplicity. Directional RSS variations can be caused by small environmental changes, affecting just a apart of the signal propagation paths. For an omni-directional antenna these four directional variations superpose additive:

$$r_{t,i}(x,y) = P_{Rx,i}(d) + \sum_{n=1}^{4} v_{d,n}\ , \quad \text{with} \quad p(v_{d,n}) \propto N(0,\sigma_\mathrm{d}) \tag{13}$$

For each directional antenna the RSS variations are calculated adding directive noise, modeled using (3):

$$r_{t,i}(x,y,\varphi) = P_{Rx,i}(d,\varphi) + (A + B\cos(\varphi_d - \varphi - \varphi_n)) \cdot v_{d,n} \tag{14}$$

In Fig. 4a the results of one simulation run for standard Wi-Fi positioning (circles) using an omni-directional antenna is presented together with the reference path and the reference attitudes (gray arrows). The positioning results are distributed around the path. In Fig. 4b the results are presented for the new attitude and position tracking (black arrows) using the setup with four directional antennas.

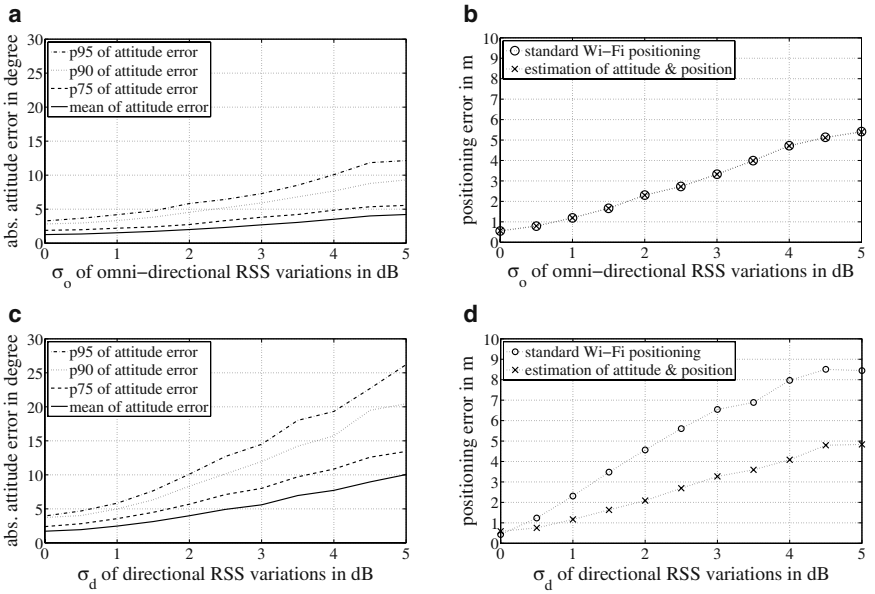



**Fig. 5** Mean values of absolute attitude errors in degree for omni-directional (**a**) and directional (**b**) RSS variations and the corresponding mean values of positioning errors using the novel approach and for comparison standard Wi-Fi positioning (**c**, **d**)

To get more general results when applying stochastic RSS variations, processing of the path has been repeated 20 times with the same values of $\sigma_o$ or $\sigma_d$. The influences of $\sigma_o$ and $\sigma_d$ are studied separately, the other parameter is set to zero. In Fig. 5a the attitude estimation results for various $\sigma_o$ and in Fig. 5c for various $\sigma_d$ are presented. The corresponding positioning results are shown in Fig. 5b and d together with the positioning results of standard Wi-Fi positioning, for the purpose of comparison.

For various $\sigma_o$ the positioning results are nearly the same and the attitude error is quite small. The attitude error increases with higher $\sigma_o$, because the error when calculating the angle between the transceiver and receiver position $\varphi_{tr}$ increases with increasing positioning errors.

On the one hand, with directional RSS variations $\sigma_d$ in Fig. 5c,d the attitude estimation errors are higher: 95% of all errors stay below 30°. So, the directional RSS variations have a bigger impact on attitude estimation.

On the other hand, the positioning errors with the novel approach are almost half of standard fingerprinting. The reason is, that the virtual omni-directional RSS values from (11) are more precise, because the measured RSS values of the directional antennas are not affected in the same way by the directional noise. So, the EKF filters the measurements and estimates a good attitude candidate. Hence, the presented approach for attitude and position estimation is more robust to directional RSS variations.

## 5 Experimental Results

In this section a proof of the introduced concept with real data is presented. The data has been collected at the "Museum Industriekultur" in Nuremberg. Results from our simulations are compared with results from the experimental setup. RSS measurements have been collected with an omni-directional and a directional antenna at 45 positions in a room for special exhibitions at the museum.

As the special antenna setup from Sect. 3.1 is not yet ready for Wi-Fi transmission a commercial directional antenna [8] has been used. It has been rotated manually at every measurement position. At first, the data for the fingerprinting database has been collected at each position with the omni-directional antenna. Secondly, RSS measurements at 33 positions on a path as depicted in Fig. 6 have been performed with both antennas. We used the omni-directional measurements for standard Wi-Fi positioning. In Fig. 6a positioning results using standard Wi-Fi fingerprinting are shown. Because the test setup is one single room without much shadowing objects, a challenging area for Wi-Fi positioning, the positioning error is higher than in Fig. 4a from simulations. The RSS measurements of the directional antennas have been used to test the attitude estimation with the EKF. The attitude estimation results of one run are depicted in Fig. 6b. The corresponding attitude and positioning errors are presented in Fig. 7.

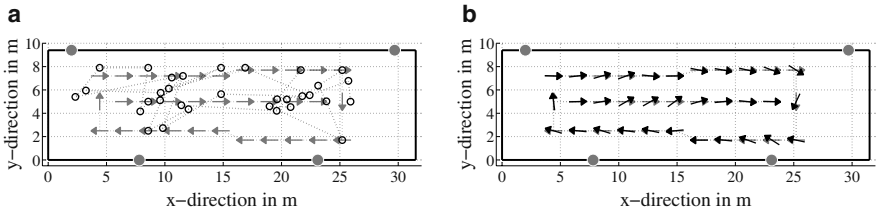



**Fig. 6** Room with access points (*gray dots*), reference path (*gray arrows*), (**a**) results from standard Wi-Fi positioning (*circles*) and (**b**) attitude estimation with the EKF (*black arrows*)

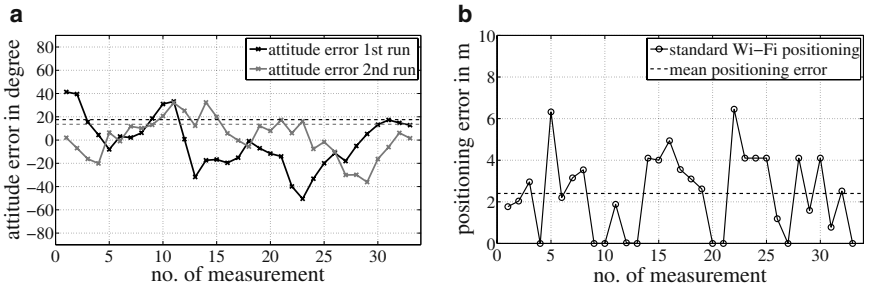



**Fig. 7** Absolute attitude errors (*solid lines*) and mean absolute attitude errors (*dotted lines*) in degree (**a**) for two measurement runs using the EKF and (**b**) the corresponding positioning errors (*solid line*) and the mean positioning error (*dotted line*) using standard Wi-Fi positioning

So far, the approach for attitude and position tracking has not been integrated in awiloc® of the Fraunhofer IIS. Therefore, attitude and position estimation have been tested separately. For attitude estimation, the reference positions instead of the estimated positions have been used to calculate $\varphi_{\text{tr}}$. So, the arrows representing the attitude in Fig. 6b are at the reference positions.

In this first test with an experimental setup the absolute attitude errors are less than 50° with mean values lower than 20°. This is promising enough for a museum visitor to find exhibits easily. The Wi-Fi positioning errors are typical for indoor environments with a mean error of 2.3 m. The possible improvement of positioning accuracy, according to Fig. 5d of the simulation results, using simultaneous attitude and position tracking will be subject of future work. Further information on RSS variations compared to fingerprinting database entries and a device calibration approach for Wi-Fi positioning can be found in [15].

## 6 Conclusions

In this work the state of the art Wi-Fi positioning in indoor environments has been expanded. RSS measurements are collected simultaneously with four directional antennas in a special antenna setup. An extended Kalman filter has been used for robust attitude estimation. For positioning standard Wi-Fi fingerprinting can be used with existing fingerprinting databases.

Simulation results demonstrate that with the novel approach the accuracy of positioning can be improved in presence of directional RSS variations. This leads to higher robustness in position tracking. The mean value of the absolute attitude errors stayed below 10° and the positioning errors below 5 m even with large RSS variations. Tests in a museum proofed the feasibility of the concept. The mean of the absolute attitude errors stayed below 20° in this case.

The presented approach for attitude and position tracking in indoor environments addresses the automation of electronic museum guides. In future work the novel approach will be integrated in awiloc® of the Fraunhofer IIS and the special antenna setup will be build for simultaneous measuring. Then, the concept will be evaluated in dynamic tests with users in museums. Furthermore, the approach will be combined with existing movement models and probabilistic concepts for pedestrian navigation, using e.g. hidden Markov models, to increase the accuracy of attitude and position tracking.

## Acknowledgements

This work is funded by the Bavarian ministry of economic affairs, infrastructure, transport and technology in the scope of the strategic program “Bayern 2020”. The authors want to thank Lucila Patiño-Studencka and Tobias Deißler for supporting and reviewing this paper.

## References


1. awiloc (2010) awiloc is a trademark of the Fraunhofer Institute for Integrated Circuits IIS, Germany. http://www.awiloc.com/
2. Bahl P, Padmanabhan V (2000) Radar: an in-building rf-based user location and tracking system. In Proceedings on INFOCOM the 19th Annual Joint Conference of the IEEE Computer and Communications Societies, Tel Aviv, Israel
3. Castro P, Chiu P, Kremenek T, Muntz R (2001) A probabilistic room location service for wireless networked environments. In Proceedings on UBICOMP the 3rd International Coference on ubiquitous computing, Atlanta, GA, USA. Springer
4. Haeberlen A, Flannery E, Ladd A, Rudys A, Wallach D, Kavraki L (2004) Practical robust localization over large-scale 802.11 wireless networks. In Proceedings on MobiCom the 10th annual international conference on mobile computing and networking, Philadelphia, PA, USA
5. Ibach P, Hübner T, Schweigert M (2004) Magicmap – kooperative Positionsbestimmung über wlan. In Proceedings on the Chaos Communication Congress, Berlin, Germany
6. Jahn J, Batzer U, Seitz J, Patiño Studencka L, Gutiérrez Boronat J (2010) Comparison and evaluation of acceleration based step length estimators for handheld devices. In Proceedings on the 13th International Conference on Indoor Positioning and Indoor Navigation (IPIN), Zürich, Switzerland
7. Kraft E (2003) A quaternion-based unscented kalman filter for orientation tracking. In Proceedings on the 6th International Conference of Information Fusion, Cairns, Queensland, Australia
8. PA-2408A patch antenna: 2.4 GHz wavelan antenna 8 dbi, WiMo Antennen und Elektronik GmbH. http://www.wimo.com
9. Skyhook Wireless: http://www.skyhookwireless.com
10. Meyer S, Vaupel T, Haimerl S (2008) Wi-fi coverage and propagation for localization purposes in permanently changing urban areas. In Proceedings on IADIS the international Conference Wireless Applications and Computing, Amsterdam, The Netherlands
11. Schilit B, LaMarca A, Borriello G, Griswold W, McDonald D, Lazowska E, Balachandran A, Hong J, Iverson V (2003) Challenge: Ubiquitous location-aware computing and the place lab initiative. In Proceedings on the 1st ACM international workshop on Wireless mobile applications and services on WLAN hotspots, San Diego, CA, USA
12. Seitz J, Vaupel T, Jahn J, Meyer S, Gutiérrez Boronat J, Thielecke J (2010) A hidden markov model for urban navigation based on fingerprinting and pedestrian dead reckoning. In Proceedings on the 13th International Conference on Information Fusion, Edinburgh, United Kingdom
13. Seitz J, Vaupel T, Meyer S, Gutiérrez Boronat J, Thielecke J (2010) A hidden markov model for pedestrian navigation. In Proceedings on WPNC the 7th Workshop on Positioning, Navigation and Communication, Dresden, Germany
14. Teuber A, Eissfeller B (2006) Wlan indoor positioning based on euclidean distances and fuzzy logic. In Proceedings on WPNC the 3rd Workshop on Positioning, Navigation and Communication, Hannover, Germany
15. Vaupel T, Seitz J, Kiefer F, Haimerl S, Thielecke J (2010) Wi-fi positioning: System considerations and device calibration. In Proceedings on the 13th International Conference on Indoor Positioning and Indoor Navigation (IPIN), Zürich, Switzerland
16. Wallbaum M (2005) Indoor geolocation using wireless local area networks. Ph.D. thesis, Department of Computer Science, RWTH Aachen University
17. Welch G, Bishop G (1995) An introduction to the kalman filter. University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
18. Wi-Fi (2003) Wi-Fi is a registered trademark of the Wi-Fi Alliance. http://www.wi-fi.org/
19. Youssef M, Agrawala A (2008) The horus location determination system. Wireless Networks 14(3):357–374

# Motion Sensing: From Single Sensors to Sensor Networks

**Martin Rulsch, Christian Arzt, Sven Feilner, Simon Jablonski, Matthias Struck, Jinghua Zhong, Daniel Tantinger, Christian Hofmann, and Christian Weigand**

**Abstract** It is well known, that regular physical activity is an important factor for preserving the health status, the challenge is how to quantify it accurately. Similarly rehabilitation programs rely on physical exercises, where the best results can be achieved through daily training. But who monitors and evaluates exercise execution at home? Micro electro mechanical systems (MEMS) based accelerometers provide a technological solution for inexpensive monitoring systems. The required number of accelerometers within the monitoring system depends on the use case. Quantifying certain activities like walking or cycling can be achieved with only one sensor, while recognizing differences in movements requires more observations and thus a network of accelerometers.
We present some typical movement related applications. For these use cases single sensor and multi-sensor systems are compared with respect to their advantages, challenges and limitations. Even signal processing requirements differ from application to application. Two approaches are explained in detail: knowledge-based and model-driven algorithms. While knowledge-based algorithms rely on feature extraction and an inference machine, which infers high level information from these features, model-driven algorithms try to describe relations between collected data and movements.

## 1 Introduction

### *1.1 Motivation*

Physical activity on a regular basis plays a key role for healthy aging. It reduces the risk to suffer from severe diseases, like hypertension, diabetes mellitus type 2 or

Martin Rulsch (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
martin.rulsch@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_18, © Springer 2011

adipositas. In addition, elderly people benefit from a lower fall risk [9]. Hence, 30 minutes of moderate physical activities five days a week or 20 minutes of vigorous physical activities three days a week is suggested [4].

Beyond the importance of activity concerning physical abilities some studies have also shown preservation and even improvement of cognitive skills induced by physical activity. Yaffe et al. have shown that women with higher baseline physical activity were less likely to suffer from cognitive decline [23]. Analyzing 18 intervention studies Colcombe and Kramer came to the conclusion that fitness training significantly increases the performance of cognitive tasks [2].

Due to advancements in the field of MEMS (***M****icro-****E****lectro-****M****echanical* ***S****ystems*) based accelerometers, inexpensive mobile monitoring systems are feasible. Such a system can be used for several purposes:

- Activity recognition
- Fall detection
- Movement reconstruction
- Energy expenditure estimation

Quantifying reliably physical activity and predicting the energy expenditure is a topic of increasing interest: for example, the knowledge of a quantitative measure of physical activity is essential in scientific studies concerning the prevention and treatment of obesity [7, 11, 20].

## *1.2 State of the Art*

Since the mid-nineties breakthrough of MEMS-based accelerometers in the automotive market a lot of new applications emerged. An example is the estimation of energy expenditure. Traditionally, two methods are used to assess energy expenditure: indirect calorimetry and doubly labeled water method. By the former $CO_2$ and $O_2$ volume are measured in the inhaled and exhaled air using a mask or a hood [8]. This approach is suitable for short term measurement but is inconvenient for assessing the energy expenditure of daily activities. By the doubly labeled water method, oxygen and hydrogen are replaced in water by uncommon isotopes of these elements [6, 14]. Because of the uncommon isotopes, it is called labeled. Oxygen leaves the body through water loss and exhaled air. Differently, hydrogen leaves the body only through water loss. Therefore, energy expenditure can be estimated with at least two samples of body water, e.g. urine or blood. The first sample is taken after the labeled water reaches equilibrium in the body and the second sample after a certain time span. The labeled water approach is expensive and only applicable for long term measurements where the average metabolic rate is of interest.

Data from accelerometers provides another indirect source for estimating energy expenditure. The use of accelerometers is less expensive than both traditional approaches. In addition, accelerometers can be used for short and long term observations and provide an energy expenditure history for the whole study. However, this

approach is not as accurate as indirect calorimetry and doubly labeled water. First accelerometer approaches used a single equation for mapping acceleration to energy expenditure. This worked for single activities, which were used for calibration (e.g. walking), but lead to over- or underestimation for different movement patterns (e.g. cycling) [3, 5].

Today most accelerometer based activity monitors recognize various classes of daily activities and use for each class a different energy expenditure estimator [1]. Common activity classes are resting, walking, running, cycling and stair climbing.

Accelerometers can also be used for movement reconstruction and for creating mobile monitoring systems. Due to the small size of MEMS-based accelerometers it is also feasible to integrate these systems into clothes. The accuracy of MEMS-based systems which rely on accelerometers, gyroscopes and magnetic field sensors can be compared with optical and electromagnetic systems [13, 19].

## 2 Single Node Sensors

With the help of a tri-axial acceleration sensor, movements can be tracked in all three spatial directions. A basic system for motion detection is based – besides the acceleration sensor – on a microcontroller, a power supply and a radio module in order to transfer the acquired data to a host. Such a sensor module can be attached to any point of the human body and provides information about the orientation.

By the use of sensors based on MEMS technology the development of very small modules is feasible. The decisive factors influencing the total size of the module are the desired duration and kind of measurement, which determine the required battery capacity. According to the use case it has to be decided whether raw data has to be continuously streamed or event triggered data packages have to be sent only if a event occurs.

An example for a single node sensor system is the activity monitor ActiSENS. Primarily, the module was developed to grasp everyday activities like *walking*, *riding a bike* and *climbing stairs*. In addition to the three-axis acceleration sensor, ActiSENS also contains a barometer in MEMS technology. The barometer complements the movement information determined by the acceleration data with information about the altitude.

A microcontroller represents the central data acquisition and processing unit, in which a sophisticated algorithm (cf. Sect. 6.2) has been implemented. Acceleration and pressure sensors are read out via a SPI-BUS. The acquired data is sent to an end device via Bluetooth.

Furthermore, the acquired data and therefore the activities of the user can be saved on a flash-memory chip, e.g. in the case of an interrupted Bluetooth-connection. To be able to assign the activities to an exact time, a real-time clock was integrated. Thereby, the user can recognize interruptions during the measurements, for example, caused by a failure of the power supply.

## 3 Sensor Networks

For a more accurate acquisition and analysis of complex movements a single sensor is in general not sufficient. By using an acceleration sensor network it is possible to acquire movement data at different points on the body at the same time. Such a network should satisfy special needs regarding the acquisition of movement:

- Reliable synchronization of network nodes
- Solid transmission of data
- High-rate sampling of movement demands a high throughput.

The last item is essential for this application. The minimal necessary throughput of the communication channel $R_b$ in bps can be calculated roughly by

$$R_b = N \cdot D \cdot f_s$$

where $N$ is the number of sensor nodes, $D$ is the maximum amount of data in bits to be transfered and $f_s$ represents how often a node is queried by the master per second.

One can distinguish between a network with sensor nodes transmitting data wirelessly and a wired network with several sensor nodes connected to a master module, which transmits the data wirelessly to a host system. Even though a wired network must face the issue of interconnection, its benefits prevail the ones of the wireless approach:

- Very easy synchronization of single sensor nodes
- Reliable transmission of data
- A central power supply
- Only one wireless transmission module necessary.

Based on this insight, the so called Fitness Assistant was realized. It consists of a wired acceleration sensor network that was completely integrated into a jacket. The basic structure of the system can be seen in Fig. 1. The design mainly consists of a master module and several sensors in the jacket. The master module includes all essential components to control the network. The communication with a host is realized by a Bluetooth module. A SD-card offers the possibility to store the data. A single acceleration sensor is already integrated into the master.

All components are controlled by a microcontroller. The communication between the master module and the acceleration sensors occurs via a SPI-Interface. One sensor node of a branch (Fig. 1, right side) basically consists of an acceleration sensor and plugs for an electrical connection to the SPI-Interface, to the address-bus and to the power supply lines. All signals and power lines are passed to the subsequent sensor node. Each node is addressed by a separate control line. The SPI-interface is clocked with a frequency of 1 MHz, therefore the theoretical throughput is 1 Mbps. Considering 11 sensors and 54 bits of data per sensor (commands, acceleration data), the theoretical sampling rate would be approximately 1600 Hz. Due to overhead and data processing activities it drops to 170 Hz in maximum.

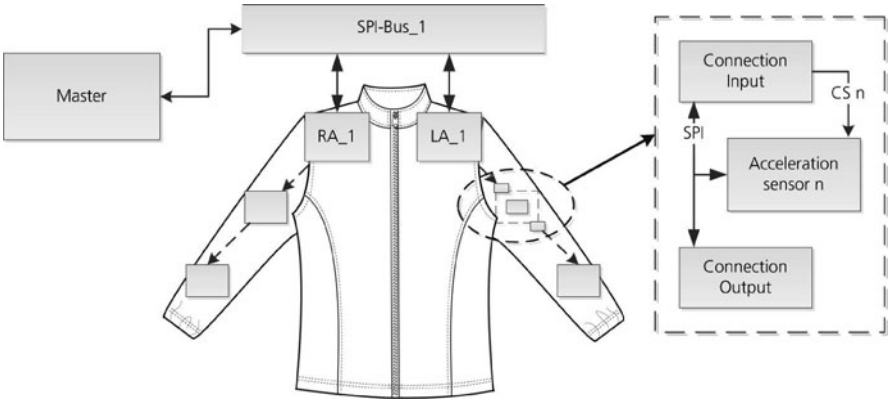



**Fig. 1** Basic structure of the activity recognition system

The textile carrier of the sensor network consists of a tight and flexible sports jacket. The master module is attached between the shoulder blades. The sensor nodes are mounted in a textile tube inside the jacket which runs from the master module along the arms to the wrists. It is possible to place the sensor nodes at almost any position along the tube. This design allows an easy replacement of broken components and the complete removal of the sensor network. Hence, the jacket can be washed easily.

## 4 Knowledge-Based Paradigm

Unlike pattern recognition algorithms, knowledge-based methods strictly separate between the so-called knowledge base declaratively describing the knowledge about the specific domain, and the so-called inference component or inference engine that tries to derive answers from the underlying knowledge base. That is to say the knowledge base can be replaced without changing the concrete inference machine. Hence, the developed activity recognition system offers a number of advantages. For instance, it is possible to use different acceleration sensors and different sensor positions on the body or adapting the system to different areas of application by just changing the knowledge base without loss of recognition accuracy. Previous works also used knowledge-based systems but focused on activity recognition systems that have to be trained for each user individually. For more details see [17].

## 5 Model-Driven Algorithms

The Kalman Filter, an optimal filter for linear systems, is the origin for numerous model-driven signal processing approaches. The most popular modification for non-

linear systems is the Extended Kalman Filter (EKF). An alternative for nonlinear systems provides the Unscented Kalman Filter (UKF) [21].

# 6 Applications

## *6.1 Fall Detection*

Automatically detecting falls with one single tri-axial accelerometer is a typical and important application. The implemented system uses the paradigm of knowledge-based methods briefly described in Sect. 4. The main part of the algorithm is based on a fuzzy-logic inference system and a neural network [15, 16]. Both described methods were integrated into the telemedical system published in [18, 22].

## *6.2 Activity Classification*

By using a single node sensor and a knowledge-based algorithm it is possible to distinguish between several activities of daily life, e.g. inactivity, walking, cycling and climbing stairs up and down.

An algorithm described in [12] requires four features for this distinction: mean acceleration (MA), change in altitude ($\Delta h$) and the energy of the acceleration signal below 3 Hz ($E_{<3}$) and above 3 Hz ($E_{\geq 3}$). The mean acceleration is computed from a high pass filtered acceleration signal, where gravity is suppressed. It was developed for the ActiSENS, comprising a tri-axial accelerometer and a barometer. The features are updated in four second intervals. While the features MA, $E_{<3}$ and $E_{\geq 3}$ are computed from the accelerometer data, the barometer is more appropriate for recognizing changes in altitude. Applying the barometric formula it is possible to compute a change in altitude from a change of atmospheric pressure.

The classification is performed using a decision tree [12]. The decision rules for the classes inactivity, walking, cycling and climbing stairs up and down are:

1. Check MA, if it is below an acceleration of 0.1 g, the person is inactive.
2. Apply cycling criterion, if it is fulfilled, the person is cycling.
3. If cycling criterion does not match, the person is walking.
4. At last check the change in altitude ($\Delta h$), if it is below $-0.75$ m or above 0.75 m in the 4 s interval, the person is walking stairs up or down.

The cycling criterion is based on the features $E_{<3}$ and $E_{\geq 3}$. While the power spectrum for walking has a peak near the walking frequency, the power spectrum for cycling is more regular. Therefore, feature $E_{<3}$ is much larger for walking than for cycling. This approach is resource efficient and can be implemented on small embedded systems like the sensor node itself.

Classification rates for the algorithm are shown in Table 1. They were obtained using a data set of 12 persons performing activities of daily life.

**Table 1** Classification rates for inactivity, walking, cycling, climbing stairs up and down using ActiSENS

| Activity | Classification Rate |
|---|---|
| Inactivity | 97.3% |
| Walking | 92.6% |
| Cycling | 82.2% |
| Upstairs | 66.8% |
| Downstairs | 65.7% |

## *6.3 Movement Reconstruction*

Using accelerometers it is feasible to reconstruct human movements, e.g. movements of the upper arm. Despite the complexity of the shoulder joint it is appropriate to model the shoulder joint with a spherical joint [10]. Neglecting tissue effects the upper arm can be modeled with a straight line. Complexity of the model is further reduced by forbidding movements of the upper body. With these constraints three angles $\Theta$, $\Phi$ and $\Psi$ are sufficient to describe upper arm movements (Fig. 2):

- Angle $\Theta$ is the angle between trunk and arm. Altering this angle raises or lowers the arm.
- Angle $\Psi$ reflects a twist of the arm.
- Angle $\Phi$ is the angle of the arm in the horizontal plane. Altering this angle moves the arm to the left or to the right.

Most accelerometers measure gravity and dynamic acceleration from movements. The angles $\Theta$ and $\Psi$ can be computed from the measured gravity. Angle $\Phi$ requires an evaluation of the dynamic accelerations and can not be reconstructed reliably using accelerometers only. Using the dynamic accelerations requires to integrate twice, which leads to a fast growing error over time. Either way raising, lowering and twisting the arm can be reconstructed. A change of the angles $\Theta$ and $\Psi$ can

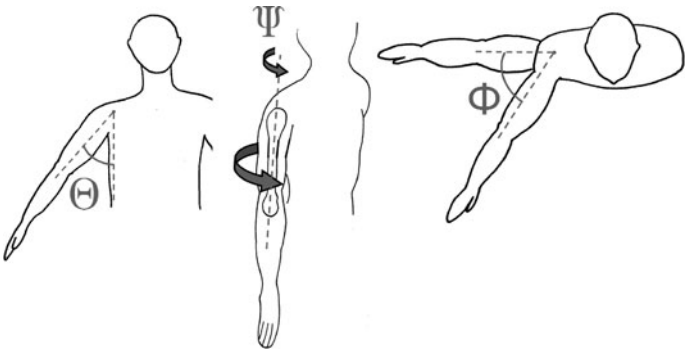



**Fig. 2** Parametrization of upper arm movements with three angles $\Theta$, $\Psi$ and $\Phi$: angle $\Theta$ rises and lowers the arm; angle $\Psi$ twists the arm and angle $\Phi$ moves the arm in the horizontal plane

be observed in the measured gravity. A sensor placed on the straight line which represents the arm with distance $r$ from the shoulder joint moves as follows:

$$\vec{p}_{\mathrm{s}} = R_{\Phi} R_{\Theta} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} .$$

Matrix $R_{\Phi}$ is a rotation by angle $\Phi$. Similarly, $R_{\Theta}$ represents a rotation by angle $\Theta$. The observed acceleration $\vec{a}_{\mathrm{w}}$ in point $\vec{p}_{\mathrm{s}}$ is the second derivation of point $\vec{p}_{\mathrm{s}}$ with respect to time:

$$\vec{a}_{\mathrm{w}} = \ddot{\vec{p}}_{\mathrm{s}} .$$

The world coordinate system, in which $\vec{a}_{\mathrm{w}}$ is measured, does not coincide with the accelerometer coordinate system due to the rotating sensor. They can be aligned using the matrices $R_{\Phi}$ and $R_{\Theta}$. The resulting acceleration measured by the accelerometer is:

$$\vec{a}_{\mathrm{s}} = R_{\Theta}^{-1} R_{\Phi}^{-1} (\vec{a}_w + \vec{g}) .$$

Vector $\vec{g}$ is the measured gravity in the initial position. A twist of the upper arm leads to a further rotation:

$$\vec{a} = R_{\Psi}^{-1} \vec{a}_{\mathrm{s}}$$

This is the acceleration measured by a sensor on the upper arm. The equation can be used as a sensor model in an extended or unscented kalman filter to compute the angles $\Theta$ and $\Psi$. Otherwise, it can be used to compute the angles directly.

## 7 Discussion and Outlook

Based on a single sensor system for motion detection, the article presented the development of a network consisting of several sensor nodes. This system has its limitations concerning the sampling rate, number of sensors and the wearing comfort.

In order to detect complex movements a high sampling rate and a large number of sensor nodes with – for example – accelerometers and gyroscopes are necessary. By using a more robust signal transmission technique, e.g differential signaling, and proper hardware design, e.g. less signal lines, the behavior regarding cross talk, reflections and signal to noise ratio could be improved, which could increase the theoretical throughput. To increase the wearing comfort as well as to ease the handling of the system it is necessary to integrate the sensor nodes in an unobtrusive way. As a consequence, the reduction of weight and size of each sensor node is inevitable to achieve this goal. Flexible printed circuit boards or implementing the components on an application-specific integrated circuits (ASIC) would be some solutions.

The model as it is described in Sect. 6.3 only addresses upper arm movement.

It is a matter of requirements to judge which movement complexity should be achievable with the model (forearm, hand, legs et cetera). Another target is to remove sensor drift which is a central problem by reconstructing horizontal arm movement. Therefore, it could be an efficient step to extend the accelerometer by additional sensors, e.g. gyroscopes or magnet sensors to get an independent source for horizontal position.

Summarized, this article presented and described several approaches for motion capturing and movement reconstruction based on commercial accelerometers. As customary for each specific application, a tradeoff between accuracy, number of sensors and computational costs has to be found. On the one hand the paradigm of knowledge-based methods, described in Sect. 4, provides a powerful, general and flexible activity classification concept, however, implementing those methods on a microcontroller is a challenging task. In order to reconstruct complex movements, e.g., for rehabilitation purposes, multiple sensor nodes and advanced reconstruction models are indispensable. The main advantages of using accelerometers instead of optical methods are relatively low costs and mobility.

## Acknowledgements

This work was partially funded by the Bavarian Research Foundation and the German Federal Ministry of Education and Research.

## References

1. Choi J, Lee J, Hwang H, Kim J, Park J, Shin K (2005) Estimation of activity energy expenditure: accelerometer approach. In IEEE-EMBS 2005. 27th Annual International Conference of the Engineering in Medicine and Biology Society, pp 3830–3833
2. Colcombe S, Kramer A (2003) Fitness effects on the cognitive function of older adults. Psychological Science 14(2):125–130
3. Crouter S, Churilla J, Bassett D (2006) Estimating energy expenditure using accelerometers. European journal of applied physiology 98(6):601–612
4. Haskell W, Lee I et al (2007) Physical activity and public health. updated recommendation for adults from the american college of sports medicine and the american heart association. Circulation 116:1081–1093
5. Hendelman D, Miller K, Baggett C, Debold E, Freedson P (2000) Validity of accelerometry for the assessment of moderate intensity physical activity in the field. Medicine & Science in Sports & Exercise 32(9):442–449
6. Lifson N, Gordon G, Visscher M, Nier A (1949) The fate of utilized molecular oxygen and the source of the oxygen of respiratory carbon dioxide, studied with the aid of heavy oxygen. Journal of Biological Chemistry 180(2):803–811
7. Lovejoy J, Champagne C, Smith S, de Jonge L, Xie H (2001) Ethnic differences in dietary intakes, physical activity, and energy expenditure in middle-aged, premenopausal women: the healthy transitions study. The American journal of clinical nutrition 74(1):90–95
8. Matarese L (1997) Indirect calorimetry: Technical aspects. Journal of the American Dietetic Association 97(10):154–S160

9. Nelson M, Rejeski W, Blair S, Duncan P, Judge J, King A, Macera C, Castaneda-Sceppa C (2007) Physical activity and public health in older adults. recommendation from the american college of sports medicine and the american heart association. Circulation 116:1094–1105
10. Prokopenko RA, Frolov AA, Biryukova EV, Roby-Brami A (2001) Assessment of the accuracy of a human arm model with seven degrees of freedom. Journal of Biomechanics 34(2):177–185
11. Reilly J, Jackson D, Montgomery C, Kelly L, Slater C, Grant S, Paton J (2004) Total energy expenditure and physical activity in young scottish children: mixed longitudinal study. The Lancet 363(9404):211–212
12. Rulsch M, Benz M, Arzt C, Podolak C, Zhong J, Couronné R (2009) A lightweight approach for activity classification on microcontrollers. In World Congress on Medical Physics and Biomedical Engineering, Munich, Germany, Springer, pp 190–193
13. Saber-Sheikh K, Bryant E, Glazzard C, Hamel A, Lee R (2010) Feasibility of using inertial sensors to assess human movement. Manual Therapy 15(1):122–125
14. Schoeller D (1988) Measurement of energy expenditure in free-living humans by using doubly labeled water. The Journal of nutrition 118(11):1278–1281
15. Struck M, Dinh C (2009) A new real-time fall detection approach using fuzzy logic and a neural network. In Proceedings of the 6th International Workshop on Wearable, Micro and Nano Technologies for the Personalised Health, pHealth
16. Struck M, Dinh C, Tantinger D (2009) Automatic emergency detection using commercial accelerometers and knowledge-based methods. In IEEE Computers in Cardiology Proceedings, pp 485–488
17. Struck M, Krassnig G, Tantinger D, Hofmann C, Wittenberg T (2010) User-friendly system for recognition of activities with an accelerometer. In Pervasive Computing Technologies for Healthcare (PervasiveHealth)
18. Struck M, Pramatarov S, Weigand C (2008) Method and system for standardized and platform independent medical data information persistence in telemedicine. In IEEE Computers in Cardiology Proceedings, pp 257–260
19. Thies SB, Tresadern P, Kenney L, Howard D, Goulermas JY, Smith C, Rigby J (2007) Comparison of linear accelerations from three measurement systems during “reach & grasp”. Medical engineering & physics 29(9):967–972
20. Trost SG, Kerr LM, Ward DS, Pate RR (2001) Physical activity and determinants of physical activity in obese and non-obese children. International Journal of Obesity 25(6):822–829
21. Wan E, Van Der Merwe R (2000) The unscented kalman filter for nonlinear estimation. In Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000, pp 153–158
22. Weigand C (2005) Use and Implementation of a Medical Communication Standard in Practice. In IEEE Computers in Cardiology Proceedings, pp 319–322
23. Yaffe K, Barnes D, Nevitt N, Lui L, Covinsky K (2001) A prospective study of physical activity and cognitive decline in elderly women. Arch Intern Med 161:1703–1708

# A Real-Time Tracking System for Football Match and Training Analysis

**Thomas von der Grün, Norbert Franke, Daniel Wolf, Nicolas Witt, and Andreas Eidloth**

**Abstract** In this paper a radio-based, real-time tracking system and its application in sports is presented. The system is capable of tracking simultaneously a large number of high dynamic objects in a pre-defined area of interest like a sports stadium. Position and sensor data are captured with high precision and at a high level of detail. Moreover, the position data is automatically analyzed and interpreted from a sports science point of view and presented in a graphical user interface to players, teams and coaches. As an example, a standardized performance test for footballers of the German football association DFB is implemented and presented here. Thus, many disciplines from radio frequency and communication technologies, media engineering, sports medicine and sports science as well as valuable input from sports professionals contributed to this research work. The paper summarizes the manifold results of the research work performed all along the closed processing chain from data capturing, analyzing, condensation and interpretation to its presentation in a 3-D graphical user interface.

## 1 Introduction and Motivation

Faster, higher, further is no longer the only maxim in sports. Sports has changed into a more dynamic, complex and very strategic competition and has become an immense branch of business, especially in football. Often crucial incidents during a game cannot be recognized neither by the human eye nor with state of the art camera techniques. Perceiving the overall situation in a team game on a tactical level is difficult because of high dynamics in the movements and due to often only minor distinctions in performance. Therefore, coaches, referees and spectators have to face these increasing challenges. Especially the world of media is looking for new ways

Thomas von der Grün (✉)
Fraunhofer IIS, Nordostpark 93, 90411 Nürnberg, Germany
thomas.vondergruen@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_19, © Springer 2011

to present sports to us and augment their content in a more attractive and exciting manner. One example is a 3-D animation in real-time to analyze the run of play or to explain certain events in a game from different points of view. In any case, sports has become an important business still increasing. Video-based tracking systems are state of the art for training and game analysis. In addition to the problems with low or blocked visibility, those systems need manual intervention and a long processing time to deliver data on a limited level of detail. The radio-based RedFIR® system of the Fraunhofer IIS is a powerful real-time analysis tool, which enables novel use cases as basis for an economic exploitation. It has been and will be generating a lot of exiting new subjects for research.

The starting point is the radio-based, real-time tracking system to capture the position and sensor data of football players and ball.

## 2 System Description

The system consists of a set of small, lightweight and miniaturized transmitters (objects to be located) and a receiving infrastructure that is set up around the area of interest, which may be the inner part of a stadium.

The functional principle of the system relies on the determination of time of arrival (ToA) values at each receiving site for every single modulated signal sequence originating from the miniature transmitters. Thus, one signal on the field leads to N ToA values in the case N receivers are installed. Using the knowledge that electromagnetic waves propagate with the speed of light, a central computing platform can then perform the second processing step in order to translate this timing information into positions according to a $xyz$-reference frame. In order to do so, N-1 time difference of arrival (TDoA) values are calculated between pairs of receivers which eliminates the need to know the time of transmission or in other words the need to use synchronized transmitters.

In the 2-dimensional case a hyperbolic curve describes the possible location of a transmitter between a pair of receiving antennas for one known TDoA value. This way, more TDoA values allow the determination of a precise location by multilateration (hyperbolic positioning). Two TDoA values by using three receivers are needed in the 2-D case and accordingly, three TDoA values by using four receivers resolve the 3-D case.

Following this functional principle, the system has been developed by Fraunhofer IIS up to the level of readiness for usage in stadium sports. This implementation has the following properties and characteristics:

The air interface employs the ISM (industrial scientific medical) band at 2.4 GHz that allows the usage of about 80 MHz while providing the advantage of few frequency regulation requirements worldwide. The miniature transmitters make use of this bandwidth by generating short but broadband signal bursts of pulse-shaped m-sequences. According to their characteristics, a number of different sequences with

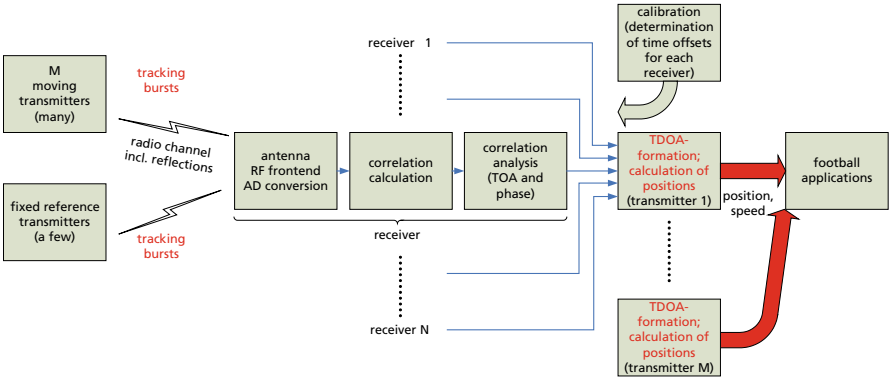



**Fig. 1** Signal processing chain including burst generation, radio channel, ToA determination and position calculation

suitable auto- and cross-correlation properties are available for a sufficient quantity of different tags to be located.

The system allows to receive overall 50 000 signal bursts per second of length 15 µs and to distinguish 144 miniature objects which operate at burst rates from below 200 to above 2000 bursts per second each.

After frequency down-conversion and complex sampling of the analog signal, the receiving side computes the ToA values by calculating correlations using the ideal form of the signals as a reference. Figure 2 shows one example of a correlation result in the form of its absolute value over time, while Fig. 3 displays its complex shape. The time resolution can be increased significantly by oversampling and interpolation. In order to find the ToA value, a characteristic point has to be found on the correlation curve that is most suitable and not prone to effects like multipath echoes. The given example shows that this may not be the case for the maximum.

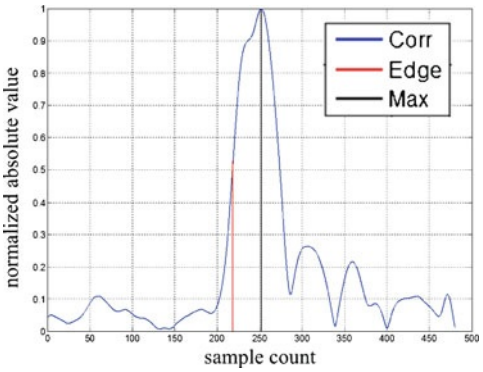



**Fig. 2** Absolute value of the complex correlation. Positions of maximum (*black*) and inflection point (*red*)

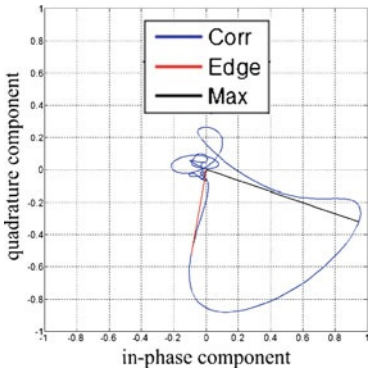



**Fig. 3** Complex correlation result depicted in IQ-representation

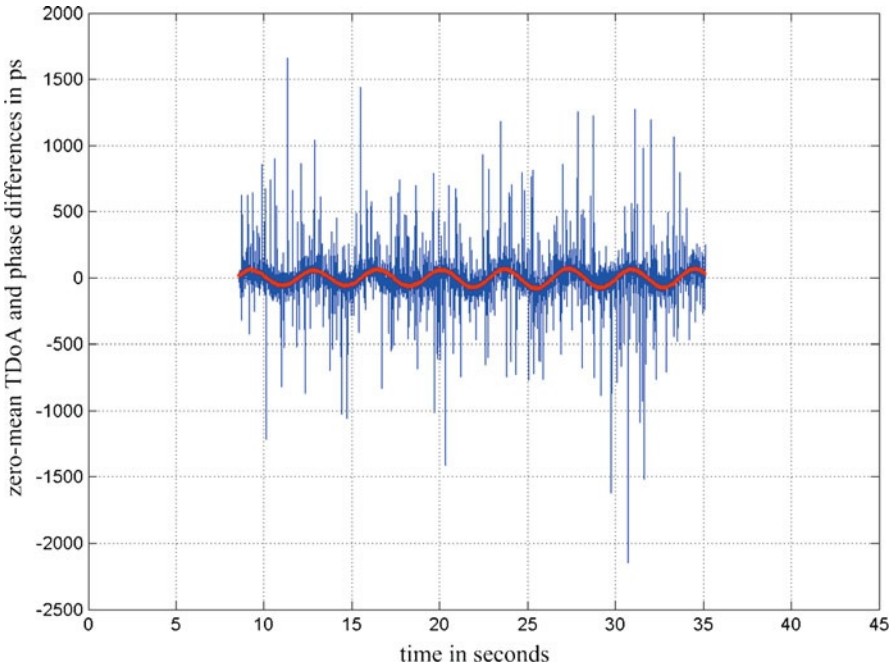



**Fig. 4** TDoA (*blue*) and carrier-phase difference (*red*) measurements of a still standing transmitter (one receiving antenna on a post slightly swaying in the wind)

Here, the maximum originated from a multipath echo and not from the direct line-of-sight path which is the only one of use. Therefore, it is advantageous to go for the inflection point on the rising edge.

Apart from ToA values, the phase of the vector leading to the characteristic point (the red line in Fig. 3) is the second valuable source of information. Between two receivers, also the phase difference of arrival values are exploited. With the receivers

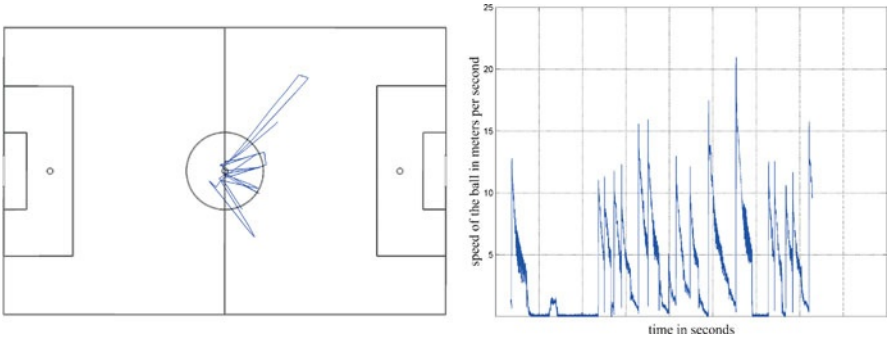



**Fig. 5** Ball track of passes between players (*left*) and absolute value of the speed of the ball (*right*)

being synchronized, a phase difference value will not change from burst to burst in the case the miniature object does not move. However, it will change if it moves towards or away from the antennas.

Therefore, changes of phase differences from burst to burst contain valuable information of the differential movement and the speed of the transmitter. Figure 4 is a practical result showing both ToA and phase (translated to timing information) results. Noise effects on the measured phases are significantly smaller compared to ToA values. Section 3.3 explains in greater detail the beneficial exploitation of phase information for position results.

In order to relate the absolute ToA values to the known position of the phase center of the receiving antennas, there are a few fixed transmitters with known position around the area of interest. These are used for the determination of calibration information, i.e. time-varying timing offsets, for each of the receiving antennas.

On the football pitch, the system achieves an absolute horizontal accuracy in the one digit centimeter range. Due to the use of phase information, the tracking of differential movements and speeds is noiseless and very precise. Figure 5 gives an example of a ball being passed between two footballers. It makes plausible that this kind of data for ball and player movements is an apt basis for automatic football event detection, statistics generation and sport scientific analysis.

## 3 Selected System Aspects

### *3.1 Miniaturized Transmitters*

In order to ensure as little disturbance as possible with the traditional way football trainings and games are conducted, the required tracking transmitters for balls and players need to be as lightweight and small as possible. This includes also that they become mechanically stable and suitable for ball integration without changing ball flight properties and player perception.

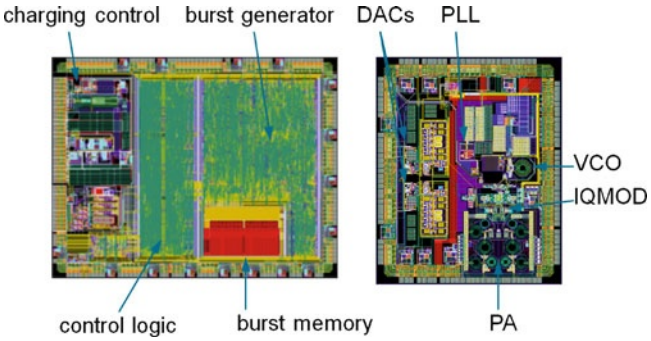



**Fig. 6** ASIC implementation of the transmitter functionality (BiCMOS 0.35 μm)

Application Specific Integrated Circuits (ASICs) have been developed in order to allow for a high level of integration of this 2.4 GHz programmable broadband signal generator functionality (Fig. 6).

Key characteristics are the ability to program arbitrary waveforms to be used as tracking signals, an application flow and power management logic that fully exploits the duty cycle to save battery capacity, and interfaces for additional sensors (e.g. inertial sensors, air pressure sensors in case of the ball, heart rate and breath rate sensors in case of the player).

The miniature transmitters have no wired contacts to the outside. Therefore, there is a 13.56 MHz wireless interface by an external coil that allows for both recharging of batteries and programming and configuring of the transmitter from the outside.

The bare dies of the ASICs are mounted using chip-on-board technology with bond wires directly from the semiconductor to the printed circuits board (PCB). Further components of the transmitter are encapsulating glue, flexible PCBs that include structures as radiating antenna elements, Lithium polymer batteries and housing elements (Fig. 7).

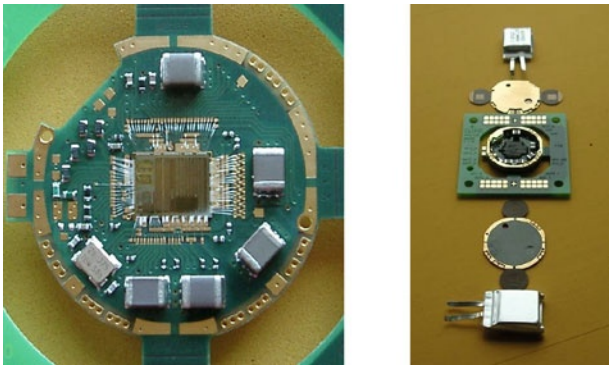


**Fig. 7** Ball transmitter packaging and build-up

The run time of the ball transmitter built in a 3-D shape is four hours at 2000 bursts per second using two 140 mAh batteries while having a weight of only 15 g. The player transmitter designed in a 2-D shape lasts three hours at 200 bursts per second powered by one 110 mAh battery.

### *3.2 Tracking-Kalman-Filter*

As in every Time-of-Flight (ToF) based localization system, multipath is potentially a major source of error [2–4]. For this real-time localization system, especially reflections of the transmitted signal on the ground or from metal advertising signs can be sources of measurement errors. In the worst case, the resulting range error could reach a few meters and might therefore have a strong influence on the result of the calculated position. Reducing the range error is one of the major areas of work within the project.

Especially the multipath error, as the predominant error source, became a well examined scientific topic. An estimator to identify the different propagation paths has been developed. Because of its membership to the tracking loop of the localization receiver, the estimator was named "Tracking-Kalman-Filter". The aim of the filter is to recognize and track all of the propagation paths, also if they are nearly vanishing behind other stronger paths. Every propagation path $i$ is modeled by five states. These are the delay time $\tau$, the amplitude $a$, the carrier phase $\phi$, and the two derivatives for delay time $\tau$ and carrier phase $\phi$.

$$x_i = [\tau_i, a_i, \varphi_i, \dot{\tau}, \dot{\varphi}]^{\mathrm{T}}. \tag{1}$$

The last two states are necessary to track the frequency offset between transmitter and receiver. For every detected propagation path one such representation of (1) is used. A complete state vector representation of $N$ paths can be written as

$$x = \left[x_1^{\mathrm{T}}, x_2^{\mathrm{T}}, \ldots, x_i^{\mathrm{T}}, \ldots, x_N^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{2}$$

From the first arrived path we directly derive the Line-of-Sight (LoS) measurement, i.e. the range between transmitter and receiver. This state model is used within an unscented Kalman filter (UKF), which models the complex valued output samples from the correlator (see Figs. 2 and 3). An estimate of one path of the measurement curve is calculated by the ideal correlation, which is delayed, scaled in amplitude and rotated by the carrier phase angle of the corresponding path estimate $x_i$. Afterwards the correlation curves from all paths of the state vector $x$ are summed up to represent the measured correlation curve.

Because propagation paths arise and vanish after a while when the transmitter is moving, a strategy for adding and dropping paths is necessary within the Tracking-Kalman-Filter. This strategy is driven by the error signal of the filter. If a not modeled path arises, it can be seen at first in the error signal. After it has proven to be

a new path, the state model is enlarged by a new five-entry path representation. If the path vanishes after a while, the amplitude of the modeled path goes below a given threshold. When this happens, the five-entry path representation is deleted from the state vector.

One of the advantages of this concept can be seen in the continuous tracking of the LoS path, which makes the system more stable and increases the reliability of the position calculations.

## *3.3 Carrier Phase Aided Position Calculation*

For position calculation not only range measurements derived from the modulation of the transmitted signals (ToA values) can be used. As already discussed, another source of information from the same measurement – the carrier phase – can be incorporated as well. This type of measurement is more accurate compared to the ToA values. The drawback is that the carrier phase measurements are not unique. Within this real-time localization system a carrier frequency $f_c$ of 2.445 GHz is used, which corresponds to a wavelength $\lambda$ of 0.123 m. Therefore the ambiguity of a measured distance derived from the carrier phase can only be given with an uncertainty of $N \cdot \lambda$, with $N$ being an integer number.

But nevertheless, there are some ways of using carrier phase measurements to calculate a more precise position result. Two cases can be distinguished:

- ToA values are used together with carrier phase values.
- Only carrier phase values are used.

For the second case a start position with an adequate quality is required.

The first case can be regarded as sensor fusion. The information of the absolute (i.e. unique) position is derived from the ToA values as usual. Two carrier phase measurements are taken from a receiver at two consecutive times. The length of this short time period of roughly 1ms or less must assure that the moving transmitter changes its position (or its distance to the receiver) at most by plus or minus half a wavelength $\lambda$. If this is the case, one can easily calculate the relative speed of the moving transmitter. If this pre-assumption does not hold, one can solve the ambiguity problem with a rough knowledge of the current speed of the moving object, but at this time with the assumption that the object does not change its speed vector very fast. Now, having the calculated speed between two time steps from every receiver, one can use this information in conjunction with the measured distances from the ToA values for calculating a carrier-phase smoothed position result. It shows a more detailed trajectory of the movements than a calculation with only ToA values.

The second case as listed above, neglects the range measurements from the ToA values within the position calculation process. As stated before, a rough start position has to be known in advance, to get the right position fix. The goal is to calculate the correct range between transmitter and receiver directly from the carrier phase measurements. Therefore, the integer ambiguity problem has to be solved.

Many ways are described in literature related to this topic [5–8]. After solving the ambiguity problem, the transmitter's position can be calculated and tracked by an extended Kalman-Filter, which uses only the distances calculated from the carrier phase measurements. The integer ambiguity solution is determined only once.

## 4 Event Detection and Sports Analysis

### *4.1 Motivation*

Providing real-time kinematic data in sports opens a wide field for automatic online and offline analysis, which can support all participators in sports like players, trainers and media. Ideal processing, combination and aggregation of position, velocity and acceleration data from all tracked objects (players, balls and even training equipment) can lead to a variety of high-level applications like optimal guidance of individual training, online match analysis enabling real-time reactions of trainers as well as supporting economic decisions through long term player efficiency analysis. The next sections give a short introduction on how the localization data captured with the system is prepared to enable applications suitable for end users.

### *4.2 Complex Event Processing (CEP) as the Key Architecture for Analysis*

The low-level kinematic data in the localization system can be transferred into a high-level representation through the combination with Complex Event Processing (CEP) to get sets of plausible and meaningful events. A comprehensive introduction to models and techniques in the field of CEP systems is done in [9].

To perform event detection, such systems are usually built on middleware in the context of event-driven architectures (EDA). Examples for events that can be detected in the football context are a pass from player A to player B or a shot on goal from player A. Even the stream of player positions can be seen as a stream of events. Certain events can be defined by making assumptions or descriptions about their emergence. These descriptions or definitions of events do not have to be written in source code of a programming language. [10] and [11] give examples of methods and systems to describe events in a loose manner by just describing situations. This could even make sport analysts without programming skills able to define reasonable events.

CEP used on a high system level is done in a hierarchical way (see Fig. 8): Low-level processing on nearly raw, kinematic data is generating so called Position Based Events (PBE). Higher-level processing is done on event level only, producing Event Based Events (EBE). Combined forms exist as Position and Event Based

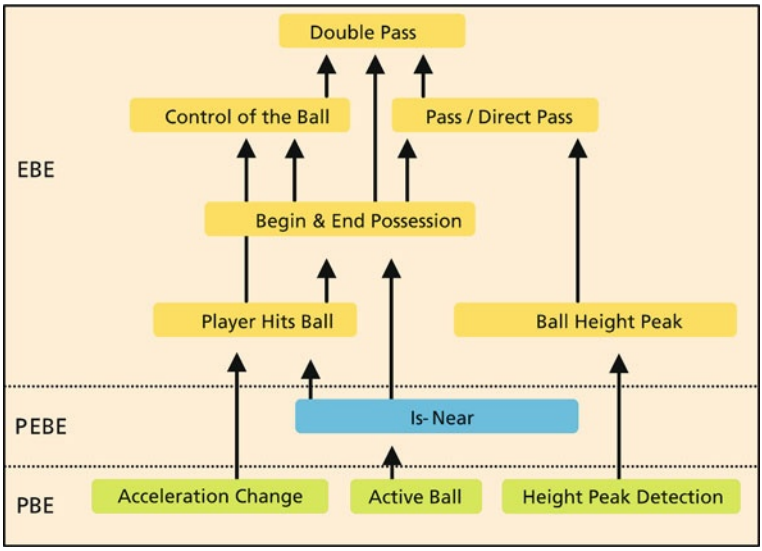



**Fig. 8** Hierarchic event structure

Events (PEBE). The low-level processing performs event detection on kinematic streams, which is usually the task that requires most of the processing power. If the delay for event detection shall be minimal for real-time ability, every new kinematic packet from any transmitter has to be handled immediately. If incoming streams are processed in such a way, there is no need to care about raw data processing any more, because we can assume that PBEs are detected and we can use these events as triggers in detectors of EBEs. These high-level events are generated by a well-defined set of rules like "event A followed by event B and then C generates some other event, if event D does not occur" using deterministic finite state automata. Figure 9 shows the corresponding automation.

Consequent aggregation of events leads to a more abstract, but also more meaningful view by discarding unneeded information. The computational challenge of generating all events needed for a football analysis system that can be used on all related fields mentioned above is solved using a distributed event processing engine

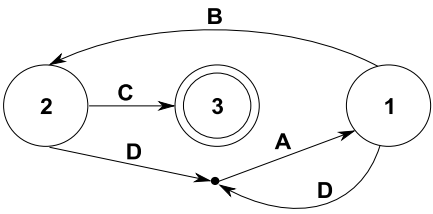



**Fig. 9** A deterministic finite state automation

(also called event-observer) which is able to grow in processing performance linearly with the number of processing nodes. Event detectors running on the nodes simple join multicast groups in a shared network and receive only subscribed events to minimize network usage. Event detectors can even move to other nodes at runtime realizing load balancing in the system. Event queuing is managed through dynamic evaluation of event-timestamps and system load at runtime. Using the described distributed environment, we are able to perform flexible, real-time event processing and analysis on heavily parallel and content rich data streams by maintaining a minimum delay at any time.

The implementation of the event-observer within the RedFIR® system deals with two modes. First, there is the analysis of regular matches and second the evaluation of training sequences. The latter is presented in the next section.

### *4.3 Standardized Performance Tests*

One of the first applications showing the benefits coaches can get realizes the automatic visualization and analysis of the so called "DFB-Testbatterie", one standardized testing method proposed by the German football association DFB. The "DFB-Testbatterie" is a set of six exercises which offer the possibility to measure the performance level of a youth player. It is a football specific talent diagnostic tool, providing information for comparison of individuals or teams. Technical, motor-driven talent attributes are measured objectively. On the basis of the individual development level and progress, performance profiles and scores can be determined. The testing exercises measure the ability to sprint (test 1), the agility of a player, the ability to run and dribble through the same course (test 2 & 3) and the ability to pass and control the ball (test 4). Also, it is measured how precise and hard a player can shoot on the goal (test 5) and how fast a player can juggle alternately with the left and the right foot (test 6).

Figure 10 gives an overview of the analysis system. While performing the "DFB-Testbatterie", transmitters are attached to players and cones defining the layout of each test.

So the distributed event processing engine mentioned above could be easily extended to operate as the analysis system where evaluation data gathering is done automatically between "start" and "finish" events for every test. These events are mainly triggered by "line crossing" events from players, where the lines are defined by the locations of the cones defining a test. "Finish" events generated by the event-observer carry all information needed for the analysis of a player's test result that can then be saved to a database and easily accessed by a Graphical User Interface (GUI) in real-time. During each test several additional test specific values for the players (e.g. intermediate times, max. velocity, etc.) are collected as well as the location of events like "ball contacts" if they are reasonable in the test. Characteristic values from each player and test are compared to those of other players in the database to rate the player's result in his corresponding age-group (see Fig. 11).

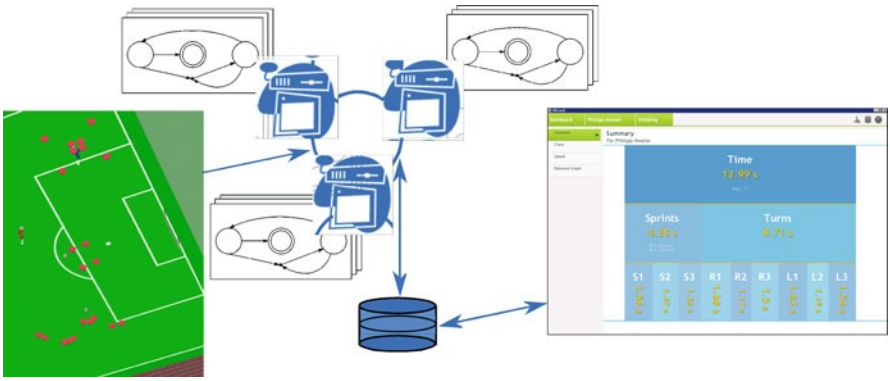



**Fig. 10** Analysis system for the "DFB Testbatterie"

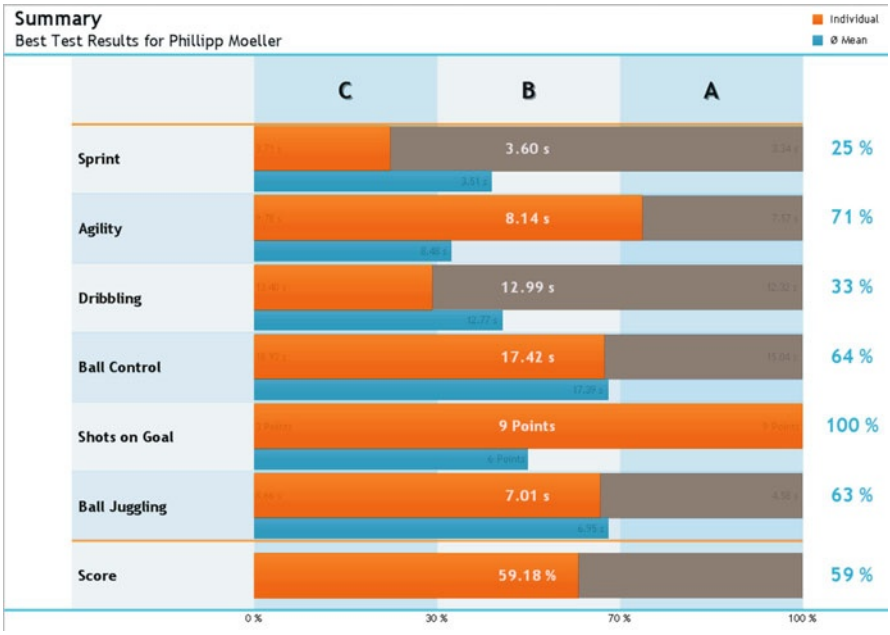



**Fig. 11** Test result presented in the GUI

The accuracy and high rate of the localization data in combination with Complex Event Processing leads to a reliable and easy to use system for performing the "DFB-Testbatterie". Instant analysis of all test results is available even with large groups of players. Therefore this approach is superior to the method used for these tests today concerning time effort and the number and detail of characteristic values that can be recorded and evaluated. At the end of the day the results of tests, trainings or game analysis can be used to control the training and to come to the right

decisions as coach of a team. Moreover, for the youth sector but also for established professionals the information gathered helps athletes to develop to a higher level of performance in a documented way.

## 5 Summary and Outlook

A radio-based real-time tracking system has been introduced by discussing the whole signal processing chain: miniaturized transmitters, correlation algorithms, Kalman filtering for line-of-sight tracking and carrier-phase aided position calculation. The application of the RedFIR® system in sports, especially in football, is motivated. The complex event processing for online analysis during matches and training is implemented based on a new architectural concept. Finally, a standardized performance test for football players proves the suitability for use as a real-life application for coaches.

Besides the up and running pilot installation in a big stadium, one more system will be installed on a training ground for regular and daily use by teams and sports scientists. Therewith, applications for commercial use of tracking data for spectators, media, federations, leagues, coaches, sports science, and sports medicine will continue to be developed gradually. They will be customized for each client and can also be adapted for other sports like American Football or track and field athletics.

Sports itself, particularly football, will not change. However, insight, objectiveness, the way of presentation and especially real-time support during training are going to be lifted to a new and exciting level.

## References

1. Rohmer G, Dünkler R, Köhler S, von der Grün T, Franke N (2005) A Microwave Based Tracking System for Football, Proceedings of the ION GNSS 2005, Long Beach, CA, USA
2. Townsend BR, van Nee DJR, Fenton PC, Dierendonck KJV (1994) Performance Evaluation of the Multipath Estimating Delay Lock Loop, Proceedings of the IEEE Position, Location and Navigation Symposium, Las Vegas, NV, USA
3. Hofmann G, Breiling M, (2004) Vorrichtung und Verfahren zum Bestimmen eines Eintrefffzeitpunktes einer Empfangsfolge, Patent DE 10,2004,059,941
4. Weill LR (1995) Achieving Theoretical Accuracy Limits for Pseudo Ranging in the Presence of Multipath, Proceedings of the 8th International Technical Meeting of the Institute of Navigation, Palm Springs, CA, USA, pp 1521–1530
5. Teunissen PJG (1996) GPS Carrier Phase Ambiguity Fixing Concepts, GPS for Geodesy, Lecture Notes in Earth Sciences, vol 60, chap 8, Springer Verlag, New York
6. Teunissen PJG (1994) A new method for fast carrier phase ambiguity estimation, IEEE Position Location and Navigation Symposium 1994, Las Vegas, NA, USA, pp 562–573
7. Euler HJ, Landau H (1992) Fast GPS ambiguity resolution on-the-fly for real-time applications. Sixth International Geodetic Symposium on Satellite Positioning, Columbus, OH, USA, pp 650–659

8. Hatch R (1991) Instantaneous ambiguity resolution, IAG International Symposium on Kinematic Systems in Geodesy, Surveying and Remote Sensing, Springer Verlag, New York, pp 299–308
9. Luckham DC (2001) The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA
10. Artikis A and Paliouras G (2009) Behaviour Recognition Using the Event Calculus, in AIAI'09, pp 469–478
11. Gyllstrom D, Wu E, Chae HJ, Diao Y, Stahlberg P, and Anderson G (2007) Sase: Complex event processing over streams, in Proceedings of the Third Biennial Conference on Innovative Data Systems Research

# Development of an Integration and Application Platform for Diverse Identification and Positioning Technologies

**Sebastian Lempert and Alexander Pflaum**

**Abstract** Diverse identification and positioning technologies like radio-frequency identification (RFID), real-time locating systems (RTLS), and wireless sensor networks (WSN) are among the enablers for the Internet of Things (IoT). Although the cost-effective integration of these enabling technologies into existing enterprise infrastructures is very important for companies, integration aspects are ignored by researches frequently. Even worse, the increasing diversity of identification and positioning technologies and corresponding applications leads to more expensive integration and application development processes. Up to now these technologies are used rather separately in different applications, but it is only a question of time before these technologies have to be used within the same application in parallel. Furthermore it could be observed, that existing commercial middleware products aiming at integration are focusing on passive RFID only. Hence there is a need to functionally extend existing solutions. This contribution introduces an integration and application platform for diverse enabling technologies for the Internet of Things that has been derived from the requirements of real world applications. An overview on applications of the Internet of Things is given and a selection of applications that have been successfully implemented in research and development projects is examined in detail. The functionalities provided by the platform and the corresponding abstract software architecture are derived from these applications.

## 1 Introduction and Motivation

Smart object technologies like RFID [4], RTLS [17] and WSN [8] are among important enabling technologies for the developing Internet of Things [18, 19]: In the context of logistics nowadays RFID is successfully used to optimize logistical processes

Sebastian Lempert (✉)
Center for Intelligent Objects ZIO and Fraunhofer IIS, Dr.-Mack-Str. 81, 90762 Fürth, Germany
sebastian.lempert@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_20, © Springer 2011

of enterprises in different industrial sectors. Besides that in logistics there is an increasing usage of other enabling technologies like RTLS and WSN. For example enterprises like DHL or Lufthansa Technik are starting to monitor selected loading units with intelligent telematics modules during transport [3] or to locate spare parts in widespread repair and overhaul centers with RFID [10]. Hospitals are introducing WSN based pilot projects in order to optimize their asset management processes for expensive medical devices and to increase transfusion safety [22]. The new possibilities of these technologies enable a wide variety of completely new applications in order to optimize processes. As depicted in Fig. 1 corresponding products differ considerably in terms of their range of functions and capabilities [20].

Nevertheless, from an application point of view different techologies might be eligible for the same application. But what technology will really catch on for what application will be decided by the application and a corresponding cost-benefit analysis. Against this backdrop today it already has become apparent that these technologies will encounter each other in different contexts and that they have to work smoothly together at the same time and at the same place. Last but not least with every single new application to be implemented and every single new technology to be used corresponding development and integration efforts rise. Thus the increasing diversity of identification and positioning technologies and corresponding applications leads to more expensive application development and integration processes.

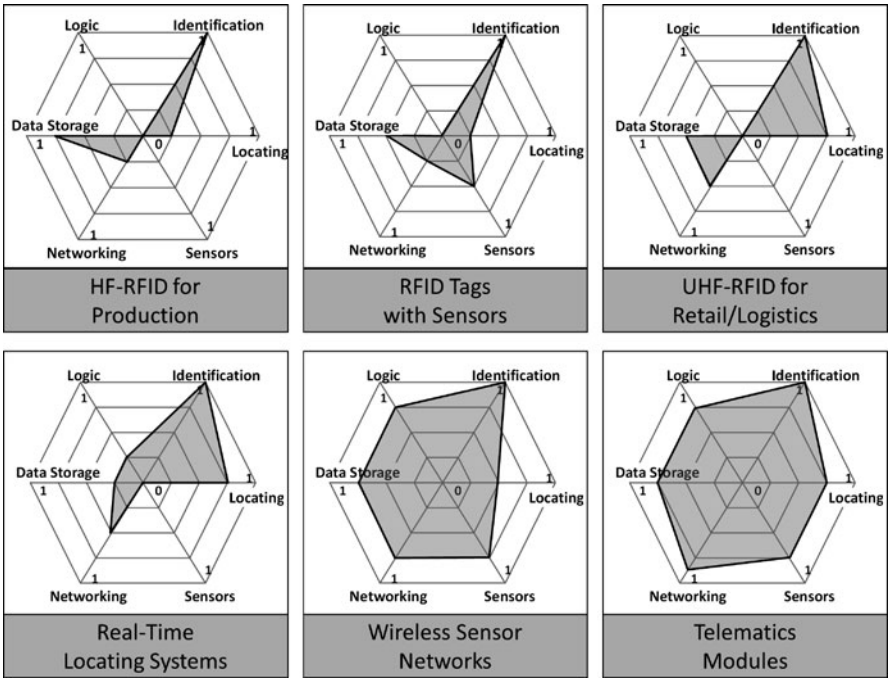



**Fig. 1** Functional profile of different smart object technologies

As a rule in existing commercial middleware solutions, aiming at the integra-tion of smart object technologies into existing IT landscapes, at the time of writing the support for smart object technologies is limited to passive RFID only. On the other hand middleware solutions for smart object technologies other than RFID are still under development and are most notably addressed in academic research, while integration aspects are frequently ignored [13]. Thus projects that involve the use of RFID in combination with other sensory devices demand new middleware architectures that fulfill additional requirements [7].

From the authors point of view an integration and application platform that claims to cover all smart object technologies has to support all functionalities as depicted in Fig. 1. Furthermore such a platform should be equipped with a set of basic applications that serve as a basis for efficient development of more sophisticated applications for the Internet of Things. Hence in this contribution the following research questions are examined:

- How can an integration and application platform for diverse identification and positioning technologies be described from a functional perspective?

In order to answer this question the following subquestions are examined:

- What are the applications that significantly influence the Internet of Things?
- Which appliations should be supported by an integration and application platform for diverse enabling technologies for the Internet of Things?
- What are the functional requirements that could be derived from these applications?
- How should a platform be designed that meets these functional requirements?

The rest of this contribution is organized as follows. Important terms that are used throughout this contribution are defined in Sect. 2. A comparision between this contribution and related work is given in Sect. 3. In Sect. 4 we describe our methodological approach and illustrate how functional requirements for an integration and application platform are derived from real world applications. In Sect. 5 this approach is applied and an abstract software architecture for such a platform is developed. We conclude this contribution with a summary and an outlook on prospective future work in Sect. 6.

## 2 Clarification of Terms

An *integration platform* denotes a middleware product or a combination of middleware products that aims at connecting different enterprise applications and supports the concept of enterprise application integration [11], where "EAI is the unrestricted sharing of data and business processes among any connected applications and data sources in the enterprise" [16].

An *integration and application platform for diverse smart object technologies* like RFID, RTLS and WSN is an integration platform, which unites these technologies with a shared technology abstraction layer, controls the interaction between

these technologies and existing enterprise infrastructures, supports intra-corporate and cross-company integration, and aims at reducing integration costs significantly. Furthermore such a platform is equipped with a set of basic applications that serve as a basis for the efficient development of more sophisticated applications for the Internet of Things.

Objects from the real world like containers, palettes and assets in general are called *intelligent objects* or *smart objects* if they are tagged with an enabling technology for the Internet of Things like RFID, RTLS or WSN. Smart objects come from a number of different *smart object technology* areas and scientific disciplines including embedded systems, ubiquitous and pervasive computing, mobile telephony, telemetry, wireless sensor networks, mobile computing, and computer networking with each area making its own imprint on the technology [29].

In research there is consensus that ideal smart objects have unique identifiers and are able to save and process information, to monitor their environment with sensors, to interact with their environment with actors, and to communicate with their environment wirelessly [21, 29]. From the authors point of view ideal smart objects are also able to detect their own position in the two- or three-dimensional space with the help of additional infrastructure.

"Today, the Internet of Things is a foundation for connecting things, sensors, actuators, and other smart technologies, thus enabling person-to-object and object-to-object communications" [28]. In addition to this definition from the others point of view the *Internet of Things* could be understood as the continuum of smart object based user-oriented solutions which are incorporating internet technologies.

## 3 State of the Art

The idea that commercial RFID middleware solutions should be further developed to support more than RFID is considered in the concepts Intelligent Network Sensor Infrastructure [24] and Bloor Sensory Middleware architecture [7].

Furthermore four RFID middleware solutions from the open source community where considered when designing the proposed integration and application platform: Global Sensor Networks [1] focuses on a sensor based abstraction for WSN and partially for RFID but does not consider RTLS or the integration of these sensors into existing infrastructures. Fosstrak [6] aims at implementing the most important standards for RFID defined by EPCglobal, while WSN and RTLS are not addressed. ASPIRE [25] focuses on RFID, supports WSN only partially and does not consider RTLS. The same holds true for Rifidi Edge [26].

Middleware solutions for RTLS and WSN are still under development and are most notably addressed in academic research, while integration aspects are frequently ignored [13]. Nevertheless altough in LocON integration into existing IT systems plays a limited role this project should be considered especially when different positioning and communication technologies should be supported over a common abstraction layer [2].

## 4 Methodological Approach

In this section we illustrate how functional requirements for an integration and application platform for diverse identification and positioning technologies are derived from real world applications.

First of all a comprehensive list of applications that significantly influence the Internet of Things is compiled from an exhaustive literature analysis. Afterwards the applications to be considered are specified: here the objective is to define a small set of applications that represents the Internet of Things. Thereby applications are chosen so that all enabling technologies for the Internet of Things are considered at least once and at least one corresponding pilot or research project exists that was successfully completed in order to strengthen practical relevance.

Therefore available publications and documents that describe corresponding projects are examined. Afterwards the most important functional requirements are identified by conducting interviews with leaders or experts from the projects to be considered. Finally based on the derived functional requirements and the current state of the art an abstract software architecture for a flexible integration and application platform is designed.

## 5 Development of an Abstract Software Architecture for a Flexible Integration and Application Platform of Tomorrow

In this section the previously presented methodological approach is applied in order to develop an abstract software architecture for a flexible integration and application platform for diverse smart object technologies.

### *5.1 Overview on Applications of the Internet of Things and Selection of Applications to be examined in Detail*

Applications of the Internet of Things can be found in many industries and in many areas of life. A comprehensive list of applications that significantly influence the Internet of Things was compiled from an exhaustive literature analysis with [4,8,17] and [5] as starting points. Due to space limitions in the following only a short and incomplete list of possible applications is presented. Nevertheless from the autors point of view this list is suitable to give an impression on the complexity of this field of research:

- *Logistics*: Supply chain event management, asset management, material flow management, electronic article surveillance, container security, hazardous materials storage, cold chain monitoring etc.

- *Healthcare*: Locate healthcare personnel, track the movements of patients, improve throughput management, track equipment, improve patient satisfaction etc.
- *Payment, ticketing and access control*: Cashless payment, tickets for public transport, ski passes etc.
- *Manufacturing*: Find products, track progress, trace parts, find tools etc.
- *Building automation*: Energy efficient humidity, ventilation and air control, assisted living systems etc.
- *Aviation*: Tracking of spare parts, runway and taxiway control on airports, tracking of air cargo etc.

In order to keep efforts low six projects successfully conducted at Fraunhofer IIS were chosen from the above mentioned list of applications for detailed examination in accordance with the previously presented methodology (see Table 1):

- *Aletheia*: The objective of this project was to ensure supply chain integrity. Thus the availability of the right product, in the right quantity, in the right condition, in the right place, at the right time, for the right customer and at the right price had to be guarenteed. Breaches of integrity like damage to products during transport or the unauthorized opening of containers were detected with the help of WSN and reported with the help of telematics modules [27].
- *Galileo*: Non-transparent and manual processes in earthworks impede the control of material flow and the documented evidence on the adherence of legal restraints. Hence the objective was to enable sustainable planning, operation and control of earth movements. Using satellite-aided positioning and sensors at the excavator boom, the digging movements of an excavator shovel were captured with high precision. These were linked to information contained in a layered digital terrain model in order to derive the type of soil in the shovel. Furthermore using satellite-aided positioning and sensors at the loading area, the movements of a dumper were captured with high precision. Failures in the process of loading and unloading were detected and reported in real-time [15].
- *LocON*: The objective of this project is to provide an airport ground control system with a complete overview of the ongoing operations and facilities to communicate with each staff member with the help of diverse positioning and communication technologies. One of the systems benefits is to predict and prevent collisions between aircrafts and ground vehicles on taxiways [2].
- *OPAL*: The objectives of this project were to enable an efficient clinical asset management and to increase transfusion safety. Expensive medical devices were equipped with sensor nodes and could be tracked and located indoor in real-time with room level precision. Blood bags and patients were also tagged with sensor nodes. The cold chain of the blood bads was monitored permanently and prior to transfusion a verification of the blood took place. Tainted or unsuitable blood was detected and reported in real-time. The system was integrated into an existing hospital information system [22].
- *OLog-PAT*: Within this project a process management system based on sensor networks has been developed in order to improve logistical, non-medical procedures in the treatment of patients in clinics and to improve the quality of care.

**Table 1** Overview on applications of the Internet of Things and corresponding enabling technologies and successfully completed projects

| Application | Enabling technology | Project |
|---|---|---|
| Logistics | WSN, Telematics Module | Aletheia |
| Logistics | RTLS, WSN | Galileo |
| Aviation | RTLS, WSN, RFID | LocON |
| Healthcare | WSN | OPAL |
| Healthcare | WSN | OLog-PAT |
| Aviation | RFID | SMMART |

Movements of staff and patients were captured with the help of sensor networks. Residential stays, were logistically collated and controlled in order to organize them more efficiently. Operationally relevant information about occupancies and utilization of resources were gathered in order to increase the transparency of the process, the coordination amongst hospital departments and, finally, to reduces waiting times and capacity overload. The system was integrated into an existing hospital information system [23].

- *SMMART*: The optimal supply of spare parts is an important factor for the availability of production resources and their cost-effective use. Therefore in this project through the global use of RFID, a comprehensive and efficient tracking and tracing system was developed in order to optimize spare parts management processes for helicopter gas turbines. This information made it possible to reduce the cost of unplanned breakdowns and the time and cost of planned servicing and maintenance. The system was integrated into an existing enterprise resource planning system [9].

## *5.2 Derivation of Functional Requirements Based on the Results of Successfully Finished Research and Development Projects*

According to the presented methodology the most important functional requirements were identified by conducting interviews with leaders or experts from the projects presented in the previous section. The results of the conducted interviews are summarized in Table 2.

## *5.3 Design of an Abstract Software Architecture for a Flexible Integration and Application Platform Based on the Derived Functional Requirements*

Based on the derived functional requirements and the current state of the art an abstract software architecture for a flexible integration and application platform was

**Table 2** Functional requirements of different real world applications

| Functional requirement | Aletheia | Galileo | LocON | OLog-PAT | OPAL | SMMART |
|---|---|---|---|---|---|---|
| Technology abstraction | ● | ● | ● | ● | ● | ● |
| Filtering & enhancement of status information | ● | ● | ● | ● | ● | ● |
| Transformation of status information into business events | ● | ● | ● | ● | ● | ● |
| Event subscription & notification | ● | ● | ● | ● | ● | ● |
| Tracking & Tracing of tags or assets | ● | ● | ● | ● | ● | ● |
| Mapping between tags and assets | ● | ● | ● | ● | ● | ● |
| Matching between objects | ● | ● | ● | ● | ● | ● |
| Dashboard for events, sensor data or KPIs | ● | ● | ○ | ○ | ● | ● |
| Display tracking & tracing information on digital maps | ○ | ● | ● | ○ | ● | ○ |
| Enterprise integration | (●) | (●) | ● | ● | ● | ● |
| Database for assets, tags and events | ● | ● | ● | ● | ● | ● |
| Conversion between coordinate systems | ○ | ● | ● | ● | ● | ● |
| Conversion between units of measurement | ● | ○ | ● | ○ | ● | ○ |
| Conversion between date and time notations | ● | ● | ● | ● | ● | ● |
| Information security | ● | (●) | ○ | (●) | ● | (●) |
| Configuration & update mechanisms for tags | (●) | ○ | ● | ○ | ○ | (●) |
| Decentralized in-network processing | ● | ○ | ○ | ○ | ● | (●) |

designed. The corresponding architecture is depicted in Fig. 2. Due to space limitations a detailed description of the platform modules and their internal modes of operation have to be omitted in this constribution. Nevertheless the interessted reader will find the corresponding information in [14].

## 6 Conclusion and Future Work

We motivated that the increasing diversity of identification and positioning technologies and corresponding applications leads to more expensive integration processes and that there is a necessity to functionally extend existing integration solutions. Against this backdrop we investigated on how an integration and application platform for diverse identification and positioning technologies can be described from a functional perspective. To this end we described our methodological approach and illustrated how functional requirements for such a platform are derived from real world applications. Furthermore we applied this approach and developed an abstract software architecture for such a platform. Currently we are implementing the proposed architecture in an internal project that focuses on the optimization of our

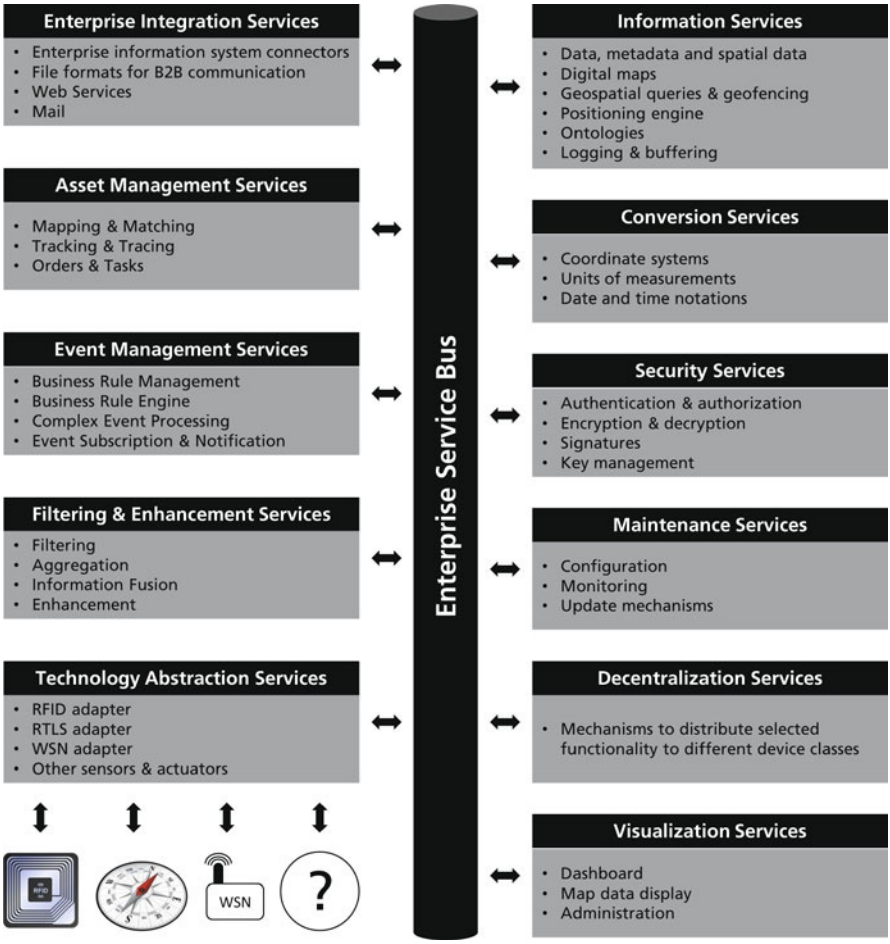



**Fig. 2** Abstract software architecture for an integration and application platform for diverse identification and positioning technologies

asset management process for highly prized measurement and testing equipment. First results on this future work are published in [12].

## Acknowledgements

This research is part of the Center for Intelligent Objects ZIO and was funded by the Bavarian Ministry of Economic Affairs, Infrastructure, Transport and Technology.

## References


1. Aberer K, Hauswirth M, Salehi A (2006) Global Sensor Networks. Technical Report LSIR-REPORT-2006-001. EPFL, Lausanne, Switzerland
2. Couronné S, Hadaschik N, Faßbinder M, von der Grün T, Weyn M, Klepal M, Widyawan, Denis T (2009) LocON – a Platform for an Inter-Working of Embedded Localisation and Communication Systems. Poster Abstract. In IEEE Secon 2009
3. Deutsche Post DHL (2009) DHL Innovation Award 2009 "Growing Ideas" Gewinnersteckbrief
4. Finkenzeller K (2008) RFID-Handbuch. Hanser, Munich, Germany
5. Fleisch E, Mattern F (eds) (2005) Das Internet der Dinge. Springer, Berlin, Germany
6. Floerkemeier C, Roduner C, Lampe M (2007) RFID Application Development With the Accada Middleware Platform. IEEE Systems Journal 1(2):82–94
7. Holloway S (2008) RFID Middleware: From RFID to Sensory Network middleware for the edge. Technical Report. Bloor Research, Towcester, Northamptonshire, UK
8. Karl H, Willig A (2007) Protocols and Architectures for Wireless Sensor Networks. John Wiley & Sons, Chichester, West Sussex, UK
9. Krupp M, Lempert S (2007) Prozesskompetenz als Basis für erfolgreiches Event Management: Bedeutung von Business Rules am Beispiel von SMMART, SCEM Forum 2007, Wiesbaden, Germany
10. Krupp M, Pflaum A, Raabe T (2010) RFID als Basis einer verbesserten Informationsgrundlage zur Steuerung logistischer Prozesse. In Krupp T; Paffrath R, Wolf J (eds) Praxishandbuch IT-Systeme in der Logistik. Deutscher Verkehrs-Verlag, Hamburg, Germany, pp 164–173
11. Laudon KC, Laudon JP, Schoder D (2006) Wirtschaftsinformatik: Eine Einführung. Pearson Studium: Wirtschaft. Pearson Studium, Munich, Germany
12. Lempert S, Pflaum A (2011) Sensornetzbasiertes Supply Chain Event Management zur Optimierung des innerbetrieblichen Asset Managements am Fraunhofer IIS. In Proceedings of LM 2011
13. Lempert S, Pflaum A (2010) Über die Notwendigkeit einer Integrationsplattform für unterschiedliche Smart Object Technologien. In Kolla R (ed) Proceedings of FGSN 2010, pp 71–74
14. Lempert S, Pflaum A (2011) Towards a Reference Architecture for an Integration Platform for Diverse Smart Object Technologies. In Höpfner H, Specht G, Ritz T, Bunse C (eds) Proceedings of MMS 2011, pp 53–66
15. Lempert S, Harrer M, Krupp M, Pflaum A (2010) Transparente und effiziente Prozesse im Erdbau durch ereignisgesteuertes Stoffstrommanagement auf Basis von Smart Objects und Business Rule Management. In Fähnrich K, Franczyk B (eds) Proceedings of INFORMATIK 2010, pp 233–244
16. Linthicum DS (1999) Enterprise Application Integration. Addison-Wesley, Boston, MA, USA
17. Malik A (2009) RTLS for Dummies. Wiley Pub. Inc., Indianapolis, IN, USA
18. Mattern F (2005) Die technische Basis für das Internet der Dinge. In Fleisch E, Mattern F (eds) Das Internet der Dinge. Springer, Berlin, Germany, pp 39–66
19. Mayordomo I, Spies P, Meier F, Otto S, Lempert S, Bernhard J, Pflaum A (2011) Emerging Technologies and Challenges for the Internet of Things. In Proceedings of MWSCAS 2011
20. Pflaum A, Hupp J (2007) Sensornetzwerke und Lokalisierungsverfahren als Schlüsseltechnologien für die intelligente logistische Umwelt von morgen. In Bullinger H, ten Hompel M (eds) Internet der Dinge. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, Germany, pp 107–118
21. Sánchez López T, Ranasinghe DC, Patkai B, McFarlane D (2009) Taxonomy, technology and applications of smart objects. ISF 13(2):281–300. doi: 10.1007/s10796-009-9218-4
22. Sedlmayr M, Becker A, Münch U, Meier F, Prokosch H, Ganslandt T (2009) Towards a Smart Object Network for Clinical Services. In Proceedings of AMIA 2009, pp 578–582
23. Seitz M, Meier F, Münch U, Ma TH, Niemann C, Kriegel J (2010) Das drahtlose Sensornetzwerk als Schattenbild der Patientenlogistik. In Proceedings of Mocomed 2010

24. Sirico L, Arteaga C, Woods T (2010) RFID Middleware is Extinct: The Intelligent Sensor Network is Born
25. Soldatos J (2009) AspireRFID Can Lower Deployment Costs. Website. Available online at http://www.rfidjournal.com/article/view/4661, visited on 2011-07-22
26. Swedberg C (2009) Pramari Launches Free Open-Source RFID Middleware. Website. Available online at http://www.rfidjournal.com/article/view/5328, visited on 2011-07-22
27. Traulsen H, Kaffenberger C, Pflaum A (2010) Senkung der Versicherungsprämien bei Überlandtransporten mittels geeigneter Smart Object Technologien. In Kolla R (ed) Proceedings of FGSN 2010, pp 33–36
28. Uckelmann D, Harrison M, Michahelles F (eds) (2011) Architecting the Internet of Things. Springer, Berlin, Germany
29. Vasseur J, Dunkels A (2010) Interconnecting smart objects with IP: The next Internet. Morgan Kaufmann/Elsevier, Burlington, MA, USA

# Part III
# Visual Computing

# Visual Computing at the IIS: From Life Sciences to Industrial Applications

**Thomas Wittenberg and Theobald Fuchs**

## 1 Introduction

Since its foundation over 25 years ago, the Fraunhofer IIS has been involved in the applied research within the field of "Visual Computing". This field of research incorporates all areas of *image acquisition*: including different kinds of imaging physics, various types of image sensors; *image formation*: describing the manipulation of the acquired image data between the sensor and the output and 3-D-image reconstruction; *image coding and transportation*: lossy and lossless coding, wire-based and wireless transmission; *image processing*: image-to-image operations such as image enhancement, panoramic image stitching, or video tracking; *image analysis*: pattern recognition, object detection, classification; and *image synthesis*: including computer graphics, image augmentation and modeling.

In the past years the image processing departments of the Fraunhofer IIS have developed competences in almost all these named subareas of visual computing and have applied their expertise to a very broad field of different applications. More specifically, the key competences of these departments have always been the fusion of all these image computing components as well as their interaction towards the goal of designing and providing systems to a broad range of different image based applications. Hence, in this contribution, an overview of the image computing activities of the Fraunhofer IIS will be provided based on examples from current applied research throughout various application fields.

## 2 Scales of Image Processing Systems

As can be seen in Fig. 1, the range of image processing applications provided by the Fraunhofer IIS in the past 25 years can be seen on various scales, namely the

Thomas Wittenberg (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
thomas.wittenberg@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_21, © Springer 2011

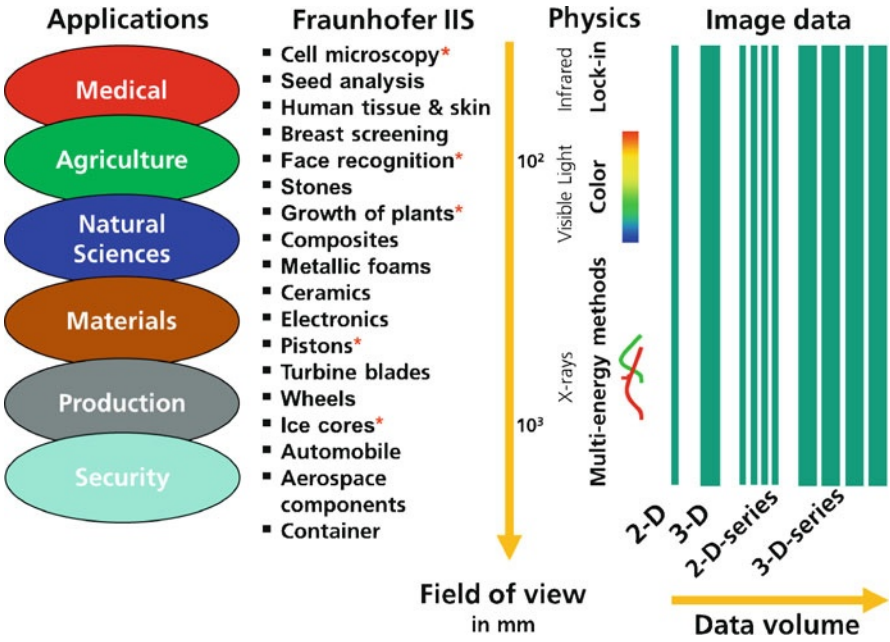



**Fig. 1** Scales of image processing systems: Application scale (*left*), physical scale (*center*), and information scale (*right*). The *red asterisks* indicate applications which will be addressed in detail in the following contributions

*application domain* scale (left), the *physics scale* (center) defining the size of objects to be captured and analyzed as well as the physics of the image modality applied, and finally the *information scale* describing the type and amount of image data (right).

The application scale shows that the image processing systems range from medicine (computer assisted microscopy, computer assisted diagnosis, and computer assisted surgery), over biological and agricultural applications (analysis of seeds, roots, leaves, whole plants, ...) and natural scenes (images and sequences of landscapes, cars, people, animals, faces, gestures, emotions, ...) to materials (foams, wood, plastic, synthetic tissues, aluminum, cast iron, ice, ...) and parts for automotive and aerospace applications such as turbine blades, break cylinders, wheels, or motors.

Regarding this list of image analysis applications it seems natural that on the physics side the image acquisition modalities applied range from optical devices such as microscopes, endoscopes, web-cams, high-speed cameras, professional movie cameras, and smart phones via ultra sound, fluorescence, multi-spectral and thermographic sensors to X-ray and computed tomography (CT) imaging systems. Hence, this spectrum of available and applied image modalities yields all possibilities to depict objects from the sub-$\mu$m range for leukocyte cell nuclei to complete cars and containers in the range of several meters, as e.g. hot steel rods.

Furthermore, on the information scale the capability can be seen, what type and amount of data must be analyzed, communicated and stored, with respect to these named applications. The image data type starts with 2-D images (X-ray images, digital mammograms, optical high-resolution still images) and the assembly of 2-D-tiles and mosaics (image stitching) to temporal image sequences (2-D+t) such as video and high-speed sequences, to volumetric (3-D) image data from CT or laser-scanning devices and ends with the acquisition of 3-D+t image data with applications such as 3-D plant growth analysis or light field coding and usage. Hence, in order to fulfill a certain image processing task under the side condition of very short (production or analysis) time slots, it can be shown that the exemplarily named systems are capable to deal with very high loads of image data up to 10 Gigabyte per second and more.

## 3 Examples for Image Processing Applications

From the described plethora of possibilities regarding the application of image processing methods in manifold application scenarios and using many different types of image acquisition modalities, six examples shall be given to show and demonstrate the variety of recent activities in the image computing domain of the Fraunhofer IIS (indicated by asterisks in Fig. 1). Using the above described physical scale, ranging between optical imaging and X-rays, these examples will be given from the domains of the life sciences, agriculture, face and object recognition, environmental analysis and car parts. Furthermore, an example shall be given regarding the possibilities of offline image enhancement.

### *3.1 Leukocyte Analysis*

In the field of optical microscopic imaging and medical image analysis, the HemaCAM system will be described by Münzenmayer et al. [1]. This system belongs to the category of computer assisted microscopy (CAM) systems and has been designed and built for the image based examination and analysis of the peripheral blood samples, being the so-called differential white blood cell count. The HemaCAM system is based on a robotic microscope which has been enhanced using an intelligent control software as well as a multi-stage image processing pipeline, including automated cell detection, cell segmentation and cell sorting. The HemaCAM system has recently been certified as a medical in vitro diagnostic device and has been in the market since fall 2010.

### *3.2 Face Detection*

Ruf et al. [3] describe a general image processing approach for the detection, tracking and classification faces, gestures and facial expression in images and image sequences obtained from a single web camera. The basic concept of the automatic and robust detection of objects with in a video stream is based on the so-called *Sophisticated Highspeed Object Recognition Engine – SHORE*, which has been designed as a flexible environment for varying detection tasks. Benchmark results using this approach have been obtained on both, standard and publicly available data sets of faces.

### *3.3 Plant Phenotyping*

Understanding plant growth and plant interaction with the environment is one aspect of modern agriculture. As a manual determination of morphological plant parameters is tedious, automatic acquisition methods are required. Specifically, 3-D scanning methods such as sheet-of-light measurements is a fast approach for the to acquisition of surface points. As an example for the domain of image processing possibilities within agriculture, Uhrman et al. [5] introduce two measurement methods based on the *sheet-of-light* principle for plant phenotyping. They discuss the required data preprocessing algorithms and finally show an example of an industrial implementation.

### *3.4 Environmental CT*

X-ray technology has went through a rapid evolution in the past 20 years. Even though modern CT systems follow the same principles of physics and technology as their precursors, the performance of today's CT machines has improved tremendously regarding spatial resolution, speed of measurement, methods of data acquisition, and 3-D evaluation of huge volumetric data sets. One application example for modern environmental CT is a system designed to acquire data from prehistoric ice cores (Voland et al. [6]): On one hand for these measurements a spatial resolution of around 12 µm or less was required, while the ice core had dimensions of 1 m length and 10 cm diameter. On the other hand, the CT system had to operate at temperatures of constantly minus 15 °C.

### *3.5 Industrial CT*

One field of research within CT imaging that is particularly important relates to data acquisition and data processing as well as automatically enhanced image evaluation

to handle huge amount of information provided by high resolution 3-D X-ray imaging. This challenge is illustrated by the contribution of Oeckl et al. [2]. Specifically, to withstand the load inside the combustion chamber, pistons for combustion motors must meet strongest requirements in terms of material quality and dimensional accuracy. Inspection of pistons using CT provides a 3-D characterization of defects with a high data volume. Hence, a CT system for casting inspection is described, covering both, hardware and software setup. One key issue within this system is the use of reference image data for defect detection.

### *3.6 Image Enhancement*

Finally, an example in the area of image formation and image enhancement will be given for professional digital camera and video systems: The contribution of Schöberl et al. [4] describes the possibilities of transferring extremely sophisticated but complex raw image enhancement methods from mobile imaging devices such as media production cameras or mobile phones towards offline processing resources. This shift allows new possibilities of image formation and image enhancement, as it opens the possibilities of applying methods with an increased need of processing power and intermediate storage.

## 4 Conclusions

The imaging activities at the Fraunhofer IIS cover a wide range of electromagnetic imaging modalities starting with infra-red light, going over visible light and soft X-rays eventually up to high energy X-ray inspection. A large variety of applications has been addressed in the recent 20 years, of which some of the most sophisticated are presented in the following reports. Based on this wide experience, research and development on the fields of medical imaging, non-destuctive testing, scientific evaluation, multimedia enterainment, and several others will play an important role in the future image processing activities of the Fraunhofer IIS.

## References

1. Münzenmayer C, Schlarb T, Steckhan D, Hasslmeyer E, Bergen T, Aschenbrenner S, Wittenberg T, Weigand C, Zerfass T (2011) HemaCAM – A Computer Assisted Microscopy System for Hematology. In Heuberger A, Elst G, Hanke R (eds) Microelectronic Systems – Circuits, Systems & Applications. Springer, Heidelberg, pp 227–235
2. Oeckl S, Gruber R, Schön W, Eberhorn M, Bauscher I, Wenzel T (2011) Process integrated inspection of motor pistons using computerized tomography. In Heuberger A, Elst G, Hanke R (eds.) Microelectronic Systems – Circuits, Systems & Applications. Springer, Heidelberg, pp 271–279

3. Ruf T, Ernst A, Küblbeck C (2011) Face and Gesture Detection with the Sophisticated High-speed Object Recognition Engine (SHORE). In Heuberger A, Elst G, Hanke R (eds) Microelectronic Systems – Circuits, Systems & Applications. Springer, Heidelberg, pp 237–246
4. Schöberl M, Keinert J, Kaup A, Foessel S (2011) Analysis of Processing Pipelines in Digital Raw Cameras. In Heuberger A, Elst G, Hanke R (eds) Microelectronic Systems – Circuits, Systems & Applications. Springer, Heidelberg, pp 281–293
5. Uhrmann F, Seifert L, Scholz O, Schmitt P, Greiner G (2011) Improving sheet-of-light based plant phenotyping with advanced 3-D simulation. In Heuberger A, Elst G, Hanke R (eds) Microelectronic Systems – Circuits, Systems & Applications. Springer, Heidelberg, pp 247–257
6. Voland V, Freitag J, Uhlmann N, Hanke R (2011) A computed tomography system for the analysis of prehistoric ice cores. In Heuberger A, Elst G, Hanke R (eds) Microelectronic Systems – Circuits, Systems & Applications. Springer, Heidelberg, pp 259–270

# HemaCAM® – A Computer Assisted Microscopy System for Hematology

**Christian Münzenmayer, Timo Schlarb, Dirk Steckhan, Erik Haßlmeyer, Tobias Bergen, Stefan Aschenbrenner, Thomas Wittenberg, Christian Weigand, and Thorsten Zerfaß**

**Abstract** Cost and competition force modern hematology laboratories to further automate their processes. To that respect the examination and analysis of the peripheral blood is of central importance as it is relevant to a large variety of diseases while on the other hand financial reimbursement is low. Over the past eight years, the HemaCAM system has been developed by the Fraunhofer IIS, which supports the assessment of peripheral blood samples and the so-called white blood differential. Since 2010, HemaCAM has been available on the market as a certified medical product, to be more specific as an in vitro diagnostic device. This contribution provides an overview of the key components of the HemaCAM system.

## 1 Introduction

### *1.1 Diagnostic Impact of Peripheral Blood*

The assessment and diagnosis of peripheral blood is of central importance for a large variety of common and hemic diseases such as inflammations, allergy, parasites, sexually transmitted diseases or all forms of anemia or leukemia. For routine examinations a blood sample obtained from the veins is typically analyzed automatically by hematology analysers or the use of Fluorescence Activated Cell Sorting (FACS) [3]. Unfortunately, these automatic methods do not deliver satisfying results in 30% to 40% of the analyses. Hence, in those cases, subsequent visual inspection of the stained blood smear under a microscope by a trained operator is essential. In addition, there are many cases where the clinical symptoms make a visual inspection of the morphology of blood cells inevitable and such require the microscopic analysis.

Christian Münzenmayer (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
christian.muenzenmayer@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_22, © Springer 2011

In preparation for visual examination under the microscope a so-called blood smear is produced. This is obtained by placing a very small amount of blood on one end of a microscope slide and dispersing it along its full length. The goal of this procedure is to yield a region where the cells are spaced far enough apart to be counted and differentiated. This region is referred to as the so-called *counting area*. Once the blood smear is completely air-dried, the slide is immersed in methanol for fixating the smear. Finally, the slide is stained in order to visually distinguish the cells from each other. A commonly used staining is the so-called May-Grünwald-Giemsa (MGG) staining, also called Pappenheim staining. Such prepared blood smears are subsequently used for cytological and hematological assessments of white (leukocytes) and red (erythrocytes) blood cells and for verifying the existence of bacteria or parasites within the blood such as nematoda or plasmodia. One of the most common examinations based on stained blood smears is the so-called white blood cell count (WBC) or white blood differential which is done by counting the various types of leukocytes [11].

Currently, several trends force modern clinical laboratories to further invest into automation of their processes. First, over the last century, a tremendous centralization of laboratory capacities and mergers has taken place which increases market pressure and competition. Hence, many small laboratories are bound to be integrated into larger laboratory chains as their leading hematologists reach retirement age and younger successors are difficult to find. This trend also leads to a specialization of laboratories to compete and generate a unique market position. Today, the market in Germany is dominated by only about five or six large laboratory chains and an end of these mergers may not be reached yet. Secondly, the pressure in terms of cost also has increased in the past years as the public health care system is cutting costs. For example, on one side currently a white blood differential in public health care is reimbursed with only about 0.70 EUR. On the other side, it takes a laboratory assistant about 2 minutes of visual inspection, not counting sample preparation, staining and handling. Therefore, there exists a still increasing motivation in the laboratories to further automate manual processes and thereby increasing cost-effectiveness and revenues.

## *1.2 Objective*

As stated in the introduction the assessment of peripheral blood is manifold and complex as there is a large variety of diseases and physiological states that can be derived from examining a stained blood smear. Hence, to this end one of the most common and important standardized examinations is the white blood differential. To automatically obtain a WBC, the HemaCAM system has been developed in the past eight years by the Fraunhofer IIS, and has been clinically certified in 2010 as an in vitro diagnostic device.

Therefore, in order to reach that goal, the workflow of the HemaCAM system starts by digitizing the relevant portion of the slide (counting area) and from that

derives the WBC, i.e. identifying a number of between 100 and 200 leukocytes, according to preceding clinical conditions. The quality of the selection and pre-classification of leukocytes is expected to be so sensitive that very few cells have to be corrected manually. Furthermore, the assessment of the red blood cell morphology and general disease states should also be possible to be interactively annotated in the system to further support full integration into the laboratory workflow. With respect to these objectives the goal of image-processing based cell analysis is to provide as much automation as possible while leaving final control as well as the final diagnostic decision to the human expert and hence enhance laboratory efficiency.

## 2 Methods

### *2.1 Hardware Components and Construction*

The HemaCAM system is based on a high-quality research microscope (Carl Zeiss, Axio Imager.Z2) with an automated height adjustable table ($z$-axis). It features a 7-fold automatic and encoded objective revolver equipped with two objectives. One objective is a Zeiss Achroplan with 10-fold magnification for fast acquisition of overview images of the slide and localization of leukocytes. The second objective is a Zeiss oil objective with 100-fold magnification for detailed high-quality images of the pre-detected leukocytes. Passing a 1-fold tube adapter the optical path ends

**Fig. 1** Hardware outfit of the HemaCAM system (Image courtesy of Horn Imaging GmbH, Aalen)

on the sensor of a digital still image camera (Allied Vision Technologies, AVT Pike F-145C, 1388 × 1038 pixel CCD-sensor). The brightfield illumination is based on the Zeiss VIS-LED light source.

Adapted to small- and medium-sized laboratories, the system is capable to handle up to eight slides in one run, fixated in an exchangeable frame which is inserted into a motorized stage manufactured by Märzhäuser.

Using two degrees of freedom ($x$- and $y$-axis) of the stage and the motorized $z$-drive of the microscope, the slide is positioned into the optical path of the microscope. Thereby the sample can be scanned field by field by use of the digital camera. Sample identification can either be obtained manually by entering the corresponding sample identification number or using the optional bar-code reader or optical character recognition (OCR) module for printed characters commonly used in laboratory automation.

The whole system (as depicted in Fig. 1) is controlled by the HemaCAM software described below running on a conventional personal computer connected to the microscope and integrated into the laboratories IT infrastructure.

## *2.2 Software Components and Algorithms*

### 2.2.1 Overview

The complete *HemaCAM Suite* consists of the *HemaCAM* software which controls the complete microscope system and provides the main user interface as well as a pure software client termed *HemaCAM Studio* allowing networked access from a distant workplace. Scanning and reviewing of slides can thereby be done in parallel as it fits to laboratory routine. After logging in to the system the scanning process of up to eight samples (loaded into the slide holder frame) can be started and the number of leukocytes to be counted can be pre-configured (in genereal 110 cells are counted). The scanning process in itself consists of several steps such as the detection of the counting area, the detection and segmentation of leukocytes with the 10-fold magnification objective, and afterwards the targeted scanning of each leukocyte with the 100-fold high magnification objective. The individual fields of view are stored in a so-called *virtual slide* that can be reviewed by the user without microscopic hardware. By means of automatic image-based feature extraction and classification each cell is assigned to the most probable class label as a suggested pre-classification for the laboratory assistant. Each of these steps will be described in more detail in the following sections.

In the classification screen (see Fig. 1) the assistant is now able to review the pre-classification results from the HemaCAM system. By an intuitive drag-and-drop graphical user interface leukocytes can be reclassified or annotated for further review. Furthermore, the configuration interface of the HemaCAM software allows the specification of additional and special cell classes for manual assignment. General disease states as well as the red blood cell (RBC) morphology can also be annotated in an intuitive and efficient way.

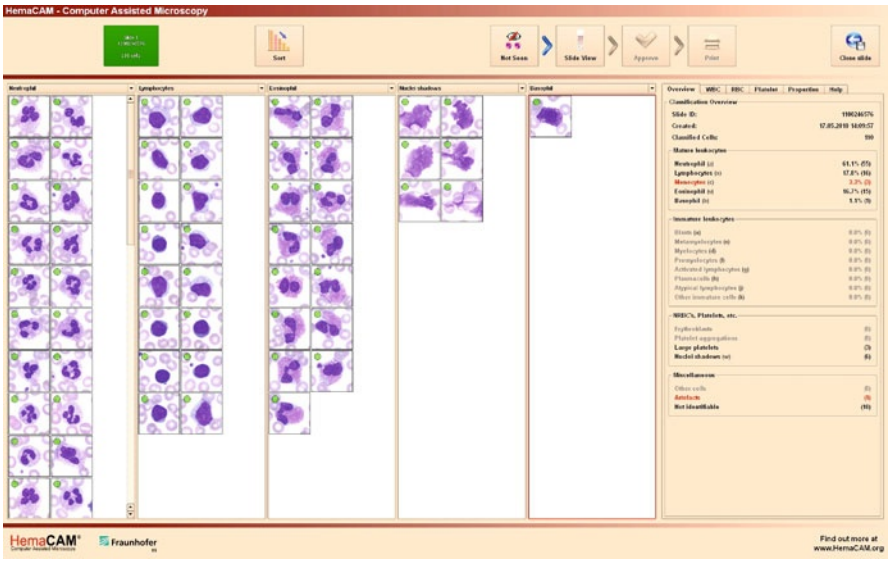



**Fig. 2** On the classification screen the HemaCAM system presents the cell pre-classification. By intuitive drag-and-drop, the operator is able to reclassify cells in an efficient manner

### 2.2.2 Virtual Slides

So-called *virtual slides* allow the storage and review of digitized blood smears or other specimen in a fully digital way with the associated advantages such as conservation against aging and bleaching, simultaneous access and remote review [5]. In a virtual slide the individual and neighborly overlapping fields of view (image tiles) that are captured sequentially during the scanning process are organized in a rectangular structure according to their global coordinates from the $xy$-stage. Within this data structure it is also possible to include further images from different magnifications, which are obtained from the 10-fold and the 100-fold objective, into one image mosaic which can then be navigated similar to commonly available applications using mouse dragging for translation and mouse wheel rotation for zooming.

However, with the high spatial resolutions involved in microscopy the tiniest inaccuracies of the scanning hardware will lead to visible misalignments even for accurately calibrated systems. Exemplary measurements in our laboratory have exposed a repeatability of 2.0 µm and an accuracy of 10.0 µm for a common hardware setting. Visually such values will lead to significant artifacts at the border between neighboring image tiles. Hence, image stitching or mosaicking algorithms are applied that match image structures in the borders of overlapping image tiles [8, 10]. These post-processing methods allow to compensate for these inaccuracies and provide a unified view of the slide. Steckhan and Paulus [9] also published an approach that uses the determined positioning accuracy of the stage as boundary conditions for a stitching algorithm based on a quadratic programming approach that is solved

according to the *Karush–Kuhn–Tucker* theorem [13]. This approach allows a consistent mosaicking with a median translation error down to 0.8 pixels.

### 2.2.3 Detection of the Counting Area

A basic prerequisite for the detection of blood cells is the reliable localization of the counting area within the blood smear, because there the cells are predominantly separated and hence do not interfere with each other. For this, multiple images of a blood smear are taken at different locations of the slide and analyzed in order to select the image corresponding to this area [7]. The analysis involves the computation of an unimodal function directly from the image content. This requires a prior segmentation of the cells that is carried out by a binarization in the HSV color space. Finally, the indicator function is derived as a combined measure from the number of cells and the average cells' surface area. Its unimodality guarantees to find a maximum value that corresponds to the counting area's image index, enabling a fast determination of the counting area on blood smears.

### 2.2.4 Leukocyte Segmentation

Segmentation of leukocytes is one major step in the process and various authors have published algorithms for this task [2,4,6]. Within the HemaCAM system there are two different steps of cell segmentation involved in the analysis process. The first step deals with the detection of the leukocyte cells in the images taken with the 10-fold objective and their coarse segmentation. This is necessary to locate the position of the single leukocytes for the recording of the images with the 100-fold magnification objective. These images and their segmentation results are used for the successive classification step [12]. As a leukocyte cell consists of two main parts, the nucleus and the plasma, the segmentation algorithm is able to segment these two parts of the cell. Furthermore, the algorithm is capable to robustly separate

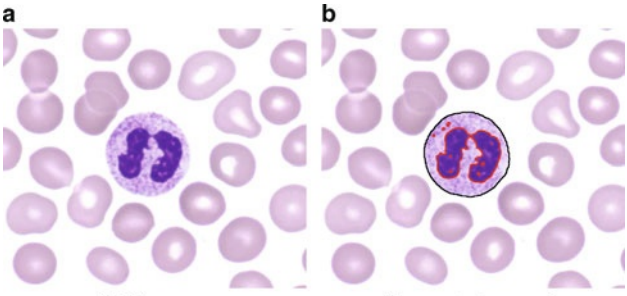



**Fig. 3** A typical leukocyte cell (Neutrophil Granulocyte, (**a**)) and the visualized result of the segmentation (**b**)

a leukocyte from surrounding and touching erytrocytes. Figure 3 shows the result of the segmentation of a typical leukocyte cell. A detailed description of the algorithm can be found in [1, 15].

### 2.2.5 Classification

Each type of leukocyte features special characteristics that are for example based on the shape of the cells nucleus, the size of the nucleus in respect to the whole cell size, the texture of the plasma, the color of the plasma and the nucleus and many more attributes. Additionally to the image data the results of the preceeding segmentation step are used as an input for the classification algorithm of the leukocytes. In the current version the considered leukocyte types are neutrophil, basophil, and eosinophil granulocytes as well as lymphocytes, monocytes, nuclei shadows, and large platelets.

This classification approach comprises a set of classifiers, where a separate classifier is applied for each class the system is able to classify into. The final decision for a certain class is obtained by comparing the classification probabilities of the different classifiers and choosing the one with the highest probability. As the set of features that could be used for classification could be very large, for each classifier and hence for every cell type a set of features was selected in an optimization process before. This was done by training each classifier on a validated reference image database and evaluating its performance on a test database, where the features that perform best are selected in the end.

## 3 Certification

According to the German Medizinproduktegesetz (MPG, German law for medical devices) § 3(2) and the European Directive 98/79/EC the HemaCAM system is considered an in vitro diagnostic device with all legal consequences. Therefore, an appropriate certification and performance evaluation process with the required documentation had to be conducted. As the Fraunhofer IIS is a publically supported research institute and not a medical device manufacturer, these work packages have been conducted at the Fraunhofer IIS in cooperation together with an industrial partner (Horn Imaging GmbH, Aalen) who acts as manufacturer and distributor of the HemaCAM system. For the performance evaluation performed in close cooperation with a specialized hematology laboratory a set of 337 blood samples containing a total of 79 072 cells was collected. These samples were then analyzed with the HemaCAM system and compared to the manually annotated results of the laboratory specialists.

There were three main topics the evaluation study focused on. The system achieved a detection rate for the counting area of 99.41%. The detection rate for

**Table 1** Pre-classification results of the evaluation study

| Cell class | Sensitivity | Specificity | Cell class | Sensitivity | Specificity |
|---|---|---|---|---|---|
| Neutrophil | 98.40% | 98.53% | Basophil | 84.63% | 99.71% |
| Lymphocyte | 98.74% | 96.32% | Nuclei Shadow | 88.09% | 99.08% |
| Monocyte | 92.67% | 99.51% | Large Platelet | 88.09% | 99.57% |
| Eosinophil | 98.56% | 99.64% | | | |

the leukocyte cells in 100-fold magnification was 99.98%. Furthermore, the pre-classification rate for seven cell classes was evaluated, the results can be found in Table 1.

## 4 Conclusion

The HemaCAM system supports hematology laboratories in the creation of white blood differentials, minimizing workload for reviewing and counting cells. Cell-based image documentation and an extensive training database lead to objective and reproducible results and enhanced diagnostic quality. Since October 1st, 2010, HemaCAM is now a certified in vitro diagnostic device and that is available on the market and installed in several laboratories in central Europe. Feedback from these laboratories is incorporated into new versions and stimulates future research and development activities at Fraunhofer IIS. The most important related activities in progress are concerned with the red blood cell morphology, an improved slide-loading system, and the automatic analysis of bone-marrow samples for which some preliminary results already have been published by Zerfaß et al. [14]. Thus, a variety of new features are expected to become available in the near future.

## Acknowledgements

This work was partially funded by the German Federal Ministry of Education and Research. The authors gratefully acknowledge the support of Dr. Thomas Nebe, Onkologikum Frankfurt, during the performance evaluation and beyond.

## References

1. Bergen T, Steckhan D, Wittenberg T, Zerfaß T (2008) Segmentation of leukocytes and erythrocytes in blood smear images. In Proc's IEEE Eng Med Biol Soc, pp 3075–3078
2. Dorini L, Minetto R, Leite N (2007) White blood cell segmentation using morphological operators and scale-space analysis. In Proc's XX Braz Symp on Comp Graphics & Image Processing, pp 294–304

3. Herzenberg LA, Parks D, Sahaf B, Perez O, Roederer M, Herzenberg LA (2002) The history and future of the fluorescence activated cell sorter and flow cytometry: A view from stanford. Clincial Chemistry 48(10):1819–1827
4. Jiang K, Liao QM, Dai SY (2003) A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering. In Proc's 2nd Int. Conf. on Machine Learning & Cybernetics, pp 2820–2825. 2.-5.11.2003, Xi'an
5. Lee S (2005) Virtual microscopy:applications to hematology. Lab Hematol 11(1):38–45
6. Liao Q, Deng Y (2002) An accurate segmentation method for white blood cell images. In IEEE Int Symposium on Biomedical Imaging (ISBI), pp 245–248
7. Rupp S, Schlarb T, Haßlmeyer E, Zerfaß T (2011) Fully automated detection of the counting area in blood smears for computer aided hematology. In Proc's IEEE Eng Med Biol Soc, vol 2011, p. accepted
8. Steckhan D, Bergen T, Wittenberg T, Rupp S (2008) Efficient large scale image stitching for virtual microscopy. In Proc's 30th IEEE Int Conf of the Engingeering in Medicine & Biology Society EMBS 2008, pp 4019–4023
9. Steckhan D, Paulus D (2010) A quadratic programming approach for the mosaicing of virtual slides that incorporates the positioning accuracy of the microscope stage. Proc's IEEE Eng Med Biol Soc 1:72–7
10. Steckhan D, Wittenberg T (2009) Optimized graph-based mosaicking for virtual microscopy. In Proc's SPIE Medical Imaging 2009, vol 7259, Orlando, FL, USA
11. Tatsumi N, Pierre RV (2002) Automated image processing: Past, present, and future of blood cell morphology identification. Clinics in Laboratory Medicine 22(1):299–315
12. Wittenberg T, Kuziela H, Mues-Hinterwäller S, Münzenmayer C, Couronné R (2005) Active Computer Vision: A necessary Paradigm for Computer Assisted Diagnosis. In Kalender W, Hahn E, Schulte A (eds) Proc's ICMP 2005 & BMT 2005, Biomed Tech, 50(1):294–296, Int Workshop
13. Wu HC (2007) The Karush-Kuhn-Tucker optimality conditions in an optimization problem with interval-valued objective function. European Journal of Operational Research 176(1):46–59
14. Zerfaß T, Haßlmeyer E, Schlarb T, Elter M (2010) Segmentation of leukocyte cells in bone marrow smears. In Proc's 23nd IEEE Int. Symp. on Computer-Based Medical Systems CMBS 2010
15. Zerfaß T, Rehn T, Wittenberg T (2008) Boundary precise segmentation of nucleus and plasma of leukocytes. SPIE Medical Imaging 2008: Image Processing 6914,69,143Q

# Face Detection with the Sophisticated High-Speed Object Recognition Engine (SHORE)

**Tobias Ruf, Andreas Ernst, and Christian Küblbeck**[†]

**Abstract** An approach enabling the detection, tracking and fine analysis (e.g. gender and facial expression classification) of faces using a single web camera is described. One focus of the contribution lies in the description of the concept of a framework (the so-called *Sophisticated High-speed Object Recognition Engine* – SHORE), designed in order to create a flexible environment for varying detection tasks. The functionality and the setup of the framework are described, and a coarse overview about the algorithms used for the classification tasks will be given. Benchmark results have been obtained on both, standard and publicly available face data sets. Even though the framework has been designed for general object recognition tasks, the focus of this contribution lies in the field of face detection and facial analysis. In addition a demonstration application based on the described framework is given to show analysis of still images, movies or video streams.

## 1 Introduction

In the past decade, several technologies enabling the detection of faces have been developed and improved, especially with respect to its performance, error rates and computational effort [11, 17, 19]. The same is valid for systems that analyze faces in order to extract information such as duration of focus, gender, age or type of reaction. Nevertheless, we believe that there are also other important features that such technology should focus on. The usage of such software in dedicated applications should be efficient, that developers can integrate such technologies with small effort. Furthermore it should provide capabilities for configuration which makes the software easy to use and offers extensive configuration possibilities.

Hence, a method for fast and robust face (and object) detection and analysis is presented, where the algorithms are embedded in a library entitled SHORE, an

Tobias Ruf (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany, tobias.ruf@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_23, © Springer 2011

acronym for *Sophisticated High-speed Object Recognition Engine*. By providing different setups, the engine can be used either for plain face detection or for more complex tasks such as gender classification, face tracking, analysis of facial features, or the analysis if the eyes are opened or closed.

In this contribution we describe the functionality and the capabilities of the SHORE system and explain the design of the framework in order to support the two side conditions mentioned above. We also describe the image processing algorithms applied for face detection and present results of benchmarks on standard data sets. Some aspects of this work already have been published in [3].

## 2 Functionality

The framework of SHORE supports and enables the detection, analysis and identification of various types of objects. In the past, we have mainly focused on the automatic detection of faces however it is also possible to detect and analyze other objects consisting of a typical structure, such as cars, hands [1] or great apes [2].

Initially, for a training phase a certain amount of annotated reference sample images is needed in order to build an adequate image model. In this context, *annotated* means that in advance a certain set of specific key features or landmarks have to be defined on the objects, and that these landmarks have to be marked manually in each image of the reference data set. Based on these annotations, a reference object model is computed beforehand which can then be applied to detect similar objects in yet unseen images.

Using such a model, we have built a system that is able to rapidly detect faces in arbitrary images and to further analyze them. Using the trained system, we are able to detect the positions of the eyes, the nose and the mouth. Additionally, we can recognize if the eyes and the mouth are opened or closed. Furthermore, we have developed and trained classification models for gender recognition and the estimation of facial expressions such as ‘happy’, ‘angry’, ‘sad’ or ‘surprised’.

## 3 The Framework

According to [7], software frameworks should be regarded as reusable design solutions for a certain domain of applications or problems. Within our work, the presented framework of SHORE addresses the problems of object detection, object analysis and object identification. To this end, we provide a so-called black-box framework [14] to the user or client, who only needs to know how to configure it and integrate it into a larger scale application. Hence, all internal details are hidden by the framework interfaces.

Figure 1 depicts an overview with the fundamental parts of the framework including three applications (seen on the top), making use of the framework in different

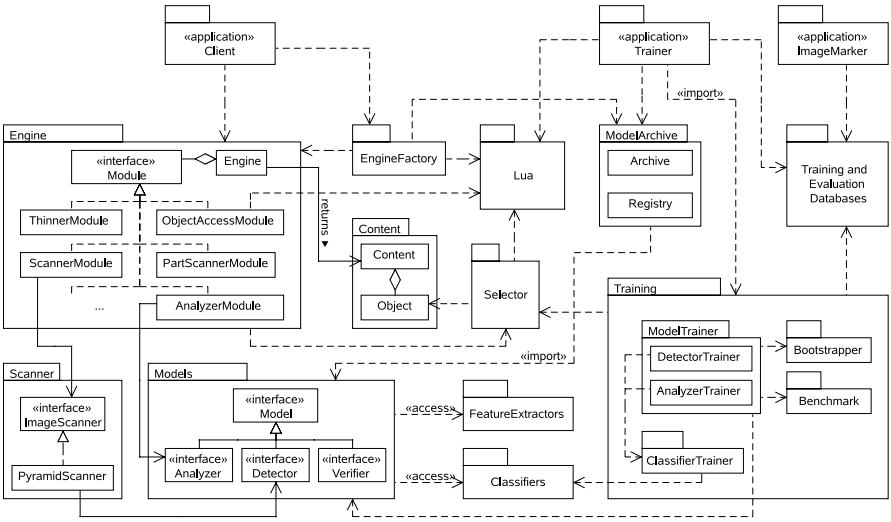



**Fig. 1** Overview with the fundamental components of the SHORE framework as an UML diagram

cases. Runtime components are located on the left and the center. The right side shows parts employed for model training and annotation of images.

With respect to this framework our design efforts were based on the following objectives:

- Support of the functionality (face detection and analysis) as described in the previous section,
- easy and flexible configuration and simple use,
- generic internal object representation,
- reusability and extensibility,
- loosely coupled training, and
- generic annotation support of reference image databases.

To cover the functionalities of object detection, analysis and recognition the concept of an object *Model* has been introduced and implemented. Models in general can be seen as an abstract representation of object characteristics, that are used for classification tasks (see left and center bottom in Fig. 1). Three interfaces for specialized models have been introduced, which insert functionalities for the *detection*, *analysis* and *verification* of the objects into the framework.

Capabilities for the configuration of the SHORE are achieved by two packages, referred to as *Engine* and *EngineFactory*. A user can set up a new *Engine* by the use of the *EngineFactory* and the assistance of the Lua scripting language [9].

The detection of certain objects in arbitrary images requires an internal representation of such objects. Using a generic representation, different types of objects (faces, eyes, hands, ...) can be handled with one single concept. Hence, an *Object* representation refers to its type, a location and region in the image, name value pairs

for ratings, attributes, marker points and other features. Using these class attributes, different types of objects can be represented. As a result of the detection and analysis process the *Engine* returns the *Content* of an image consisting of all detected and analyzed objects.

Extensibility and reusability are important issues for frameworks, these issues are addressed in several parts of the SHORE framework. One example are the modules in the package *Engine*, since new modules can easily extend the functionality.

The following sections will coarsly describe the three major use cases for the SHORE framework, namely the image annotation, the model training and the application of the engine.

## *3.1 Image Annotation*

Figure 2 depicts a screen capture of the so-called *ImageMarker* application as a part of the SHORE framework. The *ImageMarker* has been developed as a tool for flexible interactive annotation of objects in images and supports the selection and delineation of an object region in the image as well as the annotation of the object type. Furthermore, predefined and named landmarks and key attributes can be annotated for each object. The type of the objects and how these objects have to be annotated is configured in the application settings. New object types can be added any time by defining them in a preference file.

Using face detection as an example, a face in frontal position is annotated in the image by selecting both eyes, the tip of the nose and both mouth corners as landmarks. Each of these predefined markers has a key value attached, describing the marker as e.g. *'LeftEye'*. Additionally, different attributes can furthermore be assigned to the face. In the given example, the gender or facial expression can be annotated using the tool, e.g. *Gender = 'Male'* or *Expression = 'Happy'*. A screenshot of this sample annotation using the ImageMarker is provided in Fig. 2.

The output of the annotation process is an XML file containing the information about all annotated objects in the image. The annotated image databases are then used for training and benchmark of the object models within the SHORE framework as shown in Fig. 1.

## *3.2 Training the Models*

The training concept within the SHORE framework has two important properties. First the training parts are decoupled from the remaining framework. This feature allows to provide only the needed components at runtime to the user of the library. Secondly, the complete training is scriptable. Lua-based [9] training scripts are used for the training, defining both the training data and the related tasks. The SHORE *Trainer* application in Fig. 1 can parse these scripts with the help of Lua [9] and

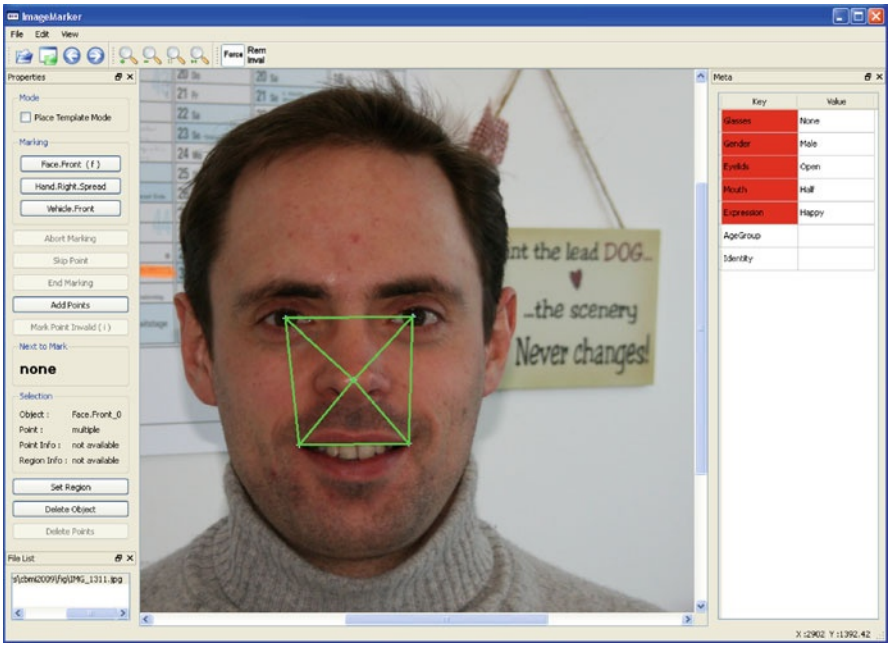



**Fig. 2** Screen capture of the *ImageMarker* tool with a sample annotation of a face in frontal position. As landmarks the center of the eyes, the tip of the nose and the corners of the mouth have been marked

uses the so-called *Training* package to carry out the training. A typical training script comprises the following workflow:

- select training images from the annotated reference databases,
- choose feature extractors,
- choose trainers for the classifiers,
- start training with a concrete model trainer,
- optimize the model,
- benchmark the trained model,
- save the model.

The training of classifiers requires the definition of two or more classes. Each class for the training can be created by selecting a subset of the training images. A concrete model trainer knows how to train a suitable model and is implemented in the package *ModelTrainer*.

## *3.3 Using the Engine*

The left side of Fig. 1 shows the usage of the framework from the viewpoint of a client application. Essentially, the client makes use of the *EngineFactory* and the

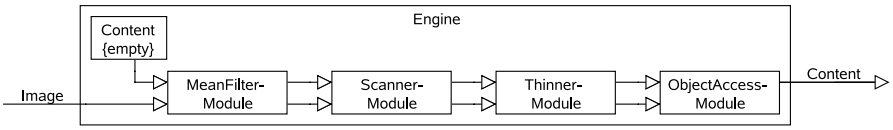



**Fig. 3** An engine setup configured for a simple face detection task

*Engine* itself. More specific, the *EngineFactory* is able to create a custom tailored *Engine* by parsing a Lua-based setup script defining the configuration of the engine.

Depending on the setup script, a concrete engine can conduct image enhancement jobs as well as complex object detection and analyzing tasks. While a simple engine is able to detect faces (objects) in the images, a more complex engine provides an additional feature analysis as well as a facial expression or gender analysis. The engine itself encapsulates a concatenation of different image processing modules as shown in Fig. 3 for a configuration example. When using the engine, the image as well as a *Content* instance is provided to the defined modules. Each module is allowed to use or modify the image data and/or access the objects in the content.

A software client may use the SHORE framework in the following way: First it provides the setup script to the *EngineFactory*. The factory parses the script to create an adequate *Engine* and returns it back to the client application. Thereafter the *Engine* is ready to process images provided by the application. For each new image the *Engine* returns the *Content* containing all objects and information gathered by the *Engine*. The client application is able to read out the content for each image and can provide further processing steps depending on the information in the content. When the *Engine* is not needed any more it is destroyed by the application.

## 4 The Algorithms

The algorithmic part of the SHORE consists of four parts, which make up the integrated pattern classification system: preprocessing, feature extraction, classification as well as training. A more detailed description of the algorithmic workflow is provided in [10].

As a preprocessing step a mean or box filter is applied which has turned out to be an efficient and fast method to smooth noise artifacts. For our face detection application no further preprocessing is needed.

As the face detection scenario should be applicable under varying illumination conditions the features providing information about the faces should be somewhat independent to illumination changes. Specifically so-called census features (also known as local binary patterns) [20] are been used. These features are defined as structure kernels of size $3 \times 3$ summarizing the local spatial image structure. Within the kernel structure information is coded as binary information $\{0, 1\}$ and the resulting binary patterns can represent oriented edges, line segments, junctions, ridges and saddle points. Furthermore, resized versions of these features are used which

**Table 1** Comparison of selected algorithms with respect to detection rate on the CMU+MIT dataset. All numbers are given in percent

| False positives | 10 | 31 | 65 |
|---|---|---|---|
| This work | 91.5 | 93.3 | 93.9 |
| Viola and Jones [16] | 83.2 | 88.4 | 92.0 |
| Garcia and Delakis [8] | 90.5 | 91.5 | 92.3 |

means that the features are not only calculated on basic $3 \times 3$ patterns but also on $3n \times 3m$ multiples of these windows. Using all these features a *feature pool* is created. Other features in the pool are edge orientation features which can be computed for each feature as $\text{atan2}(s_x, s_y)$ whereas $s_x$ and $s_y$ relate to the result of the Sobel filter applied to the respective pixel in $X$- and $Y$-direction, as described in [6]. Doing so, starting from a $24 \times 24$ window a total of approximately 30 000 features is obtained, building the set $\mathcal{C}$ of available features.

During the training phase these features are selected and weighted using the AdaBoost method. By using boosting a number of weak classifiers can be combined to form a final strong classifier. The goodness of a weak classifier is measured by its error $\epsilon$ on the training set. The final classifier then consists of a set of look-up tables $\{h_{\mathbf{x}}; \mathbf{x} \in \mathcal{C}\}$ for the features $\mathbf{x}$ chosen by the algorithm. Each look-up table holds a weight for each feature value. This method is described in more detail in [4, 10]. One important advantage using the AdaBoost method is that a very fast classifier that uses look-up tables for the classification process is obtained. The detection can be further improved by using multiple stages as described in [10] and [16]. Further speed improvements are made by carrying out a coarse-to-fine grid search as explained in [6].

## 5 Benchmark Results

The performance of our system was tested on several image data sets. None of the datasets that were used for benchmark are part of the training sets, all sets of data are disjoint. We show rates for face detection, gender classification and the classification of happy faces. The face detection rates were calculated on the CMU+MIT (consisting of 130 images with 507 faces) and the BioID data base (consisting of 1521 images with 1522 faces[1]). The outcomes are shown in Fig. 4. When examining the detection rate for certain false positives we compare to competing technologies as shown in Table 1. The numbers from Viola and Delakis are picked from the paper [15], where also other methods are compared.

The following measurements were made on images where we used pre-annotated eye-positions. The given recognition rates are based on the optimal classification thresholds chosen separately for each test data set.

[1] In the BioID data set there is one image (number 1140) showing two faces.

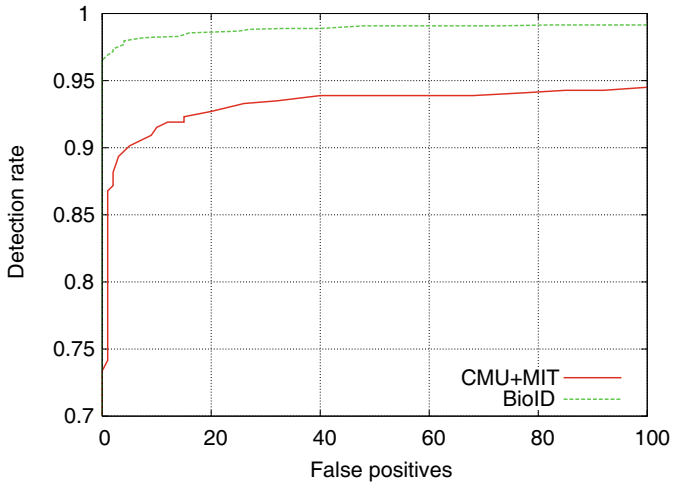



**Fig. 4** ROC-curve for the two data bases BioID [5] and CMU+MIT [18]

**Table 2** Assessment of computation time for different engine setups on a single core of a Intel Core 2 Duo 6420 CPU for an image of size 384 × 286

| | | | | |
|---|---|---|---|---|
| Face detection | × | × | × | × |
| Eye fine search | | × | × | × |
| Gender classification | | | × | × |
| Analysis of 4 expressions | | | | × |
| Time [ms] | 9.4 | 19.3 | 19.9 | 22.0 |

The performance of our gender classification module was tested on the BioID data set and on the Feret fafb data set [13]. We receive a recognition rate of 94.3% on the BioID data base and 92.4% on the FERET fafb data base.

Our happiness analyzer was evaluated on the JAFFE data base [12]. It consists of 213 images of Japanese women with 31 annotated as happy. Here the recognition rate is 95.3%.

In Table 2 we show the results of speed measurements on our system. The measurements were carried out on an Intel Core 2 Duo 6420 CPU. Although the library supports multi-threading we only used one core in this test. For evaluation we used the full BioID data set containing images of size 384 × 286 and we determined the average calculation time per image.

## 6 Conclusions

In this contribution we have described a modular framework (SHORE) to detect and analyze general objects in arbitrary images, with an application to face detection. The SHORE framework features flexibility and modularity so that the integration to

specific tasks can be carried out efficiently. We believe that due to these properties the system is well suited for object detection scenarios in real-time environments.

People interested in trying out the described system can download and evaluate the demo software from http://www.iis.fraunhofer.de/en/bf/bsy/produkte/shore/. At the moment we are involved in adding functionality as pose estimation and emotion recognition. In addition we intend to train further classifiers for example for robust hand detection. Furthermore the different components like feature extractors and classifiers will probably be extended in the future.

## Acknowledgements

This research work was partly supported by the European Commission under contract: FP6-2005-IST-5 (IMAGINATION), http://www.imagination-project.org.

## References

1. Arndt I, Ernst A, Friedl S, Kage A, Münzenmayer C, Küblbeck C, Wittenberg T (2010) Gesture based segmentation of medical imagery for sterile environments: A new approach. Proc's Int J Comput Assist Radiol Surg 5(1):407–409
2. Ernst A, Küblbeck C (2011) Fast face detection and species classification of african great apes. In 2011 IEEE International Conference on Advanced Video and Signal based Surveillance (AVSS 2011). Klagenfurt, Austria
3. Ernst A, Ruf T, Küblbeck C (2009) A modular framework to detect and analyze faces for audience measurement systems. In Fischer S, Maehle E, Reischuk R (eds) GI Jahrestagung, LNI, 154:3941–3953
4. Freund Y, Shapire RE (1999) A short introduction to boosting. Journal of Japanese Society for Artificial Intelligence 5(14):771–780
5. Frischholz R (2009) The BioID face database. Website. Available online at http://www.bioid.com/downloads/facedb/index.php, visited on January 7th
6. Fröba B, Küblbeck C(2002) Robust face detection at video frame rate based on edge orientation features. In International Conference on Automatic Face and Gesture Recognition (FG '02), pp 342–347, Washington DC
7. Gamma E, Helm R, Johnson R, Vlissides J (1994) Design Patterns – Elements of Reusable Object-Oriented Software. Addison-Wesley
8. Garcia C, Delakis M (2004) Convolutional face finder: A neural architecture for fast and robust face detection. IEEE Trans. Pattern Anal Mach Intell 26(11):1408–1423
9. Ierusalimschy R (2003) Programming in Lua, 1st edn, Roberto Ierusalimschy
10. Küblbeck C, Ernst A (2006) Face detection and tracking in video sequences using the modified census transformation. Image Vision Computing 24(6):564–572
11. Lienhart R, Maydt J (2002) An extended set of haar-like features for rapid object detection. In IEEE ICIP, 1:900–903
12. Lyons M, Akamatsu S, Kamachi M, Gyoba J (1998) Coding facial expressions with gabor wavelets. In Proc'S 3rd IEEE Int Conf on Automatic Face and Gesture Recognition, pp 200–205
13. Phillips PJ, Wechsler H, Huang J, Rauss PJ (1998) The feret database and evaluation procedure for face-recognition algorithms. Image and Vision Computing 16(5):295–306

14. Riehle D (2000) Framework design – a role modeling approach. Ph.D. thesis, Swiss Federal Institute of Technology Zurich
15. Verschae R, del Solar JR, Correa M (2008) A unified learning framework for object detection and classification using nested cascades of boosted classifiers. Machine Vision and Applications
16. Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In Proceedings IEEE Conf on Computer Vision and Pattern Recognition
17. Viola P, Jones M (2002) Robust real-time object detection. International Journal of Computer Vision 57(2):137–154
18. Wang CC (2009) CMU image data base: face. Website. Available online at http://vasc.ri.cmu.edu/idb/html/face/frontal_images, visited on January 7th
19. Wu B, Haizhou A, Chang H, Shihong L (2004) Fast rotation invariant Multi-View face detection based on Real Adaboost. In Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp 79–84
20. Zabih R, Woodfill J (1996) A non-parametric approach to visual correspondence. IEEE Trans. on Pattern Analysis and Machine Intelligence

# Improving Sheet-of-Light Based Plant Phenotyping with Advanced 3-D Simulation

**Franz Uhrmann, Lars Seifert, Oliver Scholz, Peter Schmitt, and Günther Greiner**

**Abstract** Understanding plant growth and analyzing plant interaction with the environment is an important aspect in modern agronomy and biological sciences. While measurements are often taken at field scale, current research focuses increasingly on individual plants. As a manual determination of morphological plant parameters is very time-consuming, automatic acquisition methods at high throughput are required. Optical scanning methods provide fast acquisition of surface points. However, as plants represent geometrically complex objects, planning a proper measurement setup and evaluation of the acquired data is a challenging task. This paper addresses solutions for system design and data processing for the sheet-of-light measurement method. As an example implementation a 3-D scanning system for individual plants is presented, which is amended by color images for high resolution surface and color measurements of individual plants.

## 1 Introduction

In the recent years the acquisition of morphological plant parameters has been of increasing importance in various disciplines of research: Breeders have to assess cultivars to select the best plants for breeding in order to optimize e.g. crop yield. Due to a high degree of breed optimization the morphological variations are very small and thus can hardly be rated by manual inspection alone. So an automatic method is required to determine morphological parameters objectively. Furthermore especially considering the climate change it is of importance to breed cultivars which are resistant to drought and heat stress. In this field of research automatic phenotyping systems are required in order to quantify the adaptiveness of plants to stress situations and to find correlations of phenotypic parameters with genomic properties.

Franz Uhrmann (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
franz.uhrmann@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_24, © Springer 2011

While previously the focus has been on measurements at field scale, the mentioned applications require parameter estimation of individual plants like leaf area, number of leaves and leaf parameters like slope and inclination angle. Depending on the specific application, the size, composition and architecture of plants can vary significantly, both when considering different plant species as well as individuals of a single species (for example the leaf arrangement may be steep when individuals suffer from heat stress). Furthermore the environmental circumstances can vary significantly. When it is necessary to measure plants in their natural environment, optical ascertainability can be difficult due to dense crop or tillering in late growth stages.

In this paper we present two measurement methods based on the sheet-of-light measurement principle for plant phenotyping, followed by a discussion of the required data preprocessing algorithms. As an example an implementation of a measurement system for *Arabidopsis* plants is presented. Finally present limitations and ideas for improvement like sheet-of-light simulation are discussed.

## 2 Related Work

The acquisition and measuring of morphological plant parameters is important in agronomy and biological sciences in order to assess single plants or to develop structural plant models and correlate them with functional aspects of crops. An overview of the variety of measurement methods is given in [19]. The approaches can be divided into tactile and contactless measurements.

Taking contact measurements of individual plants is very time and labor intensive, as every leaf must be touched with a test prod at various positions. Further drawbacks are the limited applicability of tactile measurements due to interference with the plant, low resolution and limited usefulness for small plants. Still it is commonly used due to standardized acquisition sequences for different plant species and the possibility to directly derive geometric parameters from the digitized 3-D coordinates [6, 13, 16].

In contrast to manually performed tactile methods, image-based contactless methods provide fast automatic acquisition of the plant's surface, generating dense point clouds as a base for further processing. The recorded data must be interpreted in a subsequent preprocessing step, i.e. the features of interest must be extracted from the captured data. Especially for complex objects this still is a challenging task, thus approaches found in literature are highly application specific. Popular optical acquisition methods include stereoscopic, time-of-flight and sheet-of-light methods.

Andersen [2] extracts depth values using a stereo camera setup to estimate geometric attributes like leaf area and height of wheat plants under laboratory conditions. A stereo imaging system with consumer cameras has been used by Biskup [3] to acquire 3-D point clouds from canopies, from which leaves are segmented to cal-

culate spatial leaf parameters. In [1] photogrammetry is used to approximate the bounding volume of tomato plants.

Recently time-of-flight systems have become commercially available, and due to the ease of handling and straightforward 3-D reconstruction algorithms, this method has also been utilized for plant measurements [9,14]. Time-of-flight sensors provide depth information derived from the time between emittance and reflectance of light pulses [11]. Presently, the data quality is significantly inferior to conventional methods due to its low resolution and 3-D-noise.

Another popular 3-D measuring method is sheet-of-light, where the object's surface is scanned by a laser line and the laser profile is captured by an optical camera. With a well-designed measurement setup, a robust and fast acquisition of dense 3-D coordinates in high resolution is possible. Thus sheet-of-light is well established for industrial applications. In the field of agricultural engineering Seatovic [15] presents a field system to automatically detect weed. Kaminuma [7] describes a scanning system for *Arabidopsis* assessment.

## 3 Methodology

### *3.1 Sheet-of-Light Measurement*

Sheet-of-light measurement is based on the triangulation principle. A laser is spread to a line, which is projected onto an object and follows the curvature of the object's surface. The camera acquires the diffuse reflection of the laser line from an angle and thus captures a single height profile of the object at the time of recording (see Fig. 1a).

In order to scan an object's surface, either the object or the sensor needs to be moved. The more height profiles are recorded during the scan, the higher the resolution in the direction of movement. Conventional video cameras only produce 30 frames per second, equivalent to only 30 profiles per second. There are smart cameras available which were specifically designed for sheet-of-light measurements, capable of generating in excess of 30 000 profiles per second. At these frequencies, it is possible to scan objects at high movement speeds and still achieve a high resolution in the direction of movement. The lateral and height resolution depend on the type of camera used and the measurement setup. The camera receives the light from the intersection of the sheet-of-light's plane with the object on its sensor's pixels. Thus the sensor resolution and the camera's field of view directly define the resolution of any given setup.

There are two related effects that can keep a camera from detecting the laser light reflected by the object though. The first effect is called shadowing (see Fig. 1b): The camera can theoretically see a part of the object intersecting the sheet-of-light plane, but another part of the object casts a shadow on this spot. Only the first intersection of the sheet-of-light can be detected. A similar effect is called occlusion

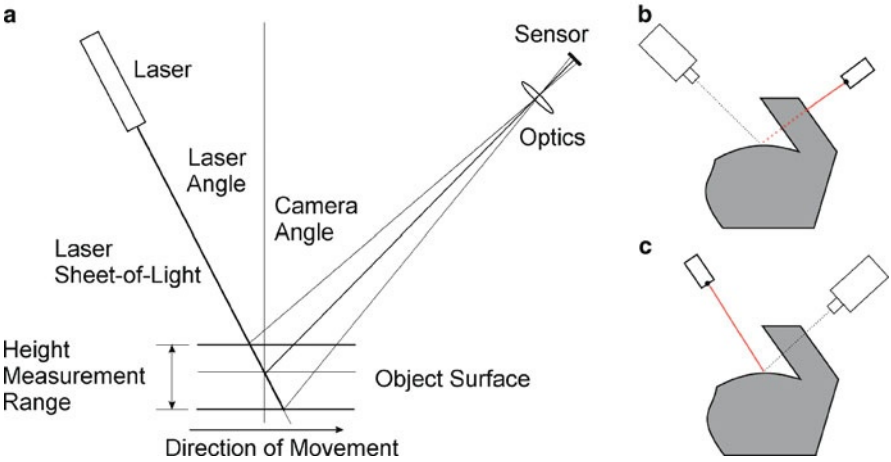



**Fig. 1** Sheet-of-light measurement: (**a**) Measurement principle. (**b**) Shadowing by laser interruption by the object. (**c**) Camera view occlusion by the object

(see Fig. 1c): The laser light hits the object, but the point of intersection is not visible for the camera. Both effects lead to an incomplete scan of the object's surface. Shadowing can be reduced using more lasers illuminating the object from different directions. For occlusion, more cameras are required to capture the laser light hitting the object.

## *3.2 Simulation-Guided System Design for Complex Objects*

For each specific scenario and application the measurement setup must be planned carefully to allow optimal acquisition in high quality. Particularly as components like smart sheet-of-light cameras are expensive it is desirable to maximize acquisition coverage with few measurement resources. As the number of possible design choices is large and measurement setup can vary a lot when being adapted to specific applications, experimental system design by trial and error quickly becomes very difficult and time-consuming.

For a fast and extensive exploration of possible configurations we implemented a simulation tool where a virtual measurement system can be set up from an arbitrary number of cameras, lasers and objects. The physical parameters for each camera (e.g. sensor resolution, optical field-of-view, position and view direction) and laser (e.g. intensity, fan angle and position) can be modified interactively. Objects can be imported as polygon mesh files (e.g. from CAD models) and arbitrarily placed and moved in the virtual scene.

After a system has been designed a virtual data acquisition can be performed. Effects like occlusion and shadowing become apparent immediately and the designer

may change individual system parameters and receives feedback quickly. This can be used as a basis for further quality measurements. The simulated data can also be analyzed with respect to different quality criteria like the coverage of the object or measurement resolution on the object's surface. Figure 4 shows the visualization of the simulation result with a *Sorghum* plant as an example for complex plant measuring.

### *3.3 Scan Data Processing*

After acquisition the scan data are being processed by subsequent processing steps in order to compute morphological plant features:

1. **Transformation to 3-D space:** Each acquired range image point $p_i = (i, j, d)^T$ at pixel position $(i, j)$ with depth value $d$ can be transformed to a world coordinate $w_i = (x, y, z)^T$ by applying a transformation $M$: $p_w = M \cdot p_i$. The parameters of $M$ can be estimated with a known set of image points and their corresponding world coordinates [10]. These point correspondences are obtained from a scan of a calibration target with defined geometry. The calibration must be performed for each sheet-of-light sensor in order to transform measured data to a common coordinate system.
2. **Sensor fusion and surface reconstruction:** In this processing step a single surface representation in terms of a polygon mesh is generated from each sensor's transformed world point set. Surface reconstruction from an unstructured point cloud is an active research topic of computer graphics, so many different approaches can be found in literature [4, 5, 8, 17, 20].
   For a sheet-of-light system where all lasers have a similar orientation, a direct fusion on the range images is possible. Therefore the range data of each sensor $i$ is transformed to the image space of an assigned reference sensor $M_0$: $p'_i = M_0^{-1} \cdot M_i \cdot p_i$. As the transformed images share the same image space, occluded areas can be reduced and the range values can be combined by averaging. Thereafter a surface reconstruction of the combined range image can be generated by creating two adjacent triangles for a neighborhood of four range image points. At the object border points may appear close in the range image while being distant in world coordinates resulting in wrongly elongated triangles. A simple method to remove those triangles is to determine the average scan resolution either analytically (if the physical parameters of the measurement setup are known) or experimentically and discard edges that are significantly longer.
3. **Color mapping:** In order to add color to the scanned 3-D surfaces, the scan system is extended by standard industrial full frame color cameras. We apply a standard calibration method in order to project points in the common world coordinate into every camera's sensor pixel coordinate system [18]. If the point was inside the field of view, the RGB data suggests a color for the world point.
4. **Plant segmentation:** In the segmentation step background vertices are removed from the reconstructed mesh so that only the surface of the plant remains. The

segmentation is based on the color value of each vertex. As suggested in [12], the $RGB$-triplet is transformed into channel $I$, which represents the 'greenness' of each vertex: $I = 0.5G - 0.25R - 0.25B$. Vertices below a manual determined greenness value are considered as soil and are removed from the mesh.

5. **Feature calculation:** With the resulting mesh representing a single plant the calculation of simple geometric features as leaf area or minimal bounding sphere is straightforward. For the computation of more complex parameters further processing steps like decomposition into single leaves have to be applied to the plant mesh.

## 4 Implementation of a Scanner for *Arabidopsis* Assessment

### *4.1 System Description*

The scanner presented is used for the growth evaluation of *Arabidopsis* plants. *Arabidopsis* is a small plant and member of the mustard (*Brassicaceae*) family. In the year 2000 the entire genome has been sequenced. The life cycle from germination to mature seed is about 6 weeks and the plant is easily cultivated in restricted space. Therefore *Arabidopsis* is one of the model organisms used for studying plant biology.

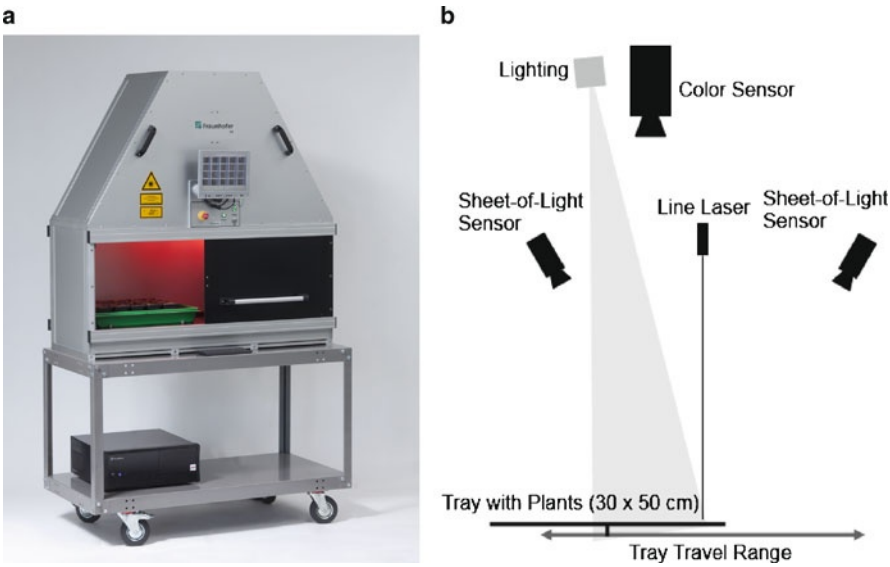



**Fig. 2** (**a**) Photo of the implemented scanner for *Arabidopsis* plants. (**b**) Schematic measurement setup

The *Arabidopsis* plant scanner is shown in Fig. 2a. A tray of for example $4 \times 5$ plants can be measured in a single measurement run. During a scan, which takes about 90 seconds, the tray is linearly moved under the sheet-of-light and color sensors. To reduce the size of the plant scanner, the measurements of shape and color are separated. First, the shape of the plants is acquired by the sheet-of-light sensor (see Sect. 3). During the return to the starting point, the color measurement takes place and the measured height data is processed in parallel. The plant scanner has a lateral resolution of 0.1 mm. We use a 10 mW class 2 line laser. The system has a safety circuit in a closed housing and the laser can only be turned on with the door closed. The plant scanner can be controlled via a touch screen mounted on the front panel.

## *4.2 Measurement Precision and Reproducibility*

To determine the measurement precision, we use a sheet of green cardboard as a reference. The cardboard reference has an area of $400 \pm 16\,\text{mm}^2$. This reference area is measured under different angles to simulate the different arrangement of the leaves. The result is shown in Fig. 3a. The measured mean area value is $396.1\,\text{mm}^2$ with a standard deviation of about $6.8\,\text{mm}^2$. In addition to the errors due to a limited resolution, segmentation errors at the edges of the measured area are the main reason for deviations.

The uncertainty rises with increasing angle. This is because of the higher influence of segmentation errors on the area calculation. Furthermore the intensity of the diffuse reflection decreases with higher angles. This can result in local measurement failure.

A repeatability measurement of twenty plants in one tray is shown in Fig. 3b. Each tray is consecutively measured ten times. The standard deviation relative to the measured leaf area is below 2%.

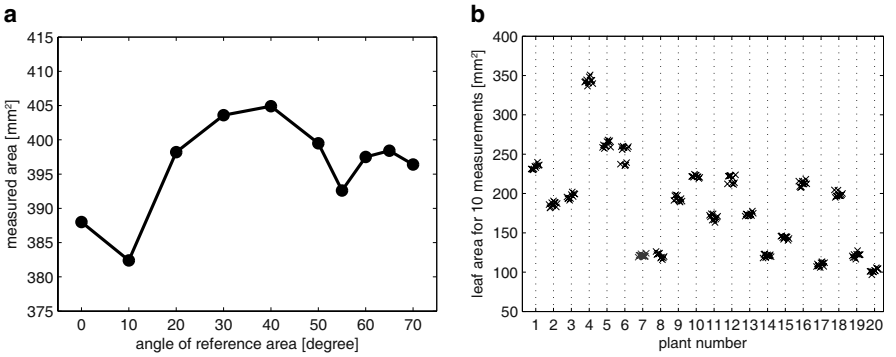



**Fig. 3** Measurement results: (**a**) Measured area of a reference cardboard at different angles. (**b**) Measured area of 20 plants for 10 consecutive scans

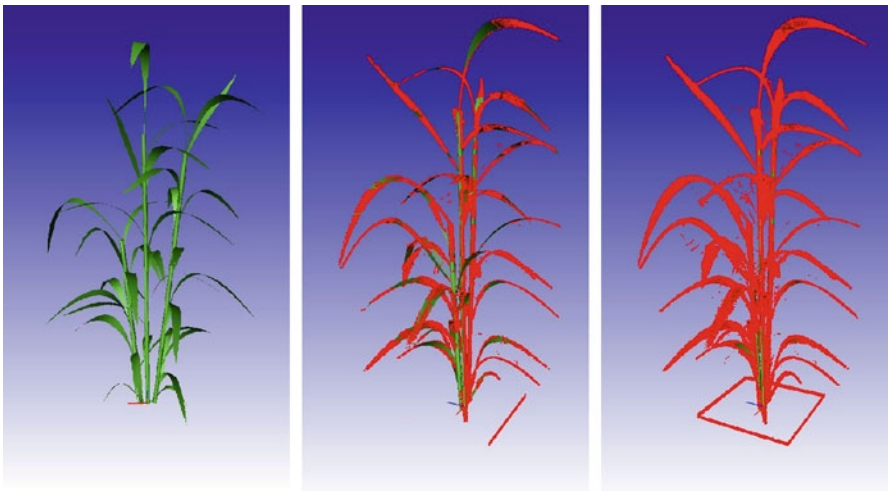


**Fig. 4** Coverage analysis for a *Sorghum* plant. *Left*: A geometric plant model used as input for the simulation. *Center*: A sheet-of-light camera and laser is virtually placed and moved vertically in order to scan the plant's surface. Scanned surfaces are marked *red*. *Right*: Visualization of the coverage when four camera-laser-pairs at different positions are used

## *4.3 Plant Assessment*

It is possible to perform scans of the same plants at multiple points of time. A separate software has been implemented to evaluate the plant evolution, i.e. to track the development of the obtained features over time. Figure 5 shows graphs for the leaf area of two individual plants. For plant 1 the leaf area increases continuously during growth as expected. At day 11 the sample has been treated with heat stress, which induces the plants to a steep leaf arrangement. This can be seen in Fig. 6 (top row), where color images of the plant at different points of times are presented: At day 11 the plant's rosette radius seems to decrease in the 2-D projection. However, the leaf area calculated from the 3-D scanner data increases as expected.

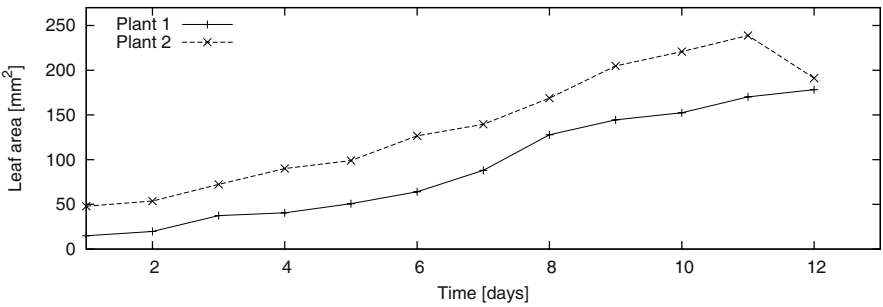



**Fig. 5** Development of the leaf area of two plants

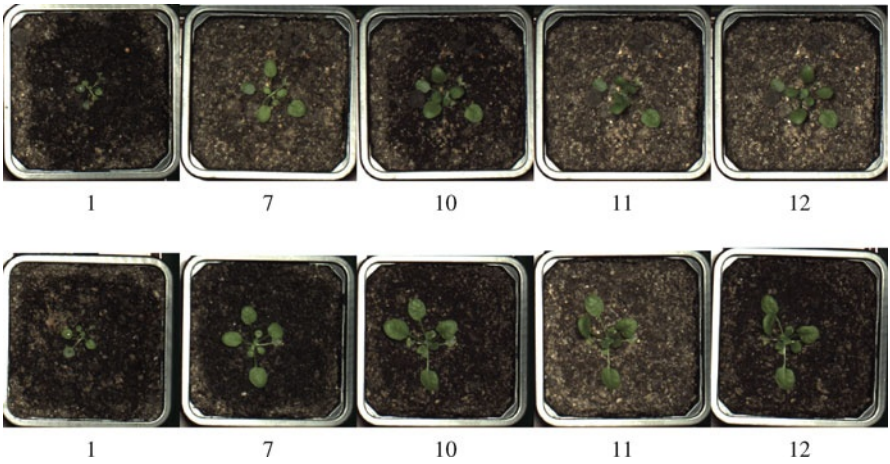



**Fig. 6** Color images of plant 1 (*top*) and plant 2 (*bottom*) shown in Fig. 5 at the specified days

For plant 2 the increment of leaf area per day is similar but the absolute value is about 50 $mm^2$ higher as for plant 1, which indicates an earlier germination date. A remarkable decline of the leaf area can be seen from day 11 to day 12. The reason for this incorrect effect becomes evident when inspecting the color images of the plant (Fig. 6, bottom row): Due to the heat stress the plant arranges the leaves that the rightmost leaf completely occludes another leaf. The occluded leaf cannot be acquired by the sheet-of-light sensor, thus it will not be considered for the calculation of the total leaf area. It is evident, that the occlusion effect due to overlapping leaves increases in later growth stages or if the architecture of the inspected plant species is more complex.

## 5 Conclusion and Future Work

The sheet-of-light method is well suited for the 3-D acquisition of plants and the automatic and accurate determination of morphological plant features. However, shadowing and occlusion lead to missing areas in the acquired data, so the requirements for a particular type of research (e.g. plant type, plant age range, etc.) must be carefully analyzed to design a good measurement setup. Even for cases with low complexity in the plant architecture, effects like overlapping leaves lead to errors in the calculated features.

In the next step a model-based approach will be implemented to supplement the acquired data realistically using a-priori knowledge of the inspected leaf shape and plant architecture. The model representation of the leaf must be flexible to allow adaption to a wide variety of applications. Additionally the geometric plant model can also be utilized in the simulation tool to simulate measurement situations even

more realistically. Interpreting the model parameters derived from a particular scan as additional plant features extends the feature set significantly, which may also lead to much better detectability of stress, diseases, or other deviations from the norm.

## References


1. Aguilar M, Pozo J, Aguilar F, Sanchez-Hermosilla J, Paez F, Negreiros J (2008) 3-D surface modelling of tomato plants using close-range photogrammetry. The Int Archives of the Photogrammetry, Remote Sensing & Spatial Inf Sciences 37:139–144
2. Andersen HJ, Reng L, Kirk K (2005) Geometric plant properties by relaxed stereo vision using simulated annealing. Computers & Electronics in Agriculture 49(2):219–232
3. Biskup B, Scharr H, Schurr U, Rascher U (2007) A stereo imaging system for measuring structural parameters of plant canopies. Plant, Cell & Environment 30(10):1299–1308
4. Curless B, Levoy M (1996) A volumetric method for building complex models from range images. In SIGGRAPH '96: Proc's 23rd Ann Conf on Comp Graphics & Interactive Techniques, ACM, New York, NY, USA, pp 303–312
5. Hoppe H, DeRose T, Duchamp T, McDonald J, Stuetzle W (1992) Surface reconstruction from unorganized points. In Proc's 19th Annual Conf on Comp Graphics & Interactive Techniques, SIGGRAPH '92, ACM, New York, NY, USA, pp 71–78
6. Kahlen K (2006) Three-dimensional architectural modelling of greenhouse cucumber (cucumis sativus l.) using l-systems. Acta Horticulturae 718:75–80
7. Kaminuma E, Heida N, Tsumoto Y, Yamamoto N, Goto N, Okamoto N, Konagaya A, Matsui M, Toyoda T (2004) Automatic quantification of morphological traits via three-dimensional measurement of arabidopsis. The Plant Journal 38:358–365
8. Kazhdan M, Bolitho M, Hoppe H (2006) Poisson surface reconstruction. In Proc's 4th Eurographics Sym on Geometry Processing, SGP '06, Eurographics Association, Aire-la-Ville, Switzerland, pp 61–70
9. Kraft M, Salomão de Freitas N, Munack A (2010) Test of a 3-D time of flight camera for shape measurements of plants. In CIGR Workshop on Image Analysis in Agriculture
10. McIvor AM (1999) Calibration of a laser stripe profiler. In Proc's 2nd Int Conf on 3-D Digital Imaging & Modeling, pp 92–98
11. Oggier T, Lehmann M, Kaufmann R, Schweizer M, Richter M, Metzler P, Lang G, Lustenberger F, Blanc N (2004) An all-solid-state optical range camera for 3-D real-time imaging with sub-centimeter depth resolution (swissranger). Soc of Photo-Optical Instrumentation Engineers (SPIE) Conf Series 5249:534–545
12. Philipp I, Rath T (2002) Improving plant discrimination in image processing by use of different colour space transformations. Computers & Electronics in Agriculture 35(1):1–15
13. Rakocevic M, Sinoquet H, Christophe A, Varlet-Grancher C (2000) Assessing the geometric structure of a white clover (trifolium repens l.) canopy using3-d digitising. Annals of Botany 86:519–526
14. Ruckelshausen A, Busemeyer L, Klose R, Linz A, Moeller K, Thiel M, Alheit K, Rahe F, Trautz D, Weiss U (2010) Sensor and system technology for individual plant crop scouting. In 10th Int Conf on Precision Agriculture
15. Seatovic D (2008) 3-D-object recognition, localization and treatment of rumex obtusifolius in its natural environment. In 1st International Conference on Machine Control & Guidance
16. Sinoquet H, Thanisawanyangkura S, Mabrouk H, Kasemsap P (1998) Characterization of the light environment in canopies using 3-D digitising and image processing. Annals of Botany 82:203–212
17. Sun Y, Dumont C, Abidi MA (2000) Mesh-based integration of range and color images. In Soc of Photo-Optical Instrumentation Eng (SPIE) Conf Series, pp 110–117

18. Tsai RY (1987) A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. IEEE J of Robotics & Automation RA-3(4):221–244
19. Vos J, Marcelis L, de Visser P, Struik P, Evers J (eds) (2007) Functional-Structural Plant Modelling in Crop Production. Springer-Verlag New York, Inc
20. Yu Y (1999) Surface reconstruction from unorganized points using self-organizing neural networks. In Proc's IEEE Visualization Conf '99, pp 61–64

# A CT System for the Analysis of Prehistoric Ice Cores

**Virginia Voland, Johannes Freitag, Norman Uhlmann, and Randolf Hanke**

**Abstract** The task was to measure ice cores with a diameter of 10 cm and a length of 1 m that were drilled out of Arctic and Antarctic glaciers down to 3000 m depth. By means of computed tomography (CT), pieces of 1 m length are measured and three-dimensional volume data with high spatial resolution are reconstructed. Complex image processing algorithms are applied to analyze the volume regarding its mean porosity as a function of depth, its total mean porosity and the volumetric distribution of the pores. Since high image quality is required to achieve precise results, a vast amount of data is acquired. The geometry of the samples increases the requirements for the X-ray components. Furthermore measurements have to be made in an environment of −15 °C.

## 1 Introduction

The polar areas covered by ice are measured by a radar system that is on board the European satellite Cryosat-2. Reliable information about ice thicknesses, however, can only be obtained by setting those data in relation with information about the porosity of the ice. Geologists obtain cylindrical drill cores with a diameter of 10 cm and a length of up to 3000 m from glaciers in polar areas like the Arctic, the Antarctic or Greenland. Cores of 1 m length have to be measured, reconstructed and analyzed in full size. Here information about inner structures is necessary to analyze the porosity of the cores. X-ray Computed Tomography (CT) is the method of choice.

For a meaningful image analysis one has to distinguish between several types of ice and several objectives: firn occurs in depths down to 100 m and is characterised by connected air structures within the core. Porous ice is obtained from depths down

Virginia Voland (✉)
Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany
virginia.voland@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_25, © Springer 2011

to 3000 m. Additionally, the transition from firn to porous ice and the occurrence of dust or volcanic ashes in certain layers are a matter of research.

These two different types of ice result in different tasks for the X-ray system. In the case of firn, where connected air structures predominate, single pores cannot be segmented and analyzed. A mean porosity per slice is calculated as a function over the length of the core. Variances of the porosity or the presence of crust allow for the determination of the occurrence of melting and freezing processes or of storms. The depth of the ice where these effects occur gives a hint on the age of the ice and hence on the period of time when climatic changes occurred.

The porous ice in contrast contains separated pores of air and gases. Their number and shape is a characteristic of the different ice layers. Therefore, not only a mean porosity per slice can be analyzed since a characterization of the number, shape and volume of the pores and their volumetric distribution contains much more information. Depending on the original depth of the ice core, several 10 000 pores can occur in segments of 1 cm height. Those pores can haven an extremely small diameter of far below 1 μm. Hence, it is desirable to increase spatial resolution in imaging in order to enable the detection and characterization of pores with a size of a few micrometers for the improvement of accuracy. A 3-D imaging system is required that achieves a low spatial resolution of a few micrometers and that at the same time is able to measure, reconstruct and analyze volumetric information of the entire drill core.

## 2 System Setup

CT, in contrast to conventional X-ray radioscopy, allows for the reconstruction of volumetric information throughout the specimen, by acquiring projective images of the object from various directions. The image quality and hence the amount of information increases with an increasing number of 2-D projections. While in medical CT the X-ray components rotate around the human body, in industrial applications usually the object is rotated while the X-ray components are not moving, because of this setup is less complex and allows for a higher precision. The firn cores are measured by means of Helical CT [3] which additionally to rotating the object applies a vertical feed. This allows for a continuous acquisition of large cylindrical objects and at the same time reduces artifacts caused by trans-versal penetration. In this application only a low spatial resolution of at least 100 μm voxel size is required which leads to comparably short scanning and reconstruction times. The low spatial resolution also leads to less than 20 GB of volume data and hence a faster analysis in comparison to the highly resolved scans of porous ice.

For an adequate analysis of the porous ice cores, a high spatial resolution is essential. Scans of down to 12.5 μm voxel size can be realized within the full diameter of the drill cores. The inner 28 mm of the cores can even be scanned with a voxel size of lower than 4 μm. The high spatial resolution in combination with the huge sample size results in various challenges: 9 TB of projection images and 2 TB of

reconstructed volume data arise with each scan and have to be handled during reconstruction and analysis. Long scanning times of up to two weeks increase the requirements on the stability and robustness of the system. At the same time, instabilities in the X-ray intensity, the detector's electronics or the software must not be the cause for a complete failure of the measurement. Continuable measurement procedures that cope with slight irregularities during the measurement are indispensible. Highly magnified projections of huge objects result in images that don't contain any information about primary intensity $I_0$ which causes reconstruction artifacts. Last but not least, the X-ray detector has to be robust enough for a 24/7 radiation exposure. Since for the analysis of inclusions pure qualitative measurements quickly reach their limits, additional quantitative physical infor-mation such as density $\rho$ or atomic number $Z$ will be determined by dual-energy measurement procedures (2X-CT, [1]). Those quantitative data allow for the detailed analysis of sedimentations i.e. if dust or volcanic ashes are present. To obtain these data, two CT data sets are necessary: one acquired with an X-ray spec-trum of low energy and one acquired with an X-ray spectrum of high energy.

In order to cope with the aforementioned requirements, a choice has to be made of the X-ray components to be used. Stability of the radiation intensity is required over a long period of time and the detector has to be radiation hard due to extensive measurement times. The realization of a Helical CT requires not only a standard system of mechanical axes, in particular the rotational axis for the object, but also requires a translational vertical object axis in order to realize the vertical feed. The supplementary manipulative device increases the measurement's uncertainty. Hence, high precision of the manipulation system is required in all degrees of freedom. Spatial resolution with cone beam imaging is limited by the geometric blurring which on its part is influenced by the X-ray components. The magnification is defined by

$$M = \frac{\mathrm{FDD}}{\mathrm{FOD}} , \tag{1}$$

whereby FDD denotes the focus-detector-distance and FOD denotes the focus-object-distance. Hence, with an increasing FDD, also the geometric blurring increases as it is defined by

$$U_{\mathrm{g}} = (M - 1) * d_{\mathrm{F}} , \tag{2}$$

with focal spot size $d_{\mathrm{F}}$.

As illustrated in Fig. 1, in high-magnification applications a small size of the focal spot is essential for image quality. At the same time an increasing FDD not only increases the geometric blurring but also reduces the solid angle covered by the detector. Consequently, the FDD is preferably reduced as much as possible, in order to use a maximum part of the X-ray flux. Finally, it is evident that with a small FDD, the FOD also has to be as small as possible in turn, in order to realize highest magnifications. Shortest distances between the ice core and the X-ray tube therefore have to be realized and an X-ray detector with a small effective pixel pitch is needed.

The following components were chosen: a 225 kV microfocus X-ray tube with an exchangeable transmission and directional tube head. Focal spot sizes are specified

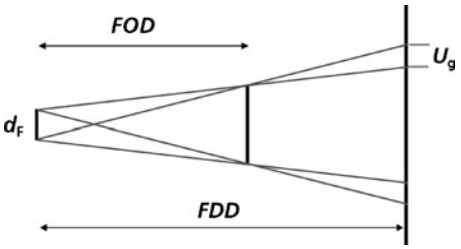



**Fig. 1** The geometric blurring increases with both, increasing focal spot size and increasing magnification. Hence, for highly magnifying applications, a small focal spot size is essential in order to minimize the geometric blurring

as 5 μm with reflection target and 1 μm with transmission target, respectively. The transmission target also allows for smaller FOD since the X-ray target is identical with the tube's exit window. On the other hand, a higher focusing of the electron beam allows for less power to be applied due to the generation of heat. In both cases pre-filters which are made of aluminum or copper are used where applicable.

The radiation imaging detector is a Fraunhofer IIS X-Eye 4020 with an active area of 40 cm by 20 cm and a pixel pitch of 50 μm. The resulting image size is 8000 by 4000 pixels. Due to the large active area, no horizontal displacement is necessary to cover the ice core in its full diameter. A small image lag below 0.1% and an external trigger mode enables fast scans.

Additionally an accurate manipulation system with a vertical wobble of 0.8 μm along 1 m traverse path allows for a high-precision alignment of the system.

## 3 Methods

### *3.1 Helical CT*

In order to cover the drill core in its full length, it is inspected by Helical CT [3]. In contrast to conventional CT the Helical CT uses a vertical feed while simultaneously rotating the object. Thereby, Feldkamp artifacts that are caused by transversal penetration of the object near the outer borders of the X-ray cone beam are avoided. The focal spot and the horizontal center row of the X-ray detector span the central plane of the CT system.

Only those parts of the object that are penetrated by the X-ray beam parallel to the central plane can be reconstructed without Feldkamp-type artifacts. Since standard 3-D-CT systems work with cone beam geometry, this part only consists of the intersection of the object with the central plane itself. A vertical feed of the object ensures artifact reduced information about the object in more than only one reconstructed slice. Limited by the mechanics only, objects of theoretically arbitrary

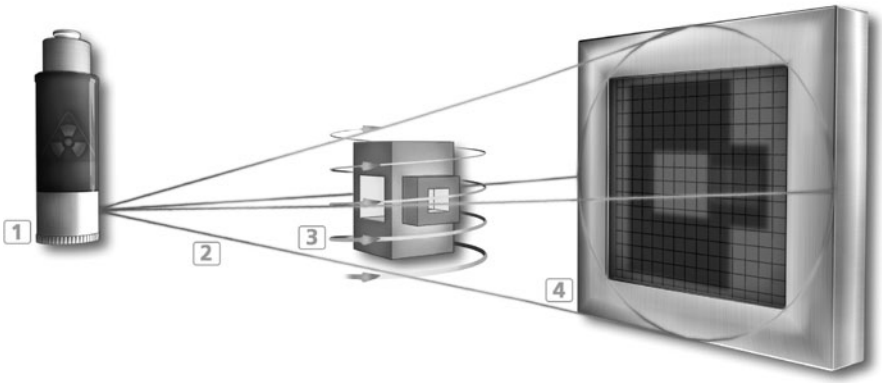



**Fig. 2** Principle of a Helical CT setup: The X-ray tube (1) emits a cone shaped X-ray beam (2) and the specimen (3) is projected onto the radiation imaging detector plane (4). Volumetric information is obtained by rotating the object. A vertical feed of the object reduces Feldkamp-type artifacts by displaying each part of the object in the central plane at least in one projection

length can be scanned with this method. In this case, the Helical CT is used to scan the entire core in one single scan.

For a precise determination of the pores' volume distribution in the case of porous ice, a high spatial resolution is indispensable. Reconstructing the core in slices of 8000 by 8000 voxels leads to a geometric resolution of 12.5 μm voxel size. This resolution is only sufficient for a reliable detection of pores with a diameter of approx. 25 μm, which is not satisfactory. The following methods are used to achieve higher spatial resolution.

## *3.2 Truncated Data and Multi-Scan Procedures*

The inner 28 mm of the ice core can be displayed on the detector in the maximum resolution of 3.5 μm voxel size with a magnification of factor 14. However, the unattenuated X-ray intensity is needed for reconstruction but not contained in the projection data since the object covers the entire detector plane. A multi-scan procedure records images of the object in various magnifications and reconstructs the inner part of the ice core with a high spatial resolution. Between every single measurement, magnification is reduced by a factor of 2, until the object can be displayed on the detector in its full width. The whole diameter of the object can only be displayed with a low spatial resolution and is used as additional information for the back projection algorithm to improve quality.

Two options for the high resolution reconstruction multiscan procedure or multi-resolution analysis region-of-interest (MRA-ROI) CT [2] can be chosen which mainly differ in the way of using the rougher scans as additional information. In comparison to a conventional CT scan, the amount of projection data is enhanced

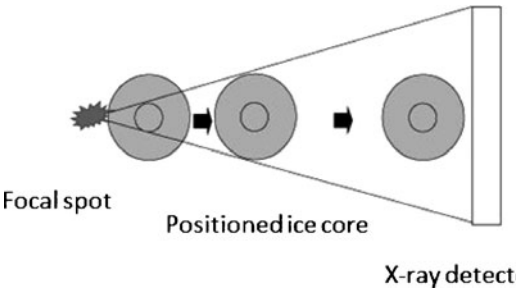



**Fig. 3** Schematic view from the top on the series of scans of the multiscan procedure or MRA ROI CT. Magnification is reduced by a factor of 2 until the object can be displayed in its full diameter plus some background intensity

by a factor $n$, where $n$ is the number of measurements used for this procedure. Figure 3 shows a schematic view of the measurement principle, where $n = 3$.

As an alternative to multi-scan methods, a single scan procedure can be used that is based on truncated data. A-priori assumptions about the geometry of the object help to reduce artifacts. The maximum penetration length is used as a-priori information. This method originally has been developed for the artifact reduced reconstruction of cylindrical objects (which ice cores are). Tests have shown that the method also works for objects with a different shape, e.g. cubic objects. In this case, the amount of measurement data is enhanced by a constant factor 3. Finally there is the option of a gradient reconstruction method (a so-called lambda reconstruction) which exclusively reconstructs the bounding surfaces between air and ice. For this procedure, the amount of measurement data does not increase in comparison to a conventional 3-D CT scan but it is very sensitive towards image noise.

### *3.3 $I_0$-Monitor*

In the case of multi-scan data acquisition as well as with truncated data reconstruction information about the primary intensity $I_0$ is not contained in all projections. This is usually coped by using a fixed value for $I_0$ during reconstruction which works well as long as the radiation intensity is stable. As soon as the dose changes during the measurement, ray sums and attenuation coefficients are set into a wrong relation and artifacts occur. Hence, an external single-pixel radiation sensor, the $I_0$-monitor, is used to measure the primary intensity. This sensor detects radiation and converts it into grey values via a calibration with respect to the behavior of the flat panel detector. Changes of the mean grey values in the detector image over the time can be traced back to their origin: either a change of penetration length has occurred or the X-ray dose itself has not been stable which is noticed by the $I_0$-monitor immediately.

Combining the methods described above allows for the inner 28 mm of the object to be scanned with a resulting voxel size of 3.5 µm.

### *3.4 Dual Energy CT*

In order to characterize inclusions within the ice, a dual energy method is applied. The objective is to obtain quantitative information about the materials present in the drill core. For this purpose the sample is scanned twice: once with a low energetic spectrum and once with a high energetic spectrum. The attenuation coefficients $\mu_1$ at the lower and $\mu_2$ at the higher energetic spectrum respectively, are known. By means of a base transform

$$\left(\frac{\mu_1(\rho, Z)}{\mu_2(\rho, Z)}\right) \rightarrow \left(\frac{\rho(\mu_1, \mu_2)}{Z(\mu_1, \mu_2)}\right) \tag{3}$$

the density $\rho$ of the materials and the kind of material itself, i.e. the effective atomic number $Z$ are obtained. This method is especially useful for materials that do not show a high contrast in radioscopy with respect to water.

## 4 Measurement Conditions

Since the ice cores are stored at a temperature of −15 °C, this is also supposed to be the ideal environmental temperature for the CT measurements. There are two possibilities to do so: both the ice cores are scanned within and together with a freezing apparatus or the entire CT system is installed in a climatic cooling chamber. The latter possibility is adequate for achieving precise results since apart from the ice no additional and possibly disturbing materials are situated within the X-ray beam that could cause scattering or beam hardening. On the other hand, cooling down the X-ray components leads to problems that are not common in conventional CT systems. Oil used in the system may change its viscosity at lower temperatures and cables may crack. To avoid unpredictable behaviour of the X-ray components at such low temperatures, the X-ray tube and the X-ray detector are situated within a climatic shielding. In contrast to a conventional CT system where X-ray tubes are often equipped with a cooling circuit, this shielding has an integrated heating that is supposed to keep a constant temperature higher than −15 °C. The shielding itself causes beam hardening to some extend that has to be corrected.

A further challenging task is the fact that the imaging geometry causes a vast amount of image data even in a binned detector mode. During a complete Helical CT scan several 10 000 projections are acquired which cause a disk space requirement of several TB. Additional TB of required disk space arises during reconstruction. The acquisition of such an amount of projection data needs some time, especially if a high image quality is desired. A robust and extremely stable measurement and

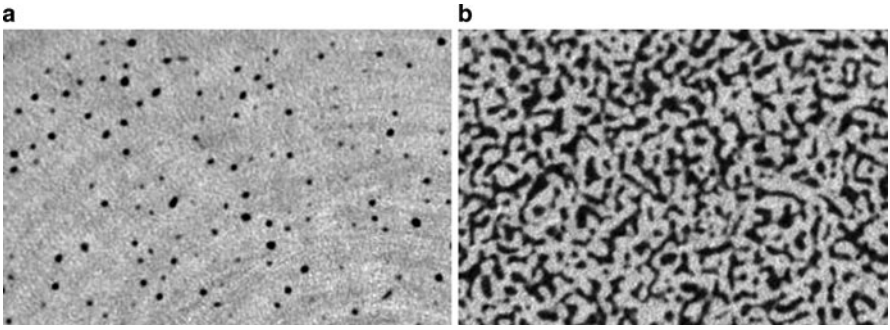



**Fig. 4** Reconstructed slices of porous ice (**a**) and firn (**b**). The single pores can be segmented while in the case of firn a mean porosity per slice is calculated

reconstruction system is hence needed. The X-ray tube must be able to produce a constant dose over a period of days, at the same time the imaging behaviour of the detector must not change. In the case of an unexpected failure of any kind it is not acceptable to repeat the entire measurement, but the projection data acquired so far has to be stored and used for a continuation of the measurement. This also applies to a complete system breakdown.

Finally a long measurement time per scan leads to a constant radiation exposure of the detector. Conventional industrial X-ray detectors are constructed for a usage of approximately 8 hours per day and would suffer from damages in their electronics soon under continuous operation. The X-Eye detector is constructed hard against radiation by an additional shielding that protects the electronics. Today it is still in use without any degradation of image quality after 1.5 years of frequent exposure.

## 5 Results

At first the $I_0$-monitor was calibrated with respect to the X-Eye detector's output signal and compared with the detector's mean grey value. The monitor's output was recorded with time. A comparison to the grey values measured with the X-Eye detector has shown a fair correlation of both datasets. Figure 5 shows the detector grey value and intensity measured with the $I_0$-monitor simultaneously. The X-ray current has been modified in regular intervals in order to generate varieties of resulting dose and grey values.

During the last 18 months the CT system was in constant use and a considerable number of measurements were made. Firn cores and cores of porous ice were scanned, reconstructed and analyzed. Some of the results can be seen in Fig. 6. The firn core was obtained from a depth of 35 m, the porous ice core is from a depth of 159 m. Structural differences between the firn and the snow are clearly visible.

A truncated data measurement of a porous ice core shows the reduction of the artifacts usually caused by ROI measurements. The data set used in Fig. 7 was ac-

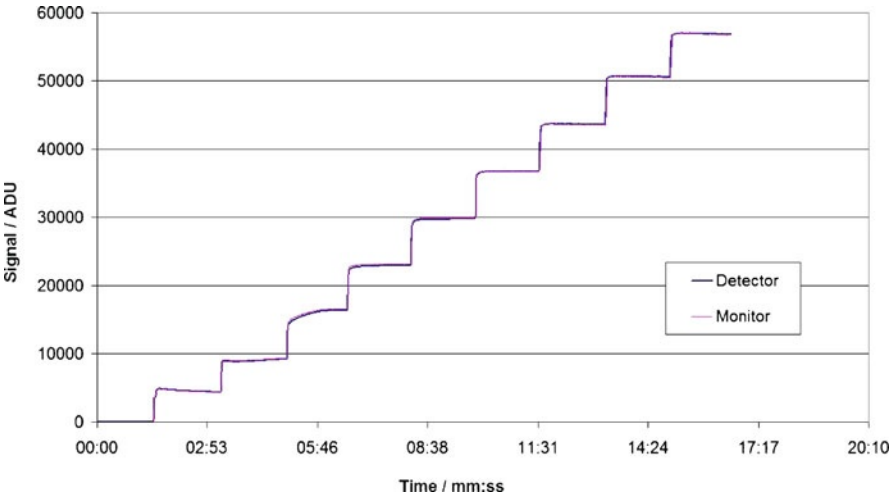



**Fig. 5** Comparison of grey- and intensity values. The tube current has been changed in regular intervals in order to change the dose

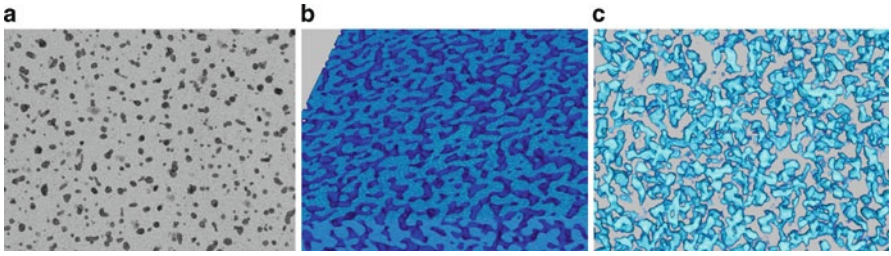



**Fig. 6** Reconstructed volumes of porous ice (**a**), firn (**b**) and superficial snow (**c**). Distinct pores can be seen clearly in the porous ice. Connected air and ice structures in the firn sample are the reason for the compact character of the firn in contrast to fluffy snow, which consists of separable snow flakes

quired without any information about the primary intensity in any of the projective images.

It was reconstructed twice: at first with a conventional ROI-CT procedure and afterwards with the artifact reducing truncated data procedure. As can be seen in the grey value profile, the irregular increase at the outer parts of the reconstruction region is completely eliminated. A homogeneous grey value level throughout the entire slice can be noticed.

One of the main applications of this CT system is the analysis of the porosity of firn cores. As an example a firn core with a diameter of 100 mm comparable to the sample of Fig. 6b was scanned over a length of 400 mm with a voxel size of 100 $\mu$m. The result of the analysis is shown in Fig. 8. As can be seen, the mean porosity per slice oscillates between 28% and 32% which is the natural fluctuation. Three abnormalities show a significantly lower porosity which can be explained by

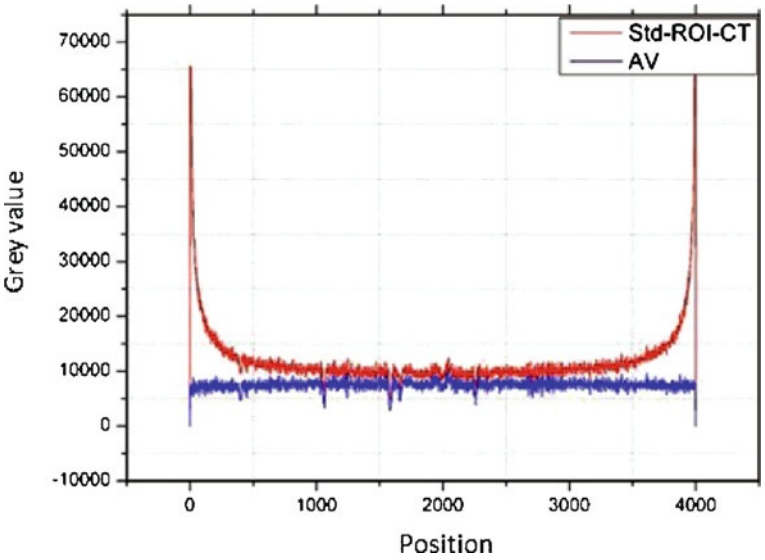



**Fig. 7** Grey value profiles of two reconstructed porous ice volumes. The *red line* shows a profile of a data set reconstructed by a conventional ROI-CT procedure, the *blue line* is a profile of the same data set reconstructed by the truncated data method. The well known artifacts especially at the outer part of the reconstruction region are considerably reduced

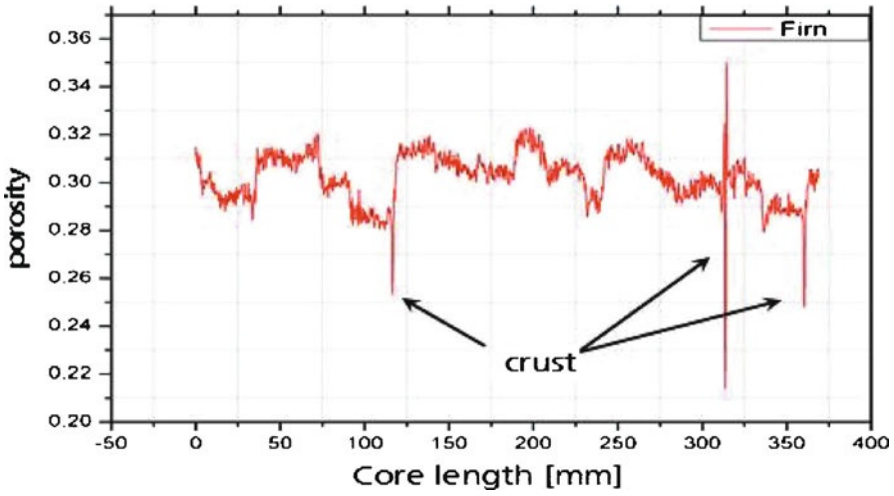



**Fig. 8** Mean porosity per slice of a firn core. Besides the slight overall variances of the porosity three abnormalities can be recognized that are caused by crusts. Those crusts can be caused by a melting and freezing process or by wind-blown dispersals

the existence of denser crusts in the core. Those crusts occur in irregular time lags and are either caused by a melting and freezing process or by storms. The effect occurs over a height of between 5 and 10 slices in the volume which corresponds to a thickness of the crust of about 0.5 mm to 1 mm.

## 6 Outlook

The reconstruction and analysis of porous ice and firn have been the main application of the presented CT system so far. Nevertheless further objectives have become a matter of research recently.

The formation of sea ice influences the gas exchange between sea water and atmosphere. It is a matter of debate how the degasification process of saline water and its contribution to the global budget of carbon dioxide in the atmosphere is controlled by sea ice. The freezing process of saline water differs from the pure consolidation of fresh water. Due to the high saline concentration, the congelation point of sea water falls below zero. The entire congelation process is accompanied by gas emissions from oversaturated water to the atmosphere which leads us to a multiphase system with three different phases. Figure 9 shows a CT slice of sea ice, where these three phases are clearly visible due two different grey values. In order to make conclusions about the congelation and degasification process of saline water, information about the temporal evolution of ice and brine structure during the freezing process is needed. This information must be brought into relation to the temperature over the height of the respective specimen. For the analysis of the mentioned dynamic processes it is planned to further the cooperation with the AWI to develop new measurement and reconstruction methods and algorithms. These developments will enable us to perform the evaluation of a wide range of dynamic processes not only ice but also i.e. crack propagation in tensile specimen.

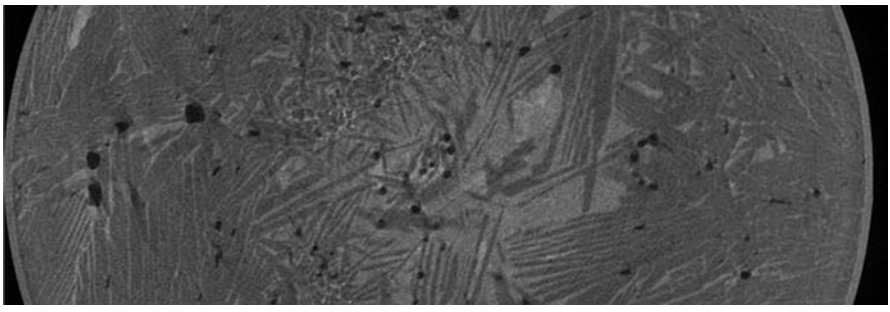


**Fig. 9** Slice image of sea ice during the formation process. *Dark regions* are caused by gas, the typical lamellar structures are caused by sea ice and the *brighter regions* show fluid brine

## References

1. Heismann BJ, Leppert J, KS (2003) Appl Phys 94:2073
2. Oeckl S (2006) Multiresolution 3D-Computerized Tomography and its Application to NDT. Proceedings of the 9th European Conference on Non-Destructive Testing (ECNDT)
3. Oeckl S (2009) Dimensionelles Messen mit Helix-Computertomographie. Fraunhofer Vision Technologietag

# Process Integrated Inspection of Motor Pistons Using Computerized Tomography

**Steven Oeckl, Roland Gruber, Werner Schön, Markus Eberhorn, Ingo Bauscher, Thomas Wenzel, and Randolf Hanke**

**Abstract** Pistons for combustion motors must meet strongest requirements in terms of material quality and dimensional accuracy to withstand the load inside the combustion chamber. The process integrated inspection of pistons using ultrasonic, eddy current and radiography is therefore state of the art. One drawback of all these methods is the lack of precise defect localisation which results in false rejects. Computerized Tomography (CT) provides a three dimensional characterization of defects and overcomes the above mentioned disadvantage but is established only for sample testing in laboratory. In this contribution we present the realisation of a CT system for process integrated casting inspection by means of combustion motor pistons. We cover the hardware setup as well as the software setup and focus on the reference data comparison which is the key method for process integrated defect detection.

## 1 Introduction

Pressure above 2000 bar and a temperature above 2700 degree Celsius are the loads a combustion motor piston has to withstand inside the combustion chamber of an engine. Therefore a combustion piston has to meet strongest requirements concerning material quality and dimensional accuracy. To ensure these requirements several non-destructive testing methods such as ultrasonic, eddy current and radiography are used for process integrated quality control [1]. Disadvantages of all these methods are the lack of precise defect localisation and the difficulties in dimensional feature extraction. Computerized Tomography (CT) provides a three dimensional description of an object and overcomes therefore the above mentioned drawbacks. But CT is an established non-destructive testing method only for sample testing in laboratory, not for process integrated quality control.

Steven Oeckl (✉)
Fraunhofer IIS/EZRT, Dr.-Mack-Str. 81, 90762 Fürth, Germany
steven.oeckl@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_26, © Springer 2011

In this contribution we present the main aspects of realising a process integrated CT system for casting inspection by means of combustion motor pistons. The realisation of the system was a joint project with the company MAHLE [2], one of the world's greatest manufacturer of combustion motor pistons. The pistons inspected in the project were rough-machined aluminium pistons for diesel motors having a maximum diameter of 10 centimetres. For stability reasons a diesel motor piston consists of an iron ring insert. A cooling gallery providing oil circulation inside the piston is used for heat dissipation. The aim of the project was the inspection of pistons concerning voids in the aluminium part and the measurement of the cooling gallery position within 30 seconds.

Following this introduction we provide a short description of the used hardware setup of the realised CT system in section two. The software setup of the CT system divided into reconstruction, cooling gallery analysis, and defect detection is content of the third section. Subsequently the results we achieved during the project and the performance of the system are presented in section four. We end the paper with a conclusion in section five.

## 2 Hardware Setup

In order to generate enough light in very short time for a sufficient signal to noise ratio in each measured projection we choose a high power tube which can be driven using 225 kV, 8 mA, and a focal spot size of 1 mm.

To speed up the data acquisition we perform a continuous rotation of the object axis instead of the classical stop and go measurement commonly used in industrial CT systems. The continuous rotation allows a fast scanning of the piston but yields motion artefact in the projection data if the exposure time is not short enough. To avoid these motion artefacts we use the Fraunhofer X-Eye detector due to its ability to deal with arbitrarily short exposure times.

**Fig. 1** Image of an aluminium piston with iron ring insert for diesel motors

**Fig. 2** Realised CT system for process integrated piston inspection within 30 seconds

Further reasons for using the Fraunhofer X-Eye detector are the advantageous image lag behaviour compared to other plat panel detectors and the resistance concerning the applied exposure rate. These aspects are of particular importance especially for a process integrated CT system where projections have to be measured continuously and permanently. More details on predestined features of the Fraunhofer X-Eye detector for process integrated CT can be found in [3].

## 3 Software Setup

Besides the hardware setup also the software of a process integrated CT system has to meet other requirements than a CT system in laboratory, especially the algorithms for automatic defect detection. We divide the software into three modules, i.e. reconstruction algorithm, cooling gallery analysis, and defect detection. The two first mentioned modules are covered only shortly and are referred to the corresponding literature, whereas the reference data comparison as the key component of defect detection is explained in detail.

### *3.1 Reconstruction*

Since we use a circular scanning geometry and rely on a fast implementation of the reconstruction algorithm we make us of the well known Feldkamp back projection algorithm [4]. The algorithm is processed on a distributed computing environment to

ensure a CT reconstruction within the required cycle time of 30 seconds. To reduce artefacts caused by the iron ring insert we apply the Iterative Artefact Reduction (IAR) as presented in [5].

### 3.2 Cooling Gallery Analysis

The minimum *radial* distance between cooling gallery and piston crown and the minimum *axial* distance between cooling gallery and piston skirt is defined as depicted in Fig. 3. Calculating these measures for every cross section of the piston that contains the central principal axis of inertia of the smallest cylinder that covers the piston yields a characterization of the cooling gallery position.

CT provides a complete three dimensional representation of a piston and therefore it is possible to calculate the above described position of the cooling gallery using image processing algorithms like registration, segmentation, and the classical Euclidean distance transform. A detailed discussion on these algorithms and how they can be used for a fast cooling gallery analysis can be found in [6].

### 3.3 Defect Detection

Software tools for offline data evaluation usually do not deal with reference information concerning the object under investigation. Rather some kind of background modelling is used to generate reference information about the object during evaluation [8]. This is a suitable approach for inspecting many different kinds of objects

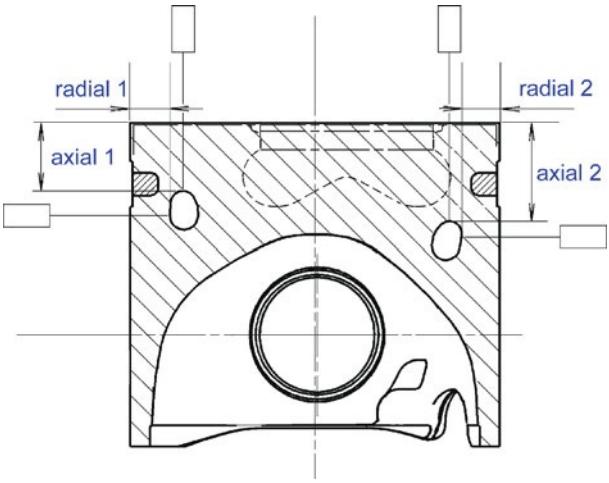



**Fig. 3** Schematic illustration of the measures that characterize the position of the cooling gallery inside the piston

because no uncomfortable teach-in process is necessary. However background modelling methods are not able to find big surface defects automatically. But this fact carries no weight, because offline tools are only used to support the staff of inspectors and the results will be usually reviewed.

In order to develop an inline evaluation method for detecting all kinds of defects, including surface defects, we have to make use of reference information in addition to background modelling. In spite of an initial registration step we need to characterize the misalignment of reference and object data caused by manufacturing tolerances to perform a suitable comparison.

CT parameters of an inline inspection system do not change during the evaluation of structurally identical objects. Therefore it is straightforward to use a CT reconstruction data set of an accurate object as reference information. In order to achieve a fast measurement of misalignments we have to pre-process the reference data set. There is no disadvantage in using time-consuming image processing at this stage, because these operations are pre-computed only once for each object type and therefore no resources are needed at inspection time.

For $r, s, t \in \mathbb{N} := \{1, 2, 3, \ldots\}$ let

$$V_{r,s,t} := \{(x, y, z) \in \mathbb{N}^3 : 1 \leq x \leq r, 1 \leq y \leq s, 1 \leq z \leq t\} \tag{1}$$

be the coordinate system of a 3-D volume data set. Since we always assume in the following that the volume data sets have same dimensions we drop the indices $r, s, t$ for convenience. Let $f_{\text{ref}}: V \to \mathbb{N}$ be the reference data set. Initially we perform an object labelling step, i.e. we determine all object representing voxels within the reference data set $f_{\text{ref}}$. Let therefore be $A \subseteq V$ the set of object indicating coordinates and $v = (x, y, z) \in V$. Then the binary label volume $l_{\text{A}}: V \to \mathbb{N}$ is defined as

$$l_{\text{A}}(v) = \begin{cases} 1, & \text{if } v \in A \\ 0, & \text{if } v \notin A \end{cases} . \tag{2}$$

Applying the Euclidean distance transform to $l_{\text{A}}$ concludes the reference data preparation step. Consider $v = (x, y, z) \in V$ and $S \subseteq V$. Then the minimal distance between $v$ and $S$ is defined as

$$\text{dist}(v, S) := \min_{s \in S} |v - s| . \tag{3}$$

Using this definition the distance transform [7] is given as

$$D(l_{\text{A}})(v) := \begin{cases} \text{dist}(v, A), & \text{if } v \notin A \\ 0, & \text{if } v \in A \end{cases} . \tag{4}$$

Summarizing the result of reference data preparation is a volume data set where all voxels which relate to object voxels in $f_{\text{ref}}$ are equal to zero, and all other voxels are coding the distance to the nearest object relating voxel.

First we define the measure for characterizing the misalignment between reference data and the current volume data set under investigation. Let $f_{\text{cur}}: V \to \mathbb{N}$ be

the current object data set and let $l_B: V \to \mathbb{N}$, $B \subseteq V$, be the corresponding binary object label volume. The misalignment measure is defined as

$$M(f_{\text{ref}}, f_{\text{cur}}) := \max\{D(l_A)(v) | l_B(v) = 1\} . \tag{5}$$

In other words, the misalignment measure calculates the maximum distance that occurs in the object area of $f_{\text{cur}}$ (see Fig. 4). This measure is suitable for inline inspection, because it can be calculated very quickly at inspection time using the prepared reference data.

The reference data comparison starts with the calculation of $M(f_{\text{ref}}, f_{\text{cur}})$. If $M(f_{\text{ref}}, f_{\text{cur}})$ is in the range of the manufacturing tolerances, no registration is necessary. If $M(f_{\text{ref}}, f_{\text{cur}})$ is bigger than the production tolerances, a straightforward registration step is performed where the inertial tensors of reference and current data are used to determine the parameters for the corresponding affine transformation. For more details see [7].

Since a simple subtraction of $f_{\text{ref}}$ and $f_{\text{cur}}$ yields many artefacts caused by manufacturing inaccuracies, we use a difference operation that takes into account the misalignment. This operation is defined as

$$S(f_{\text{ref}}, f_{\text{cur}})(v) := \min_{\substack{w \in V \\ |w - v| \leq M(f_{\text{ref}}, f_{\text{cur}})}} |f_{\text{ref}}(w) - f_{\text{cur}}(v)| . \tag{6}$$

Combining the proposed reference data comparison step and the background modelling as presented in [8] yields a defect detection method that is suitable for inline inspection.

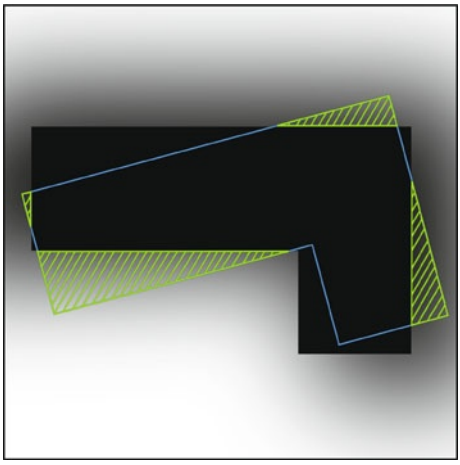


**Fig. 4** Illustration of misalignment measure calculation: The image shows the distance transform of a synthetic 2-D object; *bright pixels* indicate big distances, *dark pixels* indicate small distances, the object itself is *black*. The *hatched markings* show the misalignment to another object and indicate the distance values where the maximum defines the measure $M$

## 4 Results

Using the hardware setup as mentioned in section two a CT scan with 389 projections and an exposure time of 66 milliseconds per projection can be performed within 26 seconds. A pixel pitch of 400 μm yields a voxel edge length in the reconstructed volume data set of 330 μm. Since the object handling can be done within 4 seconds a cycle time of 30 seconds can be achieved. To get an impression of the data quality provided by the new CT system we show a comparison between a reconstructed slice of the process integrated system and a reconstructed slice of a standard system in laboratory in Fig. 5.

A typical result of the cooling gallery analysis applied to a piston where the cooling gallery has to be inclined with respect to the piston crown (see Fig. 3) can be seen in Fig. 6. The minimum distance and the maximum distance of the cooling gallery concerning the piston crown is taken from the sinus-type function that was fitted into the measured values. To validate the stability of the cooling gallery analysis we repeated the measurement of one piston 20 times and calculated some statistics of the minimum and maximum values. The standard deviation was 0.03 mm and lies in the range of the ultrasonic system which is the current method to analyse the cooling gallery. It takes 11 seconds using a state of the art dual core CPU to perform the cooling gallery analysis. A comparison between a CT measurement and the corresponding ultrasonic result is shown in Fig. 7.

A typical result of the automatic defect detection can be seen in Fig. 8. Internal defects, surface defects, and defects at the cooling gallery are automatically detected within a cycle time of 30 seconds using a distributed computing environment. The algorithm was validated using a sample of 1600 pistons with representative defects.

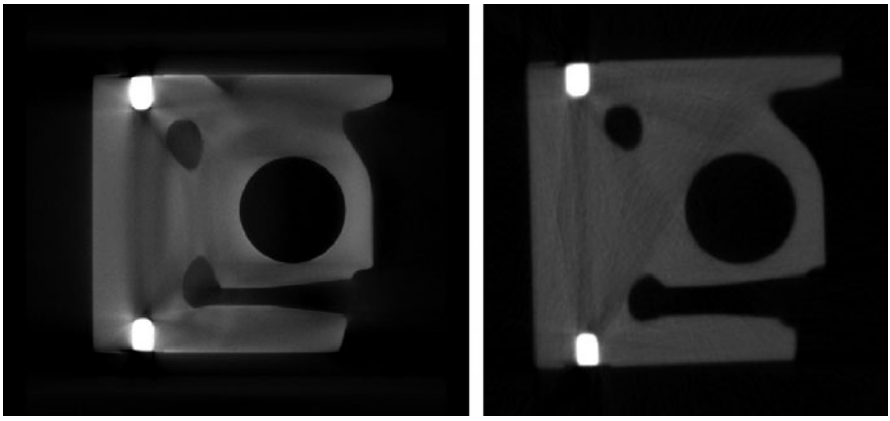


**Fig. 5** Reconstruction comparison between process integrated CT system and standard CT system for laboratory. *Left*: Result using a standard CT system with total measurement time of 20 minutes (225 kV, 570 μA, 500 ms exposure time, 800 projections, voxel edge length 110 μm). *Right*: Result using process integrated CT system with total measurement time of 0.5 minutes (225 kV, 8000 μA, 66 ms exposure time, 389 projections, voxel edge length 330 μm)

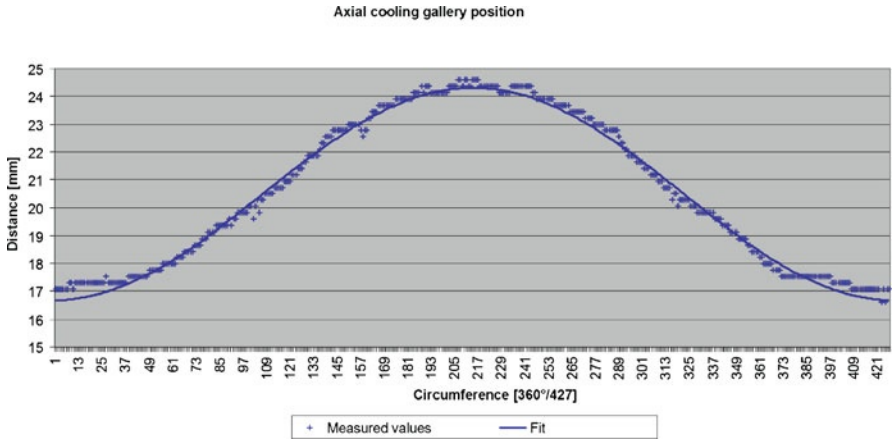



**Fig. 6** Result of the cooling gallery analysis: *Crosses* indicate measured distance values, the *line* indicates the result of the fit

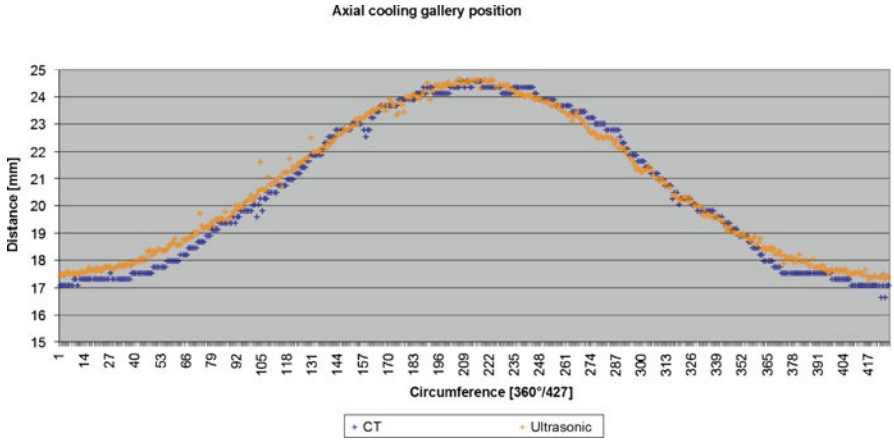



**Fig. 7** Comparison between CT and Ultrasonic methods for measuring the axial cooling gallery position: *Blue crosses* indicate the measured values using CT, *yellow crosses* indicate the measured values using Ultrasonic

A comparison between the automatic inspection results and the manual inspection of the CT data sets by an expert yields the performance of the system: Defects with an inscribed sphere of at least 1.4 mm can be detected process integrated in a stable way using the new CT system.

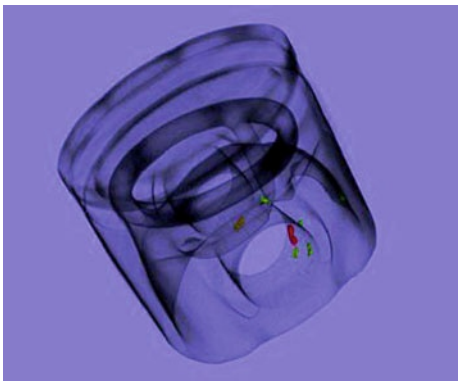


**Fig. 8** 3-D representation of an automatic defect detection result: *Coloured areas* indicate detected defects, different colours indicate different defect sizes

## 5 Conclusion

We have shown in this contribution the first process integrated CT system for casting inspection by means of combustion motor pistons. The CT scanning and reconstruction as well as the inspection concerning voids with a diameter of at least 1.4 mm and correct cooling gallery position are done process integrated within a cycle time of 30 seconds. It was shown that the stability of the cooling gallery analysis offers the possibility to replace two methods for quality control, i.e. radiography for defect detection and ultrasonic for cooling gallery analysis, by only one single process integrated CT system.

## References

1. "Mahle Kolbenkunde", www.mahle.com.
2. www.mahle.com
3. Behrendt R (2010) Röntgenkameras im industriellen Dauereinsatz, Fraunhofer Vision-Technologietag, Stuttgart
4. Feldkamp LA, Davis LC, Kress JW (1984) Practical cone-beam algorithm. J Opt Soc Am A 1:612–619
5. Kasperl S, Bauscher I, Hassler U, Markert H (2002) Reducing artefacts in industrial 3-D computed tomography (CT), presented at the 8th ECNDT, Spanish Society for Non Destructive Testing, Barcelona, Spain
6. Schön W (2008) Automatische Charakterisierung eines Zylinderkolben-Kühlkanals anhand von 3-D-Computertomographie-Daten, Studienarbeit, Friedrich-Alexander-Universität Erlangen-Nürnberg
7. Lohmann G (1998) Volumetric image analysis, NewYork, Brisbane, Toronto, Singapore, Stuttgart, Leipzig, Wiley-Teubner
8. Oeckl S, Hassler U, Wenzel T, Maisl M, Hanke R (2005) Advances in Automatic Evaluation of 3-D-CT Data, 3rd World Congress on Industrial Process Tomography, Banff, Canada

# Analysis of Processing Pipelines in Digital Raw Cameras

**Michael Schöberl, Joachim Keinert, André Kaup, and Siegfried Foessel**

**Abstract** Traditionally, image and video processing algorithms start from an RGB image. However, current image sensors deliver camera raw data that needs additional processing and interpolation for conversion into an RGB representation. While recent research delivers important improvements in image quality for processing and reconstruction of raw images, these algorithms come with a heavy computational complexity. Consequently, they are not suited for mobile solutions such as cameras for media production, mobile phones or surveillance. Offline processing on the other hand offers both higher computational power and better flexibility and is well suited for executing those algorithms. The workflow for utilizing this enhanced quality thus requires a shift from camera centric imaging to new off-camera processing strategies. This requires a novel infrastructure for transportation and interchange and enables the possibility for development of even more sophisticated algorithms for processing of camera raw data. This contribution highlights the challenges and possibilities arising from the above mentioned paradigm shift. We discuss algorithms that should stay within the camera and algorithms that benefit from offloading. Our research contributes to the increase in image quality of workflows for future video applications.

## 1 Introduction

As many media applications rely on color images of high quality, we focus on the media production market that includes high-end digital cinema applications as well as broadcast scenarios. The goal is the creation of highest image quality images together with a low camera complexity and finally low cost. In this contribution we analyze the formation of high quality images and the options for

Michael Schöberl (✉)
Chair of Multimedia Communications and Signal Processing, University of Erlangen-Nürnberg, Cauerstraße 7, 91058 Erlangen, Germany, schoeberl@lnt.de
michael.schoeberl@iis-extern.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_27, © Springer 2011

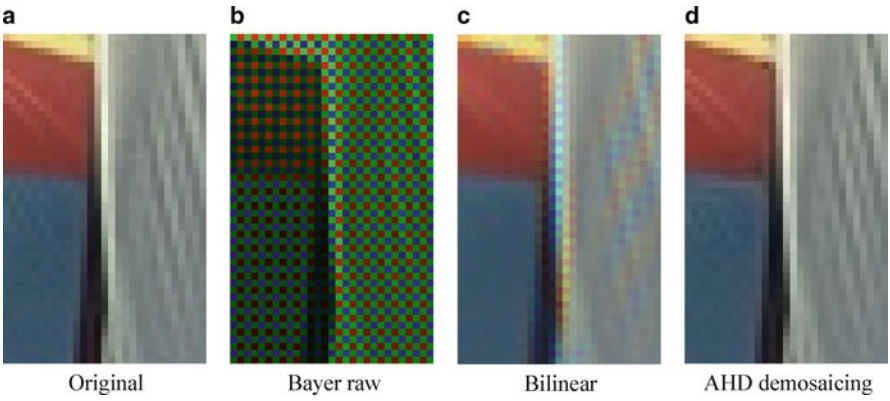



**Fig. 1** Example of simulated camera raw image: (**a**) original color image, (**b**) raw image in Bayer format, each pixel contains only one color, (**c**) after color reconstruction with simple bilinear demosaicing, and (**d**) AHD demosaicing [8]

shifting the algorithmic workload between the camera and off-camera processing stages.

Today, the majority of color camera systems utilizes a single image sensor. For the acquisition of color information a *Color Filter Array* (CFA) is placed on top of the pixels. Exemplary raw data image are shown in Fig. 1a depicts the input image of our simulation. After discarding some color information we obtain image Fig. 1b, which contains a single color value per pixel as in typical camera raw data for Bayer pattern [2] CFAs. Images with Bayer pattern sampling are not directly useful and need additional processing and interpolation of color values. For each single measured value (e.g. a pixel sensitive to green light) we need to calculate intensity values for the missing color components (red and blue). There is a wide selection of algorithms available. Figure 1c and d show the image after color reconstruction with a bilinear and the *Adaptive Homogeneity-Directed* (AHD) demosaicing algorithm [8]. Along the processing pipeline we further encounter algorithms for compensation of *Fixed Pattern Noise* (FPN) [18, 20] and interpolation of defect pixels [19]. Traditionally, these algorithms are executed inside the camera which finally outputs a fully reconstructed color image. In this contribution we look at the conditions and alternatives for shifting these algorithms from within the camera to an off-camera processing. We will show that apart from reducing the camera complexity we can also enhance the resulting image quality significantly.

## 2 Workflow Strategies

We consider two basic strategies as shown in Fig. 2: The camera centric approach performs all image processing inside the camera and delivers RGB images whereas the second workflow postpones most image processing and transports raw data.

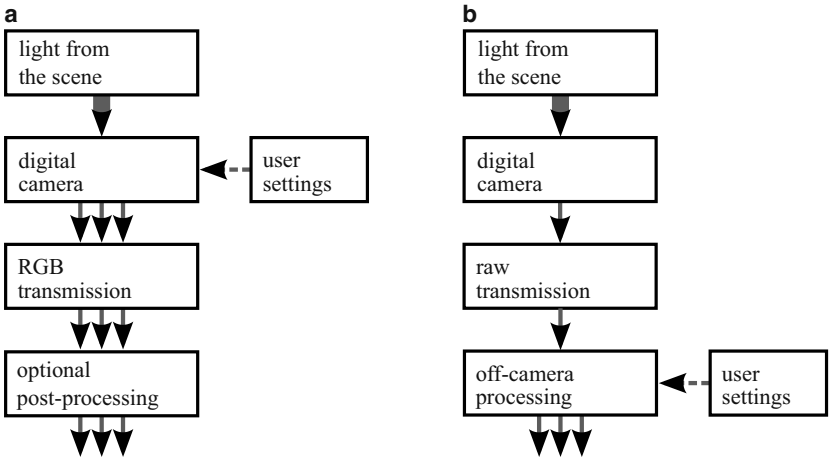



**Fig. 2** Comparison of workflow strategies: (**a**) traditional RGB workflow where all user decisions directly influence the algorithms within the camera, and (**b**) raw workflow where the camera delivers raw data and the user decisions influence the algorithms in offline processing only

## *2.1 Three Component RGB Workflow*

The traditional three component workflow in a digital camera is shown in Fig. 2a, consisting of three major steps, namely (i) image acquisition, (ii) data transport or storage, and (iii) optional post-processing. During acquisition, the camera directly creates a three component image. Typically, an RGB, XYZ or YCbCr color space is used for color representation. In the following, RGB will be used as a synonym for these color spaces, as the camera sensors typically acquire red, green, and blue pixels. In order to reduce the amount of data to transport and store, the image might be compressed. Since a large number of image processing steps, such as color reconstruction or defect pixel correction are already applied inside the camera, all decisions that influence the processing and reconstruction pipeline need to be known during recording, and are hard to change afterwards. Nevertheless, in media production it is common to apply additional off-camera processing, as e.g. color adjustments are carried out offline.

The RGB workflow depicted in Fig. 2a hints at three major drawbacks discussed below. These are rooted in the high computational effort of the algorithms necessary to achieve a high image quality: (i) the camera requires complex hardware, (ii) the flexibility in adjusting the acquisition parameters is low, and (iii) the achievable image and video quality is limited.

In more detail, algorithms such as the color reconstruction from CFA data or defect pixel correction can be computational demanding. Simpler algorithms, on the other hand, do not provide the same quality as shown below. Consequently, either image quality has to be sacrificed, or extremely high hardware effort has to be spent. While embedded systems, and thus also digital cameras, are becoming

more and more powerful, the required hardware remains expensive. With the rather low volume of high quality cameras the ASIC technology does not really solve the problem of building cheap, small and energy efficient cameras. In addition to these issues, also the flexibility is reduced, since a lot of data processing is required inside the camera. Consequently, many decisions influence the generated images which are hard to modify afterwards. For instance these decisions include the selection of the demosaicing algorithm [14]. Since both the color interpolation and the compression are not reversible, these operations cannot be undone. Consequently, optimum decisions during the image acquisition process are necessary in order to achieve best image quality. However, this is difficult to achieve because of time pressure and limited display quality on mobile devices.

## *2.2 Raw Data Workflow*

Most of the drawbacks identified in Sect. 2.1 are caused by the concentration of the video processing inside the camera. However, a full quality image is typically only required at some later stage in production. Hence, while taking the picture there is no need for having the full image quality right away. Consequently, in order to avoid named drawbacks the camera creates an almost unprocessed image when applying raw workflows as shown in Fig. 2b. Thus, decisions are not yet coupled to the data and processing and reconstruction may happen off-camera. In particular, the user is not limited to the processing chain implemented by the manufacturer, but can select an optimized variant delivering better results [12].

As a drawback, the exchange of data between different devices and institutions requires a standardized and open file format for video data. While multiple formats exist for traditional three component workflows, solutions for CFA video data are mostly proprietary. Fortunately, this starts to change with the development of the CinemaDNG format [1].

There are two major application scenarios that can profit from a consistent raw data workflow. The first one comprises video productions with high quality demands. The possibility to use the best available image processing algorithms available combined with an increased flexibility helps to increase the quality in the post-production. Typically, these applications will disclaim lossy compression and use uncompressed data for best possible quality.

The second scenario covers applications where low complexity is a major concern while still achieving good image quality. This might be because of cost, weight and battery life. Using a raw workflow provides the advantage of reduced data volume to transport. To further reduce the data volume lossy compression can be applied. This is particularly interesting since recent research demonstrated that compression before demosaicing can lead to better quality compared to traditional workflows [3, 11]. Furthermore, the computation power of general purpose computers used to process images is much less expensive then when integrating it into an embedded system such as a digital camera. Note, that live workflows also fit into the

second application scenario. The present broadcast van controlling the image acquisition can offer a much higher processing power compared to an embedded camera. Raw workflows are thus also beneficial for electronic field production and electronic news gathering type scenarios.

## 3 Analysis of Processing Pipeline Elements

In Fig. 3 the image processing pipelines for Fig. 3a an RGB workflow and Fig. 3b the proposed raw workflow are depicted. The building blocks are similar for both workflows. We chose this arrangement of the blocks based on the requirements in terms of complexity, calibration data and desired user interaction.

### *3.1 Fixed Pattern Noise Compensation*

The compensation of *Fixed Pattern Noise* (FPN) has a special role as it relies heavily on camera-specific calibration data. The compensation is therefore closely tied to

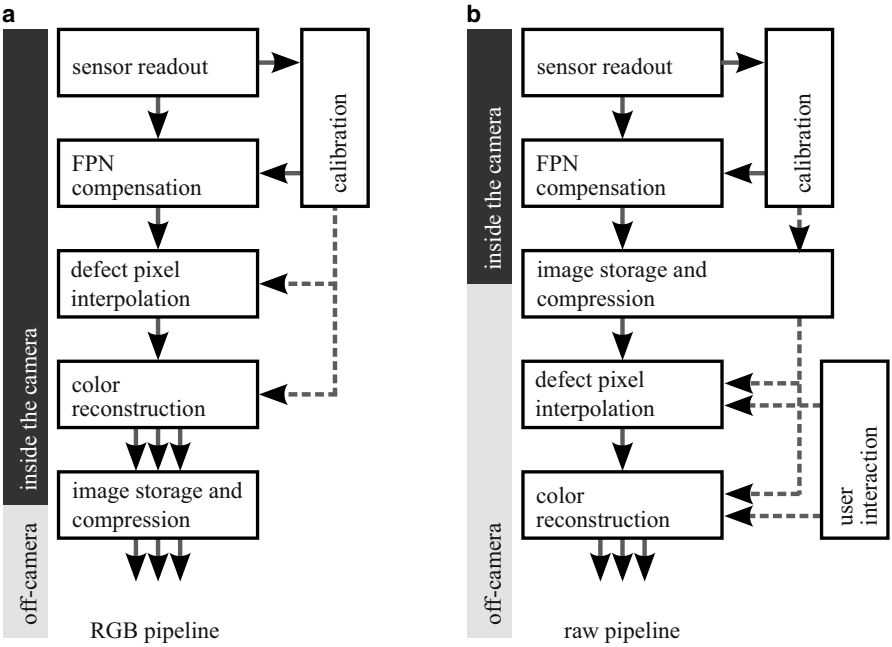



**Fig. 3** Image processing pipeline of (**a**) traditional RGB based camera, and (**b**) raw camera system. The building blocks are mostly the same but are arranged in a different order. This shifts complexity from the camera to an off-camera processing stage

the image sensor and current operating conditions of the camera. The origin of FPN can be explained as follows: Image sensors are made of many very tiny transistors and analogue circuits. Although manufacturing tolerances are tight and manufacturing is quite precise, each transistor will be slightly different from its neighbor. These tiny variations will cause the response from one pixel to the next to be quite different [6]. Fortunately, these non-uniformities are quite stable and can be calibrated and removed.

### 3.1.1 Noise Sources

We model the acquisition of the real scene information $I(x, y)$ for each pixel position $(x, y)$. The measured signal $\tilde{I}(x, y)$ is corrupted with noise $\tilde{I}(x, y) = I(x, y) + N_{\text{fixed}}(x, y) + N_{\text{dynamic}}(x, y)$. The fixed pattern noise $N_{\text{fixed}}$ can be separated into three major types of FPN distortions that are common in image sensors:

- *Dark Signal Non-Uniformity* or *Offset FPN* is one of the strongest noise sources and describes a variable black level of each pixel. Without light and exposure time $\tau_{\text{exp}} = 0$ there will still be some non-zero response. An example of typical offset noise is shown in Fig. 4a, where both a pixel-wise as well as a column wise pattern [6] can be seen.
- *Dark Current FPN* is the result of a leakage current within the pixel. Without any light a pixel will fill up with increasing exposure time $\tau_{\text{exp}} > 0$. An example of typical dark current FPN is shown in Fig. 4b. Most pixels have a uniform behavior, but few pixels have a high dark current and are highly visible as bright spots. These bright pixels are not defects and still deliver valuable image information.
- *Photo Response Non-Uniformity* describes a variable photo response per pixel. This can be caused by variable sizes, micro lens placement or even slightly different quantum efficiency for each pixel [4].

The time varying dynamic noise $N_{\text{dynamic}}$ [23] is usually much weaker for a good hardware design and is is not regarded in this contribution.

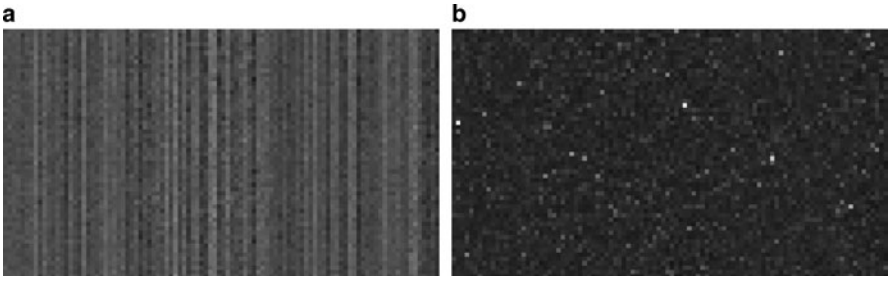



**Fig. 4** Examples of typical FPN from a CMOS camera for (**a**) offset FPN (gain ×8), and (**b**) dark current FPN (gain ×16)

### 3.1.2 Compensation Procedures

In contrast to dynamic noise, the FPN remains mostly constant from frame to frame. Unfortunately, it is not fixed either but can be assumed to change only little over time and operating conditions. This enables the estimation of the noise $\hat{N}_{\text{fixed}}$ and simplifies the actual compensation of FPN. We can compensate for FPN by subtracting an estimate $\hat{N}_{\text{fixed}}(x, y)$ for each pixel [13] $\hat{I}(x, y) = \tilde{I}(x, y) - \hat{N}_{\text{fixed}}(x, y)$.

We proposed the *Column Drift Compensation* (CDC) in [18] for estimating offset FPN $\hat{N}_{\text{offset}}$. One known problem ist that a small variation in camera temperature leads to additional noise patterns in the image. For cameras without temperature stabilization a fixed calibration frame is hardly usable. The proposed CDC uses a single fixed reference frame and adjusts for this type of temperature drift at runtime. The adjustment is performed based on drift estimates from reference rows on the sensor. As shown in [18] we achieve a high frame rate and a high image quality.

We analyzed the compensation of dark current FPN $\hat{N}_{\text{dark}}$ [20]: We found that dark current FPN is linear with exposure time $\tau_{\text{exp}}$ but can be slightly non-linear for some pixels. A calibration data set for linear compensation [17] is only suited for a limited range of exposure times. We have shown that the use of multiple sets of calibration data can extend the range of useable exposure times [20]. The resulting compensation is of the same low complexity as the linear compensation.

For some sensors a compensation method is implemented directly within the sensor. For example, the *Correlated Double Sampling* (CDS) operation mode [25] is widely used and renews the estimate $\hat{N}_{\text{offset}}$ for each frame. In many cases the camera designer has only limited access to the details of the sensor operation and hence only limited choice for the FPN compensation algorithms. We found that if the options are available, the image quality and the usable operation modes of a camera can be greatly enhanced by these algorithms. The proposed models are still of very low complexity and can be executed with just a few operations per pixel and are well suited for implementation into the camera.

### 3.1.3 Discussion

All of the FPN is sensor specific and changes with operating conditions (e.g. temperature or exposure time). This poses different challenges: The compensation requires multiple parameters per pixel. A full set of calibration parameters is often much larger than the actual image data. As can be seen from Fig. 4 the patterns are spatially random. It would be both hard and expensive to compress and transmit this calibration data along with the images. The calibration should therefore be applied on the raw camera images directly and within the camera. As we have noted above, the complexity of a FPN compensation is quite low, and even the sophisticated models require only few ADD operations per pixel. The effort for performing the compensation compared to handling all the data and performing the compensation offline is much lower. In consequence, the compensation of FPN should clearly be performed within the camera.

## *3.2 Defect Pixel Interpolation*

Another aspect of high-end cameras is the handling of defect pixels. Among the millions of pixels of a typical image sensor we can often find some that are not working properly. These pixels need to be declared as defect and need special treatment. The location of these pixels can be extracted from offset FPN [16], dark current FPN [9] and gain FPN [10] calibration data. As we have shown in [19] the compensation of defects needs to be performed before demosaicing as demosaicing would spread the error among a larger region. The interpolation of defect pixels is a difficult task. Due to the Bayer pattern [2] the direct neighbors have a different color and pixels of the same color are further away. The goal is to combine the color information of pixels of the same color. Further enhancement comes from the edge information of pixels of other colors. With multiple defects close to each other the situation becomes worse as an algorithm needs to adapt to ever changing conditions.

In [19] we have analyzed different existing algorithms. We put a special focus on analyzing the performance for situations with multiple defects close to each other. The most widely used algorithms include nearest neighbor [16], linear filtering [9] and median filtering [24]. The adaptive defect correction [22] estimates directions of edges and interpolates accordingly. We found that they all perform well in smooth image regions. Along edges they are far from perfect and the restored image often shows artifacts.

### 3.2.1 Sparsity-Based Defect Interpolation

We further proposed a new method for defect compensation in [19], based on the sparsity assumption: The transformation of an image to a transform domain can be represented with only few coefficients. These coefficients can also be found if only part of the signal is known. We found that a sparse representation can also be estimated for camera raw data and that a defect interpolation based on the *complex-valued Frequency Selective Extrapolation* (cFSE) [21] works very well.

For defect interpolation we use a small block of image data as depicted in Fig. 5a with known defect locations as shown in Fig. 5b. The cFSE algorithm iteratively updates the sparse model $g[m,n] = \sum_{k \in \mathfrak{K}} c_k \varphi_k[m,n]$ by selecting and updating a single coefficient $c_k$ in each iteration $\nu$. We use the basis functions $\varphi_k$ of the discrete Fourier transform as proposed in [21]. The set $\mathfrak{K}$ holds the indices of all basis functions used for model generation. The behavior along the iterations is visualized in Fig. 5b to k. We can see that with the adjustment of one frequency at a time, the whole region is modeled and reconstructed. The model is also valid for the missing pixels and after 100 iterations we can copy the missing pixels back to the image. In Fig. 5l we can see the original image data. The missing pixels are predicted accurately by our model.

This algorithm has the following advantages: The number and the location of additional nearby defects can directly be handled with this algorithm and no adjustment is necessary. The information from all pixels within the current block is taken

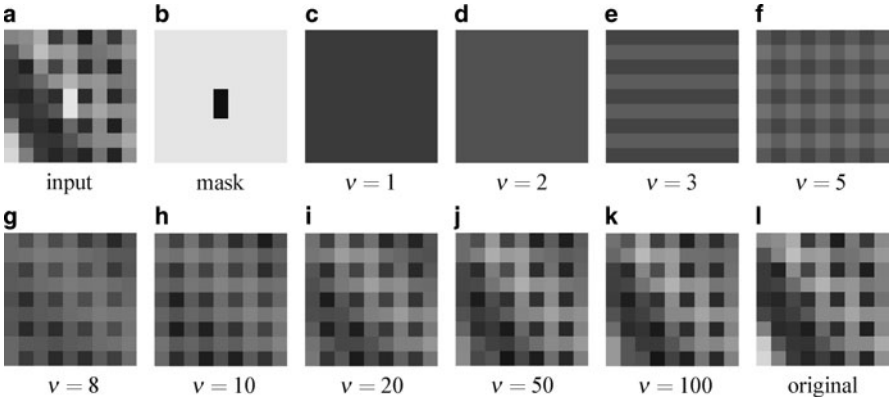



**Fig. 5** Example of defect interpolation with cFSE algorithm: (**a**) known image data with two defect pixels in the center (*white*), (**b**) defect locations, (**c**) to (**k**) model $g[m, n]$ after iteration $\nu$, and (**l**) original image block

into account and contributes to the reconstruction of the missing pixel. We do not include any assumption on the Bayer pattern. The algorithm is not tailored to the specific color of the pixel to be reconstructed and works just as well for any regular color filter array pattern.

### 3.2.2 Simulation and Results

We performed simulations for the 24 Kodak test images [5] and inserted a variable number of defects. Other than in [19] we used a defect cluster size of $3 \times 3$ pixels. The cFSE was executed with 100 iterations, $9 \times 9$ block size, decay factor $\hat{\rho} = 0.5$, orthogonality correction $\gamma = 0.5$, and DFT basis functions of size $12 \times 12$. For discussion of parameters and source code of cFSE see [21]. We measure success with the *Peak Signal to Noise Ratio* (PSNR) evaluated for the defect pixels only. The resulting Bayer-PSNR values are shown in Fig. 6. The horizontal axis shows the number of defects $K_{3\times3}$. We can see that our method outperforms the traditional methods by a large gain of 4.7 dB for isolated defects and 4.5 dB for 1–5 defects in a $3 \times 3$ region. As expected, with increasing number of defects all algorithms loose in quality. For visual results we further applied the AHD demosaicing algorithm [8] to the Bayer raw images. Figure 7 depicts some cutout of the resulting images. We can see that all traditional methods leave artifacts along edges. These distortions are highly visible after demosaicing. The proposed method does not show these artifacts. The visual inspection confirms the numeric gains.

The high gain in image quality of 4.5 dB comes at a cost. The proposed algorithm is much more complex and we can process some 5000 blocks per second in a single-threaded PC implementation [19]. This is fast enough for offline processing of images but the complexity is too large for realtime compensation of video in

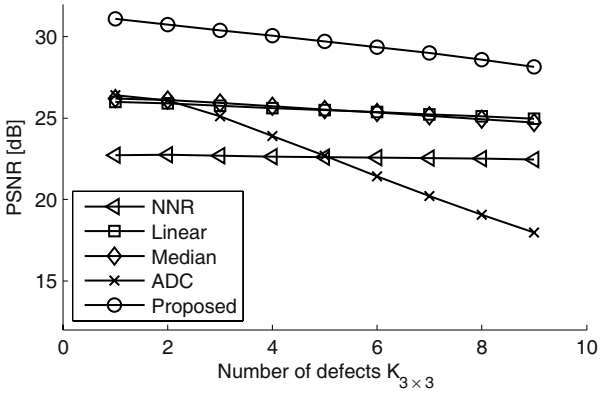



**Fig. 6** Performance of defect interpolation algorithms for a variable number of defects $K_{3\times3}$ in a region of $3 \times 3$ pixels, PSNR evaluated in Bayer image for defects only, average over all images

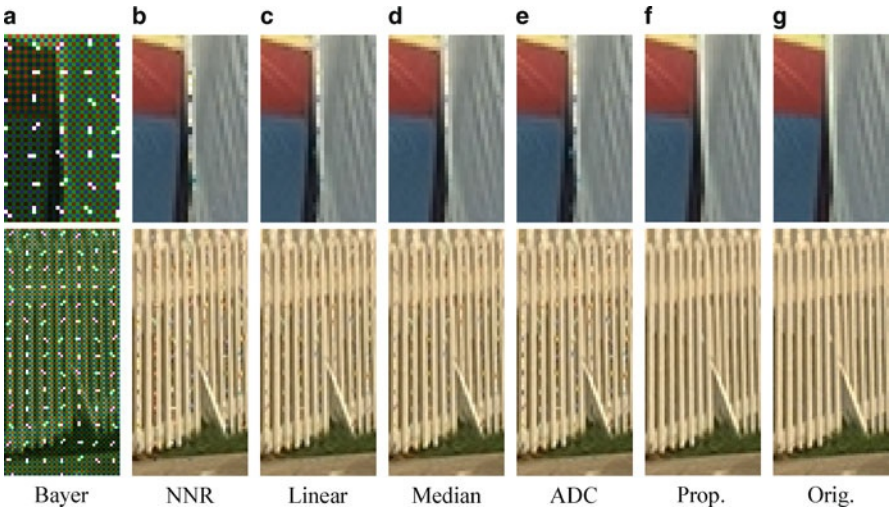



**Fig. 7** Examples from Kodak images sail (*top*) and lighthouse (*bottom*): (**a**) Bayer raw image with defect pixels (*white*), (**b**)–(**g**) defect processing and AHD demosaicing: (**b**) nearest neighbor, (**c**) linear, (**d**) median, (**e**) adaptive defect correction, (**f**) proposed cFSE method, and (**g**) original

mobile camera systems. The algorithm is therefore well suited for restoring the best possible quality in an off-camera processing step.

## *3.3 Color Reconstruction*

The interpolation of Bayer pattern raw images to a full RGB image is a widely researched topic. Applying just a linear interpolation gives a lot of unwanted dis-

tortions. In Fig. 1b we can see the result of a bi-linear interpolation. The image shows color and zipper artifacts along the edges. This is not the desired result from a high-end camera system. The interpolation with the *Adaptive Homogeneity-Directed* (AHD) demosaicing algorithm [8] does not show these artifacts as shown in Fig. 1c. Still, for some test cases even this result can be improved upon. Among other well-known algorithms such as alternating projections [7] there is still a large activity in the research community. E.g. a new approach based on compressive reconstruction has been proposed just recently [15]. This shows that there is still some improvement to be gained.

Compared to a linear interpolation these approaches are of much higher complexity and can not be used in realtime in a mobile camera system. However, in the proposed raw data workflow we are able to utilize these algorithms and spend more processing power outside the camera.

### *3.4 Calibration Data*

A camera system can have many different types of camera specific calibration data which needs to be available within the camera. The largest set is the calibration data for FPN compensation as discussed above. Multiple parameters for each pixel are common. As discussed, the FPN compensation is a data processing block that should stay within the camera for the raw data workflow.

Furthermore, we also have calibration data for defect pixels. The storage requirements are much lower as there are only few pixels within a sensor that are defect. This information can be transmitted as metadata along with the image. There is almost no change from frame to frame and calibration data could even be shared for multiple files in a sequence. Further calibration data could include color information and spectral response curves of the sensor. Again, these data sets are small compared to the image data and can easily be transmitted along with the images.

### *3.5 Image Storage and Compression*

Unlike the other elements, the storage and compression for RGB workflows differs from raw workflows. We need to distinguish two cases: First, there is the lossless or uncompressed operation. Manby high quality media production scenarios rely on best quality without any compression. This leads to fast SSD based recorders and the use of uncompressed video transmission lines based on HD-SDI. In these cases the raw workflow produces only $1/3$ of the amount of data as the demosaicing process interpolates two additional values for each measured pixel intensity.

Secondly, there are still situations where the amount of data exceeds the bandwidth of the transport channel. In these cases we need to use lossy compression. While this might be counterproductive for applications requiring utmost quality,

complexity driven scenarios can profit from such an approach as discussed in Sect. 2.2. In particular, existing literature [3, 11] reports quality gains compared to three component compression. This is an additional gain in image quality to the previously discussed raw workflows.

Since the camera outputs CFA data instead of RGB pixels when applying a raw data workflow, also the compression algorithm has to deal with this type of data. However, as adjacent pixels of CFA images belong to different colors, the traditional RGB compresion algorithms are not suitable and new algorithms need to be introduced: One possibility is to apply a special wavelet transform scheme within JPEG 2000, the so called SPACL decomposition. In essence, [26] demonstrated that an additional two dimensional wavelet transform applied to CFA data separates the high frequency pattern into different subbands. In consequence, those can be compressed similarly to ordinary images within the JPEG 2000 standard. While this approach does not lead to the best achievable compression ratios, it offers the big advantage of fully relying on a standardized and open file format, which promises quick applicability.

## 4 Conclusion

This contribution analyzed the image processing of high end camera systems and the question of handling the data. Currently, there are two major strategies, where one relys on in-camera processing in contrast to off-camera processing. The basic strategies of an in-camera workflow are compared with the off-camera raw data workflow. The raw workflow directly gives the advantages of lower in-camera complexity and better user interaction for the processing of images. We further analyzed the corresponding image processing pipelines more closely: Among the elements we found that fixed pattern noise (FPN) is one of the components that needs to be handled in-camera. Even with advanced FPN compensation algorithms the complexity is quite low. In the pipeline we also use algorithms for defect pixel interpolation. The proposed cFSE algorithm achieves superior quality but also comes at a high complexity. Along with demosaicing for color reconstruction, these algorithms should be executed off-camera.

## References

1. CinemaDNG Image Data Format Specification (2009)
2. Bayer B (1976) Color imaging array. US Patent 3,971,065
3. Doutre C, Nasiopoulos P, Plataniotis KN (2008) H.264-based compression of bayer pattern video sequences. IEEE Trans. on Circuits & Systems for Video Techn 18(6):725–734
4. Fowler B, Gamal AE, Yang D, Tian H (1998) A Method for Estimating Quantum Efficiency for CMOS Image Sensors. In SPIE Solid State Sensor Arrays: Development and Applications II, 3301:178–185

5. Franzen R (2011) Kodak lossless true color image suite. http://r0k.us/graphics/kodak/
6. Gamal AE, Fowler B, Min H, Liu X (1998) Modeling and estimation of FPN components in CMOS image sensors. In SPIE Solid State Sensor Arrays: Development and Applications II, 3301:168–177
7. Gunturk B, Altunbasak Y, Mersereau R (2002) Color plane interpolation using alternating projections. IEEE Trans. on Image Proc 11(9):997–1013
8. Hirakawa K, Parks T (2005) Adaptive homogeneity-directed demosaicing algorithm. IEEE Trans on Image Proc 14(3):360–369
9. Kovac M (1975) Removal of Dark Current Spikes from Image Sensor Output Signals. US Patent 3,904,818
10. Levine P (1986) Adaptive defect correction for solid-state imagers. US Patent 4,600,946
11. Lian NX, Chang L, Zagorodnov V, Tan YP (2006) Reversing demosaicking and compression in color filter array image processing: Performance analysis and modeling. IEEE Trans on Image Proc 15(11):3261–3278
12. Lukac R (2008) Single-Sensor Imaging: Methods and Applications for Digital Cameras. CRC Press, Inc, Boca Raton, FL, USA
13. Malueg R (1976) Detector Array Fixed-Pattern Noise Compensation. US Patent 3,949,162
14. Menon D, Calvagno G (2011) Color image demosaicking: An overview. Signal Processing: Image Communication 26(8–9):518–533
15. Moghadam A, Aghagolzadeh M, Kumar M, Radha H (2010) Compressive demosaicing. In IEEE Int Workshop on Multimedia Signal Proc, pp 105–110
16. Pape D, Reiss W (1991) Defect correction apparatus for solid state imaging devices including inoperative pixel detection. US Patent 5,047,863
17. Pillman B, Guidash R, Kelly S (2006) Fixed Pattern Noise Removal in CMOS Imagers Across Various Operational Conditions. US Patent 7,092,017
18. Schöberl M, Fößel S, Kaup A (2010) Fixed Pattern Noise Column Drift Compensation (CDC) for Digital Moving Picture Cameras. In IEEE Int Conf on Image Proc, pp 573–576
19. Schöberl M, Seiler J, Kasper B, Fößel S, Kaup A (2011) Sparsity-Based Defect Pixel Compensation for Arbitrary Camera Raw Images. In IEEE Int Conf on Acoustic, Speech, & Signal Proc, pp 1257–1260
20. Schöberl M, Senel C, Fößel S, Bloss H, Kaup A (2009) Non-linear Dark Current Fixed Pattern Noise Compensation for Variable Frame Rate Moving Picture Cameras. In Europ Signal Proc Conf, pp 268–272
21. Seiler J, Kaup A (2010) Complex-valued frequency selective extrapolation for fast image and video signal extrapolation. IEEE Signal Processing Letters 17(11):949–952
22. Tanbakuchi A, van der Sijde A, Dillen B, Theuwissen A, de Haan W (2003) Adaptive pixel defect correction. In Proc. SPIE Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications IV, 5017:360–370
23. Tian H, Fowler B, Gamal AE (1999) Analysis of temporal noise in CMOS APS. In SPIE Sensors, Cameras, & Systems for Scientific/Industrial Applications, 3649:177–185
24. Wang S, Yao S, Faurie O, Shi Z (2009) Adaptive defect correction and noise suppression module in the CIS image processing system. In Proc. SPIE Int. Symp. on Photoelectronic Detection & Imaging, 7384:73842V
25. White M, Lampe D, Blaha F, Mack I (1974) Characterization of surface channel CCD image arrays at low light levels. IEEE J of Solid-State Circuits 9(1):1–12
26. Zhang N, Wu X (2006) Lossless compression of color mosaic images. IEEE Trans on Image Proc 15(6):1379–1388

# Part IV
# Audio and Multimedia

# Audio and Multimedia

**Jürgen Herre**

## 1 Introduction

As an essential and renowned business field of Fraunhofer IIS, the audio and multimedia activities offer advanced technologies that enable a broad range of multimedia enabled applications. While this may happen in many different contexts, such as Internet transmission, consumer devices or mobile multimedia services of next-generation communication, it is always the quality experienced by the human listener/viewer that constitutes the essential criterion of merit which several generations of technology are striving to fulfill to an ever-increasing extent. Naturally, since the intended final receiver is the human auditory/visual system in this case, the quality experienced is subjective in nature and requires a profound understanding in psychometrics, psychoacoustics and psychovisual disciplines for its successful measurement and optimization.

Rooted in the successful development of perceptual audio coding schemes, such as the ISO/MPEG-1 Layer 3 specification (widely known as "mp3"), the audio and multimedia activities at Fraunhofer IIS have evolved over time to include the full range of relevant technology fields, starting from the efficient and intelligent acquisition of sound from sound fields to the efficient processing, coding, transmission and rendering of multimedia content. Creation of new technology at Fraunhofer IIS has been recently aided by the addition of the International Audio Laboratories Erlangen, that combines the strength of academic research with Fraunhofer's traditional focus on engineering and solutions to real-world problems.

This part of the book provides a set of six papers that combine science and engineering in audio signal processing and give an idea of the broad palette of topics and technologies that have been under investigation.

Jürgen Herre (✉)
International Audio Laboratories Erlangen, and Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
juergen.herre@audiolabs-erlangen.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_28, © Springer 2011

## 2 Headphone Reproduction

The first paper by Fleischmann et al. investigates a long-standing issue that is of fundamental relevance for reproducing audio over headphones. The importance of headphone playback has been increasing considerably over the recent decade in view of the vast numbers of portable audio players, PDAs, multimedia-enabled mobile phones and portable computers that are deployed today. Regardless of whether regular stereo recordings or material that has been recorded or processed for enhanced playback (e.g. binaural presentation) are presented, the ability of headphones to reproduce diffuse sound fields with the best possible accuracy has a significant impact on the perceived audio quality. The paper investigates measurements and design issues and provides results on listeners' appreciation of different approaches.

## 3 Microphone Array Processing

Located at the other end of the recording/transmission/reproduction chain, the intelligent and robust acquisition of spatial sound using microphone arrays has been a classic area of research and development for several decades. Beam-forming, signal enhancement and de-reverberation have filled many textbooks and are well-represented in academic literature. Despite the longstanding technical evolution in this field, the recent adoption of parametric concepts and time-frequency processing added a significant new viewpoint to the understanding of the field. Moreover, this new approach to spatial sound processing accounts for human auditory perception in the sense that it allows efficient analysis, representation, manipulation and re-synthesis of spatial sound using parameters that are closely related to key parameters relevant to human auditory perception, such as direction of arrival or diffuseness of a sound field. The paper by Thiergart et al. presents techniques that allow usage of parametric approaches also for linear microphone arrays instead of more traditional array geometries.

## 4 High Quality Telecommunication

Besides consumer entertainment, high-quality telecommunication may be the second most important field of application for modern audio-visual technology. After more than 100 years of being stuck with "plain old telephone service" (POTS) quality – meaning monophonic audio transmission with a frequency range limited between 300 Hz to 3.4 kHz – current state of the art technology is able to deliver a user experience that is very close to the perception of the communicating parties "being located in the same room". Current low-delay spatial sound capture, coding, IP-based transmission and spatial rendering techniques convey sound faithfully with its full audible frequency range and spatial characteristics, thus enabling "HiFi"

communication with a sound quality comparable to what has been offered to users for a long time by entertainment technology. The paper by Färber et al. provides an overview of how modern technology components can be integrated into an IP-based system to make the vision of being "together anytime and anywhere" become a reality.

## 5 Audio Coding for High Quality Speech Communication

One of the technical innovations that plays a crucial role behind such a vision of high-quality communication is the availability of a suitable source coding scheme that combines the virtues of high speech and audio quality with full bandwidth, spatial sound reproduction, high coding efficiency and low algorithmic delay, as it is a prerequisite to voice communication. The paper by Lutzky et al. introduces a novel state of the art perceptual coding scheme (AAC-ELD v2) that combines these virtues in an unprecedented way and was recently standardized within the ISO/MPEG working group. Comparative tests show that the new codec outperforms its competitors and predecessors in important respects.

## 6 Unified Speech and Audio Coding

The next paper covers one of the recent trends in source coding, namely the fact that perceptual coding of audio and coding of speech signals are no longer separate fields of technology or design philosophies for a low bitrate codec. While perceptual audio coding has an important focus on benefitting from the characteristics of the receiver (i.e. the human auditory system), speech coding has traditionally embraced the characteristics of the signal source (i.e. the human vocal tract). After many years of parallel existence, both technology worlds finally combined during a recent standardization process by the ISO/MPEG group, resulting in the Unified Speech and Audio Coding (USAC) scheme. The paper by Multrus et al. describes the new codec for the first time in its complete architecture and provides performance evaluations indicating that indeed the new scheme performs better than any of the two previous approaches. From the list of the authors it is visible that this coding scheme is the result of a major and continued effort by a sizeable group of contributors. In contrast to AAC-ELD v2, the USAC coder is not intended to be a communication codec, but serves for content distribution at very low bitrates.

## 7 Spatial Reproduction of Fine-Granular Audio Signals

Finally, the paper by Disch et al. elaborates on an innovation that has been utilized within the USAC codec for the very first time. From previous generations of high-quality spatial audio coding standards (such as ISO/MPEG HE-AAC v2, MPEG

Surround or Spatial Audio Object Coding) it has been observed as a longstanding problem that the parametric representation and re-synthesis of spatial audio (be it two channel stereo or multi-channel/surround) largely fails when trying to recreate scenes that consists of a high number of small discrete sound events, such as raindrops falling around the listener. This phenomenon is particularly disadvantageous for encoding of live recordings, since applause results from a multitude of spatially distributed clapping events by the audience. The approach introduced in the paper augments the traditional idea of decorrelation as a means of representing a wide sound image by a practicable alternative that is able to restore the necessary high temporal granularity of the signal.

## 8 Conclusions

As can be seen by the reader, the presented set of paper represents a wide area of scientific activity at Fraunhofer IIS and the International Audio Laboratories Erlangen and, in many ways, defines the forefront of technology. I sincerely hope that the reader will find them interesting and stimulating. Finally, I would like to express my particular gratitude to Robert Bleidt for his esteemed help as a native language reviewer. Other reviews and helpful comments for developing the material were provided by Bernd Edler, Emanuël Habets, Juha Vilkamo (and myself).

# Headphone Equalization – Measurement, Design and Psychoacoustic Evaluation

**Felix Fleischmann, Andreas Silzle, and Jan Plogsties**

**Abstract** Unlike loudspeakers, headphones are not designed to have a flat frequency response. Instead they should compensate for the spectral coloration caused by the acoustic transfer path when sound is travelling from a point in space to the ear. For correct headphone reproduction it is essential to control the sound pressure at the listener's ears. In the literature, there is no consensus about the optimal transfer function and equalization of headphones. In this work, several equalization strategies were tested on different commercially available headphones. Headphones were measured on an artificial head and equalization filters were designed in the frequency domain consisting of two parts: The first part of the filter is specific for each headphone and linearizes the frequency response of the headphone at the point of measurement, i.e. at the beginning of the blocked ear-channel. The second part of the filter is generic for all headphones and allows testing of different target responses on different headphones. Target responses from literature and the ISO 11904-1 diffuse-field equalization were tested. Further, target responses were designed by expert-listeners by directly comparing headphone and loudspeakers playback. A listening test showed a general preference for the equalization curves adjusted by experts and the diffuse-field equalization. The experiment also verified that the equalization process enhances the perceptual quality irrespective of the headphone model.

## 1 Introduction

Music and most of other sound material is produced and mixed for reproduction over loudspeakers. The sound engineer designs the material with certain timbral and spatial properties. Loudspeakers with a flat on-axis frequency response in a neutral sound studio are used. The listener is typically positioned at the *sweet* spot. There

Felix Fleischmann (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
felix.fleischmann@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_29, © Springer 2011

are standards for stereo and multi-channel (5.1) loudspeaker set-ups and reproduction rooms [4, 11, 12]. The listening scenario is not only defined by the frequency response of the loudspeaker, it includes also the acoustic transfer function from the loudspeaker to the listener's ears. The transfer path comprises the direct sound and multiple reflections at the boundaries of the room. Additionally, the body of the listener changes the acoustic path to the eardrum. The sound of one loudspeaker is perceived – with time and level differences – at both ears. The summation of the sound fields of two or more loudspeakers results in the intended perception and forms the auditory events between or behind the loudspeakers, called summing localization [3].

Reproducing a stereo signal over headphones, the acoustic situation and thus also the perception changes completely. There are no room reflections and there is no crosstalk between the channels. The left audio signal is fed directly to the left ear, and the right audio signal to the right ear. As a result of this unnatural situation, the whole auditory scene is perceived inside the head – between the two ears.

Spatial hearing relies on time-, level- and spectral cues of the left and right ear signals [2]. Binaural rendering makes use of head-related transfer functions (HRTFs) or binaural room transfer functions (BRTFs) to create virtual sound sources in the corresponding locations. Loudspeaker signals can be processed in a headphone compatible fashion by using such binaural processing. The auditory accuracy or plausibility of binaural reproduction depends on many factors, such as the selected HRTFs [19, 24, 25] or BRTFs [16] or head-tracking [15].

An important contribution to the accuracy of binaural signals is to properly control the frequency response of the ear signals. For that purpose the transfer function of the headphones can be measured at the same microphone positions as for the HRTF/BRTF measurements. In this case, the equalization method to control the sound pressure at the ear canal is well described and understood [17].

However, in most practical cases it is desirable that the rendering of the binaural signals can be done without prior knowledge of the specific headphones used, i.e. the binaural rendering needs to assume a specific frequency response of the headphone. The standard for diffuse-field equalization could serve as a convention, but only a small number of headphones show a transfer function that complies with the standard [18]. Therefore, an additional step is needed to equalize the headphone transfer function to follow a specific design goal.

In the following section headphone design goals and the target response are reviewed followed by an overview of the required measurement techniques. In Sects. 4 and 5 a novel equalization strategy and its perceptual evaluation are presented.

## 2 Headphone Design Goal

To adapt the loudspeaker signals for headphone reproduction in terms of timbre, a proper equalization should be used, see [18, 27, 29]. The free-field equalization will reproduce a flat frequency response, when a frontal sound wave passes

the microphone in an anechoic environment. In the case of a human head as receiver, the frequency response is not flat when the wave length is smaller than the size of the head, i.e. for frequencies above 2 kHz, [6] and depends on the position of the source relative to the head [1]. The diffuse-field equalization reproduces a flat frequency response in the diffuse sound field of a reverberation chamber when sound waves from all directions hit the transducer. In literature it is assumed that the diffuse-field equalization is the adequate interface for reproducing loudspeaker signals on headphones [27, 28]. However, in most loudspeaker listening situations the listener is positioned somewhere in-between a free-field and a complete diffuse-field. Thus, the timbre of the diffuse-field equalization does not exactly match the loudspeaker reproduction. Therefore, the headphone diffuse-field standard is very rarely applied and every headphone manufacturer uses its own equalization [18].

## *2.1 Diffuse-Field Equalization for Headphones*

In ITU-R BS.708 [10], instructions are given how to measure the diffuse-field transfer function of humans in a reverberant chamber. Møller [19] showed a different approach to measure the diffuse-field transfer function of humans by measuring HRTFs in an anechoic chamber. The diffuse-field transfer function is then gained by energy averaging of the single HRTF measurements. The results of this research were incorporated to ISO 11904-1 [5]. In this standard, a transfer function for the diffuse sound field is given. It shows a gain of approx. +10 dB between 4 kHz and 5 kHz.

Other studies use a perceptually motivated approach to identify the equalization curve. Experiments by Lorho [14] indicate that people tend to prefer a more subtle form of equalization – a gain of +3 dB at 3 kHz is rated best. These transfer functions are presented in Fig. 1.

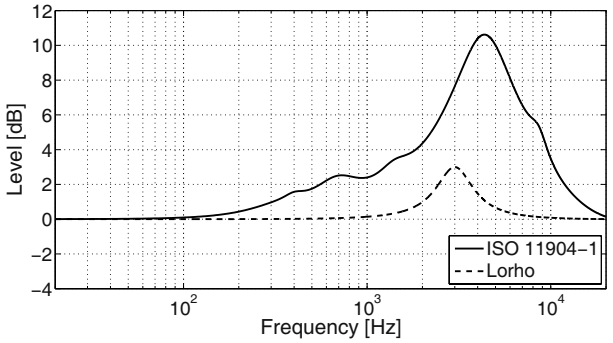



**Fig. 1** Diffuse-field equalization according ISO 11904-1 [5] and perceptual motivated equalization, Lorho [14]

In this investigation new equalization curves for headphones are found by directly comparing loudspeaker with headphone reproduction. Therefore, the measurement of the headphone transfer function and its inversion is necessary.

## 3 Measurement of Headphone Transfer Function

The basic characteristics of a headphone can be determined by its frequency response. It is important to choose a standardized and reproducible measurement point [13]. Nevertheless, headphone measurements on humans show high variances [18]. Therefore the measurement on a standardized artificial head is preferred. For checking the reproducibility, the measurement has to be repeated several times, where in between the headphone is removed and reattached again. All headphones in this investigation are measured on an artificial head [30] which is compliant to IEC 60959 [9].

In order to gain all relevant information about the headphone, the combination 'headphone – ear – microphone' is excited by a determined stimulus, in this case an exponential sine sweep. This excitation signal offers certain advantages: small crest-factor, short measuring time, robustness against surrounding noise, for details see [8, 20]. The transfer function can be derived by dividing the response spectrum by the stimulus spectrum.

## 4 Headphone Equalization

The new approach in this investigation divides the necessary headphone equalization into two parts. The first filter equalizes the specific headphone to a flat frequency response measured on the blocked ear canal of a dummy head. The second filter adds the target response, which is the focus of this study.

### *4.1 Inversion of the Headphone to Flat Frequency Response*

The measurement of headphone transfer functions (HpTFs or PTFs) has been described in Sect. 3. After removing outlier measurements, the absolute square of the transfer functions of the repeated measurements of left and right transducers are averaged. The resulting transfer function is then smoothed in the frequency domain (1/12 octave).

Because a headphone transducer is limited in linearly reproducing very low frequencies, the inversion should be done with care by defining some frequency limits. A realistic low frequency limit for most headphones is 50 Hz, e.g. [23]. Because the measurement is not accurate and reproducible above 12 kHz, the frequencies in the

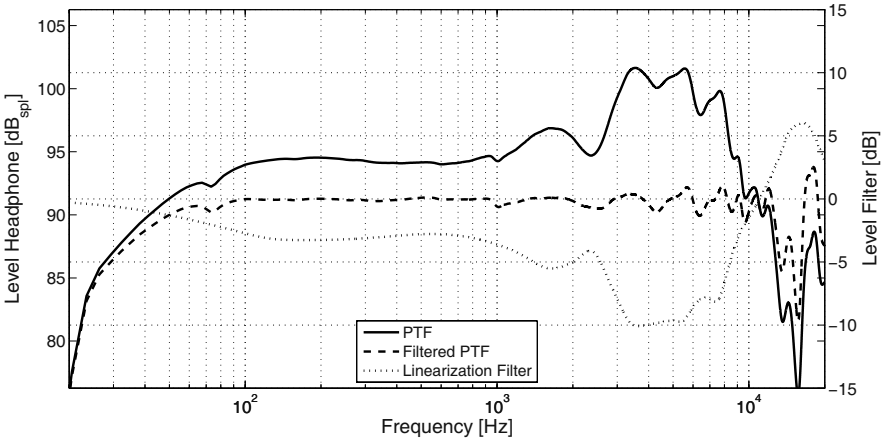



**Fig. 2** Frequency response of a headphone (PTF), inverse filter (15 SOS) and filtered response

region from 12–20 kHz are only roughly equalized by adjusting the mean energy, see Fig. 2. Then this transfer function of the linearization goal is inverted.

For the filter implementation FIR or IIR filters can be used [21]. IIR filters have clear advantages in terms of low-frequency performance and adjustability. They can be efficiently realized as stable second order sections (SOS).

To minimize the error between the linearization goal (flat frequency response) and the actual measured frequency response of the headphone, a parametric filter design method described in [22] is used. The filter was restricted to 15 SOS. The error values are calculated considering the double logarithmic behaviour of the ear in frequency and sound pressure. The result of the inversion is depicted in Fig. 2.

## *4.2 How to Find the Target Headphone Response*

To find the correct equalization for headphones which match the timbre of loudspeaker reproduction a new approach of direct comparison between loudspeaker and headphone reproduction is used. To keep the reproduction between these two cases as similar as possible, binaural room impulse responses (BRIR) of five loudspeakers are measured at the listening point. The reproduction room follows the

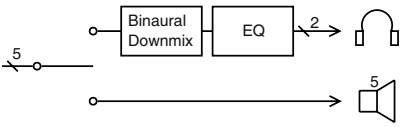



**Fig. 3** Principal signal flow for the expert-tuning

standard ITU-R BS.1116 [12] and is described in [26]. With the measured impulse responses a binaural downmix is formed by filtering with BRIRs and summation. This headphone signal is directly compared by three expert listeners to a 5.1 loudspeaker reproduction in the same room. They used parametric equalizers (EQ) to adjust the timbre of the headphone playback, see Fig. 3.

## *4.3 Equalization Overview*

Figure 4 gives a general overview of the different filters involved and lists the different approaches for headphone equalization. The first approach (Fig. 4(1)) shows the theoretical case with an amplifier and a perfectly linear headphone. In the second case (Fig. 4(2)) a normal headphone with a certain PTF is used with its inversion filter. This approach can only be used when the inversion of the headphone transfer function on the human head is known. The third equalization (Fig. 4(3)) is the normally used one, with a standard diffuse-field equalized signal (EQ BRIR) and the assumption that the used headphone is diffuse-field equalized as well. In the fourth approach (Fig. 4(4)) the two-step equalization is shown that is presented in this investigation. The headphone inversion filter from the dummy-head measurement and the newly found equalization (EQ HP) found by expert listeners.

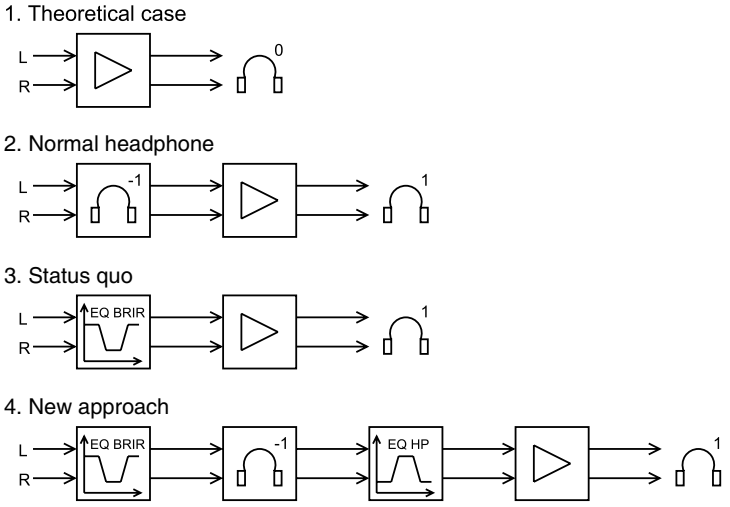



**Fig. 4** Schematic presentation of the different headphone equalizations

# 5 Perceptual Evaluation

## *5.1 Listening Test Design*

To evaluate the new equalization curves found by the experts and compare them to other possible equalization curves, a listening test was designed. For the listening test, four 5-channel tracks of different musical style used in EBU multi-channel tests [7] were used, see Table 1.

For every track, eight different equalization filters were investigated, as shown in Table 2. As the different equalization filters change the amplitude of the signal in different ways (compare *Linear* to e.g. *Low Anchor*), the overall amplitude of the signals has to be normalized for similar loudness. The listening test is conducted on three different headphones, see Table 3. For each headphone, every track was tested with each equalization filter. A listener was not allowed to listen to more than one headphone per day. $KH_{el.stat.}$ was tested by 16 subjects. Thereafter, the group was divided into 8 subjects for $KH_{dyn.A}$ and 8 subjects for $KH_{dyn.B}$. A MUSHRA [12] like test approach was used without open reference. The subjects were asked to judge overall quality especially focussing on timbre and localization on an absolute scale from 0 to 100.

**Table 1** Overview of the different listening test items

| Item | Artist/Composer | Style | Type | Duration [s] |
|---|---|---|---|---|
| *A* | Led Zeppelin | Rock | Rockband | 20 |
| *B* | Sedambonjou | Latin | Bigband | 16 |
| *C* | Tschaikowsky | Classic | Orchestra | 16 |
| *D* | Mozart | Radio Play | Male and female speaker | 13 |

**Table 2** Equalization curves used in the test

| Filter | Information |
|---|---|
| *Linear* | Linearization filter active, equalization filter not active |
| *Lorho* | +3 dB at 3 kHz (see Fig. 1) |
| *Expert 1-3* | Expert equalization filters from comparing to loudspeakers |
| *ISO 11904–1* | +10 dB at 4–5 kHz (see Fig. 1) |
| *Low anchor* | +20 dB at 4 kHz (purposely poor and sharp sounding item) |
| *Unfiltered* | Linearization filter not active, equalization filter not active |

**Table 3** Selection of headphones

| Abbreviation | Model | Manufacturer | Class | Nominal EQ |
|---|---|---|---|---|
| $KH_{el.stat.}$ | SR-404 | Stax | circumaural | diffuse-field |
| $KH_{dyn.A}$ | HD-600 | Sennheiser | circumaural | diffuse-field |
| $KH_{dyn.B}$ | PX-200 | Sennheiser | supraaural | n/a |

## 5.2 Results and Discussion

The listening test results with mean values and 95% confidence intervals are presented in Figs. 5, 6 and 7. The preference of different equalization curves is dependent on the sound items, but the general rank order is similar. The confidence intervals for the dynamic headphones are higher as for $KH_{el.stat.}$, because the number of listeners is smaller. The average values for $KH_{dyn.B}$ are lower than for the two other types. Interestingly, the unfiltered version gets a lower ranking than the low anchor in this case. This shows the poor frequency response of this low-cost headphone, which improves drastically by filtering. The *Linear*, *Lorho* and *Expert 1* equalization only show mid-range ratings. The unfiltered $KH_{dyn.A}$ is rated nearly as good as with the new equalization curves showing the good design of this headphone. The electrostatic headphone $KH_{el.stat.}$ can be approximately improved by 15 points on the scale by the proper equalization. Over all sound items the equalization curves from

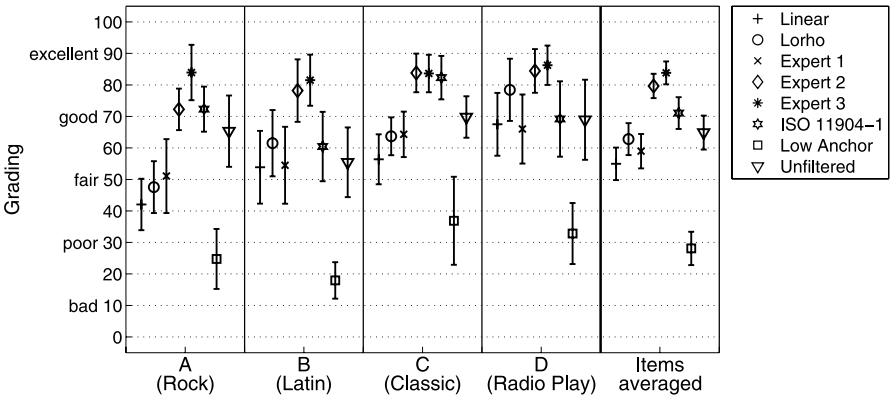



**Fig. 5** Listening test results for headphone $KH_{el.stat.}$, 16 listeners, 8 equalizations

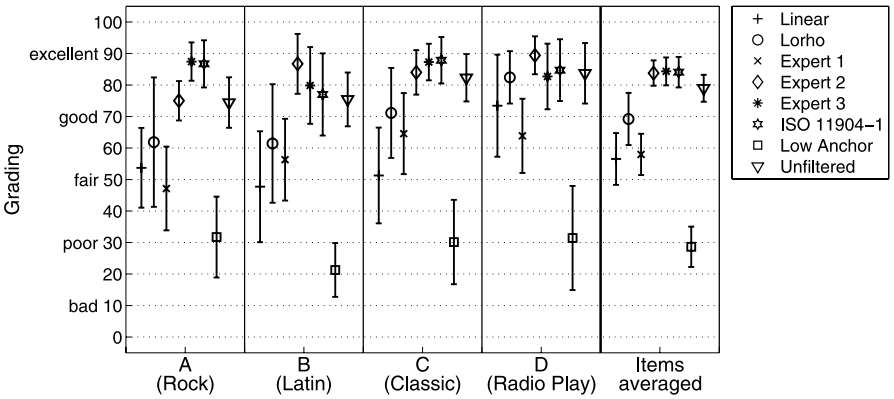



**Fig. 6** Listening test results for headphone $KH_{dyn.A}$, 7 listeners, 8 equalizations

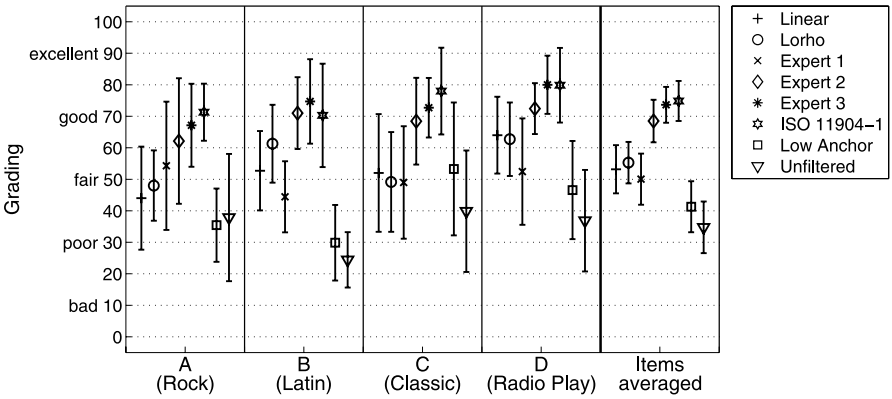



**Fig. 7** Listening test results for headphone $KH_{dyn.B}$, 7 listeners, 8 equalizations

expert 2 and 3 and the ISO curve get the best rating, independent of the headphone type.

## 6 Conclusions

This study focuses on the design of headphone equalization and its evaluation. The described measurement techniques result in reproducible headphone responses at the ear-canal entrance. It is shown that careful design of an equalization filter results in a linearized magnitude response. However, a linear response is only the first step for a good headphone reproduction, especially for binaural signals. As there is no consensus about the target response for headphones, different equalization filters were implemented and evaluated. The results show that the equalization provides a benefit in terms of perceptual quality for all tested headphones. Highest quality ratings are seen with equalization curves adjusted by expert listeners, while linear responses are not preferred. The preferred curves result from directly comparing loudspeaker reproduction with binaural headphone reproduction and adjusting the headphone equalization to the same perception of timbre.

The main contribution of this study is that the complete filter design can be seen as a two stage process: the linearization of the headphone response and then the implementation of the target response. In this way it is possible to improve the perceptual quality of headphone reproduction independent of the headphone model.

## References

1. Algazi VR, Duda RO, Satarzadeh P (2007) Physical and filter pinna models based on anthropometry. In Audio Engineering Society Convention 122, preprint #7098. Vienna Austria

2. Blauert J (1972) Die Schallausbreitung im äußeren Ohr und Konsequenzen für das räumliche Hören. In Audio Engineering Society Convention. Munich, Germany
3. Blauert J (1997) Spatial hearing, the psychophysics of human sound localization, 2 edn. MIT Press, Cambridge, Massachusetts
4. Borwick J (2001) Loudspeaker and headphone handbook, 3 edn. Focal Press, Oxford
5. DIN EN ISO 11904-1 (2003) Bestimmung der Schallimmission von ohrnahen Schallquellen Teil 1: Verfahren mit Mikrofonen in menschlichen Ohren (MIRE-Verfahren). Deutsches Institut für Normung e.V.; Beuth Verlag, Berlin
6. Duda R, Martens W (1998) Range dependence of the response of a spherical head model. J Audio Eng Soc 104(5):3048–3058
7. European Broadcasting Union Project Group (2007) Evaluations of multichannel audio codecs. Tech 3324
8. Farina A (2000) Simultaneous measurement of impulse response and distortion with a swept-sine technique. In Audio Engineering Society Convention 108, preprint #5099. Paris, France
9. IEC TR0 60959 (1990) Provisional head and torso simulator for acoustic measurements on air conduction hearing aids. Swedish Standards Institute
10. ITU (1990) Recommendation ITU-R BS.708; Determination of the electro-acoustial properties of studio monitor headphones. Intern. Telecom Union. Geneva, Switzerland
11. ITU (1992–1994) Recommendation ITU-R BS.775-1; Multichannel stereophonic sound system with and without accompanying picture. Intern. Telecom Union. Geneva, Switzerland
12. ITU (1994–1997) Recommendation ITU-R BS.1116-1; Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems. Intern. Telecom Union. Geneva, Switzerland
13. ITU (1996) ITU-T P.57; Series P: Telephone transmission quality; Objective measuring apparatus; Artificial ears. Intern. Telecom Union. Geneva, Switzerland
14. Lorho G (2009) Subjective evaluation of headphone target frequency responses. In Audio Engineering Society Convention 126, preprint #7770. Munich, Germany
15. Minnaar P, Olesen Søren K, Christensen F, Møller H (2001) The importance of head movements for binaural room synthesis. In Proceedings of the ICAD. Espoo
16. Minnaar P, Olesen Søren K, Christensen F, Møller H (2001) Localization with binaural recordings from artificial and human heads. J Audio Eng Soc 49(5):323–326
17. Møller H (1992) Fundamentals of binaural technology. Applied Acoustics 36(3–4):171–218
18. Møller H, Hammershøi D, Jensen CB, Sørensen MF (1995) Transfer characteristics of headphones measured on human ears. J Audio Eng Soc 43(4):203–217
19. Møller H, Sørensen MF, Hammershøi D, Jensen CB (1995) Head-related transfer functions of human subjects. J Audio Eng Soc 43(5):300–321
20. Müller S, Massarani P (2001) Transfer-function measurement with sweeps. J Audio Eng Soc 49(6):443–471
21. Oppenheim AV, Schafer RW (1975) Digital Signal Processing. Prentice Hall, London
22. Ramos G, López JJ (2006) Filter design method for loudspeaker equalization based on IIR parametric filters. J Audio Eng Soc 54(12):1162–1178
23. Schärer Z, Lindau A (2009) Evaluation of equalization methods for binaural signals. In Audio Engineering Society Convention 126, preprint #7721. Munich, Germany
24. Silzle A (2002) Selection and tuning of HRTFs. In Audio Engineering Society Convention 112, preprint #5595. Munich, Germany
25. Silzle A (2003) Quality of head-related transfer functions – some practical remarks. In First ISCA Tutorial & Research Workshop on Auditory Quality of Systems. Akademie Mont-Cenis, Germany
26. Silzle A, Geyersberger S, Brohasga G, Weninger D, Leistner M (2009) Vision and technique behind the new studios and listening rooms of the Fraunhofer IIS Audio Laboratory. In 126th AES Convention, preprint #7672. Munich, Germany
27. Spikofski G (1988) Das Diffusfeldsonden-Übertragungsmass eines Studiokopfhörers. In Rundfunktechnische Mitteilung, 32(3):101–115

28. Theile G (1986) On the standardization of the frequency response of high quality studio headphones. J Audio Eng Soc 34(12):956–969
29. Toole FE (1984) The acoustics and psychoacoustics of headphones. In Audio Engineering Society Conference: 2nd International Conference: The Art and Technology of Recording. Ottawa, Canada
30. Zollner M (1995) Der neue IEC-959 Kunstkopf. Tech. rep., CORTEX Instruments GmbH. http://homepages.fh-regensburg.de/~elektrogitarre/pdfs/neuerieckuko.pdf

# Parametric Spatial Sound Processing Using Linear Microphone Arrays

**Oliver Thiergart, Markus Kallinger, Giovanni Del Galdo, and Fabian Kuech**

**Abstract** The recording and reproduction of spatial audio gains increasing attention since multichannel sound systems have been established in modern audio applications. Directional Audio Coding (DirAC) represents an efficient description of spatial sound in terms of a few audio downmix signals and parametric side information in frequency sub-bands, namely the direction-of-arrival and diffuseness of the sound. Traditionally, the directional parameters are derived based on the active sound intensity vector, which can be efficiently estimated via two or three-dimensional microphone grids. However, due to form factor constraints, linear microphone arrays are often more suitable in practice and thus, alternative parameter estimation techniques are required. In this contribution, a framework for parametric spatial audio processing based on linear microphone arrays is presented allowing for an efficient recording, enhancement, and reproduction of spatial sound. Different application examples are discussed including parametric directional filtering and de-reverberation.

## 1 Introduction

Spatial audio processing is becoming more important as the variety of commercial applications for multichannel audio is constantly increasing. A common scenario is the use of home entertainment systems for listening to multi-channel music or watching movies featuring surround sound. Furthermore, there is growing interest in making spatial audio available for high-quality teleconferencing systems. Clearly, the latter application not only requires a realistic reproduction of the spatial audio, but also appropriate methods for its recording and audio enhancement.

An efficient approach to the recording and reproduction of spatial sound is the Directional Audio Coding (DirAC) technique [4]. DirAC uses a parametric rep-

Oliver Thiergart (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
oliver.thiergart@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_30, © Springer 2011

resentation of the sound field based on features which are relevant for the perception of spatial sound, namely the Direction-Of-Arrival (DOA) and the diffuseness of the sound field in frequency sub-bands. On the reproduction side, the signals of an arbitrary loudspeaker setup are determined as function of these parameters so that an accurate spatial rendering is achieved at a desired listening position. The DirAC parameterization assures that the interaural time differences, interaural level differences, and the interaural coherence are perceived correctly when the DOA of the recorded sound and the diffuseness is accurately reproduced. Since practical loudspeaker configurations in consumer applications are usually limited to two-dimensional (2-D) reproduction, we restrict ourselves to this scenario.

Originally, the author of [4] proposes to determine the DirAC parameters based on the active sound intensity vector which, in practice, can be determined using (comparatively cheap) microphone arrays [6]. Usually, cross-shaped planar microphone arrays are utilized for this task, allowing to carry out an accurate 2-D DirAC analysis. In many applications, however, it is not feasible to employ 2-D microphone arrays due to geometrical constraints, e.g., as in teleconferencing applications when the microphones have to be integrated into a flat TV screen.

To overcome this problem, an alternative method for estimating the DirAC parameters is presented employing a linear microphone array for the sound recording. More specifically, the DOA is determined based on the Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [5]. In contrast to the traditional ESPRIT, however, the proposed algorithm does not require computationally expensive eigenvalue decompositions allowing for a low-complexity integration into consumer products. Moreover, the mathematical framework enables an efficient estimation of the diffuseness parameter based on the Magnitude Squared Coherence (MSC) [1] between the array microphones. In the context of spatial audio teleconferencing, the estimated DirAC parameters are further exploited for enhancing the recorded speech. For this purpose, the concept of parametric directional filtering in order to attenuate sound originating from undesired directions is discussed. Moreover, it is shown that the parametric description of spatial sound additionally allows the recording or the reproducing system to efficiently reduce the diffuse sound energy caused by room reverberation on the recording side.

## 2 Sound Field Model and Parametric Framework

The parametric description of spatial sound in DirAC assumes that the total sound pressure $S(\omega, t, \vec{r})$ in an arbitrary point $\vec{r}$ at angular frequency $\omega$ and time instance $t$ results from a superposition of directional sound $S_{\text{dir}}(\omega, t, \vec{r})$ (e.g. a plane wave), arriving with angle $\varphi(\omega, t, \vec{r})$, and diffuse sound $S_{\text{diff}}(\omega, t, \vec{r})$, i.e.,

$$S(\omega, t, \vec{r}) = S_{\text{dir}}(\omega, t, \vec{r}) + S_{\text{diff}}(\omega, t, \vec{r}) \; . \tag{1}$$

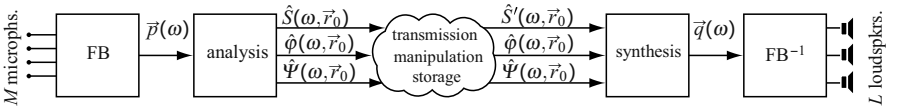



**Fig. 1** Parametric framework for the analysis, manipulation, and synthesis of spatial sound

The energy ratio between both sound components can be expressed by the Direct-to-Diffuse-Ratio (DDR) $\Gamma(\omega, t, \vec{r})$, i.e.,

$$\Gamma(\omega, t, \vec{r}) = \frac{|S_{\text{dir}}(\omega, t, \vec{r})|^2}{|S_{\text{diff}}(\omega, t, \vec{r})|^2} \,. \tag{2}$$

The spatial sound at the measurement location $\vec{r}_0$ is described by the sound pressure $S(\omega, t, \vec{r}_0)$, the DOA $\varphi(\omega, t, \vec{r}_0)$, and the so-called *diffuseness* $\Psi(\omega, t, \vec{r}_0) \in [0, 1]$ being 0 when only directional sound arrives, 1 when the sound is entirely diffuse, and 0.5 when both sound components possess equal energy. In fact, $\Psi(\omega, t, \vec{r}_0)$ represents an alternative expression to the DDR $\Gamma(\omega, t, \vec{r}_0)$, i.e.,

$$\Psi(\omega, t, \vec{r}_0) = \frac{1}{1 + \Gamma(\omega, t, \vec{r}_0)} \,. \tag{3}$$

Notice that all dependencies on time are omitted in the following for simplicity.

The parametric description of spatial sound by means of $S(\omega, \vec{r}_0)$, $\varphi(\omega, \vec{r}_0)$, and $\Psi(\omega, \vec{r}_0)$ is obtained and exploited as depicted in Fig. 1 showing a block diagram of the overall system. Input to the system is a microphone array of $M$ omnidirectional capsules where the array center is located in $\vec{r}_0$. The microphone signals are transformed into the time-frequency domain using a filterbank whose resolution is chosen such that the single wave model in (1) holds reasonably well. The transformed microphone signals, contained in vector $\vec{p}(\omega)$, are passed to the analysis block for deriving estimates $\hat{\varphi}(\omega, \vec{r}_0)$ and $\hat{\Psi}(\omega, \vec{r}_0)$ of the true DOA and diffuseness, respectively. Both parameters together with the sound pressure $\hat{S}(\omega, \vec{r}_0)$ in the array center form a *DirAC stream*, which can be stored, transmitted, or manipulated. On the reproduction side, the DirAC stream is passed to the synthesis stage generating a vector $\vec{q}(\omega)$ with $L$ loudspeaker signals, where $L$ can even be larger than $M$. The inverse filterbank finally yields the loudspeaker signals for sound reproduction.

## 3 Sound Analysis

The parametric representation of spatial sound introduced in the previous section is obtained in practice using a linear array of $M \geq 2$ omnidirectional microphones aligned with the horizontal plane. The sound pressure $\hat{S}(\omega, \vec{r}_0)$ in Fig. 1 is found from the microphone closest to the array center $\vec{r}_0$, or, alternatively, via linear combinations of the microphone signals $\vec{p}(\omega)$.

### 3.1 Direction-Of-Arrival (DOA) Estimation

The DOA $\hat{\varphi}(\omega, \vec{r}_0)$ of the directional sound $S_{\text{dir}}(\omega, \vec{r}_0)$ in (1) is determined similarly to [5] by employing a linear microphone array which can be separated into two identical but displaced subarrays with spacing $\Delta$ (cf. Fig. 2). The microphone signals $\vec{p}(\omega)$ are written as

$$\vec{p}(\omega) = \vec{s}(\omega) + \vec{n}(\omega) \ , \tag{4}$$

where $\vec{s}(\omega) = [S(\omega, \vec{r}_1), S(\omega, \vec{r}_2), \ldots, S(\omega, \vec{r}_M)]^{\mathrm{T}}$ is the total sound pressures defined in (1) at the microphone positions $\vec{r}_{1\ldots M}$ and $\vec{n}(\omega)$ contains independent zero-mean microphone noise with $\vec{n}(\omega) = [N_1(\omega), N_2(\omega), \ldots, N_M(\omega)]^{\mathrm{T}}$.

Without loss of generality, a Uniform Linear Array (ULA) with equal spacing $\Delta$ between $\vec{r}_{1\ldots M}$ is considered for the derivation. It is assumed that the directional sound $S_{\text{dir}}(\omega, \vec{r}_{1\ldots M})$ results from a single plane wave arriving from $\varphi(\omega)$ with unit magnitude. Moreover, no diffuse sound is present, i.e., $S_{\text{diff}}(\omega, \vec{r}) = 0$. Therefore, the sound field in $\vec{r}_{1\ldots M}$ can be expressed as

$$\vec{s}(\omega) = e^{j\phi_o}\, \vec{a}(\omega) \ , \tag{5}$$

where $\phi_0$ is the phase of the wave in $\vec{r}_1$, $\vec{a}(\omega) = [1, e^{j\mu}, e^{j2\mu}, \ldots, e^{j(M-1)\mu}]^{\mathrm{T}}$, and $\mu$ is the spatial frequency which can be written as

$$\mu = \frac{\omega}{c} \Delta \sin \varphi(\omega) \ , \tag{6}$$

where $c$ is the speed of sound. Equation (5) requires the narrowband assumption, for which a temporal delay equals a phase shift of the wave. When diffuse sound is present, i.e., when the model underlying to (5) is violated, the diffuseness $\hat{\Psi}(\omega, \vec{r}_0)$ in Fig. 1 increases (as discussed later) indicating the unreliable DOA estimation.

From the microphone signals $\vec{p}(\omega)$ in (4), the covariance matrix $\boldsymbol{R} \in \mathbb{C}^{M \times M}$ is computed which, using (4) and (5), can be expressed as

$$\boldsymbol{R} = \mathrm{E}\left\{\vec{p}(\omega)\vec{p}^{\mathrm{H}}(\omega)\right\} = \begin{bmatrix} 1+|N_1|^2 & e^{-j\mu} & \ldots & e^{-j(M-1)\mu} \\ e^{j\mu} & 1+|N_2|^2 & \ldots & e^{-j(M-2)\mu} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(M-1)\mu} & e^{j(M-2)\mu} & \ldots & 1+|N_M|^2 \end{bmatrix} = \begin{bmatrix} \vec{d}_1 \\ \boldsymbol{D} \\ \vec{d}_2 \end{bmatrix} \ , \tag{7}$$

where $\mathrm{E}\{\cdot\}$ denotes the expectation, $\vec{d}_1, \vec{d}_2 \in \mathbb{C}^{1\times M}$, and $\boldsymbol{D} \in \mathbb{C}^{(M-2)\times M}$. Assuming no microphone noise for the moment, i.e., $N_{1\ldots M}(\omega) = 0$, it follows from (7) that

$$\begin{bmatrix} \vec{d}_1 \\ \boldsymbol{D} \end{bmatrix} \Theta = \begin{bmatrix} \boldsymbol{D} \\ \vec{d}_2 \end{bmatrix} = \boldsymbol{A}\Theta = \boldsymbol{B} \ , \tag{8}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are the subarray matrices and $\Theta = e^{j\mu}$. Equation (8) provides a way for determining $\Theta$, for instance via the Least Squares (LS) method. Finally,

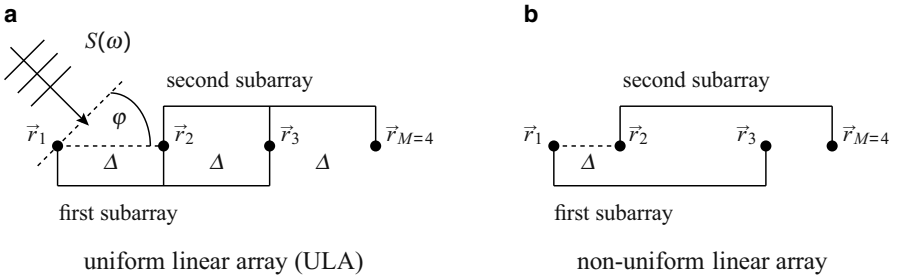



**Fig. 2** Exemplary microphone arrays used for the sound field analysis

using (6),

$$\hat{\varphi}(\omega) = \arcsin \frac{c \angle \Theta_{\mathrm{LS}}}{\omega \Delta} \,. \tag{9}$$

It can be shown that (9) is independent from the noise energies $|N_{1\ldots M}(\omega)|^2$. Therefore, the initial assumption $N_{1\ldots M}(\omega) = 0$ becomes superfluous. Moreover, when using no ULA, the subarray matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ have to be chosen accordingly to the two subarrays. For instance, for the right-hand array in Fig. 2, $\boldsymbol{A}$ contains the first and third row of $\boldsymbol{R}$, while $\boldsymbol{B}$ contains the second and fourth row. Clearly, the DOA estimates resulting from (9) are limited to the interval $-90° < \hat{\varphi}(\omega) < 90°$ meaning that front-back discrimination is not possible using the linear array.

## *3.2 Diffuseness Estimation*

The diffuseness estimator assumes for the diffuse sound $S_{\mathrm{diff}}(\omega, \vec{r})$ a spherically isotropic field. The diffuseness is estimated via the Magnitude Squared Coherence (MSC) $C_{1M}(\omega)$ between the two outer microphones located in $\vec{r}_1$ and $\vec{r}_M$, respectively. The MSC, defined e.g. in [1], can directly be computed from the covariance matrix $\boldsymbol{R}$ in (7), i.e.,

$$C_{1M}(\omega) \equiv \frac{\Phi_{1M}(\omega)\Phi_{M1}(\omega)}{\Phi_{11}(\omega)\Phi_{MM}(\omega)} = \frac{r_{1,M} \cdot r_{M,1}}{r_{1,1} \cdot r_{M,M}} \,, \tag{10}$$

where $\Phi_{1M}(\omega)$ and $\Phi_{M1}(\omega)$ are the cross Power Spectral Densities (PSDs) between the two microphone signals, $\Phi_{11}(\omega)$ and $\Phi_{MM}(\omega)$ are the corresponding auto PSDs, and $r_{i,k}$ is the $(i,k)$-th element of $\boldsymbol{R}$. The diffuseness $\Psi(k,n)$ is then estimated with

$$\hat{\Psi}(\omega) = 1 - \left[ \frac{C_{1M}(\omega) - C_{\mathrm{diff}}(\omega, d)}{1 - C_{\mathrm{diff}}(\omega, d)} \right], \tag{11}$$

where $d$ is the distance between $\vec{r}_1$ and $\vec{r}_M$ and $C_{\text{diff}}(\omega, d)$ is the MSC in spherical isotropic sound fields which, for omnidirectional microphones, is [1]

$$C_{\text{diff}}(\omega, d) = \frac{\sin^2(d\omega/c)}{(d\omega/c)^2} \,. \tag{12}$$

The term in brackets in (11) represents a linear mapping of $C_{1M}(\omega)$ to the interval $[0, 1]$. When the sound is purely diffuse, the MSC becomes $C_{1M}(\omega) = C_{\text{diff}}(\omega, d)$ resulting in $\hat{\Psi}(\omega) = 1$. For only directional sound, where $C_{1M}(\omega) = 1$, zero diffuseness is obtained as desired. In general, when the directional sound $S_{\text{dir}}(\omega, \vec{r})$ arrives from broadside direction at the linear array, i.e., $\varphi(\omega) \to 0$, the diffuseness estimator (11) represents a close approximation to the ideal diffuseness (3) [7].

## 4 Sound Manipulation and Synthesis

The parametric representation of spatial sound introduced in Sect. 2 and depicted in Fig. 1 allows for flexible processing approaches to manipulate and reproduce spatial sound. All the following approaches have in common that the directional sound $S_{\text{dir}}(\omega, \vec{r})$ and the diffuse sound $S_{\text{diff}}(\omega, \vec{r})$ in the signal model in (1) are treated individually. For this issue, it is assumed that both sound components are mutually uncorrelated so that

$$|S_{\text{dir}}(\omega, \vec{r})|^2 + |S_{\text{diff}}(\omega, \vec{r})|^2 = |S(\omega, \vec{r})|^2 \,. \tag{13}$$

The individual energies of the directional and diffuse sound can be determined from to overall sound field energy by inserting (2) and (3), i.e.,

$$|S_{\text{dir}}(\omega, \vec{r})|^2 = \big(1 - \Psi(\omega, \vec{r})\big)\, |S(\omega, \vec{r})|^2 \,, \tag{14}$$

$$|S_{\text{diff}}(\omega, \vec{r})|^2 = \Psi(\omega, \vec{r}) |S(\omega, \vec{r})|^2 \,. \tag{15}$$

Thus, the directional sound $S_{\text{dir}}(\omega, \vec{r}_0)$ and diffuse sound $S_{\text{diff}}(\omega, \vec{r}_0)$ in the measurement position $\vec{r}_0$ can be reconstructed from the measured total sound pressure $\hat{S}(\omega, \vec{r}_0)$ via

$$\hat{S}_{\text{dir}}(\omega, \vec{r}_0) = \sqrt{1 - \hat{\Psi}(\omega, \vec{r}_0)}\, \hat{S}(\omega, \vec{r}_0) \,, \tag{16}$$

$$\hat{S}_{\text{diff}}(\omega, \vec{r}_0) = \sqrt{\hat{\Psi}(\omega, \vec{r}_0)}\, \hat{S}(\omega, \vec{r}_0) \,. \tag{17}$$

In fact, the two equations provide estimates of the individual sound components possessing the correct magnitude but equal phases. The equal phases are only relevant for the spatial sound synthesis as discussed later.

### 4.1 Directional Filtering

As an example for spatial sound manipulation, directional filtering [3] implements the attenuation of interfering sound from undesired directions while the source signals from desired directions are preserved. Figure 3a depicts the general processing structure for the spatial sound manipulation. Directional filtering is carried out in the upper signal path. After estimating the directional sound $\hat{S}_{\mathrm{dir}}(\omega, \vec{r}_0)$ using (16), directional filtering is achieved by means of spectral attenuation, i.e., by multiplying $\hat{S}_{\mathrm{dir}}(\omega, \vec{r}_0)$ with a time and frequency-dependent gain function $f_{\mathrm{difi}}(\omega)$. The gain function $f_{\mathrm{difi}}(\omega)$ depends on the DOA $\hat{\varphi}(\omega, \vec{r}_0)$ of the directional sound, i.e., $f_{\mathrm{difi}}(\omega)$ is close to zero if $\hat{\varphi}(\omega, \vec{r}_0)$ corresponds to an undesired direction, and close to one otherwise. Notice that directional filtering modifies only the directional sound $\hat{S}_{\mathrm{dir}}(\omega, \vec{r}_0)$ since for diffuse sound, the estimated DOAs $\hat{\varphi}(\omega, \vec{r}_0)$ contain no useful information.

### 4.2 Diffuse Noise Suppression

The suppression of ambient noise, e.g., reverberation, is carried out in the lower signal path in Fig. 3a. Clearly, the estimated diffuse sound portion $\hat{S}_{\mathrm{diff}}(\omega, \vec{r}_0)$ contains the undesired ambient noise, which is attenuated by multiplying a scalar factor $0 \leq \beta < 1$. The reader is referred to [2] for further details on de-reverberation within the DirAC framework.

### 4.3 Spatial Sound Synthesis

One of the major advantages of the parametric representation of spatial sound introduced in Sect. 2 is the support of almost any loudspeaker setup for the sound reproduction. However, since the DOA $\hat{\varphi}(\omega, \vec{r}_0)$ in Fig. 1 can only be resolved within

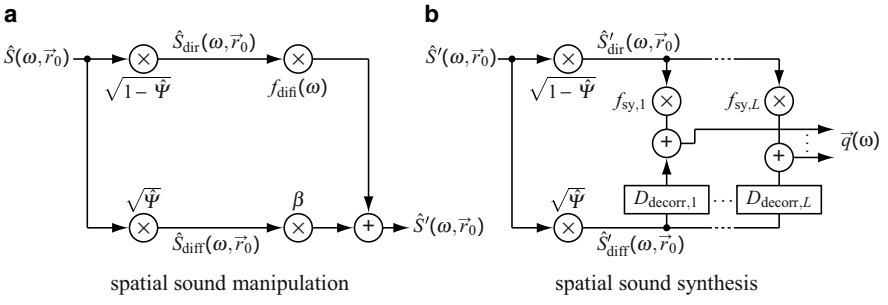



**Fig. 3a,b** General description of the manipulation and synthesis procedures

a range of 180° due to the linear microphone array processing, only sound reproduction setups with all $L$ loudspeakers within $-90° \ldots 90°$ are considered in the following.

The different steps of the spatial sound synthesis are depicted in Fig. 3b. The directional sound $\hat{S}'_{\text{dir}}(\omega, \vec{r}_0)$ and diffuse sound $\hat{S}'_{\text{diff}}(\omega, \vec{r}_0)$ are computed using (16) and (17), respectively. The directional portions of the loudspeaker signals are then obtained by multiplying $\hat{S}'_{\text{dir}}(\omega, \vec{r}_0)$ with a corresponding synthesis function $f_{\text{sy},i}(\omega)$, which is determined via Vector Base Amplitude Panning (VBAP) [4]. The functions $f_{\text{sy},i}(\omega)$ represent time and frequency-dependent gain functions depending on the time and frequency-dependent DOA $\hat{\varphi}(\omega, \vec{r}_0)$ and dedicated, frequency-independent panning law-functions. The panning laws are designed in a way that gains larger than zero are assigned to only the two loudspeakers adjacent to the current DOA $\hat{\varphi}(\omega, \vec{r}_0)$. The closer the DOA approaches a loudspeaker position, the larger is the gain factor assigned to the respective loudspeaker channel. A further prerequisite for VBAP is the preservation of energy, i.e., the sum of all squared panning law-functions equals one for any $\hat{\varphi}(\omega, \vec{r}_0)$.

The diffuse sound $\hat{S}'_{\text{diff}}(\omega, \vec{r}_0)$ is fed to all loudspeakers. Thus, the diffuse sound portions in the loudspeaker channels are completely coherent which may lead to inaccurate subjective spatial cues, especially if loudspeakers are only placed within $-90° \ldots 90°$. In the latter case, the diffuse sound might be perceived as coming from the center of the loudspeaker arrangement. Consequently, the different copies of $\hat{S}'_{\text{diff}}(\omega, \vec{r}_0)$ are mutually decorrelated in the processing stages denoted as $D_{\text{decorr},i}$. This ensures that the spatial cue of *interaural coherence* matches the perception of a diffuse sound field.

Finally, the loudspeaker signals $\vec{q}(\omega) = [Q_1(\omega), Q_2(\omega), \ldots, Q_L(\omega)]^{\text{T}}$ depicted in Fig. 1 are computed by summing the directional and diffuse portions for each loudspeaker channel, i.e.,

$$Q_i(\omega) = f_{\text{sy},i}(\omega)\hat{S}'_{\text{dir}}(\omega, \vec{r}_0) + \frac{1}{\sqrt{L}} D_{\text{decorr},i} \left\{ \hat{S}'_{\text{diff}}(\omega, \vec{r}_0) \right\}, \tag{18}$$

where $1 \leq i \leq L$ is the loudspeaker index and the factor $1/\sqrt{L}$ ensures that the diffuse sound is perceived with correct strength.

## 5 Summary

Directional Audio Coding (DirAC) is an efficient technique for the recording, manipulation, and reproduction of spatial audio. Its main advantage lies in the description of the sound field by means of two parameters, namely Direction-Of-Arrival (DOA) and diffuseness of the sound. This contribution proposes two methods for the estimation of these parameters when using a linear array of omnidirectional microphones. For the DOA, an adaptation of the ESPRIT algorithm is introduced, whereas for the diffuseness, an estimator based on the magnitude squared coherence

is derived. The possibility to have estimators for linear arrays characterized by a low computational complexity allows to integrate DirAC in real-time consumer products in which a traditional 2-D microphone array cannot be employed.

## References

1. Brandstein M, Ward D (eds) (2001) Microphone Arrays: Signal Processing Techniques and Applications. Springer
2. Kallinger M, Del Galdo G, Kuech F, Thiergart O (2011) Dereverberation in the spatial audio coding domain. In Audio Engineering Society Convention 130, London UK
3. Kallinger M, Ochsenfeld H, Del Galdo G, Kuech F, Mahne D, Schultz-Amling R, Thiergart O (2009) A spatial filtering approach for directional audio coding. In Audio Engineering Society Convention 126. Munich, Germany
4. Pulkki V (2007) Spatial sound reproduction with directional audio coding. J Audio Eng Soc 55(6):503–516
5. Roy R, Kailath T (1989) ESPRIT-estimation of signal parameters via rotational invariance techniques. Acoustics, Speech and Signal Processing, IEEE Transactions on 37(7):984–995
6. Schultz-Amling R, Kuech F, Kallinger M, Del Galdo G, Ahonen J, Pulkki V (2008) Planar microphone array processing for the analysis and reproduction of spatial audio using directional audio coding. In Audio Engineering Society Convention 124. Amsterdam, The Netherlands
7. Thiergart O, Kratschmer M, Kallinger M, Del Galdo G (2011) Parameter estimation in directional audio coding using linear microphone arrays. In Audio Engineering Society Convention 130. London UK

# High-Definition Audio for Group-to-Group Communication

**Nikolaus Färber, Manfred Lutzky, and Fabian Kuech**

**Abstract** The European research project "Together anywhere, together anytime" (TA2) explores how technology can make communication and engagement easier among groups of people separated in space and time, e.g. families living in different cities. It combines high quality communication with shared group activities using distributed multimedia applications and broadband technology in the living room. This paper provides an overview of the TA2-project with a focus on high-definition audio communication, which has the goal of providing an experience "as if speaking to someone in the same room". The most important aspects for making audio communication a natural experience are high audio bandwidth, robust transmission with low delay, hands-free operation, and spatial audio (surround sound) reproduction. These requirements are addressed by the Audio Communication Engine (ACE), which includes audio coding, acoustic processing and IP transport in a single software module. Based on concept demonstrators developed in the TA2-project, results from user experience tests are presented which indicate that more than 90% of users rate the audio quality as *natural* when using high-definition audio, compared to less than 10% when standard audio quality is used.

## 1 Introduction

Many of our enduring experiences, such as holidays, celebrations and moments of fun and laughter, occur as group events. These experiences are events that current technology does not address well. Modern media and communications serve individuals best. Phones, computers, and electronic gaming devices are usually owned and used by individuals and provide individual experiences. For example, consider a teenager when checking her Facebook account on an iPad in the living room. She

Nikolaus Färber (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
nikolaus.faerber@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_31, © Springer 2011

would certainly be most annoyed when a parent is looking over her shoulder or even trying to touch her screen. Groups of people sitting in a coffee bar can often be observed sitting at a table, each twittering on their Blackberry smartphone without taking much notice of each other. Hence, social interactions within the real environment of a group or family are inhibited rather than supported by such devices and services. Contrast this to the situation of playing a conventional board game at the table. Here, the family members and invited friends talk to each other and interact together with the board game. Other examples of group-based engagement that fill our lives are sport events, music concerts, sharing pictures from a past vacation, or simply having a chat at the dinner table.

Since families and friends are often forced to live in different cities, the question arises how technology can help to support similar group experiences despite geographic distance. This is the central question addressed by the European research project "Together Anywhere, Together Anytime" (TA2), which started in February 2008 [1]. Based on social science theory and interviews with users [2], TA2 identified two components which are essential for applications aiming at making communication and engagement easier among groups of people separated in space and time:

1. High quality communication,
2. Shared activity.

While the former is required for natural and intense communication during group events, the latter is equally important for triggering regular events and building a group experience. Following this basic design guideline, TA2 developed several *concept demonstrators* which address different application areas and use a diverse set of advanced technologies. In the context of this paper, we focus on the communication aspect and, more specifically, on the audio part. However, it should be noted, that the overall TA2 project has a much bigger scope covering aspects of audiovisual analysis [3], dynamic video composition [4], or and artificial intelligence for automatic camera control [5].

## 2 Audio Communication Technologies

For natural communication, audio is particularly important and still challenging to implement when full audio bandwidth, low delay, low bit-rate transmission, hands-free operation and multichannel capability are considered. The goal of audio communication should be to allow a user experience *as if speaking to someone in the same room*.

Considering that the quality of current phone calls is still limited to mono at less than 4 kHz audio bandwidth, it becomes obvious that many technical challenges have to be addressed. Over the past three years we have worked on various aspects of this vision and integrated those into the Audio Communication Engine (ACE). The ACE covers the complete audio chain from the microphone to IP transport. Its

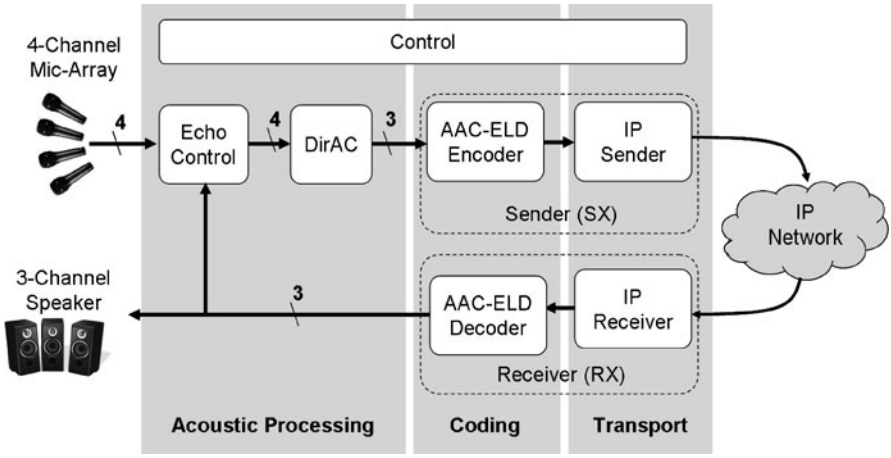



**Fig. 1** Components of the Audio Communication Engine

functionality is separated into three main categories: acoustic processing, coding and transport, as illustrated in Fig. 1 and described in the following sections.

## *2.1 Acoustic Processing*

Because the conversation should be as natural as possible, clip-on microphones and headsets shall be avoided in the living room and audio signals should be captured unobtrusively with hands-free operation. Nevertheless, sufficient speech quality should also be provided in noisy and reverberant environments. Furthermore, it is desirable to be able to capture complex acoustic scenes in one location and to realistically reproduce them at the other end of the communication chain. All of the above requirements and challenges have to be dealt with by acoustic processing, i.e. signal processing steps somewhere in between the microphone and the audio encoder. In the following we discuss associated technical challenges and briefly present solutions as implemented in ACE and the TA2 concept demonstrators.

### 2.1.1 Spatial Audio Recording and Reproduction

For reasons of flexibility, the recording technique should not put any constraints on the loudspeaker configuration used for reproduction. Thus, common spaced-microphone techniques are not suitable, as they assume a priori knowledge of the loudspeaker systems [6]. Although coincident-microphone approaches, such as Ambisonics, are independent of the loudspeaker set-up, they do not yield the localization accuracy for talkers that is desired for communication applications. Therefore,

the acoustic interface of the ACE is based on a parametric description of sound fields taking into account human perception of spatial sound, namely Directional Audio Coding (DirAC) [7]. The features to be determined from the microphone signals are the direction-of-arrival (DOA) of sound and the diffuseness of the sound field. These parameters can be estimated via an energy analysis of the sound field based on the active sound intensity vector. As discussed in [8], the DirAC analysis can be performed using small two-dimensional microphone configurations extending only a few centimetres and allowing unobtrusive integration into the living room. The actual signals of the loudspeaker channels for surround playback are determined from the microphone signals as a function of the DirAC parameters so that an accurate spatial rendering can be achieved at a desired listening position. As discussed in [7], the loudspeaker configuration can be chosen almost arbitrarily and can be designed to match application-specific requirements. It is also worth mentioning that the directional information concerning the sound field can further be exploited to efficiently implement speech enhancement functionality, such as spatial filtering [9] or de-reverberation [10] as described in other papers in this book.

As illustrated in Fig. 1, the TA2 concept demonstrator uses a configuration consisting of a 4-channel microphone array and a 3-channel speaker setup (left, center, right). The array consists of four omnidirectional boundary condenser microphone capsules arranged at the corners of a 4.4 cm square [8]. The microphone array is mounted in a horizontal plane, integrated into a floor lamp with an arm extending over a table in front of the participants. For the given configuration, DirAC can be seen as a pre-processing step before encoding; it is used to calculate 3-channel speaker signal from the 4-channel microphone array in a flexible way. No interaction with subsequent encoding by AAC-ELD is needed as the output of DirAC is encoded as three audio channels.

### 2.1.2 Acoustic Echo Control

Acoustic echoes are a serious problem for hands-free operation in communication systems. They arise from an acoustic coupling between the loudspeakers and the microphones of the acoustic interface. The signal emitted from the loudspeaker travels to the microphone directly and via reflections in the room and is thus transmitted back to the far-end side. Acoustic echoes represent an annoying disturbance and can severely impede natural conversation. The most common solution to this problem is to insert adaptive filters in parallel to the acoustic echo path and to subtract an estimate of the echo from the microphone signal. In the context of wideband speech communication and, more importantly, multi-channel acoustic interfaces, this approach imposes very high computational complexity, which typically requires dedicated hardware solutions. Instead, the TA2 concept demonstrators employ an echo control technique based on a simple parametric model of the acoustic echo path [11]. The model includes a direct propagation path between the loudspeakers and the microphones, coloration effects of early reflections, and the contribution of late reverberation. Based on this model, an estimate of the echo power spectrum included

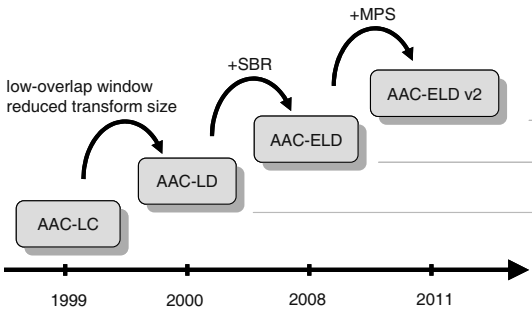
in the microphone spectra is obtained from the corresponding spectra of the loudspeaker signals. The echo control algorithm then attenuates those spectral portions in the microphone signals that mainly include echo components. In typical application scenarios including multi-channel configurations, robust attenuation of the echo by 60 dB can be expected and is reliably achieved.

In Fig. 1, all four microphone signals are echo-controlled based on the input of the three speaker channels before further processing by DirAC. Even though all audio signals are sampled at 48 kHz, the ACE can still be implemented to run in real-time on a general purpose CPU of a standard PC together with all other components of the system, demonstrating the low computational complexity.

## *2.2 Audio Coding*

After obtaining clean and undisturbed PCM samples through acoustic processing, it is the task of audio coding to reduce the amount of data while preserving audio quality as much as possible. ISO/MPEG Advanced Audio Coding (AAC) [12] is already in widespread use today, e.g. most prominently the Low Complexity version (AAC-LC) in Apple's iPod music player. Though AAC offers excellent audio quality at stereo bit-rates of 128 kbps, it features a high algorithmic delay (about 100 ms). While this is not a disadvantage for offline music playback, interactive conversation suffers from too long delay. For example, the frequency of unintended interruptions is increased as conversation partners realize too late that the remote person already started to talk. For this reason, MPEG has standardized variations of AAC with reduced delay as illustrated in Fig. 2 and described below.

The development towards low delay started in the year 2000 with the standardization of AAC-LD, featuring an excellent audio quality for speech and music at bitrates comparable to those of MP3. A high conversational quality is achieved by providing full audio bandwidth at an algorithmic delay of only 20 ms. In 2008 this codec was enhanced to create AAC Enhanced Low Delay (AAC-ELD) [13] by the incorporation of a low-delay Spectral Band Replication tool (SBR) that reduces the




| algorithmic delay [ms] | typical bit rate [kbit/s] |
|---|---|
| 19-39 | 24 – 48 (stereo) |
| 15-32 | 24 (mono) – 128 (stereo) |
| 20-40 | 32 (mono) – 128 (stereo) |

**Fig. 2** Evolution of MPEG AAC conversational codecs

**Table 1** Min. bit-rates required for excellent quality of mono signals [14]

| | Codec | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| [kbps] | AAC-ELD | AAC-LD | CELT | G.718 | G.719 | G722.1-C | G722.2 | G.722 | SILK | Speex |
| Audio item 1 | 32 | 48 | 64 | 48 | 32 | 48 | – | – | 40 | – |
| 2 | 24 | 48 | 48 | – | 32 | 32 | – | – | – | – |
| 3 | 48 | 32 | 64 | – | 48 | – | – | – | 40 | – |
| 4 | 32 | 48 | 48 | – | 48 | 24 | – | – | – | – |
| 5 | 32 | 48 | 48 | – | 32 | 32 | – | – | – | – |
| 6 | 32 | 48 | 48 | 32 | 48 | – | – | – | 40 | – |

operating point to 24 kbps per channel. A further bit-rate reduction for stereo content has been achieved by adding the low-delay MPEG Surround tool (MPS) in 2011 which is denoted as AAC-ELD v2 and described in more detail in other papers in this book. Figure 2 summarizes the relationship among the MPEG AAC-LD/ELD codecs along with their key features. The listed bit rates indicate the typical operation points. Since more recent versions include previous versions as a subset, higher bit rates can always be used if desired. All codecs support the full audio bandwidth, i.e. up to 48 kHz sampling rate.

Evidence for the quality of AAC-ELD compared to other state of the art super-wideband codecs has been provided by Deutsche Telekom AG in an independent listening test [14]. The test was conducted to gain insights into the required bit-rate for excellent speech and audio quality in mobile conversational applications. The test set comprises MPEG codecs (AAC-LD, AAC-ELD), ITU codecs (G.718, G.719, G.722.1-C, G.722.2), the audio codec used by Skype (Silk) as well as open source codecs (CELT, Speex). The listening panel consisted of 26 non-expert listeners. The report concludes that "The AAC-ELD seems to perform most efficient [sic] for the given contents". An excellent quality (MUSHRA score $>80$) can only be achieved with a subset of the codecs under test (AAC-ELD, AAC-LD, CELT, G722.1-C and G.719). Table 1 shows the minimum bit-rate at which excellent audio quality can be achieved for mono signals. Only AAC-LD/ELD, CELT and G.719 achieve excellent quality consistently over all 6 test items. AAC-ELD provides this quality at an average bit-rate of 32 kbps, which is significantly lower than the bit-rates required by CELT and G.719.

Stereo performance has also been tested in this listening test. It shows that AAC-ELD is the only codec that achieves excellent audio quality at all items at 64 kbps. This is especially relevant as spatial audio is one of the main features of the ACE. For the implementation of the TA2 demonstrators, as illustrated in Fig. 1, the AAC-ELD codec is used to encode three channels (left, center, right) at a total bit-rate of 192 kbps.

In summary, AAC-ELD is a very competitive codec compared to other state of the art super-wideband codecs. It can assure excellent audio quality at 32 kbps/channel while offering a very low delay of 15–32 ms.

## 2.3 IP Transport

Giving the rising importance of the Internet as a ubiquitous network infrastructure, the ACE focuses on transport over the Internet Protocol (IP). Besides practical issues, such as session initialization, NAT traversal, and packetization in RTP/UDP protocols, the transport module of the ACE has to cope with the delay variation of IP-packets, usually referred to as *jitter*.

One of the most challenging tasks in low-delay IP-streaming is to set the correct size of the de-jitter buffer. Normally, a *fixed playout* deadline is chosen, meaning that packets undergo a fixed end-to-end delay. However, if this fixed delay is set too low, then packets may arrive too late to be played out in time. Since these packets are not lost on the network but arriving late, the term *late loss* is used. When the fixed delay is increased, all packets are received in time but with the drawback of an unnecessary high *buffering delay*, i.e. the amount of time that a packet is kept in the buffer before playout. Because the variations in network delay are not easy to predict, it is very difficult to find a good trade-off between late loss and buffering delay when using a fixed playout. This is illustrated in Fig. 3 (right, dashed line), where unnecessary buffering delay is introduced in the beginning, while severe late loss has to be accepted around $t = 70\,\mathrm{s}$.

As a solution to this problem, *adaptive playout* offers an algorithm which estimates the jitter on the network and adapts the size of the de-jitter buffer in order to minimize buffering delay and late loss. For example, if the delay and/or jitter increases on the network, then the playout time is increased to reduce late loss. By contrast, if the delay and/or jitter decreases at a certain point, the buffer is reduced to minimize buffering delay. Hence, the playout follows effectively the network delay as illustrated in Fig. 3 (right, solid line).

Note that an adaptation of the playout time implies shrinking and stretching the audio signal in time. The required time modification is realized with low complexity and good quality by exploiting the overlap-add structure of the AAC-ELD codec.

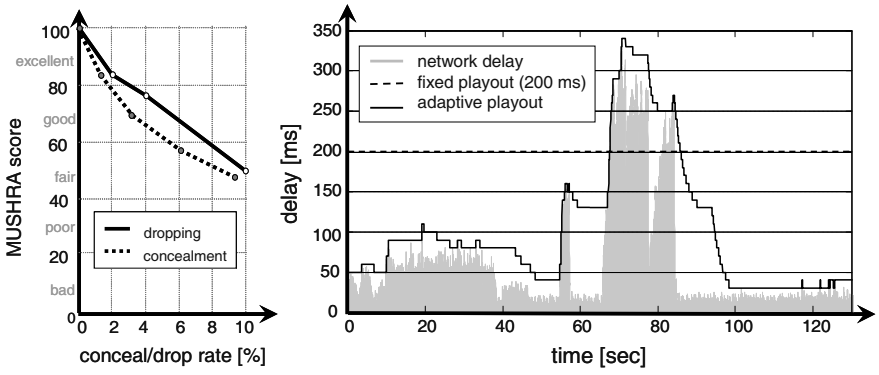



**Fig. 3** Subjective quality for time scaling using AAC-ELD concealment and dropping (*left*) and example of fixed vs. adaptive playout (*right*)

To reduce the buffer size, the ACE simply skips single audio frames at the decoder. I.e. instead of decoding the frame sequence (1, 2, 3, 4), the ACE drops frame 3 and decodes (1, 2, 4) and thus implements time-shrinking. For time stretching, the AAC-ELD decoder generates an additional frame through a concealment operation (C). I.e. instead of decoding frames (1, 2, 3) the ACE decodes (1, 2, C, 3) and therefore extends the timeline by one frame interval. The left plot in Fig. 3 shows the subjective quality of time-modification using the MUSHRA test methodology. As can be seen, this approach of time-scale modification based on dropping and concealment results in excellent quality as long as the modification rate is kept below 2%. Hence, this number is used as a design target for optimizing the parameters of the Jitter Buffer Management (JBM) within the ACE.

In summary, the adaptive playout algorithm implemented in the transport module of the ACE achieves the lowest possible delay while still maintaining good audio quality. The achieved end-to-end ("mouth to ear") delay is below 80 ms on a Local Area Network (LAN), leaving enough margin for additional delay on the Internet. Further information on ACE adaptive playout is given in [15].

## 3 User Experience Evaluation

While some of the above results already evaluate specific aspects of audio quality in isolated test conditions, it is still a valid question if the demonstrated benefits are also perceived in a scenario such as envisioned in TA2. Therefore, user experience (UX) tests have been conducted with families and friends when being involved in an interactive group activity. For this purpose, the ACE was installed in two rooms for computer-mediated communication. In addition, an HD-videoconference and a coffee table with a touch-screen where installed. The latter allowed playing of simple board-games such as "Memory" or "Ludo". This setup resulted in a casual group conversation, where people started to chat and laugh spontaneously, triggered through the game as a framing experience. The left side of Fig. 4 shows a screenshot from the HD-videoconference and provides an impression of the scenario and atmosphere as experienced by the users.

The group was exposed to "high" and "low" audio quality for a duration of about 5 min. each and then asked to express their agreement to a statement such as "The communication was natural and without problems" on a 5-point rating scale ranging from 1 (strongly agree) to 5 (strongly disagree). While the "high" audio quality is characterized by high bandwidth (44.1 kHz sampling rate), spatial audio (left, center, right), and low delay (80 ms), the "low" audio quality is characterized by narrowband (8 kHz sampling rate), mono (routed to all three speakers), and high delay (600 ms). The low quality settings have been derived from an analysis of communication solutions in current game consoles (Wii Speak, Xbox Live, Teamspeak) and can therefore be considered as "standard". From the 25 people who participated in the test, only 12 were above the age of 12 and allowed to vote. Since each test condition was presented twice, the results are based on 24 votes.

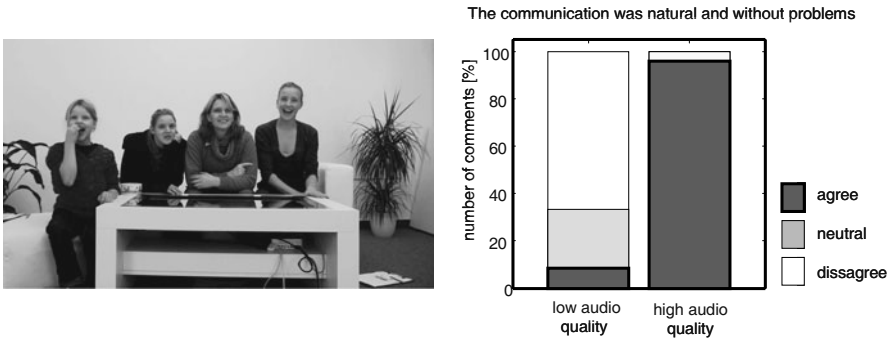



**Fig. 4** User experience during interactive group activity for low and high audio quality

Figure 4 (right) illustrates the results when mapping the scores (1, 2) to a general "disagree" and scores (4, 5) to a general "agree". A score of (3) is interpreted as "neutral". As can be seen, the agreement raises from below 10% to above 90% when improving the audio quality from low to high. Though further tests are needed to validate these results and to investigate which parameters of the audio chain influence the user experience most (audio bandwidth, number of channels, or delay), the initial results provides a good indication for the importance of audio quality in group-to-group communication.

## 4 Summary and Conclusions

Despite innovative devices and services in the social media landscape, we believe that there is a gap to fill in between the user experiences of iPad and Facebook. This gap includes geographically distributed group experiences combining high quality communication with a shared activity, as investigated in the EU project TA2.

Audio and particularly speech is an important aspect for social interaction and it is the goal of the Audio Communication Engine (ACE) to provide an experience *as if speaking to someone in the same room*. This vision requires significant progress in all areas of the transmission chain including audio capture, acoustic processing, coding, and transport. The specific technology components within the ACE, which address these areas include AAC-ELD, DirAC, echo control and adaptive playout.

The performance of the above technologies in a group-to-group communication scenario has been assessed through user experience tests involving the remote playing of a board game while communicating through the ACE. More than 90% of users rate the audio quality as natural when using high-definition audio, compared to less than 10% when standard audio quality is used. This indicates the importance of audio in making communication and engagement easier among groups of people separated in space.

## Acknowledgments

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. ICT-2007-214793

## References

1. Homepage of TA2 project, www.ta2-project.eu
2. Williams D, Ursu MF, Meenowa J, Cesar P, Kegel I, Bergstrom K (2011) Video Mediated Social Interaction Between Groups. System Requirements and Technology Challenges. Elsevier Telematics and Informatics (Elsevier T&I), 28(4):251–270
3. Korchagin D et al (2011) Just-in-Time Multimodal Association and Fusion for Home Entertainment. In Proc IEEE Int Conf on Multimedia & Expo (ICME), Barcelona, to be published
4. Jansen J et al Enabling Composition-Based Video-Conferencing for the Home. IEEE Transactions on Multimedia, accepted for publication
5. Falelakis M et al (2011) Reasoning for Video-mediated Group Communication. In Proc IEEE Int Conf on Multimedia & Expo (ICME), Barcelona, to be published
6. Eargle J (2001) The microphone book, Boston: Focal Press
7. Pulkki V (2007) Spatial Sound Reproduction with Directional Audio Coding. Journal of the Audio Engineering Society, 55(6):503–516
8. Kallinger M, Kuech F, Schultz-Amling R, Galdo GD, Ahonen J, Pulkki V (2008) Analysis and Adjustment of Planar Microphone Arrays for Application in Directional Audio Coding. 124th AES Convention, Paper 7374, Amsterdam, The Netherlands
9. Kallinger M, Galdo GD, Kuech F, Mahne D, Schultz-Amling R (2009) Spatial filtering using directional audio coding parameters. In Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Taipei
10. Kallinger M, Galdo GD, Kuech F, Thiergart O (2011) Dereverberation in the spatial audio coding domain. 130th AES Convention, Paper 8429, London, UK
11. Kuech F, Kallinger M, Schmidt M, Favrot A, Faller C (2008) Acoustic Echo Suppression Based on Separation of Stationary and Non–Stationary Echo Components. In Proc Int Workshop on Acoustic Echo and Noise (IWAENC), Seattle
12. ISO/IEC 13818-7 (1997) Information Technology – Generic Coding of Moving Pictures and Associated Audio, Part 7: Advanced Audio Coding
13. Schnell M et al (2007) Enhanced MPEG-4 Low Delay AAC – Low Bitrate High Quality Communication. 122th AES Convention, Paper 6998, Vienna
14. 3GPP Document S4-100479 (2010) Listening tests concerning reference codecs for EVS, from Deutsche Telekom AG, TSG-SA4#59 meeting, Prague, Czech Republic, available at ftp://ftp.3gpp.org/tsg_sa/WG4_CODEC/TSGS4_59/Docs/S4-100479.zip
15. Issing J et al (2008) Adaptive Playout for VoIP based on the Enhanced Low Delay AAC Audio Codec. 124th AES Convention, Paper 7395, Amsterdam

# MPEG-4 AAC-ELD v2 – The New State of the Art in High Quality Communication Audio Coding

**Manfred Lutzky, Markus Schnell, María Luis Valero, and Johannes Hilpert**

**Abstract** Recently MPEG finished the standardization of a Low Delay MPEG Surround tool that is tailored for enhancing the widely adopted AAC-ELD low delay codec for high-quality audio communication into AAC-ELD v2. In combination with the Low Delay MPEG Surround tool, the coding efficiency for stereo content outperforms competing low delay audio codecs. This paper describes the technical challenges and solutions for designing a low delay codec that delivers a performance which is comparable to that of existing state of the art compression schemes. It provides a comparison to competing proprietary and ITU-T codecs, as well as a guideline for how to select the best possible points of operation. Applications facilitated by AAC-ELD v2 in the area of broadcasting and mobile video conferencing are discussed.

## 1 Introduction

Given the new applications of Web 2.0, the user behavior in modern communication scenarios has changed dramatically towards more interactivity and a more diverse use of media. Examples for this evolution can be found in the success of Facebook or mobile video conferences provided, for instance, by Apple's "Facetime". This change in user experience comes along with new expectations driving the development of audio codecs that are able to fulfil novel requirements. A need for low delay, low bitrate and high quality communication codecs has arisen, for which stereo transmission is a key feature. In this paper the recently standardized Low Delay MPEG Surround is presented as a parametric stereo tool for AAC-ELD [5], which provides good audio quality for bitrates as low as 24 kbps. The combination of both technologies is also known as AAC-ELD v2. Its advantages and operation points will be described in the course of this paper.

Manfred Lutzky (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
manfred.lutzky@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_32, © Springer 2011

## 2 Stereo Communication Scenarios

Most traditional communication scenarios, such as plain old telephone, VoIP or video- and teleconferencing, are still based on mono audio connections, mainly due to hardware restrictions. Considering the field of emerging applications, those restrictions of the devices will disappear and stereo will become a mandatory feature. One of these applications is *Telepresence at home*, an upcoming key application for IP network enabled TVs and set top boxes. Stereo enhances also services such as *Live Broadcasting over Next generation Networks* where the reporter in the field interacts with the director in a broadcasting studio [5]. *Spatially enhanced audio conferencing* is another novel application which aims at recreating a conferencing room by spatially rendering all online participants. Each participant is considered as an independent audio object which is rendered at a dedicated spatial position and level. Hereby, the mono streams from each participant are mixed into a single stereo stream processed at a centralized unit and sent back to each participant. This spatial conditioning of the signal allows the recipient to exploit the *cocktail party effect*, to enhance speech intelligibility.

While the demand for high audio quality is obvious for those scenarios, low bitrates are necessary to achieve robust transmissions over wireless channels as well. Furthermore, lower data rates result in reduced radio activity of the transmitter which leads to a significantly increased battery lifetime of the mobile device. Finally, low delay is needed for applications where synchronization with external video is demanded, for example the transmission of TV stereo sound to wireless headphones. If the audio delay is too high, sufficient lip synchronization to the video presentation of the talker can not be maintained.

## 3 Communication Codecs

The high audio quality demanded by modern communication scenarios calls for codecs which provide an audio bandwidth significantly higher than the 3.5 kHz of standard telephone systems. Such codecs are known as super wideband (SWB) or full band (FB) codecs. By SWB codecs, an audio bandwidth of up to 14 kHz can be transmitted while FB codecs extend this range even further.

SWB and FB communication codecs have been released by standardization bodies such as ITU-T (G.722.1 Annex C, G.719, G.718), MPEG (AAC-ELD) and other bodies such as xiph.org (CELT).

### *3.1 ITU-T*

With G.722.1 Annex C [11], the ITU-T defined their first SWB codec in 2005. The codec is based on an MDCT transform coding scheme utilizing a sine window.

Later in 2008, the ITU-T finalized the work on the successor of G.722.1 Annex C, G.719 [10] which is also the first full band codec of the ITU-T. This codec is based on transform coding as well but improves this approach by a signal adaptive time-frequency filter bank. The latest SWB development from the ITU-T was finalized in 2010, namely G.718 Annex B [4, 8]. The core coder features a layered scalable coding approach where the lower layers contain ACELP coding techniques while the higher layers are based on transform coding. The transform comprises a delay optimized asymmetric window. Such technologies are also known as low delay filterbanks.

As this publication examines stereo capability it has to be mentioned that the codecs G.722.1 Annex C, G.718 and G.719 do not support joint coding of both channels of stereo signals. This means to code stereo signals two codec instances have to process each channel separately. For G.718, a stereo extension layer is planned although yet not released.

### *3.2 Xiph*

The xiph.org organization also worked on a SWB communication codec which was first released in 2008 under the name CELT [9]. This codec is based on the transform coding scheme as well but has symmetric transform windows which are delay optimized by reducing their overlap.

### *3.3 MPEG*

With AAC-LD [1], the Moving Picture Experts Group (MPEG) introduced the worldwide first low delay codec for high quality audio at low bitrates already back in 2000. This codec is derived from the general MPEG music codec AAC-LC by optimizing its tools towards low delay operation. This results in a codec that keeps the high audio quality of its parent while reducing the delay down to 20 ms, almost without an increase in data rate. In 2008 AAC-ELD [7] was standardized and included into the MPEG set of communication codecs. It was the first codec using low delay filter banks. This further reduced the delay down to 15 ms while maintaining the high audio quality. Moreover, its combination with a low delay version of the spectral band replication (SBR) tool expanded the codec's operating range towards lower bit rates [6]. Finally, as other codecs of the AAC family, it included dedicated joint stereo coding tools. The latest step to AAC-ELD v2 is the incorporation of low delay MPEG surround [5] to enhance the stereo performance at low bitrates, which is described in more detail in Sect. 4.2.

# 4 Stereo Coding

This chapter explains the stereo coding algorithms of AAC-ELD as well as the parametric stereo coding extension of AAC-ELD v2.

## *4.1 Discrete Stereo Coding*

AAC-ELD has a set of encoder controlled stereo coding tools or modes designed for different operation points:

For totally independent coding of two audio channels, a dual-mono operation may be used. Dual mono is only recommended when the two channels contain two different mono programs, as for instance bilingual material. Usually only the content of one of the two channels is presented at the receiver side. The data rate is almost doubled with respect to mono operation to achieve the same quality.

For transmission of dedicated stereo material, e.g. stereo microphone signals or stereo music content, AAC-ELD offers the option to use Mid-Side (MS) and intensity stereo (IS) coding to exploit inter-channel redundancies and to avoid stereo unmasking effects for highly correlated audio channels. These techniques are used in well-known codecs like MP3 or AAC as well. For typical material this type of joint stereo coding can help to reduce the bit rate slightly and increase the performance for critical material, when compared to dual-mono coding. Vice versa, higher audio quality can be delivered at the same data rate.

## *4.2 Parametric Stereo Coding with Low Delay MPEG Surround*

To achieve stereo performance at bit rates close to monophonic operation, a parametric stereo extension has been integrated into AAC-ELD v2. This parametric extension is based on a 2-channel version of Low Delay MPEG Surround (LD-MPS) [5]. Its main principle is to resynthesize the stereo image from a mono downmix of the two input signals by using a description of frequency- and time-variant spatial parameters, such as inter-channel level differences (ILD) and inter-channel cross-correlation (ICC). To achieve only a negligible increase in the codec's processing delay, the stereo-to-mono downmix is applied in the time representation of the input signal. The LD-MPS parameter estimation is done in parallel to the AAC-ELD coding of the downmix signal. The LD-MPS upmix is carried out as a postprocessing step to the AAC-ELD decoder. Spatial parameters are supplied as a multiplex in the bitstream of the AAC-ELD v2 compressed data.

When the codec operates in LD-MPS mode, the spatial processing shares a newly designed low delay Quadrature Mirror Filterbank (LD-QMF) with the Spectral Band

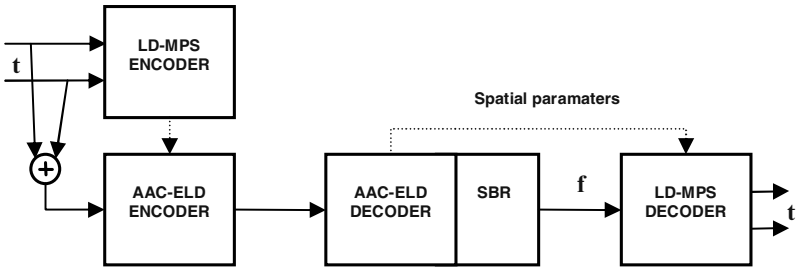



**Fig. 1** AAC-ELD v2 Block Diagram: combination of Low Delay MPEG Surround and AAC-ELD

Replication (SBR)-module. This filterbank has an asymmetric prototype that combines low delay properties with sufficiently high frequency separation. The frequency domain interface with SBR reduces the computational complexity as well as the overall delay for the processing, since no additional analysis and synthesis filterbanks are needed. The spatial parameters are quantized and coded efficiently by entropy coding to achieve a very compact representation of the stereo image. Typically the bit rate overhead for the stereo parameters is around 3 kbps at 48 kHz. This allows AAC-ELD v2 to code stereo signals at bit rates significantly lower than with its discrete stereo coding tools without introducing significant additional delay. In the subsequent section, the optimum cross-over point between these modes is discussed and validated by subjective listening tests.

## 5 Performance

In the following subsections, the results of two listening tests are provided. Firstly, Fig. 2 shows the results of an independent test performed by Deutsche Telekom [2] comparing the stereo operation of AAC-ELD with respect to competitive codecs. Secondly, a comparison of the efficiency of different stereo coding modes for AAC-ELD v2 is given in Fig. 3. Both listening tests were conducted using the MUSHRA test methodology [3], with a hidden reference and low-pass filtered anchor conditions.

### *5.1 Performance Evaluation for AAC-ELD vs. ITU vs. XIPH*

The listening test conducted by Deutsche Telekom [2] provides a good overview of the stereo performance of SWB communication codecs.

Here, the codecs AAC-ELD, AAC-LD, G.718, G.719 and CELT were compared with 5 audio items, these items are described in Table 1. The results, which are depicted in Fig. 2, include only relevant codecs evaluated in the original test and

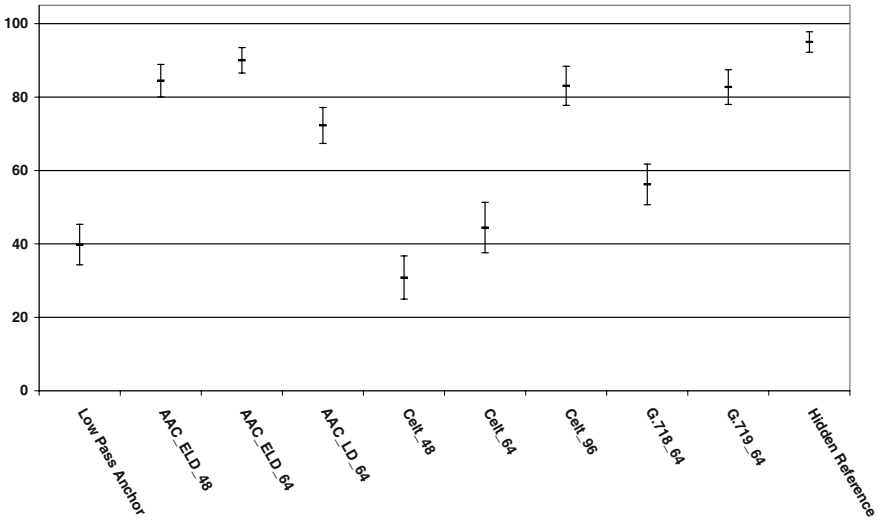



**Fig. 2** MUSHRA Listening test results for the comparison of different communication codecs

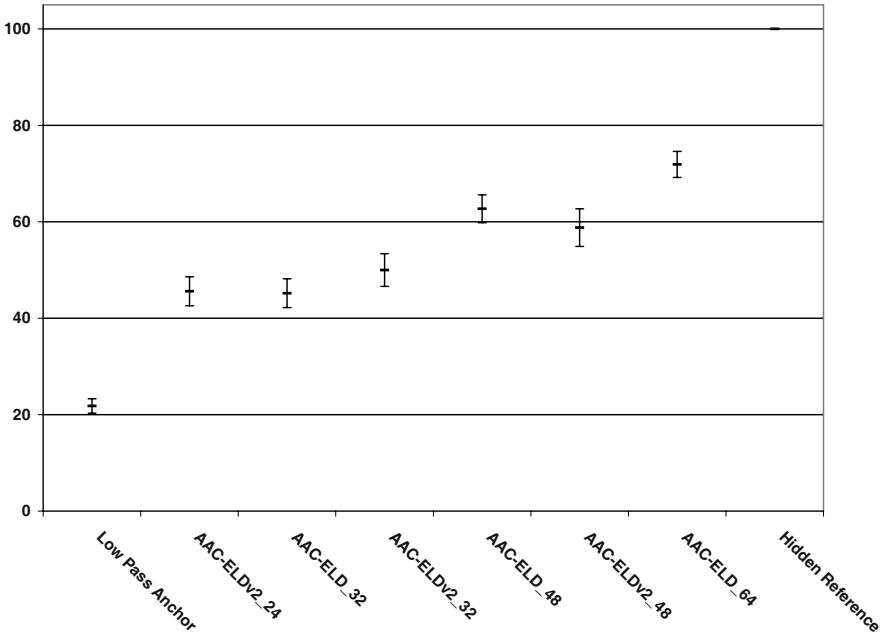



**Fig. 3** MUSHRA Listening test results for the comparison of AAC-ELD and AAC-ELDv2 at different bitrates

**Table 1** List of items used in the stereo listening test conducted by Deutsche Telekom in [2]

| Item name | Description |
|---|---|
| Chapman | Tracy Chapman part from Mountains of things |
| Dialog | Male on right and female on left speaker talking simultaneously |
| Hockey | Ice hockey report with announcement and noise of the auditory |
| Klassik | Bach trumpet |
| Neues | Male speaker with background music |

therefore exclude the Speex codec which showed a poor audio quality at all tested bitrates. From this test it was concluded that "*for stereo, AAC-ELD achieves excellent quality for all contents at 64 kbit/s*" [2], while no other codec was able to achieve this performance.

## *5.2 Performance Evaluation for AAC-ELD v2*

The test results presented in Sect. 5.1 show the rate-distortion advantage of AAC-ELD at intermediate bit rates. For this test, AAC-ELD operated in joint stereo mode at 48 and 64 kbps. However, this test does not reveal the bit rate range where the use of a parametric stereo extension with LD-MPS is advantageous over joint stereo coding and what quality gain can be achieved. To rank the performance of AAC-ELD v2, a listening test was set up, in which the following conditions were evaluated:

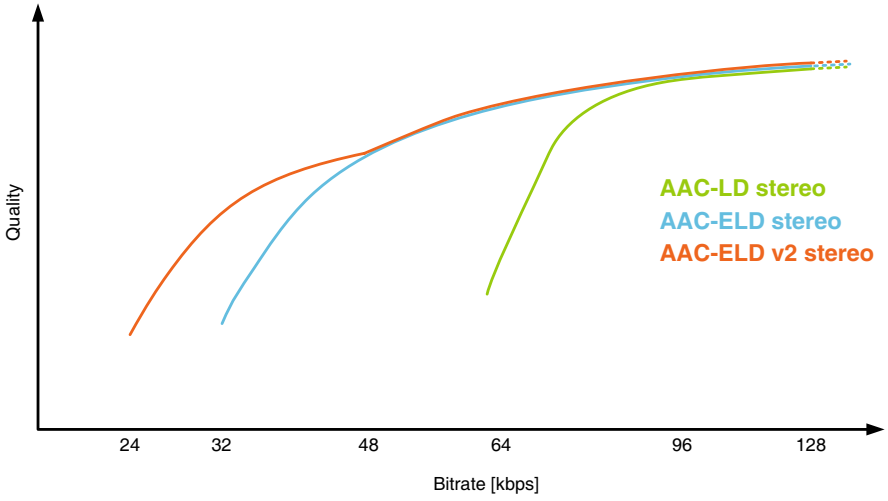



**Fig. 4** Graphic representation of the overall AAC-ELD v2 quality compared to AAC-ELD and AAC-LD

- AAC-ELD stereo at 64 kbps, 48 kbps and 32 kbps,
- AAC-ELD v2 parametric stereo at 48 kbps, 32 kbps and 24 kbps,
- Hidden Reference and 3.5 kHz Low Pass Anchor.

The same 5 pre-processed test items as in [2] were used for the evaluation. The stimuli were presented over Stax "Lambda Pro" headphones in an acoustically isolated laboratory. 8 listeners, all experienced in the field of audio coding, took part in the test. The results are shown in Fig. 3.

From the results it may be concluded that the optimum threshold between the operation modes is between 32 and 48 kbps. For lower rates, LD-MPS gives a quality advantage over discrete stereo coding. AAC-ELD v2 at 24 kbps is basically on a par with AAC-ELD at 32 kbps. In Fig. 4 the relative relationship between quality and bitrate of AAC-ELD v2 compared to AAC-ELD is illustrated, using AAC-LD as a quality reference.

## 6 Conclusions

This paper introduces the new Low Delay MPEG Surround parametric stereo coding tool of AAC-ELD v2 which upgrades the AAC-ELD full bandwidth communication codec. As it is fully backwards compatible to its predecessor, only the extended operation points are evaluated. Listening tests on stereo content show an improved audio quality below 48 kbps. The useful range of operation is extended down to 24 kbps where its perceived quality is comparable to AAC-ELD at 32 kbps for stereo signals.

## Acknowledgements

The authors would like to thank the entire Low Delay AAC and MPEG Surround teams at Fraunhofer IIS for their help in evolving the technology, in particular Prof. Jürgen Herre for his valuable contributions and his support in the standardization activities.

## References

1. Allamanche E, Geiger R, Herre J, Sporer T (1999) MPEG-4 Low Delay Audio Coding Based on the AAC Codec. In 106th AES Convention. Munich, Germany. Preprint 4929
2. Feiten B, Kroll J, Raake A, Wältermann M, Wüstenhagen U (2010) Evaluation of superwideband speech and audio codecs. In Audio Engineering Society Convention 129
3. International Telecommunication Union (2001) Method for the subjective assessment of intermediate sound quality (MUSHRA). ITU-R, Recommendation BS 1534-1, Geneva, Switzerland

4. Laaksonen L, Tammi M, Malenovsky V, Vaillancourt T, Lee MS, Yamanashi T, Oshikiri M, Lamblin C, Kovesi B, Miao L, Zhang D, Gibbs J, Francois H (2010) Superwideband Extension of G.718 and G.729.1 Speech Codec. In INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association. Makuhari, Japan
5. Luis Valero M, Hölzer A, Schnell M, Hilpert J, Lutzky M, Engdegård J, Purnhagen H, Ekstrand P, Kjörling K, Oomen W (2010) A New Parametric Stereo and Multichannel Extension for MPEG-4 Enhanced Low Delay AAC (AAC-ELD). In 128th AES Convention. London, UK
6. Schnell M, Schmidt M, Ekstrand P, Albert T, Przioda D, Lutzky M, Geiger R, Ruoppila V, Henn F, Tårnes E (2008) Delayless mixing – on the benefits of MPEG-4 AAC-ELD in high quality communication systems. In 124th AES Convention. Amsterdam, The Netherlands. Preprint 7337
7. Schnell M, Schmidt M, Jander M, Albert T, Geiger R, Ruoppila V, Ekstrand P, Lutzky M, Grill B (2008) MPEG-4 Enhanced Low Delay AAC – a new standard for high quality communication. In 125th AES Convention. San Francisco, CA, USA. Preprint 7503
8. Vaillancourt T, Jelinek M, Ertan AE, Stachurski J, Rämö, A., Laaksonen L, Gibbs J, Mittal U, Bruhn S, Grancharov V, Oshikiri M, Ehara H, Zhang D, Ma F, Virette D, Ragot S (2008) ITU-T EV-VBR: A robust 8–32 kbit/s scalable coder for error prone telecommunication channels. In 16th European Signal Processing Conference. Lausanne, Switzerland
9. Valin JM, Terriberry TB, Maxwell G (2009) A full-bandwidth audio codec with low complexity and very low delay. In 17th European Signal Processing Conference. Glasgow, Scotland
10. Xie M, Chu P, Taleb A, Briand M (2009) ITU-T G.719: A new low-complexity full-band (20kHz) audio coding standard for high quality conversational applications. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp 265–268
11. Xie M, Lindbergh D, Chu P (2006) From itu-t g.722.1 to itu-t g.722.1 annex c: A new low-complexity 14khz bandwidth audio coding standard

# MPEG Unified Speech and Audio Coding – Bridging the Gap

**Markus Multrus, Max Neuendorf, Jérémie Lecomte, Guillaume Fuchs, Stefan Bayer, Julien Robilliard, Frederik Nagel, Stephan Wilde, Daniel Fischer, Johannes Hilpert, Nikolaus Rettelbach, Christian Helmrich, Sascha Disch, Ralf Geiger, and Bernhard Grill**

**Abstract** Speech and audio coding schemes originate from different worlds. Speech coding schemes typically assume a source model i.e. the human vocal tract. General audio coding schemes primarily rely on a sink model i.e. the human auditory system. While speech coding schemes work well for the signal class they were designed for at very low rates, they are known to fail for general audio signals even at higher rates. In contrast, general audio coders work well for any content at higher rates, but typically have limited performance especially for speech signals at very low rates. Recently the ISO/MPEG group started a standardization activity to develop a new Unified Speech and Audio Coding scheme. A state of the art AAC based general audio coder, featuring transform coding, parametric bandwidth extension and parametric stereo coding, was extended by source model coding tools. All codec modules were further improved and revised for enhanced performance in particular at very low bitrates. The new unified coding scheme outperforms dedicated speech and general audio coding schemes and bridges the gap between both worlds. This paper describes the new codec in detail and shows how the goal of consistent high quality for all signal types is reached.

## 1 Introduction

With the advent of devices which unite a multitude of functionalities, the industry has an increased demand for an audio codec which can deal equally well with all types of audio content including speech at low bitrates. In many use cases e.g. broadcasting, movies or audio books the audio content is not limited to only speech or music. Instead, a wide variety of content is to be processed. Hence, a unified audio codec that can deal equally well with all types of audio content is highly desired.

Markus Multrus (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
markus.multrus@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_33, © Springer 2011

Audio coding schemes, such as MPEG-4 High Efficiency Advanced Audio Coding (HE-AAC) [10, 19], are advantageous in that they show a high subjective quality at low bitrates for music signals. However, the spectral domain models used in such audio coding schemes do not perform well on speech signals at low bitrates.

Speech coding schemes, such as Algebraic Code Excited Linear Prediction (ACELP) [14], are well suited for representing speech at low bitrates. The time domain source-filter model of these coders closely follows the speech production process. State of the art speech coders such as the 3GPP Adaptive Multi-Rate Wideband (AMR-WB) [1, 3] perform very well for speech even at low bitrates, but show a poor quality for music. Therefore, the source-filter model of AMR-WB was extended by transform coding elements in the 3GPP AMR-WB+ [2, 16]. Still, for music signals AMR-WB+ is by far not able to provide a similar audio quality like HE-AACv2.

Addressing the obvious need for an audio codec that can code speech and music equally well, ISO/MPEG issued a Call for Proposal (CfP) on Unified Speech and Audio Coding within MPEG-D [12]. The joint contribution from Fraunhofer IIS and VoiceAge Corp. performed best in the CfP listening tests and was selected as reference model zero (RM0) at the 85th MPEG meeting in Hanover, Germany, in summer 2008 [18]. Already at that time, the system fulfilled all requirements for the new technology listed in the CfP [13]. Nevertheless, subsequent collaborative refinement based on RM0 was carried out within the MPEG Audio Subgroup until early 2011 when the technical development was essentially finished. This paper outlines the main functionalities and features of the new MPEG-D Unified Speech and Audio Coder (USAC).

## 2 State of the Art

### *2.1 General Codec Structure*

A general codec structure of common modern audio codecs is shown in Fig. 1. This scheme, which is also employed in HE-AACv2 and AMR-WB+, consists of three main components: (1) a core-coder (i.e. transform or speech coder) which provides a high quality and largely wave-form preserving representation of low frequency signal components; (2) a parametric bandwidth extension such as spectral band replication (SBR) [19], which reconstructs the high frequency band from replicated low frequency portions through the control of additional parameters; and (3) a parametric stereo extension such as “Parametric Stereo” [4, 10], which represents stereo signals with the help of a mono downmix and a corresponding set of inter-channel level, phase and correlation parameters. For low bitrates the parametric tools are able to reach much higher coding efficiency with a good quality/bitrate trade-off. At higher bitrates, when the core coder is able to handle a wider bandwidth and coding of multiple channels, the parametric tools can be selectively disabled.

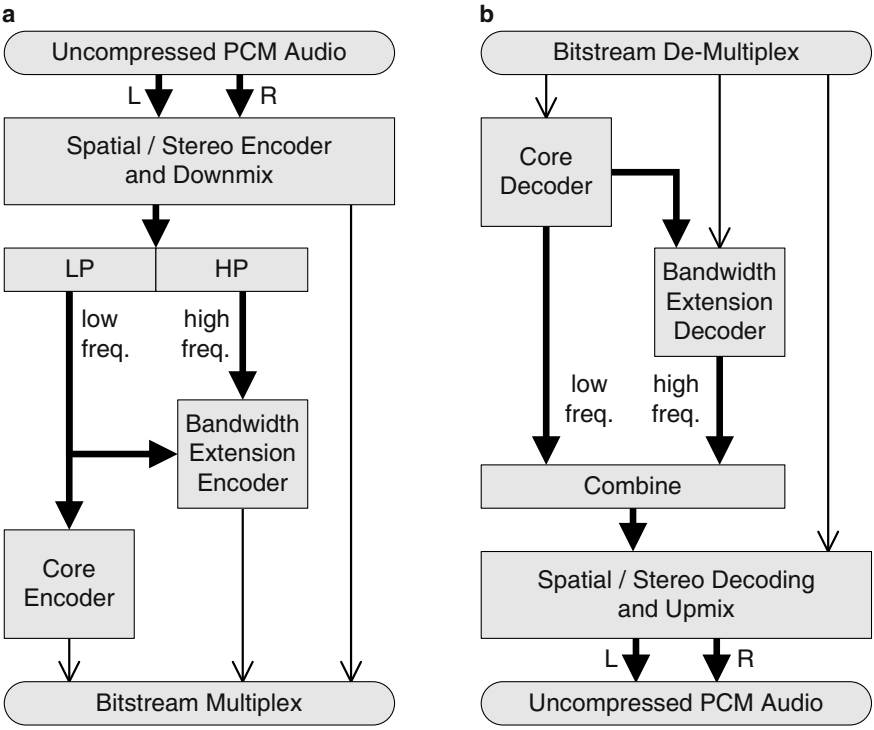



**Fig. 1** General structure of a modern audio codec with a core codec accompanied by parametric tools for coding of bandwidth extension and stereo signals. USAC closely follows this coding paradigm. (**a**) Shows the encoder. (**b**) Shows the corresponding decoder structure. *Bold arrows* indicate audio signal flow. *Thin arrows* indicate side information and control data

## *2.2 Transform Coding*

General transform coding schemes, such as AAC [10] rely on a sink model motivated by the human auditory system. By means of this psychoacoustic model, temporal and simultaneous masking is exploited for irrelevance removal. The resulting audio coding scheme is based on three main steps: (1) a time/frequency conversion; (2) a subsequent quantization stage, in which the quantization error is controlled using information from a psychoacoustic model; and (3) an encoding stage, in which the quantized spectral coefficients and corresponding side information are entropy-encoded. The result is a highly flexible coding scheme, which adapts well to all types of input signals at various operating points.

### 2.3 Speech Coding

Speech coding schemes, such as AMR-WB rely on a source model motivated by the mechanism of human speech production. Efficient speech coding schemes, such as AMR-WB, typically have three major components: (1) a short-term linear predictive coding scheme (LPC), which models the vocal tract; (2) a long-term prediction (LTP) filter, which models the periodicity in the excitation signal from the vocal chords; and (3) an innovation codebook, which encodes the non-predictive part of the speech signal. AMR-WB follows the ACELP approach which uses an algebraic representation for the innovative codebook: a short block of excitation signal is encoded as a sparse set of pulses and associated gain for the block. The pulse codebook is represented in algebraic form. The encoded parameters in a speech coder are thus: the LPC coefficients, the LTP lag and gain, and the innovative excitation. This coding scheme can provide high quality for speech signals even at low bitrates.

## 3 Technical Description

### 3.1 System Overview

USAC preserves the same overall structure of HE-AACv2 as depicted in Fig. 1. An enhanced SBR tool serves as a bandwidth extension module, while MPEG Surround 2-1-2 supplies parametric stereo coding functionality. The core coder consists of an AAC based transform coder enhanced by speech coding technology. All these components have been upgraded, as described in the following sections.

### 3.2 Description of USAC Core Coder

Figure 2 depicts the layout of the USAC core decoder. In the MPEG philosophy, the encoder is not normatively specified. Implementers are free to choose their own encoder architecture as long as it produces valid bitstreams. In that respect, USAC provides a wide range of flexibility and permits continuous performance improvement.

USAC retains all capabilities of AAC. In Fig. 2 the left signal path resembles the AAC coding scheme as described in Sect. 2.2 (“Transform Coding”). It comprises the function of entropy decoding (arithmetic decoder), inverse quantization, scaling of the spectral coefficients by the means of scale factors and inverse MDCT transform. With respect to the MDCT, all flexibility inherited from AAC regarding the choice of the transform window, such as length, shape and dynamic switching is maintained. All AAC tools for discrete stereo or multi-channel operation are included in USAC. As a consequence, USAC can be operated in a mode equivalent to AAC.

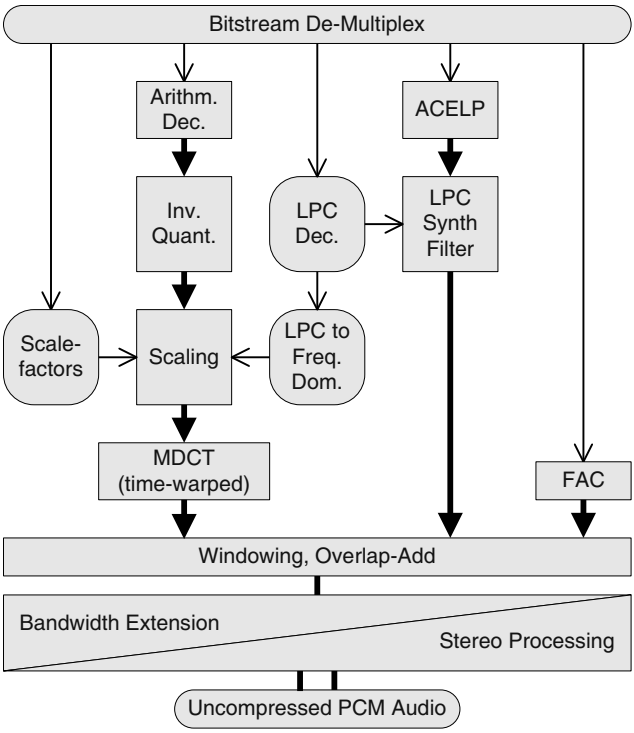



**Fig. 2** Overview of USAC core decoder modules. The main decoding path features a Modified Discrete Cosine Transform (MDCT) domain coding part with scalefactor based or LPC based noise shaping. An ACELP path provides speech coder functionality. The Forward Aliasing Cancellation (FAC) enables smooth and flawless transitions between transform coder and ACELP. Following the core decoder, bandwidth extension and stereo processing is provided. *Bold black lines* indicate audio signal flow. *Thin arrows* indicate side information and control data

In addition USAC introduces new technologies, which offer increased flexibility and enhanced efficiency. The AAC Huffman decoder was replaced by a more efficient context-adaptive arithmetic decoder. The scale factor mechanism as known from AAC can control the quantization noise shaping with a fine spectral granularity. If appropriate, it can be substituted by a Frequency Domain LPC Noise Shaping (FDNS) mechanism which consumes fewer bits. The USAC MDCT features a larger set of window lengths. The 512 and 256 MDCT block sizes complement the AAC 1024 and 128 sizes, providing a more suitable time-frequency decomposition of many signals.

The MDCT time warping mechanism and the ACELP module relieve one of AAC's most critical problems, tonal signals with varying pitch within an MDCT window. Both techniques are particularly well suited to deal with signals such as speech. As a result, USAC gains considerable perceived audio quality over AAC at lower bitrates and especially for speech signals.

## *3.3 Coding Tools and Further Enhancements*

### 3.3.1 Context Adaptive Arithmetic Coding

Entropy coding of spectral coefficients is performed by a context-adaptive arithmetic coding scheme. The cumulative frequency tables, as used for the arithmetic coding, are selected based on a context, which is deduced from the neighboring spectral coefficients. This algorithm allows a saving from 3 to more than 6% of the overall bitrate over AAC Huffman coding while showing about the same complexity requirements [6].

### 3.3.2 Noise Filling

Prior to scaling and inverse MDCT transform, a noise filling procedure is applied at the decoder side: in place of spectral coefficients that have been quantized to zero, comfort noise is injected and scaled to match a transmitted energy. This conceals spectral hole artifacts and improves perceptual quality.

### 3.3.3 LPC Coefficient Decoder

If the ACELP core is active, the LPC coefficients are utilized in the standard way employed by many speech coders. Additionally, in the USAC coder the LPC filter coefficients represent an alternative to the scale factor mechanism inherited from AAC. Like these they can convey the spectral shaping of the MDCT domain spectral coefficients. The number of coefficients and their line spectral pair (LSP) representation which is used for transmission are adopted from AMR-WB+. The application of the LPC filter coefficients to shape the spectrum are described in the next paragraph.

### 3.3.4 Frequency Domain LPC Noise Shaping

The Frequency Domain LPC tool is an additional possibility for USAC to control the quantization of its spectral coefficients. It allows a LPC-based perceptual shaping of the quantization noise in the MDCT domain, which is more efficient than transmitting scale factors especially at low bitrates. Moreover, it facilitates the transitions between ACELP and transform-based coding: the LPC-based filters of ACELP can be properly initialized without sending overhead information.

### 3.3.5 Time Warped Modified Discrete Cosine Transform (TW-MDCT)

The TW-MDCT increases the coding gain for harmonic signals with a varying fundamental frequency. This variation is compensated for, and thereby the spectral

smearing of partial tones caused by frequency modulation is reduced. The reduction is achieved by time-varying resampling within one block. A careful calculation of the resampling and an adaption of the window shapes retain the perfect reconstruction property and the constant framing of the classic MDCT [5].

### 3.3.6 Algebraic Code Excitation Linear Predictive Coder

The ACELP coder is still the state of the art speech coder and is particularly well suited to code sequences of voiced speech [14]. The key components are a time-domain based coder that mimics the pulse like excitation of the human glottis and a linear prediction filter which shapes the spectral envelope of the signal, thus modeling the human vocal tract. USAC employs the ACELP technology of AMR-WB+ [2].

### 3.3.7 ACELP/MDCT Transitions

Coding of successive MDCT and ACELP frames requires smooth and efficient handling of the transitions [15]. While MDCT frames apply a non-rectangular, overlapping windowing scheme, ACELP uses a rectangular, non-overlapping window. Moreover, MDCT frames are afflicted with a time-domain aliasing component which needs to be canceled by adjacent MDCT frames. ACELP frames do not provide a corresponding aliasing component. To allow for proper transitions between adjacent MDCT and ACELP frames, additional information for canceling the time-domain aliasing and windowing effects, called “Forward Aliasing Cancelation” (FAC) is transmitted.

### 3.3.8 Harmonic Bandwidth Extension

HE-AAC employs SBR [10, 19] for reproducing high frequency (HF) content by a spectral shift of the low frequency (LF) content. Thereby, the harmonic relations of tonal components of the signal are not maintained. As a consequence, if tonal peaks in the boundary region between LF and HF are placed in close spectral vicinity to each other, undesired beating effects can occur. These are perceived as unpleasant auditory roughness. The USAC harmonic bandwidth extension (HBE) [17] avoids such problems by stretching the LF spectrum using multiple phase vocoders. This intrinsically ensures harmonic continuation of the LF band in the HF band.

### 3.3.9 MPEG Surround 2–1–2

The USAC parametric stereo principles have been adopted from MPEG Surround (MPS) [11]. MPS is a bitrate-efficient method for coding multichannel signals (2

or more audio channels) using a core-coded downmix signal and an associated parametric representation of the human's auditory cues for spatial perception. For USAC, an additional stereo via mono (2–1–2) operation mode has been introduced. The upmixed channels are obtained by level- and phase weighting of the downmix and a decorrelated version of itself.

#### 3.3.10 Unified Stereo Coding

Due to the synthetic nature of the decorrelated signal, the perceived quality of the parametric stereo model saturates towards higher bitrates. To allow for seamless scaling towards transparent audio quality, the parametric stereo concept has been extended to waveform coding for the higher rates. Unified stereo technology serves this purpose by application of residual coding in combination with phase information.

#### 3.3.11 Complex Prediction Stereo Coding

MPEG Surround 2–1–2 and unified stereo coding employ quadrature mirror filterbanks (QMF), which are shared with the bandwidth extension tools. At very high rates where parametric coding and ACELP are typically not active, stereo coding can be performed exclusively in the MDCT domain by means of complex-valued stereo prediction [7]. This eliminates the need for QMF-based processing and thus allows for very low codec complexity.

## 4 Listening Test Results

As mentioned above, the RM0 technology met all performance requirements listed in the CfP. The RM0 performed at least as good as the better of two reference codecs (HE-AACv2 and AMR-WB+) over a large bitrate range. At the end of the collaborative standardization phase, the performance of the improved codec was benchmarked against the same state of the art speech and audio codecs at selected bitrates.

### *4.1 Listening Test Procedure and Setup*

The listening test was conducted according to MUSHRA methodology [9]. For each item, all stimuli including bandlimited anchor signals (3.5 kHz and 7 kHz) and an unprocessed hidden reference were presented to 19 experienced listeners from Fraunhofer IIS, VoiceAge Corp., Dolby and Philips. The test was carried out in ded-

icated listening rooms with reference equipment [8]. The stereo test material consisted of 24 different items of three signal categories: music, mixed speech/music, and speech. Three additional items (one per category) were presented in a preparatory training session. USAC was tested at bitrates of 16, 20 and 24 kbps against HE-AACv2 and AMR-WB+ at 16 and 24 kbps.

## *4.2 Listening Test Results*

Figure 3 shows the test results averaged per signal category and over all items. The performance of the two state of the art codecs depends on the signal type. For music HE-AACv2 is graded "Fair" to "Good" and outperforms AMR-WB+ which is rated "Bad" to "Fair". For speech AMR-WB+ is graded "Fair" to "Good" and outperforms HE-AACv2 which is rated "Bad" to "Fair". For mixed speech/music signals both reference codecs show similar performance, ranging in between their speech and music results. Due to the unbalanced behavior the overall quality of both codecs is only "Fair" at the tested bitrates.

In contrast, USAC shows a balanced performance for any signal type. It is graded "Good" for all signal categories over all bitrates. For each test point it outperforms the better of the two reference codecs at all signal classes by 5 to 20 MUSHRA points. Due to its consistently good performance the overall quality is also graded "Good" at the tested bitrates.

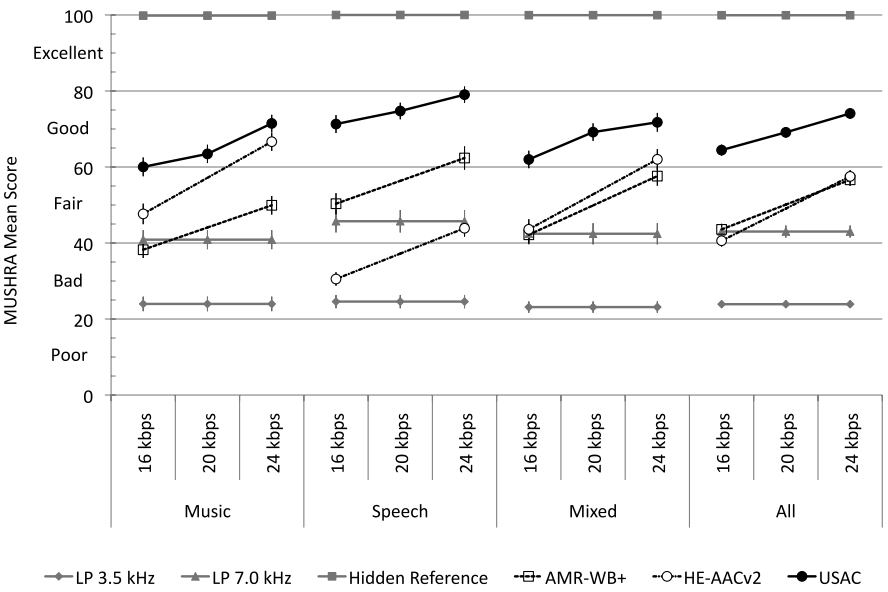



**Fig. 3** Listening test results for music, speech, mixed speech/music and overall results. Mean values and according confidence intervals (95% level of significance, Student-*t* distribution)

## 5 Conclusions

In the past 15 years the AAC family has always represented the state of the art in audio coding. The basic AAC codec, standardized in 1996, and its advancements High Efficiency AAC Version 1 in 2002 and Version 2 in 2005, were always able to set new reference quality standards at even lower bitrates. USAC now further extends the working envelope by evolutionarily upgrading the AAC codec's capabilities for music and, at the same time, considerably improving its performance for speech with and without background signals. While for the current AAC family speech represents one of the most critical signals, with USAC speech signals are now coded at a quality level better than the best available speech codecs even at very low bitrates. The speech quality of USAC surpasses AMR-WB+ which provided the basic modules for the speech improvements in USAC and which has been widely considered to offer the best low-rate speech performance to date. With USAC it is possible to implement services in the bitrate range of 8 to 24 kbps in mono or stereo. Until now, such systems have had to revert to speech codecs, which in turn failed for non-speech material. USAC provides a general audio coding scheme which, above all, sets a new reference for low rate coding of speech signals. Even for the new wireless transmission technologies like 3GPP Long Term Evolution (LTE) or more advanced broadcasting systems bandwidth will remain expensive in many application areas. USAC will allow to packing more audio channels into the same transmission bandwidth, and therefore provide significant cost saving for many applications.

## Acknowledgements

The authors would like to thank VoiceAge Corporation for their long-standing commitment during the preparation of the CfP submission and productive and successful cooperation during the collaborative phase of the MPEG standardization. Furthermore we would like to thank Dolby Laboratories, Philips, Samsung, and Panasonic for their collaboration and support in preparing joint core experiments. Additionally, we would like to thank the numerous further companies which have contributed to the USAC standard by supplying valuable technology and by participating in constructive discussions at the MPEG meetings. Finally, the authors would like to thank VoiceAge Corporation, Dolby Laboratories and Philips for sharing listening test results that could be used in this publication.

## References

1. 3GPP (2002) Adaptive Multi-Rate – Wideband (AMR-WB) speech codec; General description. 3GPP TS 26.171
2. 3GPP (2004) Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Transcoding functions. 3GPP TS 26.290

3. Bessette B, Salami R, Lefebvre R, Jelinek M, Rotola-Pukkila J, Vainio J, Mikkola H, Jarvinen K (2002) The adaptive multirate wideband speech codec (AMR-WB). IEEE Transactions on Speech and Audio Processing 10(8):620–636
4. Breebaart J, Faller C (2007) Spatial Audio Processing: MPEG Surround and Other Applications. John Wiley & Sons Ltd, West Sussex, England
5. Edler B, Disch S, Bayer S, Fuchs G, Geiger R (2009) A time-warped MDCT approach to speech transform coding. In 126th Convention, 7710. Audio Eng Soc, Munich
6. Fuchs G, Subbaraman V, Multrus M (2011) Efficient Context Adaptive Entropy Coding for Real-Time Applications. In Proc. IEEE ICASSP 2011, pp 493–496
7. Helmrich CR, Carlsson P, Disch S, Edler B, Hilpert J, Neusinger M, Purnhagen H, Rettelbach N, Robilliard J, Villemoes L (2011) Efficient Transform Coding of Two-Channel Audio Signals by Means of Complex-Valued Stereo Prediction. In Proc. IEEE ICASSP 2011, pp 497–500
8. International Telecommunication Union (1994) Methods for the Subjective Assessment of Small Impairments in Audio Systems including Multichannel Sound Systems. Recommendation ITU-R BS.1116-1, Geneva, Switzerland
9. International Telecommunication Union (2001) Method for the subjective assessment of intermediate sound quality (MUSHRA). ITU-R, Recommendation BS 1543-1, Geneva, Switzerland
10. ISO/IEC 14496-3:2009 (2009) Coding of Audio-Visual Objects, Part 3: Audio
11. ISO/IEC 23003-1:2007 (2007) MPEG-D (MPEG audio technologies), Part 1: MPEG Surround
12. ISO/IEC JTC1/SC29/WG11 (2007) Call for Proposals on Unified Speech and Audio Coding. Shenzhen, China. MPEG2007/N9519
13. ISO/IEC JTC1/SC29/WG11 (2008) MPEG press release. Hannover, Germany. MPEG2008/N9965
14. Laflamme C, Adoul JP, Salami R, Morissette S, Mabilleau P (1991) 16 kbps wideband speech coding technique based on algebraic celp. In IEEE Int Conf on Acoustics, Speech and Signal Proc, vol 1, pp 13–16, Inst of Electrical and Electronics Eng, Toronto, Ontario, Canada
15. Lecomte J, Gournay P, Geiger R, Bessette B, Neuendorf M (2009) Efficient cross-fade windows for transitions between lpc-based and non-lpc based audio coding. In 126th Convention, 7712. Audio Eng Soc, Munich
16. Makinen J, Bessette B, Bruhn S, Ojala P, Salami R, Taleb A (2005) AMR-WB+: a new audio coding standard for 3rd generation mobile audio services. In Proc IEEE ICASSP'05, 2:1109–1112
17. Nagel F, Disch S (2009) A harmonic bandwidth extension method for audio codecs. In IEEE Int Conf on Acoustics, Speech and Signal Proc, pp 145–148
18. Neuendorf M, Gournay P, Multrus M, Lecomte J, Bessette B, Geiger R, Bayer S, Fuchs G, Hilpert J, Rettelbach N, Nagel F, Robilliard J, Salami R, Schuller G, Lefebvre R, Grill B (2009) A novel scheme for low bitrate unified speech and audio coding – MPEG RM0. In 126th Convention, 7713. Audio Eng Soc, Munich
19. Wolters M, Kjörling K, Homm D, Purnhagen H (2003) A closer look into MPEG-4 High Efficiency AAC. In 115th Convention. Audio Eng Soc, New York, NY, USA. Preprint 5871

# A Dedicated Decorrelator for Parametric Spatial Coding of Applause-Like Audio Signals

**Sascha Disch and Achim Kuntz**

**Abstract** Parametric spatial coding of multichannel content has acquired widespread use in today's audio codecs to enable efficient operation in the medium bitrate range. However, applause signals are still challenging to code with good perceptual quality using these techniques. To obtain a sufficiently low side information bitrate, the spatial parameter update rate is usually chosen in the range of several tens of milliseconds which is too low to restore a convincing spatial listener envelopment of applause-like signals. Moreover, state of the art decorrelators that are mandatory in parametric multichannel decoders inevitably deteriorate the signal quality of transient sound events, like handclaps, by dispersion. Therefore, we propose a novel decorrelation and parametrization technique for efficient coding of applause sounds: transient events are separated from the core decoder output and a dedicated decorrelator algorithm distributes the transients in the spatial image according to parametric guiding information transmitted in the bitstream. Listening tests show a substantial improvement in subjective quality compared to standard methods.

## 1 Introduction

Contemporary audio codecs often utilize parametric spatial coding techniques to efficiently compress a multichannel signal at medium bitrates around 32 kbps. These techniques encode and transmit a downmix of the multichannel signal and additional parametric side information that guides the upmix at the decoder side. Prominent examples for such audio codecs are High Efficiency-Advanced Audio Coding (HE-AAC) [5] including the parametric stereo (PS) tool [14], MP3 Surround [6] and MPEG Surround (MPS) [7, 8]. The latter is a versatile multichannel format that can be combined with different so-called core coders for compression of the down-

Sascha Disch (✉)
Fraunhofer IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany
sascha.disch@iis.fraunhofer.de

A. Heuberger, *Microelectronic Systems*.
DOI 10.1007/978-3-642-23070-7_34, © Springer 2011

mix [2]. In parametric spatial audio decoders, decorrelators are needed in the upmix process to restore specific correlation properties between two or more signals that are reconstructed from one or several downmix signals [13]. The application of decorrelators significantly improves the perceptual quality of the output signal when compared to previous approaches, e.g. intensity stereo [4]. Specifically, the use of decorrelators enables the proper synthesis of spatial sound with a wide sound image, several concurrent sound objects and/or ambience [1, 3]. However, decorrelators are also known to introduce artifacts like changes in temporal signal structure, timbre, etc. [11]. Since the kind and severity of decorrelator artifacts strongly depend on the processed signal, we propose to apply in a signal-adaptive manner dedicated decorrelators that are tailored to the specific characteristics of the input signal. Therefore, we present a decorrelator especially suited for reproduction of applause signals.

## 2 State of the Art

### *2.1 Parametric Spatial Audio Coding*

At the encoder, parametric spatial coding calculates a downmix signal by applying a downmix matrix $\mathbf{H}_{\mathrm{DMX}}$ to the input signals. For stereo input signals this reads as

$$\begin{bmatrix} M \\ res \end{bmatrix} = \mathbf{H}_{\mathrm{DMX}} \begin{bmatrix} L_{\mathrm{in}} \\ R_{\mathrm{in}} \end{bmatrix} . \tag{1}$$

Instead of transmitting both the downmix signal $M$ and the 'residual' signal *res*, only $M$ and the spatial properties of the input channels are transmitted. Alternatively, a bandlimited residual signal can be transmitted. In that case, only those spectral regions that are not contained in the residual signal are reconstructed by parametric spatial coding.

Since the spatial parameters, e.g. the inter-channel level differences (ICLDs) and inter-channel cross-correlations (ICCs), only have to be transmitted at the relatively low time-frequency resolution of the human auditory perception of spatial phenomena, a considerable bitrate saving is obtained.

At the decoder, the downmix signal is fed to an upmixing structure that calculates the stereo or multichannel output based on the parametric spatial data. Figure 1 shows the structure of a mono-to-stereo decoder. From the downmixed input signal $M$ a single decorrelator generates a decorrelated signal $D$ that is fed into the upmixing matrix $\mathbf{H}$ along with signal $M$. The coefficients of $\mathbf{H}$ are determined by the parametric spatial side information generated in the encoder.

The decorrelator allows for synthesizing output signals with adapted cross-correlation properties. The characteristics of commonly used decorrelators are described in the following section. Additionally, we identify typical problems caused by such decorrelators when applied to applause-like signals.

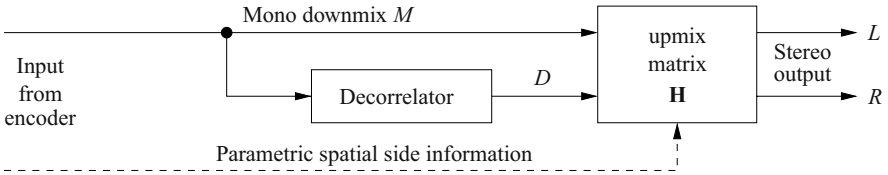



**Fig. 1** Typical structure of a mono to stereo upmixer used in a parametric spatial audio decoder. Audio signal paths are depicted as *solid lines*, the parametric data path is indicated as a *dashed line*

### *2.2 Decorrelation*

The decorrelators employed in many spatial audio decoders (including MPS) contain IIR allpass structures. These act as artificial late reverb generators and are consequently well suited for generating homogeneous, smooth, noise-like, immersive sounds. However, there are examples of sound fields with a non-homogeneous spatio-temporal structure that are still immersing the listener: one prominent example are applause-like sound fields. Applause signal recordings can be modeled as a composition of many single, distinct nearby claps that are temporally separated by a few milliseconds and a superimposed noise-like ambience originating from very dense far-off claps. Such signals create listener-envelopment not only by homogeneous noise-like fields, but also by rather dense sequences of single claps from different directions. Hence, the non-homogeneous component of applause sound fields may be characterized by a spatially distributed mixture of transients. Due to their reverb-like behavior, IIR allpass decorrelators are incapable of generating immersive sound fields with the characteristics e.g. of applause. Instead, they tend to temporally smear transients which results in a noise-like immersive sound field lacking the distinctive, original spatio-temporal structure.

## 3 Applause Decorrelator

### *3.1 Processing Principle*

A good decorrelator for applause signals should therefore

1. create an output signal that is sufficiently decorrelated from the input.
2. not alter the temporal structure of the single claps/transients.
3. reinstate a spatial distribution of the transient signal components similar to the original.
4. allow for bitrate vs. quality trade-offs (e.g. from random spatial transient distribution at low bitrates to near-transparency at high bitrates).
5. exhibit a low computational complexity.

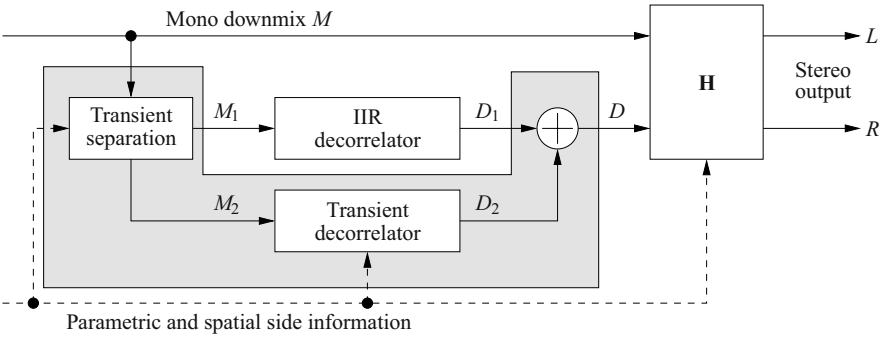



**Fig. 2** Dedicated transient decorrelator within a mono-to-stereo upmix system. The processing blocks highlighted by *gray shading* perform the transient handling

We propose a decorrelating signal processing structure that allows for a higher quality decorrelation of applause-like audio signals through the inclusion of a dedicated transient decorrelator. Its application in the upmix-process of spatial audio coders is depicted in Fig. 2. It involves three processing steps:

Firstly, a separation of the input signals in two components $M_i$ is performed: $M_2$ contains the transients, $M_1 = M - M_2$ the remaining (non-transient) part. The non-transient component $M_1$ is processed like in conventional systems, e.g. by feeding $M_1$ into an IIR allpass decorrelator.

Secondly, $M_2$ is fed to a dedicated transient decorrelator. The decorrelation of the transient stream is carried out by applying phase terms at a high temporal resolution:

$$D_2[n,k] = M_2[n,k] \cdot e^{j\Delta\varphi[n]} \ , \tag{2}$$

where $n$ and $k$ are the time and frequency indices of the short-time frequency domain signal representations $D_2[n,k]$ and $M_2[n,k]$, respectively. $\Delta\varphi$ ideally reflects the phase difference between downmix and residual. Hence, the transient residuals are effectively replaced by phase modified transients from the downmix. Applying the phase information inherently results in a panning of the transients since the spatial parameters (ICC, ICLD) are also effective for the transient signals through the subsequent upmix. For example, the ICC controls the width of the rendered transient distribution.

Lastly, the outputs of both the conventional decorrelator and the transient decorrelator are added to form the decorrelated signal $D = D_1 + D_2$ that is passed to the upmix-process.

The signal separation into the transient and non-transient component can be controlled by parameters that are either generated in the encoder or decoder. Similarly, the transient decorrelator can be steered by parameters (phase terms $\Delta\varphi[n]$) that are either transmitted from the encoder or generated in the decoder. Three variants of obtaining $\Delta\varphi[n]$ are discussed in the following.

### *3.2 Variants*

There are several possibilities to obtain the phase information for the transient decorrelator.

1. Guided operation: the phase difference between residual and downmix signal is calculated in the encoder.
2. Unguided operation: the phase terms are derived from a pseudo-random process in the decoder.
3. Residual guided operation: phase estimates are obtained in the decoder from non-fullband residuals.

The first method restores the actual fine-grain signal properties at the expense of additional side information, whereas the second method aims at only preserving the long term integral signal properties. The latter might already yield perceptually satisfactory results in low bitrate scenarios, since a convincing envelopment is much more important than the restoration of the precise location of single clapping events. The third method exploits phase information that is available in the decoder if residual signals are transmitted for a frequency range. At the decoder, the phase information is to be applied to transient signal segments only.

For guided operation, the transmission of the necessary phase information can thus be limited to segments, for which transients have been detected at the encoder. Moreover, transient sound components often exhibit little frequency dependency of the ICLDs, ICCs, and transient phase values. The parameters can therefore be transmitted as broadband values to reduce the required parameter bitrate. Additionally the transient segment positions within an audio frame need to be transmitted to the decoder efficiently.

For unguided operation, the simplest method is to generate phase values with a uniform random distribution in the range $[-180°, 180°]$. More advanced methods include the measurement of the statistical properties of the phase distribution in the encoder. These properties are coded and transmitted at a low time resolution. In the decoder, random phase values are generated subject to the statistical properties.

For the residual guided operation, residual signals have to be accessible at the decoder for at least some frequency bands. The phase difference between the downmix and the residual signals is measured in these frequency bands and is extrapolated to bands where no residuals are available. Thus, no phase terms have to be transmitted to the decoder. The correctness of the extrapolated phase values depends on the consistency of the measured phase difference along the frequency axis. For transient signals, a high consistency is usually encountered.

## 4 Application

The upcoming MPEG standard for Unified Speech and Audio Coding (USAC) [12] includes a decorrelating structure named the Transient Steering Decorrelator (TSD).

The TSD consists of a fully encoder-steered variant of our newly proposed decorrelator: transient positions within each time frame are losslessly encoded and transmitted alongside with their corresponding broadband phase values that are quantized at 3 bits/value. For example, USAC may be operated at a sampling frequency $f_s = 44.1$ kHz in a 32 kbps total bitrate configuration using its parametric stereo mode. Parametric stereo for USAC operates on the 64 subbands of a Quadrature Mirror Filterbank (QMF). The maximum time resolution that can be obtained in the subband domain is $64/f_s \approx 1.4$ ms. This resolution is fully utilized by TSD. In contrast, one spatial parametric frame amounts to a much longer duration, specifically 32 subband samples $2048/f_s \approx 46$ ms. Up to 2 spatial parameter sets can be transmitted within a frame.

## 5 Results

During MPEG standardization, listening tests have been conducted evaluating USAC with and without TSD at a 32 kbps stereo configuration.

The listening test setup was based on a standard MUltiple Stimuli with Hidden Reference and Anchor (MUSHRA) test according to the ITU recommendation BS.1534 [10]. For each item, the test presents the labeled reference and all test conditions, alongside with the hidden reference and a hidden 3.5 kHz cut-off lowpass filtered anchor (lp35) to the listener. All conditions are time-aligned. Hidden reference and lower anchor are included to check the listener's reliability.

Only one person at a time is subject to the test. For each listener, the sequence of test items is randomly ordered and the order of the conditions of each item is randomized as well. Individual switching between conditions while listening is permitted and so is setting a loop on arbitrarily selected partitions of the item, as suggested in the BS.1116-1 recommendation [9] and applicable to MUSHRA tests as well. There is no limit of the number of repetitions the test subjects may listen to before finally rating the test item and proceeding to the next. This allows for a very close comparison and thorough examination of the different conditions. The perceptual quality of the items is rated on a scale ranging from 100 points to 0 points, labeled "excellent" through "good", "fair", "poor" and down to "bad".

Nine expert listeners participated in the tests. The test set comprised five applause test items each of which exhibiting different spatio-temporal properties (different clap densities, different direct-to-diffuse ratios). The tests were performed in a dedicated listening test environment using high-quality electrostatic headphones.

Figures 3 and 4 show the TSD evaluation listening test results as mean MUSHRA scores along with their 95% confidence intervals, assuming a Student's $t$-distribution. The absolute scores in Fig. 3 of USAC with TSD consistently show a higher mean score than without TSD for all items. No item was degraded versus the plain USAC condition. The difference scores for USAC+TSD with respect to plain USAC are plotted in Fig. 4. Here, we observe a statistically significant improvement for all items since none of the confidence intervals include zero.

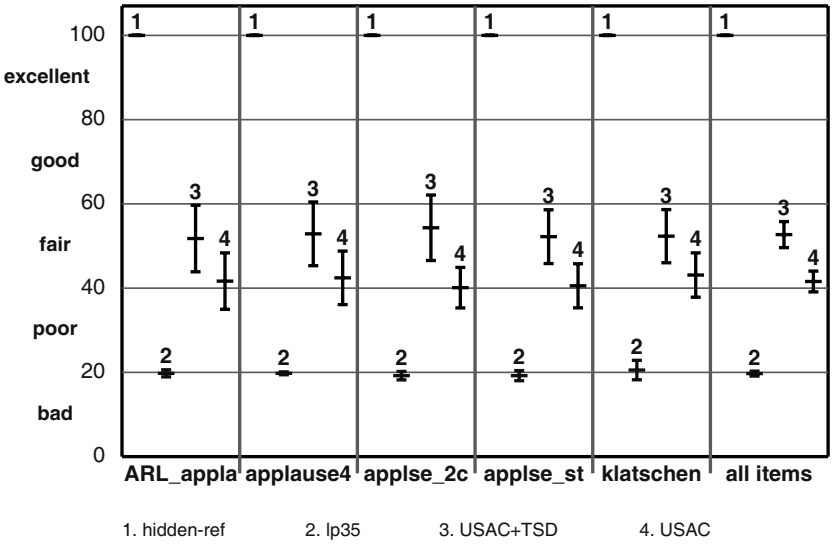



**Fig. 3** Listening test results for 32 kbps stereo coding. Average scores and the 95% confidence intervals are shown

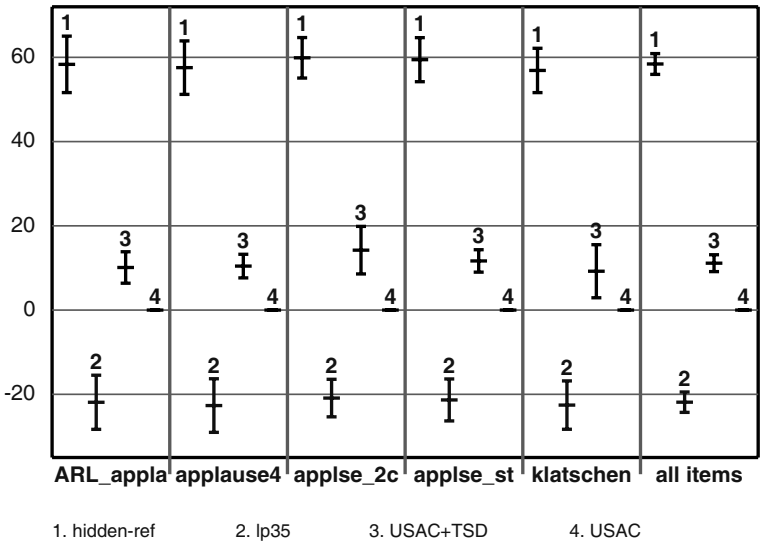



**Fig. 4** Listening test results for 32 kbps stereo coding. Difference scores relative to "4. USAC" (coding without the proposed decorrelator structure) and the 95% confidence intervals are shown

## 6 Conclusion

Applause-like signals are known to be critical items for parametric spatial audio coders. In state of the art spatial decoders, an insufficient time resolution of the

spatial parameters due to bitrate constraints and the usage of IIR reverb-type allpass structures for decorrelation severely compromise the perceptual quality of transient mixtures such as applause. To remedy these problems, we have presented a novel decorrelating structure that adds a second decorrelator path in parallel to the existing one. This parallel decorrelator is tailored to the special characteristics of transient mixtures. The benefits of the proposed structure have been proven in subjective listening tests. Our dedicated decorrelator for parametric spatial coding of applause-like audio signals in its fully encoder-steered variant has already been adopted into the upcoming MPEG Unified Speech and Audio Coding (USAC) standard as the Transient Steering Decorrelator (TSD) tool.

## Acknowledgements

The authors like to thank the following persons who contributed to development and MPEG standardization of TSD (in alphabetical order): Tom Bäckström, Jürgen Herre, Johannes Hilpert, Markus Multrus, Max Neuendorf, Julien Robilliard, Christian Uhle.

## References

1. Baumgarte F, Faller C (2003) Binaural cue coding-part i: Psychoacoustic fundamentals and design principles. Speech and Audio Processing, IEEE Transactions on 11(6):509–519
2. Breebaart J, Faller C (2007) Spatial Audio Processing: MPEG Surround and Other Applications. John Wiley & Sons Ltd, West Sussex, England
3. Faller C, Baumgarte F (2003) Binaural cue coding-part ii: Schemes and applications. Speech and Audio Processing, IEEE Transactions on 11(6):520–531
4. Herre J, Brandenburg K, Lederer D (1994) Intensity stereo coding. In 96th Convention. Audio Eng Soc
5. Herre J, Dietz M (2008) Mpeg-4 high-efficiency aac coding [standards in a nutshell]. Signal Processing Magazine, IEEE 25(3):137–142
6. Herre J, Faller C, Ertel C, Hilpert J, Hoelzer A, Spenger C (2004) Mp3 surround: Efficient and compatible coding of multi-channel audio. In 116th Convention Audio Eng Soc. Preprint 6049
7. Herre J, Kjoerling K, Jeroen B, Faller C, Disch S, Purnhagen H, Koppens J, Hilpert J, Roeden J, Oomen W, Linzmeier K, Chong KS (2007) Mpeg surround; the iso/mpeg standard for efficient and compatible multi-channel audio coding. In 122nd Convention Audio Eng Soc
8. Hilpert J, Disch S (2009) The mpeg surround audio coding standard [standards in a nutshell]. Signal Processing Magazine, IEEE 26(1):148–152
9. ITU-R BS.1116-1 (1997) Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems. ITU, Geneva, Switzerland
10. ITU-R BS.1534-1 (2003) Method for the subjective assessment of intermediate quality levels of coding systems. ITU, Geneva, Switzerland
11. Laitinen MV, Kuech F, Disch S, Pulkki V (2011) Reproducing applause-type signals with directional audio coding. J Audio Eng Soc 59(1/2):29–43

12. Neuendorf M, Gournay P, Multrus M, Lecomte J, Bessette B, Geiger R, Bayer S, Fuchs G, Hilpert J, Rettelbach N, Salami R, Schuller G, Lefebvre R, Grill B (2009) Unified speech and audio coding scheme for high quality at low bitrates. In IEEE Int Conf on Acoustics, Speech and Signal Proc Inst of Electrical and Electronics Eng, Taipei, Taiwan
13. Purnhagen H, Engdegard J, Roden J, Liljeryd L (2004) Synthetic ambience in parametric stereo coding. In 116th Convention Audio Eng Soc
14. Schuijers E, Breebaart J, Purnhagen H, Engdegard J (2004) Low complexity parametric stereo coding. In 116th Convention Audio Eng Soc