

MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE

<http://www.comsoc.org/~mmc>

MMTC Communications - Frontiers

Vol. 11, No. 1, January 2016

CONTENTS

Message from MMTC Chair.....	5
SPECIAL ISSUE ON MULTIMEDIA COMMUNICATIONS in TERAHERTZ and WIRELESS OPTICAL NETWORKS.....	6
<i>Guest Editors: Tuncer Baykas¹ and Thomas Kürner².....</i>	<i>6</i>
<i>¹Istanbul Medipol Univeristy, Turkey, ²Technische Universität Braunschweig, Germany;</i>	<i>6</i>
<i>tbaykas@medipol.edu.tr; kuerner@ifn.ing.tu-bs.de</i>	<i>6</i>
TERAPAN – Towards a 100 Gbit/s Point-to-Point Link with Electronic Beam Steering Operating at 300 GHz	8
<i>Sebastian Rey¹, Iulia Dan², Thomas Merkle³, Axel Tessmann³, Ingmar Kallfass², Thomas Kürner¹</i>	<i>8</i>
<i>¹Technische Universität Braunschweig, Germany, {rey kuerner}@ifn.ing.tu-bs.de</i>	<i>8</i>
<i>²University of Stuttgart, Germany, {iulia.dan ingmar.kallfass}@ilh.uni-stuttgart.de</i>	<i>8</i>
<i>³Fraunhofer Institute for Applied Solid State Physics, Germany, {thomas.merkle axel.tessmann}@iaf.fraunhofer.de</i>	<i>8</i>
iBROW – Innovative ultra-BROadband ubiquitous Wireless communications through terahertz transceivers	12
<i>Luis Pessoa¹, Bile Peng², Jue Wang³, Loïc Sanchez⁴, Abdullah Al-Khalidi³, Edward Wasige³, Amy Liu⁵, Horacio Cantu⁶, Ian Oxtoby⁷, Bruce Napier⁸, Anthony Kelly³, José Figueiredo⁹, Mohand Achouche¹⁰, Henrique Salgado¹, Thomas Kürner², Moz Fisher⁴, Andreas Pascht¹¹ and Wolfgang Tempel¹¹</i>	<i>12</i>
<i>¹INESC TEC, Portugal; ²TU Braunschweig, Germany; ³University of Glasgow, UK; ⁴CEA-LETI, France; ⁵IQE Silicon, UK; ⁶CST Global, UK; ⁷Opticap, UK; ⁸Vivid, Germany; ⁹University of Algarve, Portugal; ¹⁰III-V Lab, France, ¹¹Nokia Bell Labs, Germany.....</i>	<i>12</i>
<i>luis.m.pessoa@inesctec.pt; edward.wasige@glasgow.ac.uk</i>	<i>12</i>
Terahertz wireless communication activities in China.....	18
<i>Guangkai Li, Bo Ai, Ke Guan, Zhangdui Zhong.....</i>	<i>18</i>
<i>State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, 100044, Beijing, China</i>	<i>18</i>
<i>{gkli.rcs, boai, kguan, zhdzhong}@bjtu.edu.cn</i>	<i>18</i>

IEEE P802.15.3d – Towards the First Wireless Communications Standard

Operating in the 300 GHz Band	23
<i>Thomas Kürner¹, Iwao Hosako², Yumi Ueda², Alexander Fricke¹, Sebastian Rey¹, Akifumi Kasamatsu², Norihiko Sekine²</i>	23
¹ Technische Universität Braunschweig, Germany; ² NICT, Japan;	23
<i>{kuerner fricke rey}@ifn.ing.tu-bs.de; {hosako ueda kasa nsekine}@nict.go.jp</i>	23
Free Space Optical Communications:	27
Fundamentals, Applications and Recent Achievements	27
<i>Murat Uysal</i>	27
<i>Ozyegin University, Istanbul, Turkey, 34794</i>	27
<i>murat.uysal@ozyegin.edu.tr</i>	27
Research on Visible Light Communication in Korea	32
<i>Nam-Tuan Le and Yeong Min Jang</i>	32
<i>Kookmin University</i>	32
<i>Seoul, Korea</i>	32
<i>namtuanlnt@gmail.com, yjang@kookmin.ac.kr</i>	32
Let there be Light Again!	36
An Amendment to IEEE 802 Visible Light Standard is in Progress	36
<i>Tunçer Baykaş¹, Morteza Soltani¹, Mohamed Siala², Trang Nguyen³, Yeong Min Jang³</i>	36
¹ Istanbul Medipol University, Turkey, ² SUP'Com,tunusia, ³ Kookmin University, Korea	36
<i>tbaykas@medipol.edu.tr</i>	36
LiFi – Technological and Commercial Considerations	42
<i>Nikola Serafimovski¹, Dobroslav Tsonev¹, Harald Burchardt¹, Stefan Videv¹, Harald Haas^{1,2}</i>	42
¹ pureLiFi Ltd, ² The Li-Fi R&D Center University of Edinburgh	42
<i>Edinburgh, UK</i>	42
<i>{nikola.serafimovski, dobroslav.tsonev, harald.burcahrdt, harald.haas}@purelifi.com</i>	42
<i>{s.videv, h.haas}@ed.ac.uk</i>	42
SPECIAL ISSUE ON ULTRA-HIGH DEFINITION VIDEO COMMUNICATIONS	48
<i>Guest Editors: Pedro Amado Assuncao¹ and Rahul Vanam²</i>	
¹ Instituto de Telecomunicações/IPLeiria, Portugal; ² InterDigital Communications, Inc., USA, <i>amado@co.it.pt, rahul.vanam@interdigital.com</i>	
Recent advances in UHD video coding technology: High Dynamic Range and Wide Color Gamut	50
<i>Dragorad Milovanovic and Dragan Kukolj</i>	50
<i>Faculty of Technical Sciences, University of Novi Sad, Republic of Serbia</i>	50
<i>RT-RK Institute of Computer based Systems, Republic of Serbia</i>	50
<i>dragoam@gmail.com, dragan.kukolj@rt-rk.uns.ac.rs</i>	50
Analytical evaluation of high frame-rate video signal for encoder design	56
<i>Yukihiro Bandoh, Seishi Takamura and Atsushi Shimizu</i>	56
<i>Nippon Telegraph and Telephone Corporation, Japan</i>	56
<i>{bandou.yukihiro, takamura.seishi, shimizu.atsushi}@lab.ntt.co.jp</i>	56

From HD to UHD video: implications for embedded systems.....	61
implementations of software-based HEVC video encoders	61
Eduarda Monteiro ¹ , Guilherme Corrêa ² , Luis A. da Silva Cruz ³ , and Sergio Bampi ¹	61
¹ Universidade Federal do Rio Grande do Sul, ² Universidade Federal de Pelotas, Brazil	61
³ Instituto de Telecomunicações/Universidade de Coimbra-DEEC, Portugal	61
ermonteiro@inf.ufrgs.br, gcorrea@inf.ufpel.edu.br, lcruz@deec.uc.pt, bampi@in.ufrgs.br	61
Contouring artefacts prevention in compressed UHD video sequences:.....	67
tools and analysis of their performance	67
André Seixas Dias, Matteo Naccari and Marta Mrak	67
British Broadcasting Corporation - R&D Department, London, UK	67
Andre.SeixasDias@bbc.co.uk, Matteo.Naccari@bbc.co.uk, marta.mrak@bbc.co.uk	67
DASH Segment Scheduling for Scalable UHD Video Streaming	73
Eduardo Asbun ¹ , Ahmed Hamza ² , Srinivas Gudumasu ³ , and Yong He ¹	73
¹ InterDigital Communications, Inc., U.S.A., ² Simon Fraser University, Canada, ³ Aricent Technologies Ltd., U.S.A.	73
{ed.asbun, yong.he}@interdigital.com, srinivas.gudumasu@aricent.com, aah10@sfu.ca	73
UHD Video Retargeting based on Visual Attention Models and Temporal Consistency	79
Lino Ferreira ^{1,2,3} and Luís A. da Silva Cruz ^{1,2} and Pedro Assuncao ^{1,3}	79
¹ Instituto de Telecomunicações, Portugal	79
² Universidade de Coimbra/DEEC, ³ Intituto Politecnico de Leiria/ESTG, Portugal	79
lino.ferreira@ipleiria.pt; {lcruz, amado}@co.it.pt	79
MMTC OFFICERS (Term 2014 — 2016)	85

Message from MMTC Chair

Dear MMTC colleagues:

How time flies by! It is the beginning of 2016 and the current executive team for MMTC are proceeding toward the last quarter of our term. It has been a vibrating journey for the team. In the process we have achieved huge successes and come across obstacles too. In this letter, it is with my greatest pleasure to reflect on some lessons and call for volunteers for future MMTC leadership.

First, the few flagship activities pioneered by MMTC have been highly touted by the research community. Our signature E-letter/R-letter have been so successful that the authors consider them as a regular publication venue. Indeed, the citations to the articles published on the MMTC E-letter have been rising over the years. Bravo to the leadership and the volunteer team who have made all of us proud! Our interest group framework has also become an iconic practice that is followed by several other TCs in IEEE ComSoc.

Second, suggested by IEEE ComSoc Publication VP, we are working on further positioning the MMTC E-Letter as a community communication platform. Specifically, we are discussing with IEEE ComSoc to finalize a better template and format that would clearly demonstrate its designed purpose. With this new development, we expect that its reach will be broadened. Therefore, I would like to call for more contributions to the E-letter and more recommendations to the R-Letter.

Third, it is time to plan for our next team of executive leadership for MMTC. Our official call for nominations will start in March 2016 and the selection will be confirmed in June 2016 at ICC 2016. The selection committee is made up of previous MMTC chairs (excluding the outgoing chair). I would like to personally extend an invitation to all the MMTC members to consider the opportunity of serving our community. The more active members volunteer, the better the MMTC will be!

Finally, I would like to thank all the existing IG chairs and other volunteers for the work they have already done and will be doing for the success of MMTC and hope that any of you will find the proper IG of interest to get involved in our community!



Yonggang Wen
Chair, Multimedia Communications TC of IEEE ComSoc

***SPECIAL ISSUE ON
MULTIMEDIA COMMUNICATIONS in TERAHERTZ and WIRELESS OPTICAL
NETWORKS***

Guest Editors: Tuncer Baykas¹ Thomas Kürner²

¹*Istanbul Medipol Univeristy, Turkey,* ²*Technische Universität Braunschweig, Germany;*

tbaykas@medipol.edu.tr; kuerner@ifn.ing.tu-bs.de

The thirst for multimedia consumption almost drained traditional spectrum bands used for cellular and local area network cellular bands. Starting with the millimeter wave bands, researchers enable new technologies at higher frequencies for faster data rates. This special issue of E-Letter focuses on the recent progresses of two of those emerging technologies. The first technology, terahertz communication aims to utilize bands around 300GHz to achieve multimedia wireless communications with data rates of 100 Gbit/s and beyond. Whereas the second technology of this special issue, the Optical wireless communication (OWC), uses optical carriers, i.e., visible, infrared, or ultraviolet bands, for wireless transmission. Various research groups all around the globe are currently working in both areas, which have recently also attracted the interest of the industry including the launch of dedicated standardization activities.

It is the great honor of the editorial team to have eight articles, from academia and industry laboratories and standardization bodies.

Terahertz Communication Networks

In this special issue of the E-Letter four THz communication networks related papers from research groups in Europe, China and Japan are describing the ongoing activities.

The first article titled “TERAPAN – Towards a 100 Gbit/s Point-to-Point Link with Electronic Beam Steering Operating at 300 GHz” by Rey et al describes the preliminary outcome of a German collaboration research project, which is aiming at demonstrating a 300 GHz wireless communications system with a targeted data rate of 100 Gbit/s. The approach is an all-electronic one based on 35 nm mHEMT technology. The interim results that have been demonstrated so far are a data rate of 64 Gbit/s in conjunction with mechanical beam steering.

Another technological approach is described in the second article “iBROW- Innovative ultra-BROadband ubiquitous Wireless communications through terahertz transceivers” by Pessoa et al., where a Resonant Tunneling Diode (RTD) based transceiver technology is used. This European project targets at the development of both an all-electronic RTD approach for portable devices and an optoelectronic RTD approach for the integration of femto cell base stations connected to a fibre-optic network at 90 and 300 GHz. The paper gives an overview on the ongoing activities in channel modeling, system design, RTD technology development as well as on manufacturing III-V materials on silicon.

An overview on the ongoing research activities in China in the area of THz communications is presented in the third article titled “Terahertz wireless communication activities in China” by Li, Ai, Guan and Zhong from Beijing Jiaotong University. Four different systems operating between 140 GHz and 4.13 THz developed by three Chinese institutions are presented together with an outlook on future research topics.

The fourth article titled “IEEE P802.15.3d – Towards the First Wireless Communications Standard Operating in the 300 GHz Band” by Kürner et al. describes the ongoing standardization and regulation activities for THz communications. IEEE 802 is currently working at the first 300 GHz wireless communications standard and the allocation of spectrum beyond 275 GHz is an agenda item at the next world radio conference in 2019.

Optical Wireless Communication Networks

OWC systems can be classified into short and long range systems, according to the distance between the transmitter and receiver. Short range systems are mainly for indoor use while long range systems are used for outdoor and space applications. The focus of the first article is outdoor OWC links which are widely referred to as free space optical (FSO) communication. Fundamentals, applications and recent achievements of FSO communication are described by Uysal.

Korea is one of the leading countries in visible light communication (VLC) technologies, which is part of OWC. The second paper of this section is by Le and Jang from Kookmin University and details the status of VLC research in Korea. They provide information about specific research projects such as “the proposed system of LED-ID from “Development of Home Network Technology based on LED-ID” and general research topics.

One of the first standards solely dedicated on VLC is 802.15.7, which is published in 2011. Currently a task group in IEEE 802 LMSC is working on an amendment to this standard. The third article entitled “Let there be Light Again! An Amendment to IEEE 802 Visible Light Standard is in Progress” by Baykas et. al. provides the status of this work and its future.

Last article “LiFi – Technological and Commercial Consideration” by Serafimovski et. al. of this issue is about one of the most important commercialization attempts in VLC technology. LiFi takes VLC further by using light emitting diodes (LEDs) to realize fully networked wireless systems. Synergies are harnessed as luminaries become LiFi attocells resulting in enhanced wireless capacity providing the necessary connectivity.

While we are far from delivering a complete coverage on these two exciting research areas, we hope that the articles provide an overview of recent activities in THz and Optical Wireless Communication. Finally, we would like to thank all the authors for their great contribution and the E-Letter Board for making this special issue possible.



Tuncer Baykas works as an assistant professor at Istanbul Medipol University. He was the chair of IEEE 802.19.1 TG. He served as co-editor and secretary for 802.15 TG3c and contributed to many standardization projects, including 802.22, 802.11af and 1900.7. Currently he is the vice director of the “Centre of Excellence in Optical Wireless Communication Technologies (OKATEM)”.



Thomas Kürner is a full Professor for Mobile Radio Systems at Institut für Nachrichtentechnik at Technische Universität Braunschweig and is working, amongst other topics, on THz communication systems. He is chair of IEEE 802.15 TG3d and IEEE 802.15 IG THz and the coordinator of the German TERAPAN project.

TERAPAN – Towards a 100 Gbit/s Point-to-Point Link with Electronic Beam Steering Operating at 300 GHz

Sebastian Rey¹, Iulia Dan², Thomas Merkle³, Axel Tessmann³, Ingmar Kallfass², Thomas Kürner¹

¹*Technische Universität Braunschweig, Germany, {rey|kuerner}@ifn.ing.tu-bs.de*

²*University of Stuttgart, Germany, {iulia.dan|ingmar.kallfass}@ilh.uni-stuttgart.de*

³*Fraunhofer Institute for Applied Solid State Physics, Germany,
{thomas.merkle|axel.tessmann}@iaf.fraunhofer.de*

1. Introduction

The objective of the TERAPAN project [1] is to demonstrate an all-electronics approach for 300 GHz communications with beam steering in an indoor environment. The targeted data rate is 100 Gbit/s for a distance of up to ten meters. The transmitter and the receiver are MMICs realized in a 35 nm mHEMT technology [2]. After a very short introduction to THz communications in general, the remainder of this letter will give an overview of the TERAPAN project and the current status.

Basic propagation conditions

The idea of THz communications is to use lower-order modulation schemes with huge absolute bandwidths in the order of 50 GHz to achieve data rates of 100 Gbit/s. The frequency range around and above 300 GHz is very promising, since this part of the spectrum is still not allocated to any services. However, passive services, like earth exploration services and radio astronomy, are operating in this frequency range. Still for indoor communications, no interference with these services is expected due to the high free space path loss and the losses through building materials.

Actually, the free space path loss is one of the disadvantages at these high frequencies. Over a distance of just 10 m it is already around 100 dB at 300 GHz. Any link at these frequencies will require highly directive antennas (e.g. 25 dBi at the receiver and the transmitter) to cope with the free space path loss [3]. For short ranges the atmospheric attenuation at THz frequencies can be neglected. For longer transmission distances there are atmospheric windows with bandwidths of several tens of GHz with a rather low attenuation. Another challenge are the reduced link budget margins due to lower available transmit power and higher receiver noise compared to transceivers operating in the microwave and low millimeter-wave spectrum.

Envisioned Applications

The TERAPAN project does not directly demonstrate a solution tailored to one specific application but demonstrates the feasibility of important parts for THz communications. Due to the required directive antennas, especially applications with point-to-point links are currently discussed. For instance, within the IEEE 802 project a first standard for THz communications is developed with a focus on intra-device, close proximity kiosk downloading, wireless data centers and front-/backhaul-links as applications [4]. These applications all have in common that they focus on point-to-point links and only require beam steering in terms of switching from one link to another. The drastically varying transmission distance can basically be achieved by applying antennas with a suitable antenna gain.

In addition to the previously mentioned application the TERAPAN project keeps one more application in mind: Wireless Personal Area Networks (WPAN). In such an application with a directive antenna, electronic beam steering is additionally required in order to react in time to a movement of the transmitter or the receiver.

2. System Overview

In this section, the enabling technology and the design of the TERAPAN hardware are briefly summarized. In the second part some measured figures of merit are presented for a Single-Input Single-Output (SISO) transmission. A much more detailed description of the underlying technology, the design of the RF-frontend and the measurements can be found in [2].

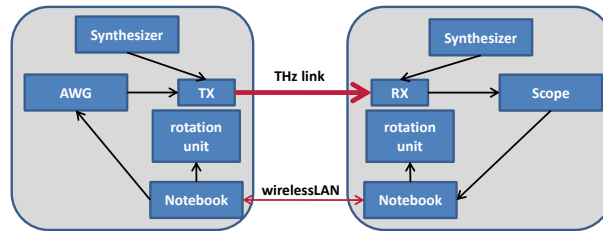


Figure 1: Setup of the mechanical beam steering demonstration.

In Figure 1 the setup for the mechanical beam steering demonstration is illustrated. As a source for I-/Q-data an Arbitrary Waveform Generator (AWG) is utilized. At the receiver a high speed oscilloscope (Scope) records the transmitted signals. Two synthesizers provide the local oscillator signals at 8.333 GHz which is multiplied by twelve to 100 GHz.

Technology and Chip Design of the RF-frontends

The key components are the TERAPAN RF-frontend MMICs. Within the transmitter the 100 GHz local oscillator is multiplied by three to 300 GHz. The I- and Q-Data is then modulated on this carrier frequency. Together with some amplifiers, all these components are integrated on a $0.75 \times 3.25 \text{ mm}^2$ chip, cf. Figure 2. The architecture of the receiver is almost identical except for an LNA instead of the output power amplifier. The chips are manufactured in a 35 nm InGaAs/GaAs mHEMT technology, which enables transit frequencies above 500 GHz in combination with low noise figures. The MMICs are mounted in waveguide modules.

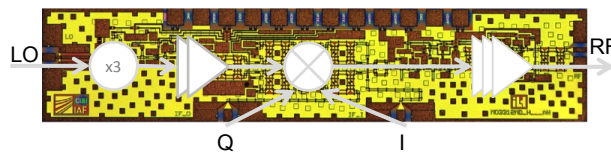


Figure 2: TERAPAN Transmitter MMIC.

SISO Capabilities

After fabrication the MMICs have been characterized and in a measurement campaign the performance of the SISO link has been estimated. The transmitter provides a transmitted power of -7.0 dBm. The estimated noise figure of the receiver is in the order of 7 dB. In order to cope with the free space path loss of 82.0 dB for a transmission distance of one meter at 300 GHz, a horn antenna with ~24 dBi is applied at the receiver and the transmitter, respectively. With a QPSK modulation and a symbol rate of 32 Gbaud a data rate of 64 Gbit/s has successfully been demonstrated. In Figure 3 the constellation diagram after some simple equalization of this transmission is depicted. The four constellation points are clearly separated from each other, obviously allowing for quasi error free decoding in terms of a low bit error rate even without a forward error correction. Since the AWG does not support higher symbol rates, the achievable data rate with a QPSK is likely not limited to 64 Gbit/s by the RF-frontends.

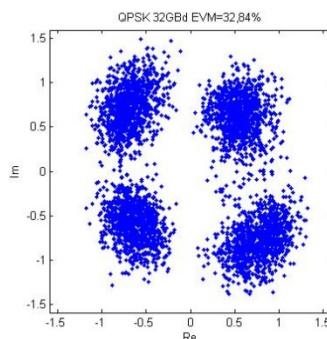


Figure 3: Constellation diagram of the 64 Gbit/s transmission with QPSK at a distance of 1 m.

In addition, a transmission has also been feasible for an increased distance of two meters. Possibly, in a future setup even a further increase can be realized with longer RF cables attached. Higher order modulation schemes have more demanding requirements e.g. in terms of the linearity of the amplifiers or the maximum allowable phase noise. Their performance still has to be evaluated.

3. Demonstration of Steerable Beams

In the first phase beam steering has been realized by mounting the SISO link with the attached horn antennas on mechanical rotation units, Figure 4 (left). For an impression of the dimensions on the right side of Figure 4 a detail of the transmitter is depicted. The multiplier and the amplifier are not integrated in the transmitter module for flexibility reasons. Therefore the size can be further reduced in the future.

With a very simple algorithm, the feasibility of tracking a slowly moving transmitter over a small transmission distance of 1 m has successfully been demonstrated at the NGMN Industry Conference and Exhibition 2015 [5]. At that time, a data rate of 12 Gbit/s was demonstrated in the first beam steering experiments. A brief video of this demonstration can be found on the TERAPAN website [1].

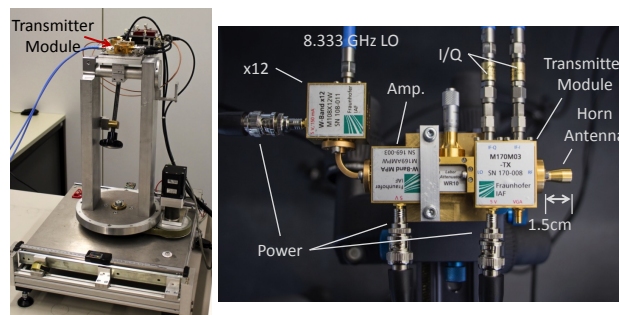


Figure 4: Transmitter modules on a mechanical rotation unit (left); detail of the transmitter modules (right).

The overall setup for the demonstration is shown in Figure 1. The same AWG and a high speed scope as in the other experiments are used to generate and to receive the transmitted data. Besides the 300 GHz connection, there is a second link which represents a WLAN at a lower frequency (Figure 1). This link is used to exchange signaling information in order to coordinate the beam steering at the receiver and the transmitter. Even in a future fully integrated THz communications system, this second link at a lower frequency enables THz devices to easily discover that they are within communication range without the need to already align the highly directive THz antennas.

In the second phase of the TERAPAN project electronic beam steering will be demonstrated with a phased array of four transmitters and four receivers. Currently, a suitable antenna module and the receiver and transmitter modules are manufactured.

4. Conclusion

This letter provided an overview over the ongoing TERAPAN project which aims at the demonstration of a 100 Gbit/s link with steerable antennas. A SISO transmission with mechanical beam steering was presented. For the remainder of the project, electronic beam steering will be realized and demonstrated till the end of 2016.

Acknowledgment

TERAPAN is funded by the German Federal Ministry of Research and Education (BMBF) in the frame of the VIP (Validating the Innovative Potential) under grant number 03V0411. The authors thank Dr. Peter Meissner for his valuable input as the Innovation Mentor of the TERAPAN project.

References

- [1] "TERAPAN Project Website," [Online]. Available: www.terapan.de.
- [2] I. Kallfass, I. Dan, S. Rey, P. Harati, J. Antes, A. Tessmann, S. Wagner, M. Kuri, R. Weber, H. Massler, A. Leuther, T. Merkle and T. Kürner, "Towards mmic-based 300 ghz indoor wireless communication systems," *Trans. Institute of Electronics, Information and Communication Engineers*

IEICE, Vols. E98-C, no. 12, Dec 2015.

- [3] T. Kürner and S. Priebe, "Towards THz Communications - Status in Research, Standardization and Regulation," *Journal of Infrared, Millimeter and Terahertz Waves*, vol. 35, pp. 53-62, 2014.
- [4] IEEE 802, "P802.15.3d Project Authorization Request (PAR)," [Online]. Available: http://ieee802.org/PARs/2015_11/15-15-0682-01-003d-3d-par-change.pdf.
- [5] NGMN Alliance, "NGMN Industry Conference & Exhibition 2015," [Online]. Available: <http://ice2015.ngmn.org/>.



Sebastian Rey (Dipl.-Ing.) is pursuing his Ph.D. as a research assistant with the Institut für Nachrichtentechnik at Technische Universität Braunschweig. His work mainly focuses on propagation and system measurements and modelling at 300 GHz. He is a contributor to IEEE 802.15.3d and currently the editor of the evaluation criteria document of this group.



Iulia Georgiana Dan received the M.Sc. Degree in Electrical Engineering from the university of Stuttgart in 2015. Currently, she is pursuing her Ph.D. as a research assistant with the Institute of Robust Power Semiconductor Systems at the University of Stuttgart, with a focus on monolithic integrated circuit design for 300 GHz transceivers.



Thomas Merkle (Dr.-Ing.) is the project leader of the TERAPAN research project at the Fraunhofer IAF. His research work is focused on MMIC design and packaging technologies for millimeter-wave radar and wireless communication applications at 240 GHz and 300 GHz.



Axel Tessmann is group leader of the packaging and sub-system group at the Fraunhofer IAF. His main research areas are the design and packaging of millimeter-wave and submillimeter-wave ICs using high electron-mobility transistors on GaAs, GaN, and InP, as well as circuit simulation and linear and nonlinear device modeling.



Ingmar Kallfass holds the chair for Robust Power Semiconductor Systems at the University of Stuttgart, where his major fields of research are compound semiconductor based circuits and systems for microwave and power electronics



Thomas Kürner is a full Professor for Mobile Radio Systems at Institut für Nachrichtentechnik at Technische Universität Braunschweig and is working, amongst other topics, on THz communication systems. He is chair of IEEE 802.15 TG3d and IEEE 802.15 IG THz and the coordinator of the German TERAPAN project.

iBROW – Innovative ultra-BROadband ubiquitous Wireless communications through terahertz transceivers

Luis Pessoa¹, Bile Peng², Jue Wang³, Loïc Sanchez⁴, Abdullah Al-Khalidi³, Edward Wasige³, Amy Liu⁵, Horacio Cantu⁶, Ian Oxtoby⁷, Bruce Napier⁸, Anthony Kelly³, José Figueiredo⁹, Mohand Achouche¹⁰, Henrique Salgado¹, Thomas Kürner², Moz Fisher⁴, Andreas Pascht¹¹ and Wolfgang Templ¹¹

¹INESC TEC, Portugal; ²TU Braunschweig, Germany; ³University of Glasgow, UK; ⁴CEA-LETI, France;

⁵IQE Silicon, UK; ⁶CST Global, UK; ⁷Opticap, UK; ⁸Vivid, Germany; ⁹University of Algarve, Portugal;

¹⁰III-V Lab, France, ¹¹Nokia Bell Labs, Germany

luis.m.pessoa@inesctec.pt; edward.wasige@glasgow.ac.uk

1. Introduction

The demand for bandwidth in wireless short-range communications has doubled every 18 months since 1980 [1], and it is expected that wireless data-rates of multiple tens of Gbit/s will be required by the year 2020 [2]. Broadband internet with high bandwidth services and applications such as high definition multimedia, web-TV with multi-casting capabilities, real-time video and fast peer-to-peer applications are in continuous evolution and demanding ever higher data rates. In parallel, the increasing penetration of the population with mobile devices like tablets and smart phones, as well as the general trend to mobile broadband connectivity along with the “smartification” of our technical environment (“Internet of Everything”) will contribute decisively to this trend.

To address the predicted future network usage requirements, the iBROW project aims at the development of energy-efficient and compact ultra-broadband short-range wireless communication transceiver technology. The envisaged usage and deployment scenarios such as ultra-broadband services for indoor- and campus- or stadium-like environments have in common that they address end users and may involve a high number of detached RF-frontend nodes in order to provide good coverage. Furthermore, high bitrate communicating wireless nodes operating in the range of up to 100 Gbit/s as targeted by iBROW involve extremely high signal processing effort in the front- or backhaul, driving costs and requiring high power consumption. This problem can be to certain extent addressed by radio-over-fibre (RoF) concepts in centralised radio access network-like architectures (CRAN) or distributed antenna systems (DAS). Both have in common that the bulk of signal processing tasks is shifted into a central node, which may run the high performance signal processing and hence can be more expensive and power consuming than the remote RF-frontend nodes which can be held relatively primitive and efficient needing only low power.

iBROW is pursuing its objective through the exploitation of Resonant Tunnelling Diode (RTD) based transceiver technology, and targets the development of (1) an all-electronic RTD suitable for integration into cost-effective wireless portable devices and (2) an optoelectronic RTD, consisting of the integration between a RTD and a photodetector and a laser diode, which will be suitable for integration into mm-wave/THz femtocell base stations connected to high-speed 40/100Gbps fibre-optic networks.

To achieve the targeted 100 Gbit/s wireless transmission speeds it is indispensable to rely on the exploitation of the mm-wave and THz frequency bands [3], above 60 GHz and up to 1 THz, since the frequency spectrum currently in use (in the 1-6 GHz range) is not expected to be suitable to accommodate the predicted future data-rate requirements, in spite of the significant and continuous progress that has been achieved in spectrum efficiency techniques, including spatial re-use methods such as beam-forming and MIMO. The proposed RTD technology can serve both source and detector functions.

This paper will describe several results already achieved on iBROW and preview upcoming work including a brief description of the target application scenario, and the modelling of the mm-wave/THz channels at the target operation frequencies of 90 and 300 GHz. In addition, a few results will be presented on the high performance RTD oscillators in the 75-200 GHz range as well as first results on the integration of InP RTD epitaxial wafers on a silicon host substrate through wafer bonding.

2. iBROW target application scenario

Figure 1 illustrates the iBROW target scenario, identifying both downlink (DL) and uplink (UL) communication directions as well as the key technology blocks being developed, namely (1) the all-electronic RTD suitable for integration into cost-effective wireless portable devices, and (2) the optoelectronic RTD, consisting of the monolithic integration between a RTD and a photodetector (an RTD-PD) and hybrid integration with a laser diode (RTD-PD-LD), which will be suitable for integration into mm-wave/THz femtocell basestations connected to high-speed 40/100 Gbit/s fibre-optic networks.

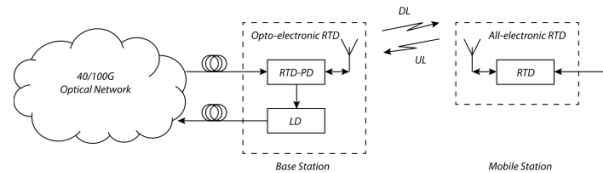


Figure 1 – Schematic representation of the iBROW target application scenario.

3. Channel modelling for THz indoor communications

A channel model is an essential component for the system and link level simulations. It is based on channel measurements with Vector Network Analyser (VNA) or channel sounder (CS) and describes the channel characteristics. Channel models are generally classified into deterministic and statistical models. The former requires a detailed description of application scenario, such as the ray launching channel model [4]. The latter generates a channel randomly based on the statistical channel characteristics. Both models depend on the choice of application scenario and the statistical channel model provides more sense of generality and requires less computational effort. Besides the static channel model, another important aspect of the indoor wireless channel is the human blockage problem of the propagation path since the human body can no longer be assumed transparent as at several GHz carrier frequencies. To model this phenomenon, a human blockage model was developed [5] as well as a realistic human movement model [6]. The generation and application of channel models are depicted in Fig. 2.

Within iBROW, three representative application scenarios are considered: a small office scenario, a meeting room and an auditorium. The single office scenario accounts for 1-2 users while the meeting room scenario has an increased numbers of users where in addition to the static situation, a dynamic component plays an additional role. For the third scenario, the number of users is higher and the concept of distributed antennas can be investigated. Fig. 3 shows a photo of the second scenario and the corresponding ray tracing model.

Preliminary channel measurements have been carried out in the small office scenario. The results are presented in [7]. Compared to the channel with conventional lower frequencies, the investigated channels suffer from a much higher free space path loss and the distributions of the angles of departure/arrival are more specular. Since the bandwidth is very broad, the symbol duration is usually shorter than the delay spread, which suggests serious inter-symbol interference.

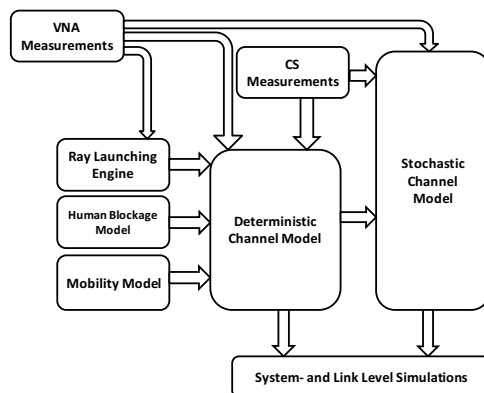


Figure 2 – Generation & application of channel models.



Figure 3 – Meeting room scenario and ray launching model.

4. The RTD technology

An RTD device consists of a narrow band gap semiconductor material sandwiched between two thin wide band gap materials. These devices lend themselves to an efficient integration with optoelectronic components (photodetector and laser) since they are made from the same material system [10]. RTDs also require very low power consumption and unlike traditional transistor technologies, they possess an intrinsic gain (provided by its negative differential conductance region) which allows for a simple and energy efficient implementation, avoiding the employment of a number of complex transceiver building blocks, such as power or low noise amplifiers which can be rather inefficient [10]. RTD devices exhibit the highest oscillation frequency among traditional electronic devices. The published highest frequency of single device RTD oscillator is 1.55 THz with 0.4 μ W output power [11].

iBROW targets the realisation of indium phosphide (InP) based RTDs on a silicon host substrate with devices and circuits realised using conventional integrated circuit technology in order to achieve a potentially very high-speed and low-cost solution. The RTD technology being developed at University of Glasgow aims to raise the power level up to several milliwatt (mW) for frequencies over 100 GHz. To date on iBROW, 76 GHz / 0.95 mW, 125 GHz / 0.34 mW, 156 GHz / 0.24 mW, 166 GHz / 0.17 mW RTD oscillators have been realized [12][13].

The epitaxial heterostructure of RTDs fabricated in iBROW (Fig. 4a) are grown by molecular beam epitaxy (MBE) on a semi-insulating InP substrate by IQE Ltd. The devices are fabricated using optical lithography. A micrograph picture of the completed RTD device is shown in Fig. 4b. The maximum

available power (P_{\max}) of single RTD oscillator can be estimated from $(3/16)\Delta V\Delta I$, where ΔV is the peak-valley voltage difference, and ΔI is the peak-valley current difference. By optimizing the layer structure, it is possible to maximize the ΔV and ΔI to further increase the obtainable power [14].

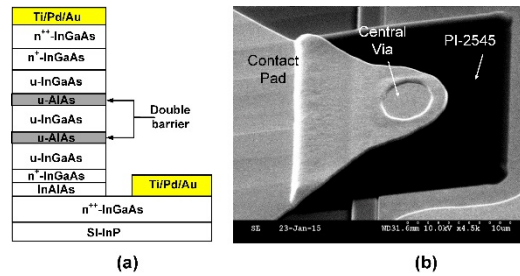


Fig. 4 (a) The epitaxial layer structure of a RTD device. (b) Micrograph of the fabricated $5 \times 5 \mu\text{m}^2$ RTD device.

Measurement results

The oscillators were characterized on-wafer by using an Agilent spectrum analyser with relevant external mixers. The results are, summarized in Table 1. RTD oscillators presented here demonstrated significantly high performance in terms of higher output power, especially in W-band frequency range. In these designs, no intent was made to match the impedance. Future designs will employ optimized epi-structures and integrated impedance matching networks, in particular integrated antenna loads designed to match the RTD impedances.

TABLE I. SUMMARY OF RTD OSCILLATOR PERFORMANCE

Device size (μm^2)	CP W length (μm)	Freq. (GHz)	Power (dBm/mW)	DC Power (mW)
16	120	75	-0.2/0.95	422
25	30	125	-1.7/0.68	415
15	30	156	-6.3/0.24	374
15	20	166	-7.7/0.17	191

5. III-V on silicon technology

In order to introduce III-V materials in the low cost silicon platform manufacturing, the wafer bonding approach is of great interest due to the limitations of growing compound III-V hetero-epitaxial layers directly onto silicon. Direct wafer bonding refers to a process by which two mirror-polished wafers are put into contact and held together at room temperature by adhesive forces, without any additional materials. This technique has become a technology of choice for materials integration in various areas of microelectronics, microelectromechanical systems and optoelectronics [15]. In iBROW, characterizations of III-V epitaxial layers on InP wafers required for the RTD manufacturing show that surface defectivity is low enough to expect a correct bonding. Furthermore, a plasma surface activation before bonding was used to maximize the bonding strength after a low temperature annealing imposed by the difference of thermal expansion coefficient between InP and Si. Epitaxial layers on 50mm InP wafers were successfully bonded on oxidized 200 mm Si wafers. Acoustic characterization of the bonding interface shows a correct bonding yield which will allow us to develop post bonding RTD manufacturing step.

6. System considerations

In the emerging mass markets of fibre to the x (FTTx) and ubiquitous high bandwidth wireless access all the benefits of integration will be important, particularly low power and low cost. The technology must

therefore be capable of addressing end-consumer markets, which means that it will face extremely high cost pressure and must support ease of deployment, i.e. our solutions need to be low cost, or at least must have the potential to become low cost when the technology matures during the coming years. In addition, back- or front-haul technologies for the wireless nodes need to support the above mentioned ease of deployment and must not involve high cabling effort. Especially in indoor environments, cabling can involve high effort and costs; under such conditions, unsuitable system architectures may become showstoppers. For instance, star-like network topologies would allow for highest data rates, but involve very high cabling effort (i.e. a fibre to each node). For this reason, advanced multiplexing techniques (TDM, WDM, T-WDM) and bus or daisy chain topologies are favourable. Last but not least, the system needs to be very energy efficient in order to avoid expensive and noisy active and noisy cooling solutions and to allow for low operation costs.

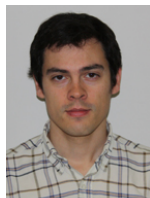
7. Conclusion

This paper described several aspects of the work being carried out on the European project iBROW towards realising a mm-wave/THz transceiver technology in RTD technology. Several achievements were highlighted and future directions pointed out. In the 300 GHz band iBROW is already well positioned to make impact both on Standardization of THz communications systems at IEEE 802 and spectrum regulation at the ITU-R level [16].

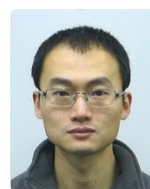
References

- [1] S. Cherry, "Edholm's Law of Bandwidth," *IEEE Spectrum*, vol. 41, no. 7, pp. 58–60, 2004.
- [2] T. Kleine-Ostmann and T. Nagatsuma, "A review on terahertz communications research," *Journal of Infrared, Millimeter, and Terahertz Waves*, 2011, 32.2: 143-171.
- [3] T. Kürner, "Towards future THz communications systems," *Int. Journal of THz Science and Technology*, vol. 5, no. 1, pp. 11–17, 2012.
- [4] D. M. Rose and T. Kürner, "An analytical 3D ray-launching method using arbitrary polygonal shapes for wireless propagation prediction," in *IEEE 80th Vehicular Technology Conference (VTC Fall)*, Vancouver, 2014.
- [5] M. Jacob et al. "A ray tracing based stochastic human blockage model for the IEEE 802.11 ad 60 GHz channel model." *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*. 2011.
- [6] D.M. Rose et al. "Impact of realistic indoor mobility modelling in the context of propagation modelling on the user and network experience." *7th IEEE European Conf. on Antennas and Propagation (EuCAP)*, 2013.
- [7] B. Peng et al. "Channel characteristics study for future indoor millimeter and submillimeter wireless communications." *10th IEEE European Conf. on Antennas and Propagation (EuCAP)*, 2016.
- [8] B. Romeira, L.M. Pessoa et al, "Photo-detectors Integrated with Resonant Tunneling Diodes," *Sensors* 13, no. 7, pp. 9464-9482, 2013.
- [9] H. Mizuta and T. Tanoue, "The Physics and Applications of Resonant Tunneling Diodes," Cambridge University Press, 2006.
- [10] J.M.L. Figueiredo et al, "Resonant Tunneling Optoelectronic Circuits, Advances in Optical and Photonic Devices," Ki Young Kim (Ed.), ISBN: 978-953-7619-76-3, InTech, DOI: 10.5772/7146
- [11] T. Maekawa, H. Kanaya, S. Suzuki and M. Asada, "Frequency increase in terahertz oscillation of resonant tunnelling diode up to 1.55 THz by reduced slot-antenna length," *Electronics Letters*, 50(17), pp.1214-16, 2014.
- [12] J. Wang et al, "W-band InP-based Resonant Tunneling Diode Oscillator with Milliwatt Output Power," *Indium Phosphide & Related Materials Conference*, 2014.
- [13] J. Wang et al, "High performance resonant tunnelling diode oscillators for THz applications," *IEEE Compound Semi. Integrated Circuit Symp. (CSICS)*, pp.1-4, 2015.
- [14] J. Wang et al. "G-Band MMIC resonant tunnelling diode oscillators" *Indium Phosphide & Related Materials Conference*, 2016. (Submitted).

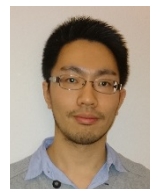
- [15] L. Sanchez et al, "Chip to wafer direct bonding technologies for high density 3D integration," 62nd IEEE Elec. Components & Techn. Conf., pp. 1960-1964, 2012.
- [16] T. Kürner, B. Napier, A. Pascht, B. Peng, L. Pessoa, W. Templ, and E. Wasige, "*Input from the Horizon 2020 iBROW Project to the TG3d Call for Contributions to the Response on the Liaison Statement from ITU-R WP1A,*" IEEE 802.15 Document 15-16-0034-00-003d, Atlanta/USA, Jan. 2016.



Luís Pessoa is a senior researcher at INESC TEC, mainly involved in the conception and management of R&D projects. His research interests include: coherent optical systems, fibre supported microwave systems, RF and microwave devices and antennas.



Bile Peng is a PhD candidate at Technische Universität Braunschweig. His research interests are THz channel measurement, modelling and estimation.



Jue Wang is a Postdoc. Researcher at the University of Glasgow working on RTD devices and circuits.



Edward Wasige is a Senior Lecturer at the University of Glasgow. His research interests include RTD devices and circuits for THz electronics and optoelectronics.



Thomas Kürner is full Professor for Mobile Radio Systems at Institut für Nachrichtentechnik at Technische Universität Braunschweig and is working, amongst other topics, on THz communication systems. He is chair of IEEE 802.15 TG3d and IEEE 802.15 IG THz and the coordinator of the TERAPAN project.



Wolfgang Templ heads the department for radio transceiver research at Nokia Bell Labs where the investigation and design of new disruptive concepts for future communication systems based on new enabling device technologies is an important area of work.

Terahertz wireless communications activities in China

Guangkai Li, Bo Ai, Ke Guan, Zhangdui Zhong

State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, 100044, Beijing, China

{gkli.rcs, boai, kguan, zhdzhong}@bjtu.edu.cn

1. Introduction

The Terahertz (frequencies ranging from 0.1 THz to 10 THz) wireless communication has received great popularity in satisfying very high speed wireless data rate and exploding the restriction of limited bandwidth in conventional wireless communication systems. It is also the matchless research direction for achieving large data transmission via wireless up to now.

Owing to the active encouragement of pioneers for Terahertz wireless communication, the National High Technology Research and Development Program ("863" Program) of China started a program of "Technologies Development for Millimeter Wave and Terahertz Wireless Communication" at the end of 2011. The final report of the program was delivered in August 2015, which declares that the first important program in terms of Terahertz wireless communication in China has been completed. Furthermore, the program convened a number of Chinese research institutions and further established National Cooperative Innovation Centre of Terahertz Science (NCICTZ) in 2012.

In recent years, the Chinese research institutions, e.g., China Academy Of Engineering Physics [1]~[7], Chinese Academy of Sciences [8][9], University of Electronic Science and Technology of China [10], Nanjing University [11], Shanghai Aerospace Electronics Equipment Institute [12][13], etc., have developed the researches on Terahertz wireless communication.

2. Typical examples of Terahertz wireless communication systems in China

In this section, several representative Terahertz wireless communication systems in China will be briefly introduced.

0.14 THz wireless communication system developed by China Academy of Engineering Physics [4]:

A 0.14 THz wireless communication system has been testified to transmit a maximum 10 Gbit/s signal over 1.5 Km with bits error rate (BER) lower than $1e-6$. As shown in Fig. 1, the system is constructed by sub-harmonic mixer and multiplier based on Schottky barrier diodes, Cassegrain antennas, local oscillator multiplier and other high performance elements. Based on 16 quadrature amplitude modulation (16QAM), the spectrum efficiency of the system can reach to 2.86 bit/s/Hz in generating 10 Gbit/s signal within only 3.6 GHz bandwidth. Moreover, the system supports two demodulation models. In non-real-time mode, high speed analog digital converter (ADC) and software demodulator in computer are developed to support up to 10 Gbit/s data. In real-time mode, as Field-Programmable Gate Array (FPGA) demodulator embedded into hardware, a 2Gbit/s signal at 140 GHz can be transmitted over 1.5 Km with BER of $1.8e-11$.

It is noteworthy that, for 1.5 Km wireless links, the output power of the 140 GHz front end is -5dBm which can be further pushed to level of tens of watts for longer distance of wireless links by vacuum traveling-wave tubes cascaded with solid-state power amplifiers.

0.34 THz wireless communication system developed by China Academy of Engineering Physics [5][6][7]:

The 0.34 THz wireless communication system is based on 16QAM modulation. With two Cassegrain antennas (48.4 dBi) on both Tx and Rx. It can support 3 Gbit/s real time data transmission with lowest BER of $1.784e-10$ over 50 m line of sight channel. Fig. 2 (which is from Fig. 1 in [5]) shows a sketch of the system. Moreover, the 0.34 THz wireless communication system was further enhanced to develop 0.34 THz WLAN system based on IEEE 802.11 b/g protocol. According to the experiment reported in

[5][7], during transmission time of 30 min, the average transmission data rate of 6.54 Mbit/s has been achieved within 1.15 m link distance.

3.27 THz and 4.13 THz wireless communication system developed by Chinese Academy of Sciences [8][9]:

These two wireless communication systems are developed based on THz quantum cascade laser (QCL) processes in transmitter and spectrally matched THz quantum-well photo-detector (QWP) in receiver (see Fig. 3 which includes Fig. 1 and Fig. 4 from [9]). For 4.13 THz wireless communication system, 1 Mbit/s data rate has been demonstrated within 2.2 m distance. 5 Mbit/s data rate has been obtained with deeper optimization of the system. For 3.27 THz wireless communication system, a low BER (lower than $1e-8$) transmission is obtained up to 20 Mbit/s within 2.2 m distance, which is shown in Fig. 3. The further enhancement of the 4.13 THz system will depend on improvement of THz QWP and lower capacitance of cable link.

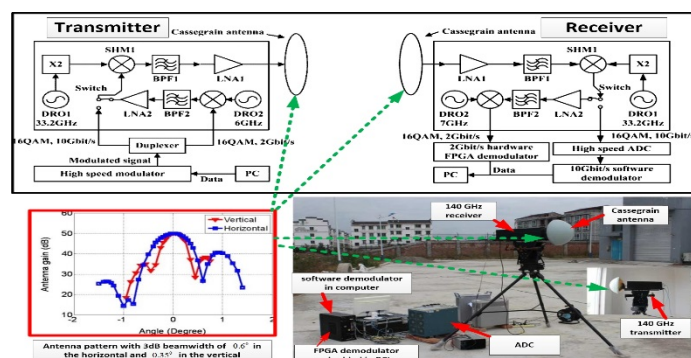


Figure 5 Schematic and photograph of 0.14 THz wireless communication system with measured pattern of Cassegrain antenna.

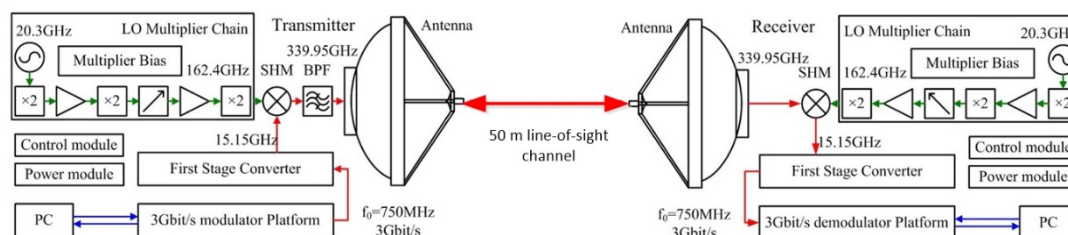


Figure 6 Schematic of 0.34 THz wireless communication system

0.3 THz wireless communication system developed by Shanghai Aerospace Electronics Equipment Institute [12][13]:

The applications of terahertz wireless communication suffer the atmosphere attenuation on the ground, which is contrary to the applications in space. Shanghai Aerospace Electronics Equipment Institute in China designed a 0.3 THz wireless communication system for potential space applications. Based on Schottky diode mixer technology, the system has been tested and verified to transmit high definition video signal with 1.5 Gbit/s data rate over 14 m distance.

3. Conclusions

In this letter, we survey the state of the art of the researches on developing THz wireless communication systems in China. China Academy of Engineering Physics developed 0.14 THz and 0.34 THz system; Chinese Academy of Sciences developed 3.27 THz and 4.13 THz systems; University of Electronic Science and Technology of China developed a 0.3 THz system. Apart from these THz fixed wireless communications, some researchers from Beijing Jiaotong University, e.g., Dr. Ke Guan, propose to develop THz communications for enabling dynamic smart rail mobility by 2020.

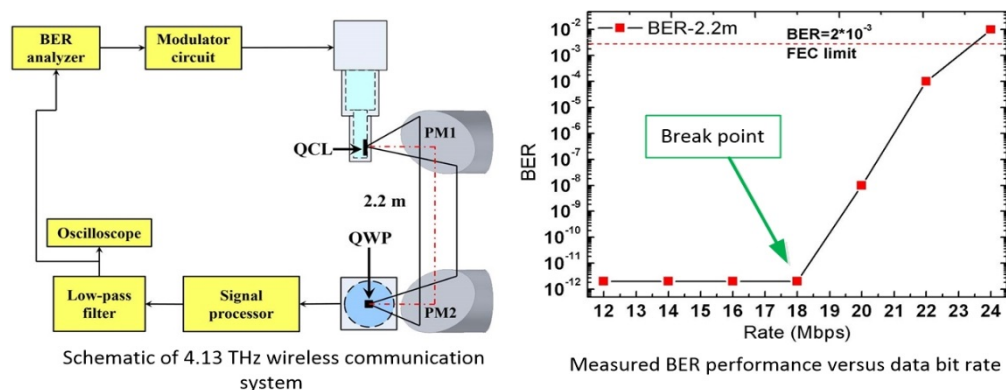


Figure 7 Sketch and performance of 4.13 THz wireless communication system

References

- [17] X. Deng, C. Wang, C. Lin, et al., "Experimental research on 0.14 THz super high speed wireless communication system". *High Power Laser & Particle Beams*, 2011, 23(6):1430-1432.
- [18] C. Wang, C. Lin, X. Deng, et al., "140 GHz Data Rate Wireless Communication Technology Research". *Journal of Electronics & Information Technology*, 2011, 33(9):2263-2267.
- [19] C. Wang, C. Lin, Q. Chen, et al., "0.14 THz high speed data communication over 1.5 kilometers". *International Conference on Infrared*. IEEE, 2012:1-2.
- [20] C. Wang, C. Lin, Q. Chen, et al., "A 10-Gbit/s Wireless Communication Link Using 16-QAM Modulation in 140-GHz Band". *IEEE Transactions on Microwave Theory & Techniques*, 2013, 61(7):2737-2746.
- [21] C. Wang, B. Lu, C. Lin, et al., "0.34-THz Wireless Link Based on High-Order Modulation for Future Wireless Local Area Network Applications". *IEEE Transactions on Terahertz Science & Technology*, 2014, 4(1):75-85.
- [22] C. Wang, B. Lu, L. Miao, et al. "0.34 THz T/R front-end for wireless communication". *High Power Laser & Particle Beams*, 2013, 25(6):1530-1533.
- [23] C. Lin, L. Bin, C. Wang, et al., "0.34 THz Wireless Local Area Network demonstration system based on 802.11 protocol". *Journal of Terahertz Science & Electronic Information Technology*, 2013, 11(1):12-15.
- [24] Y. Tan, Z. Chen, J. Cao, et al., "Wireless terahertz light transmission based on digitally-modulated terahertz quantum-cascade laser". *Chinese Optics Letters*, 2013, 11(3):38-40.
- [25] L. Gu, Z. Tan, Q. Wu, et al., "20 Mbps wireless communication demonstration using terahertz quantum devices". *Chinese Optics Letters*, 2015, 13(8):59-61.
- [26] Z. Chen, B. Zhang, Y. Fan, "120GHz real-time wireless data link based on Schottky technology". *Infrared, Millimeter, and Terahertz waves (IRMMW-THz)*, 2014 39th International Conference on. IEEE, 2014: 1-2.
- [27] X. Deng, Y. Li, W. Wu, et al., "A fundamental OOK modulator for 340 GHz communication using 0.13 μm CMOS technology". *Microwave & Optical Technology Letters*, 2015, 57(5):1166-1168.
- [28] W. Yuan, H. Wang, H. Li, et al. "0.3THz wireless communication systems for space applications". *Selected Proceedings of the Photo electronic Technology Committee Conferences held June-July 2015*. International Society for Optics and Photonics, 2015.
- [29] H. Wang, W. Yuan, B. Zhang, et al. "The design, test, and application of the front end in 0.3 THz wireless communication systems". *Selected Proceedings of the Photo electronic Technology Committee Conferences held June-July 2015*. International Society for Optics and Photonics, 2015: 97953H-97953H-5.



Guangkai Li was born in Hefei, China on May 22, 1992. He was recruited as Ph.D. student to join State Key Laboratory of Rail Traffic Control and Safety in Beijing Jiaotong University, Beijing, China, in 2015. His research includes optimizing ray tracing algorithm which based on measurements and deterministic modeling for mm-Wave and terahertz channel in dynamic scenarios.



Bo Ai received his Master and Ph.D. degree from Xidian University in 2002 and 2004 in China, respectively. He graduated in 2007 with great honors of Excellent Postdoctoral Research Fellow in Tsinghua University. He is now working in Beijing Jiaotong University as a professor and advisor of Ph.D. candidates. He is a deputy director of State Key Lab of Rail Traffic Control and Safety. He is an associate editor for IEEE Trans. on Consumer Electronics and an editorial committee member of journal of Wireless Personal Communications. He has authored/co-authored 6 books, 140 scientific research papers and 26 invention patents in his research area till now. His current interests are the research and applications OFDM techniques, HPA linearization techniques, radio propagation and channel modeling, GSM for railway systems, and LTE for railway systems. He is an IET Fellow and an IEEE Senior member.



Ke Guan received B.E. degree and Ph.D. degree from Beijing Jiaotong University in 2006 and 2014, respectively. He is an Associate Professor in State Key Laboratory of Rail Traffic Control and Safety & School of Electronic and Information Engineering, Beijing Jiaotong University. In 2015, he has been awarded a Humboldt Research Fellowship for Postdoctoral Researchers. He was the recipient of a 2014 International Union of Radio Science (URSI) Young Scientist Award. His paper received the honorable mention in the third International URSI student prize paper competition in 2014 URSI GASS. In 2009, he was a visiting scholar in Universidad Politecnica de Madrid, Spain. From 2011 to 2013, he has been a research scholar at the Institut fur Nachrichtentechnik (IfN) at Technische Universitat Braunschweig, Germany. From September 2013 to January 2014, he was invited to conduct joint research in Universidad Politecnica de Madrid, Spain.

His current research interests are in the field of measurement and modeling of wireless propagation channels, high-speed railway communications, vehicle-to-x channel characterization, and indoor channel characterization for high-speed short-range systems including future terahertz communication systems. He has authored/co-authored over 70 research papers in international journals and conferences. He received the Huawei Excellent Student Award of China in 2013 and the First National Scholarship for Ph. D Candidates in 2012. He serves as a Technical Program Committee (TPC) member for the IEEE VTC 2014 Fall, IEEE VTC 2015 Spring, and IEEE ICC'15 DVC Workshop. He has been a member of the IC1004 initiative.



Zhangdui Zhong is a professor and advisor of Ph.D. candidates in Beijing Jiaotong University. He is now a director of School of Computer and Information Technology and a Chief Scientist of State Key Laboratory of Rail Traffic Control and Safety in Beijing Jiaotong University. He is also a director of the Innovative Research Team of Ministry of Education, and a Chief Scientist of Ministry of Railways in China. He is an executive council member of Radio Association of China, and a deputy director of Radio Association of Beijing. His interests are wireless communications for railways, control theory and techniques for railways, and GSM-R system. His research has been widely used in the railway engineering, such as Qinghai-Xizang railway, Datong-Qinhuangdao Heavy Haul railway, and many high-speed railway lines of China. He has authored/co-authored 7 books, 5 invention patents, and over 200 scientific research papers in his research area. He received Mao Yi Sheng Scientific Award of China, Zhan Tian You Railway Honorary Award of China, and Top 10 Science/Technology Achievements Award of Chinese Universities.

IEEE P802.15.3d – Towards the First Wireless Communications Standard Operating in the 300 GHz Band

Thomas Kürner¹, Iwao Hosako², Yumi Ueda², Alexander Fricke¹, Sebastian Rey¹, Akifumi Kasamatsu², Norihiko Sekine²

¹Technische Universität Braunschweig, Germany; ²NICT, Japan;
{kuerner|fricke|rey}@ifn.ing.tu-bs.de; {hosako|ueda|kasa|nsekine}@nict.go.jp

1. Introduction

Already a couple of years ago THz communications have not only become an attractive new research area but also triggered discussions and activities in standardization and regulation [1]. In March 2014, IEEE 802 established its project P802.15.3d with the task to develop the worldwide first wireless communications standard operating in the 300 GHz frequency band [2,3]. In parallel, the standardization process is accompanied by activities at the ITU-R level targeting on the provision of an appropriate regulatory framework [4,5]. This letter gives a brief overview on the status of the standardization activities at IEEE 802 and the regulatory processes towards the World Radio Conference 2019 (WRC-2019).

2. Standardization at IEEE 802

According to its granted project authorization request (PAR) [3], the Task Group 3d of the Working Group IEEE 802.15 develops an amendment to the IEEE Standard 802.15.3.

Scope of IEEE P802.15.3d

The amendment to be produced within the project “defines a wireless switched point-to-point physical layer to the IEEE standard 802.15.3 operating at a nominal PHY data rate of 100 Gbit/s with fallbacks to lower data rates as needed. Operation is considered in the bands from 252 GHz to 325 GHz at ranges as short as a few centimeters and up to several 100 m.”[3] The PAR also allows modifications to the Medium Access Control (MAC) layer, if those are necessary to support the new 300 GHz physical layer (PHY).

Applications

The targeted applications are emerging wireless switched point-to-point links in data centers, in wireless back- and fronthauling, in intra-device communications and in close proximity communications (eg. kiosk downloading, file exchange, touch less gate systems). While on the one hand these applications seem to be quite diverse, both with respect to their communication range and their operational environments, on the other hand the commonality of all these applications lies in their point-to-point character with known positions of transmit and receive antennas and the option to switch between different links. In all four fields of application, this new wireless standard will enable the replacement of cables by wireless links while keeping the capability of transmitting data rates of multiple 10s of Gbit/s. This is not only attractive at relatively short distances appearing in intra-device communication and close-proximity point-to-point links but also for larger communication distances occurring in data centers or with backhaul/fronthaul links. Wireless links in these environments will make frequent reconfiguration easier and more cost-effective and will reduce costs for the case when installing a fiber network is not cost-effective, respectively.

Current Status

To prepare a call for proposals, IEEE 802.15 TG3d has been working on four supporting documents:

- Applications Requirements Document [6];
- Technical Requirements Document [7];
- Channel Modeling Document [8];
- Evaluation Criteria Document [9].

These documents have the purpose to assist both in preparing and evaluating the proposals by setting the expectations, defining the requirements and providing defined scenarios, criteria and standardized channel models. The TG3d Channel Model [8] will be the first comprehensive channel model for the 300 GHz band. The current plans are to finish this preparatory work in March 2016 and issue a call for proposals targeting to get the amendment to the standard approved until beginning of 2018 at the latest.

5. Regulatory Issues

The operational frequency band considered in TG3d can be split into two bands based on the regulatory status according to the radio regulations (RR) [10]. While the band 252-275 GHz is already allocated to the mobile and fixed services, the band 275-325 GHz is identified for the passive services by footnote No. 5.565 of the RR. Administrations may make available this band available for other services, e. g. THz communications, as long as the passive services are protected from interference from these other services. In order to make this band generally available for THz communications on a worldwide basis, the only possibility is to envisage a shared use by THz communications and passive services [11]. Therefore, further studies are needed to review RR footnote No. 5.565 for use of these bands by active services in the future. A separate agenda item has been defined for the upcoming WRC-2019.

WRC-19 Agenda Item 1.15

At the World Radiocommunication Conference 2015 (WRC-15), the identification of the frequency range 275-475 or 275-1000 GHz to the land mobile and fixed services was considered to be one of the agenda items for the next conference (WRC-19). After the fruitful debates among the delegates of concerned countries, WRC-15 approved this agenda item to be the WRC-19 Agenda Item 1.15 [5]:

“1.15 to consider identification of frequency bands for use by administrations for the land-mobile and fixed services applications operating in the frequency range 275-450 GHz, in accordance with Resolution 767 [COM6/14] (WRC 15);”

The first session of conference preparatory meeting (CPM 19-1) was held right after WRC-15, and the procedure for conducting the conference preparatory work was also decided. The technical and operational characteristics of the above services and the passive services already identified in some of the above frequency range should be studied first by the relevant Working Parties. After that, the sharing studies among those services should be conducted by Working Party 1A which is a responsible group for the WRC-19 Agenda Item 1.15. Finally, the candidate frequency band(s) in the above frequency range should be decided and provided by the CPM Report.

6. Conclusion

This letter provides a short status on the development of an amendment to IEEE Standard 802.15.3 targeting at a wireless switched point-to-point physical layer operating at a nominal PHY data rate of 100 Gb/s in the frequency band 252 – 325 GHz. The goal is to finalize this standard until the beginning of 2018. In parallel, regulatory activities have started in order to provide stable regulatory conditions for the frequency bands up to 450 GHz at WRC-19.

References

- [30] Kürner, T.; Priebe, S.: Towards THz Communications – Status in Research, Standardization and Regulation. Journal of Infrared, Millimeter, and Terahertz Waves, (2014) Vol. 35, pp. 53-62 <http://link.springer.com/content/pdf/10.1007%2Fs10762-013-0014-3.pdf>
- [31] Home Page IEEE 802.15 Task Group 3d http://www.ieee802.org/15/pub/index_TG3d.html
- [32] P802.15.3d Project Authorization Request (PAR): http://ieee802.org/PARs/2015_11/15-15-0682-01-003d-3d-par-change.pdf
- [33] Report ITU-R SM.2352-0- Technology trends of active services in the frequency range 275-3000 GHz; Geneva, June 2015 http://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-SM.2352-2015-PDF-E.pdf

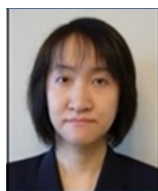
- [34] Provisional Final Acts WRC-15, Geneva 2015; https://www.itu.int/dms_pub/itu-r/opb/act/R-ACT-WRC.11-2015-PDF-E.pdf
- [35] TG3d - Applications Requirements Document (ARD): <https://mentor.ieee.org/802.15/dcn/14/15-14-0304-16-003d-applications-requirement-document-ard.docx>
- [36] TG3d - Technical Requirements Document (TRD): <https://mentor.ieee.org/802.15/dcn/14/15-14-0309-17-003d-technical-requirements-document.doc>
- [37] TG3d – Channel Modeling Document (CMD): <https://mentor.ieee.org/802.15/dcn/14/15-14-0310-16-003d-channel-modeling-document.docx>
- [38] TG3d – Evaluation Criteria Document (ECD): <https://mentor.ieee.org/802.15/dcn/15/15-15-0412-10-003d-evaluation-criteria-document.docx>
- [39] ITU-R: Radio Regulations, Edition of 2012 <https://www.itu.int/pub/R-REG-RR>
- [40] Priebe, S.; Britz, D. M.; Jacob, M.; Sarkozy, S.; Leong, K. M. K.; Logan, J. E.; Gorospe, B.; Kürner, T.: Interference Investigations of Active Communications and Passive Earth Exploration Services in the THz Frequency Range. IEEE Transactions on Terahertz Science and Technology, Vol. 2, No. 5, S. 525– 537, 2012.



Thomas Kürner is a full Professor for Mobile Radio Systems at Institut für Nachrichtentechnik at Technische Universität Braunschweig and is working, amongst other topics, on THz communication systems. He is chair of IEEE 802.15 TG3d and IEEE 802.15 IG THz and the coordinator of the German TERAPAN project.



Iwao Hosako is currently Director General of Advanced ICT Research Institute at the National Institute of Information and Communications Technology (NICT), and is working on other THz technologies (sensing, imaging, and etc.). He is vice chair of IEEE 802.15 TG3d and IEEE 802.15 IG THz.



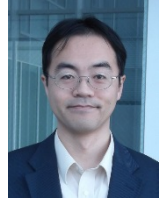
Yumi Ueda is working as a technical expert at the National Institute of Information and Communications Technology. Her job is the standardization for terahertz wave radiocommunication systems and devices mainly in ITU-R and Asia-Pacific Telecommunity.



Alexander Fricke is working as a research assistant at the Institute for Communications Technology at Technische Universität Braunschweig. Currently, he is dealing with propagation and channel modeling in the field of Terahertz intra-device communications. He had several contributions to and is the editor of the CMD of IEEE 802.15 TG3d.



Sebastian Rey is pursuing his Ph.D. as a research assistant with the Institut für Nachrichtentechnik at Technische Universität Braunschweig. His work mainly focuses on propagation and system measurements and modelling at 300 GHz. He is a contributor to IEEE 802.15.3d and currently the editor of the evaluation criteria document of this group.



Akifumi Kasamatsu is a director of the terahertz and millimeter wave ICT laboratory at the National Institute of Information and Communications Technology (NICT), Tokyo and his research interests are in wireless communication technology such as wireless transceivers and nano-scale semiconductor devices for millimeter wave and terahertz communications. He is a contributor to IEEE 802.15.3d.



Norihiko Sekine is a Research Manager at the National Institute of Information and Communications Technology (NICT), Tokyo and his research interests include the physical properties of semiconductor nanostructures in the terahertz regime and their application to terahertz devices and systems. He is a contributor to IEEE 802.15.3d.

**Free Space Optical Communications:
Fundamentals, Applications and Recent Achievements**

Murat Uysal

*Ozyegin University, Istanbul, Turkey, 34794
murat.uysal@ozyegin.edu.tr*

1. Introduction

Optical wireless communication (OWC) [1-4] uses optical carriers, i.e., visible, infrared, or ultraviolet bands, for wireless transmission in unguided propagation media. OWC systems can be classified into short and long range systems, according to the distance between the transmitter and receiver. Short range systems are mainly for indoor use while long range systems are used for outdoor and space applications. Indoor and outdoor propagation environments are inherently differently from each other and pose unique challenges which need to be carefully taken into account for OWC system design. The focus of this paper is outdoor OWC links which are widely referred to as free space optical (FSO) communication in the literature.

FSO products with transmission rates of 10 Gb/s are already in the market [5] and the speeds of recent experimental FSO systems are competing with fiber optic [6, 7]. FSO systems use very narrow laser beams. This spatial confinement provides a high reuse factor, an inherent security against jamming, and robustness to electromagnetic interference. Furthermore, the frequency in use by the FSO technology is unlicensed worldwide. Therefore, FSO systems do not require license fees providing a cost-effective alternative to microwave counterparts. FSO systems are easily deployable and can be reinstalled without the cost of dedicated fiber optic connections. With such advantages, FSO can be used in a wide range of application areas. In this paper, we first provide an overview of FSO communications and then present application areas of this promising technology with huge potential. We also discuss some recent achievements including the ongoing large-scale projects by Facebook and Google using FSO technology.

II. Fundamentals of FSO Communications

In an FSO communication system, a source produces waveforms which are then modulated onto an infrared (IR) carrier. The generated optical field is radiated through the atmosphere towards a remote destination. In terrestrial applications, this can be several kilometers while in space applications the range might be on the order of thousands of kilometers. At the receiver, a photodetector transforms the optical field to an electrical current. The receiver processes the detected electrical current to recover the original transmitted information.

Current FSO systems typically operate in the near-IR wavelengths ranging from 750 to 1600 nm. In the near-IR band, certain wavelengths experience severe absorption due to the presence of different molecules in the atmosphere [8]. There are some atmospheric windows, located around 850, 1060, 1250, and 1550 nm where an attenuation of less than 0.2 dB/km is experienced [9]. It is worth noting that 850 and 1550 nm wavelengths coincide with the standard transmission windows of fiber communication systems. Commercially available FSO systems also prefer these two wavelengths to take advantage of the already available off-the-shelf components.

A typical FSO transmitter consists of an optical source, a modulator, an optical amplifier, and beamforming optics. The input bits are first modulated with an intensity modulation technique such as on-off keying (OOK) or pulse position modulation (PPM). Channel coding can be optionally used before modulation if required. To boost the optical intensity, the modulated laser beam is passed through the optical amplifier. FSO systems typically use semiconductor laser diodes (LDs) [10] as light sources. Vertical-cavity surface-emitting lasers (VCSEL) are mostly used for operation at 850 nm while Fabry-Perot (FP) and distributed feedback (DFB) lasers are preferred for operation at 1550 nm. As alternative to

LDs, high power light emitting diodes (LEDs) with beam collimators can be also used [11].

The use of laser transmitters is regulated by safety standards. These standards limit the transmitted optical power for eye-safety [12]. In fact, only certain wavelengths in the near-IR range can penetrate the eye with enough intensity resulting in retina damage. Other wavelengths are absorbed by the front part of the eye. It is known that the absorption coefficient at the front part of the eye is much higher for longer wavelengths (> 1400 nm). Therefore, the allowable transmission power for lasers operating at 1550 nm is higher, and hence, they are considered for longer distance transmissions.

FSO receivers commonly use non-coherent detection. The receiver front-end consists of optical filters and a lens which has the role of collecting and focusing the received beam onto the photodiode (PD). The PD output current is converted to a voltage by means of a transimpedance circuit. The output of the transimpedance circuit is low-pass filtered in order to limit the thermal and background noise levels.

P-i-N (PIN) diode or an avalanche photodiode (APD) are used in practical FSO systems. PIN diodes are usually used for FSO links with ranges up to a few kilometers [13]. In FSO receivers with PIN PDs, the performance becomes limited by the thermal noise. For long range links, APDs are mostly preferred with their high current gains [14]. The advantage of APD comes at the expense of increased implementation complexity. In particular, a relatively high voltage is required for APD reverse biasing that necessitates the use of special electronic circuits. This also results in an increase in the receiver power consumption. In long range links, optical pre-amplifiers are also employed [15]. Erbium-doped fiber amplifier (EDFA) is a common choice at the 1550 nm wavelength. In direct detection receivers, an optical pre-amplifier can degrade the SNR by at least 3 dB [16]. Nevertheless, when the receiver performance is limited by the electronic noise, optical pre-amplification can be highly beneficial.

III. FSO Applications

FSO systems can be used for a wide range of applications [17–20]. The major application area is enterprise/campus connectivity. FSO systems can bridge multiple buildings in corporate and campus networks supporting ultra-high speeds without the cost of dedicated fiber optic connections. With their high-bandwidth capacity, they are also appealing for video surveillance and monitoring. Similarly, FSO links are capable of satisfying the most demanding throughput requirements of today's high definition television (HDTV) broadcasting applications. In broadcasting of live events such as sports and ceremonies or television reporting from remote areas and war zones, signals from the camera need to be sent to the broadcasting vehicle which is connected to a central office via satellite uplink. The required high-quality transmission between the cameras and the vehicle can be provided by an FSO link.

The potential use of FSO links for backhauling/fronthauling in cellular links has also attracted recent attention. Wireline connections such as T1/E1 leased lines and microwave links are typically deployed between the base stations and the mobile switching center in a cellular system. The growing number of bandwidth-intensive mobile phone services now requires the deployment of technologies such as FSO which allow much higher throughputs [21].

FSO links are particularly useful for deployment as redundant links in disaster recovery and emergency situations. Temporary FSO links can be readily deployed within hours in such disaster situations in which local infrastructure could be damaged or unreliable.

While earlier uses of FSO links were mainly for fixed installations, it is possible to establish such links in mobile applications given that reliable pointing-acquisition-tracking algorithms are designed. This enables the deployment of FSO links for aircraft-to-aircraft, aircraft-to-ground, aircraft-to-high altitude platforms (HAPs). Such uses of FSO are particularly useful in tactical field and research is pursued in this

direction by military organizations and defence companies.

FSO can be used as a powerful ultra-long connectivity solution (for more than 10000 kilometers) for ground-to-satellite and satellite-to-satellite communications as well as intraplanet communication. One of the major milestones in this area took place in 2001 when a 50 Mb/s FSO link was successfully established between ARTEMIS geostationary satellite and the SPOT-4 French Earth observation satellite in sun-synchronous low earth orbit. With the introduction of coherent modulation techniques, data rates on the order of Gbps were further achieved. It is expected that OWC will continue to be a major enabling technology in space and satellite links.

III. Recent Achievements

There are several companies which are currently active in FSO market including Canon (Japan), Cassidian (Germany), fSONA (Canada), GeoDesy (Hungary), Laser ITC (Russia), LightPointe Communications (USA), MRV (USA), Northern Hi-Tec (UK), Novasol (USA), Omnitek (Turkey), Plaintiff Systems (Canada), and Wireless Excellence (UK) among others. The commercially available FSO systems offer transmission rates up to 10 Gb/s. Recent experimental demonstrations target transmission rates on the order of Tb/sec. For example, a rate of 1.6 Tb/s was reported in [22]. The implemented system exploits polarization multiplexing and wavelength division multiplexing (WDM). QPSK modulation and coherent detection are used. The system allows enough power budget to support the record transmission of 16 channels, operating each at 100 Gb/s over 40 km of fiber and 80 m of FSO between two buildings. In another demonstration [23], 100 Tb/s was achieved over a short distance (1 m). Multiplexing/demultiplexing of 1008 data channels using orbital angular momentum (OAM) mode-division-multiplexing combined with WDM were used in this experimental link. 24 OAM modes, each carrying 42 WDM channels, and each wavelength transporting a 100 Gb/s QPSK signal, were multiplexed, providing an aggregated capacity of 100.8 Tb/s.

In 2014, another record was achieved by NASA [24] which was able to transmit data across the 384,633km distance between Earth and the LADEE satellite orbiting the Moon. Record breaking data download rate was reported at 622 Mb/s. The system uses four separated laser transmitters resulting in 40 W of transmitter power.

FSO technology is also deployed in ongoing projects *internet.org* [25] and *Loon* [26] led by Facebook and Google. Internet.org aims to provide affordable access to less developed countries. The project plans to use different vehicles to deliver Internet to different types of locals. Less populated areas will be served by low-Earth orbit and geosynchronous satellites. In suburban areas it will use solar-powered drones. Facebook plans to use FSO among drones and satellite links. This will basically form a mesh network in the sky which is able to distribute high capacity data streams, similar to fibre optic networks through the air and between aerial platforms. In Loon project led by Google, high-altitude balloons, instead of drones, are used to build a network 20 miles above the earth. It was reported that a 155 Mb/s optical connection between two balloons more than 100 kilometers apart was established. Ongoing work in the same project aims to further improve the data rate to the range of Gb/s.

4. Conclusions

In this paper, we have provided an overview of FSO communications and discussed several potential areas. FSO has been treated as a niche technology in the past. Thanks to the recently launched projects from Facebook and Google, FSO is now making headlines and its advantages of this promising technology will be better understood leading to a wider deployment in near future.

Acknowledgements

This work is carried out as an activity of “Centre of Excellence in Optical Wireless Communication Technologies (OKATEM)” funded by Istanbul Development Agency (ISTKA) under Innovative Istanbul Financial Support Program, 2015. The statements made herein are solely the responsibility of the authors and do not reflect the views of ISTKA and/or T.R. Ministry of Development.

References

- [41] S. Arnon, J. R. Barry, G. K. Karagiannidis, R. Schober, and M. Uysal (Eds.), *Advanced Optical Wireless Communication*, Cambridge University Press, June 2012.
- [42] M. Uysal, C. Capsoni, Z. Ghassemlooy, A.C. Boucouvalas, E. G. Udvarý (Eds.), *Optical Wireless Communications – An Emerging Technology*, Springer, 2016.
- [43] M. Uysal and H. Nouri, “Optical Wireless Communications – An Emerging Technology”, Invited Paper, 16th International Conference on Transparent Optical Networks (ICTON), Graz, Austria, July 2014.
- [44] Z. Ghassemlooy, S. Arnon, M. Uysal, Z. Xu, and J. Cheng, “Emerging Optical Wireless Communications-Advances and Challenges,” *IEEE Journal on Selected Areas in Communications*, vol.33, no.9, pp.1738-1749, September 2015.
- [45] D. Rodewald, “MRV introduces industry’s first 10G ethernet wireless point-to-point system,” MRV Communications, Inc., Chatsworth, CA, USA, 2008.
- [46] M.-C. Jeung et al., “8 × 10-Gb/s terrestrial optical free-space transmission over 3.4 km using an optical repeater,” *IEEE Photon. Technol. Lett.*, vol. 15, no. 1, pp. 171–173, Jan. 2003.
- [47] E. Ciaramella et al., “1.28-Tb/s (32 × 40 Gb/s) free-space optical WDM transmission system,” *IEEE Photon. Technol. Lett.*, vol. 21, no. 16, pp. 1121–1123, Aug. 2009.
- [48] S. Bloom, E. Korevaar, J. Schuster, and H. Willebrand, “Understanding the performance of free-space optics,” *J. Opt. Netw.*, vol. 2, no. 6, pp. 178–200, Jun. 2003.
- [49] A Comparison of FSO Wavelength System Designs, LightPointe White Paper, 2002. [Online]. Available: <http://www.lightpointe.com/support/whitepapers.html>
- [50] J. M. Kahn and J. R. Barry, “Wireless infrared communications,” *Proc. IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.
- [51] Plaintree Systems Inc., *Free Space Optics, Point-to-Point Wireless Connectivity*, Annapolis, ON, Canada. [Online]. Available: <http://freespaceoptics.ca/about.html>
- [52] G. Clark, H. Willebrand, and B. Willson, *Free Space Optical Laser Safety*, LightPointe White Paper. [Online]. Available: <http://www.lightpointe.com/support/whitepapers.html>
- [53] Terescope Product Series, MRV website. [Online]. Available: <http://www.mrv.com>
- [54] K. Kiasaleh, “Performance of APD-based, PPM free-space optical communication systems in atmospheric turbulence,” *IEEE Trans. Commun.*, vol. 53, no. 9, pp. 1455–1461, Sep. 2005.
- [55] M. Razavi and J. H. Shapiro, “Wireless optical communications via diversity reception and optical preamplification,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 975–983, May 2005.
- [56] S. B. Alexander, *Optical Communication Receiver Design* Bellingham, WA, USA: SPIE, 1997.
- [57] V. W. S. Chan, “Free-space optical communications,” *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4750–4762, Dec. 2006.
- [58] H. A. Willebrand and B. S. Ghuman, “Fiber optics without fiber,” *IEEE Spectr.*, vol. 40, no. 8, pp. 41–45, Aug. 2001.
- [59] E. Leitgeb et al., “Current optical technologies for wireless access,” in *Proc. Int. ConTEL*, Zagreb, Croatia, Jun. 2009, pp. 7–17.
- [60] D. Kedar and S. Arnon, “Urban optical wireless communication networks: The main challenges and possible solutions,” *IEEE Commun. Mag.*, vol. 42, no. 5, pp. 2–7, May 2004.
- [61] M. Alzenad, M. Z. Shakir, H. Yanikomeroglu, and M.-S. Alouini, “FSO-based vertical backhaul/fronthaul framework for 5G+ wireless networks”, to appear in *IEEE Communications Magazine*

- [62] G. Parca, A. Shahpari, V. Carrozzo, G. M. Tosi Beleffi, and A. L. J. Teixeira, "Optical wireless transmission at 1.6-Tbit/s (16×100 Gbit/s) for next-generation convergent urban infrastructures", *Opt. Eng.* 52 (11), 116102 (Nov 05, 2013).
- [63] H. Huang, G. Xie, Y. Yan, N. Ahmed, Y. Ren, Y. Yue, D. Rogawski, M. Tur, B. Erkmen, K. Birnbaum, S. Dolinar, M. Lavery, M. Padgett, and A. E. Willner, "100 Tbit/s free-space data link using orbital angular momentum mode division multiplexing combined with wavelength division multiplexing, " in *Optical Fiber Communication Conference/National Fiber Optical Engineers Conference 2013*, Anaheim - USA, paper OTh4G.5, 17-21 September 2013
- [64] NASA Lunar Laser Communications Demonstration (LLCD) webpage [Online]
<http://esc.gsfc.nasa.gov/267/271.html>
- [65] <https://info.internet.org/en/>
- [66] <http://www.google.com/loon/>



Murat Uysal is currently a Full Professor and Chair of the Department of Electrical and Electronics Engineering at Ozyegin University, Istanbul, Turkey. Prior to joining Ozyegin University, he was a tenured Associate Professor at the University of Waterloo, Canada, where he still holds an adjunct faculty position. Dr. Uysal's research interests are in the broad areas of communication theory and signal processing with a particular emphasis on the physical layer aspects of wireless communication systems in radio, acoustic and optical frequency bands. He has authored some 250 journal and conference papers on these topics and received more than 5000 citations.

Prof. Uysal currently serves as the Chair of IEEE Turkey Section. He lead the EU COST Action OPTICWISE which is a European scientific network for interdisciplinary research activities in the area of optical wireless communications. He serves on the editorial boards of IEEE Transactions on Communications and IEEE Transactions on Wireless Communications. In the past, he served as an Editor for IEEE Communications Letters, IEEE Transactions on Vehicular Technology, Wiley Wireless Communications and Mobile Computing (WCMC) Journal, Wiley Transactions on Emerging Telecommunications Technologies (ETT), and Guest Editor of IEEE JSAC Special Issues on Optical Wireless Communication (2009 and 2015). He was involved in the organization of several IEEE conferences at various levels. He served as the Chair of the Communication Theory Symposium of IEEE ICC 2007, Chair of the Communications and Networking Symposium of IEEE CCECE 2008, Chair of the Communication and Information Theory Symposium of IWCMC 2011, TPC Co-Chair of IEEE WCNC 2014 and General Chair of IEEE IWOW 2015. Over the years, he has served on the technical program committee of more than 100 international conferences and workshops in the communications area. His distinctions include the Marsland Faculty Fellowship in 2004, NSERC Discovery Accelerator Supplement Award in 2008, University of Waterloo Engineering Research Excellence Award in 2010, Turkish Academy of Sciences Distinguished Young Scientist Award in 2011 and Ozyegin University Best Researcher Award in 2014 among others.

Research on Visible Light Communication in Korea

Nam-Tuan Le and Yeong Min Jang

Kookmin University

Seoul, Korea

namtuanlnt@gmail.com, yjang@kookmin.ac.kr

1. Introduction

The development of LED with the possibility of high-speed switching is one of important factor of Visible Light Communication (VLC) success. During the last decade it has been achieving a big change in wireless communication. The research on VLC is spread out from material design to application service. The specification of VLC which is standardized as IEEE 802.15.7 [1] in 2011 is one full reference of VLC application deployment. However the research issues of VLC still going on especially on the performance enhancement. To the best of our knowledge, this article provides an overview of the activities of Visible Light Communication research in Korea. The research issues will be discussed briefly on the target goal of Korea VLC. Based on the provided information, we will discuss some trend issues of visible light communication on standardization and business trend in the last section.

2. Research on Visible light communication in Korea

From early of the 2000s, Visible Light Communication is one of the challenges issues in Korea. The research on visible light communication has focused on application service, localization, modulation, dimming control, MIMO diversity, and access control. The research is classified by two directions based on receiver's architecture: photodiode and image sensor.

With photodiode visible light communication system, the research issues include:

- For dimming control, it is one important issue of VLC with illumination advantage comparison with RF system. This issue which is analyzed from [18] is one of the main parts of IEEE 802.15.7 specification. The proposed scheme from "Color and Brightness Control of RGB LEDs" [2] used Pulse width modulation and variable pulse position modulation on RGB component to control the brightness. Every color is perceived by the human eye by a ratio of three basic RGB color. By applying PWM on controlling on/off period of each RGB light source, the proposed scheme can generate modulated color. The proposed dimming control in [17] was a new forward error correction (FEC) coding method using a coset constructed from a bent function and a Reed-Muller (RM) code. It can provide accurate dimming control in on-off keying-based visible light communication systems.
- For MIMO technique, the research of "Spatial Multiplexing in MIMO" [3], [4] represents the performance of MIMO visible light communication system on image sensor array. The spatial multiplexing of MIMO technique is applied to improve the throughput performance.
- For application service, the proposed system of LED-ID from "Development of Home Network Technology based on LED-ID" project [5], [6] was a deployment of IEEE 802.15.7 specification for ID applications. The operation of LED-ID includes two phases: First one is the identification of LED-ID tag; second is data transmission. The research also considered link switching and link recovery for movement topology of Tag and Reader. The scenario topology of LED-ID service is mainly focused on spatial diversity so the collision control is considered by a new protocol. The receiver will initial the connection and error control.
- For localization issue, the proposed scheme of "Carrier allocation VLC" [9], [10] used received signal strength indication (RSSI) for distance estimation and reference coordinator [11]. The reference coordinators are three LEDs array with different subcarriers. The scheme can prevent flickering and inter-cell interference from different transmitters.

- For error correction issue, the suggested scheme of “soft-input soft-output run-length limited (RLL) decoder” [12] for on-off keying (OOK) modulation and Reed–Solomon (RS) codes can enhance the performance of bit error rate (BER) and frame error rate (FER) performances VLC systems.
- For link connection control, the work on link recovery and link switching [13], [14] focus on the efficient mechanism for connection management. By predicting the link connection based on movement direction or link handshake protocol between a mobile terminal and VLC Aps, the result of proposed schemes showed good performance on delay connection or link QoS.
- For modulation issue, the research of VLC OFDM technique [15], [16] shows the performance of OFDM on VLC system in different configuration and interference factor. With the possibility of inter-symbol interference avoidance, VLC OFDM will be a promising technique for high-speed VLC system.

About Visible Light Communication based image sensor, it is a new issue which is considered as an extension of IEEE 802.15.7 specification. Up to the current development, image sensor technique is classified as rolling shutter and global shutter technique. Image sensor Visible light communication is also called as Optical camera communication (OCC) because of the development of camera and smart devices. These are different from photodiode case. The research mainly focuses on application, modulation, synchronization and localization.

- For modulation approach, the proposed schemes on frequency subcarrier and color shift modulation are core research for OCC [19], [20], [21]. Frequency subcarrier modulation is applied for a non-flickering scenario which is usually for indoor full illumination application. Color shift modulation is for flickering scenario in vehicle communication, data broadcasting scenario. The performance of two modulation schemes on rolling shutter and global shutter is different. Rolling shutter image sensor can support high data rate compare with global shutter. However global shutter has the advantage of distance and MIMO topology.
- For synchronization issue, the research on the carrier frequency, OOK and color shift [22], [20] specify the limitation of image sensor characteristics for OCC: shutter speed, frame variation and exposure time.
- For localization issue, the research of “Localization based Image sensor communication” project [7], [8], [9] suggested photogrammetry calculation from image data using computer vision technique and reference coordinate.
- For application service, the proposed scenario and topology of vehicle communication, seaside communication or data broadcasting [21], [8], [23]. Depending on application scenario, the image sensor type should be considered based on data rate, communication distance or MIMO technique.

3. Conclusion and Discussion

With the development of image sensor technology, human demand and mobile system, Optical Camera Communication will be a promising service in near future. It should be considered as one candidate for revision specification of IEEE visible light communication specification.

References

- [1] IEEE, IEEE standard for local and metropolitan area networks - Part 15.7: Short-Range Wireless Optical Communcation Using Visible Light, IEEE 802.15.7, Sep. 2011.
- [2] Kyungmook Choi, Yunseon Jang, MinChul Ju, and Youngil Park, “Visible Light Communication with Color and Brightness Control of RGB LEDs,” ETRI Journal, vol. 35 no. 5, pp. 927-930, Oct. 2013.
- [3] Ratan Kumar Mondal, Nirzhar Saha, Yeong Min Jang, “Performance Analysis of Spatial Multiplexing in MIMO Based Visible Light Communication System,” The Journal of Korea Information and Communications Society, vol. 38C no. 09, pp. 797-801, Sep. 2013.

- [4] Atul Sewaiwar, Phyu Phyu Han, Yeon Ho Chung, "3-Gbit/s Indoor Visible Light Communications Using Optical Diversity Schemes," IEEE Photonics Journal, vol. 7 no. 6, Dec. 2015.
- [5] Nam-Tuan Le and Yeong Min Jang, "Broadcasting MAC Protocol for IEEE 802.15.7 Visible Light Communication," in Proceeding of 2013 Fifth International Conference on Ubiquitous and Future Networks, pp. 667-671, July 2013.
- [6] Nam-Tuan Le and Yeong Min Jang, "Non-Link Switching Issue for Broadcasting Visible Light Communication," in Proceeding of 2015 International Conference on Information Networking, pp. 238-241, Jan. 2015.
- [7] Md Shareef Ifthekhar, Ratan Kumar Mondal, Nam-Tuan Le, Yeong Min Jang, "Simple method for indoor localization in OCC using smart phone image sensor," in Proceeding of 2014 Sixth International Conf on Ubiquitous and Future Networks, pp. 55-58, July 2014.
- [8] Shareef Ifthekhar, Nirzhar Saha, Yeong Min Jang, "Stereo-vision-based cooperative-vehicle positioning using OCC and neural networks," Optics Communications, vol.352, pp. 166-180, Oct. 2015.
- [9] H.-S. Kim, D.-R. Kim, S.-H. Yang, Y.-H. Son, and S.-K. Han, "Inter-cell interference mitigation and indoor positioning system based on carrier allocation visible light communication," 5th International Conference Signal Processing Communication System, pp. 1-7, Hawaii, U.S.A., Dec. 2011.
- [10] S. Jung, S. Hann, and C. Park, "TDOA-based optical wireless indoor localization using LED ceiling lamps," IEEE Transactions on Consumer Electronic, vol. 57 no. 4, pp. 1592-1597, Nov. 2011.
- [11] Tuan Nguyen and Yeong Min Jang, "Highly Accurate Indoor Three-Dimensional Localization Technique in Visible Light Communication Systems," The Journal of Korea Information and Communications Society, vol. 38C no. 09, pp. 775-780.
- [12] He Wang and Sunghwan Kim, "Soft-Input Soft-Output Run-Length Limited Decoding for Visible Light Communication," IEEE Photonics Technology Letters, vol. 28 no. 3, Feb. 2016.
- [13] Tuan Nguyen, Yeong Min Jang, Mostafa Zaman Chowdhury, "A pre-scanning-based link switching scheme in visible light communication networks," in Proceeding of 2013 Fifth International Conference on Ubiquitous and Future Networks, pp. 366-369, July 2013.
- [14] Le Nam-Tuan, Choi Sun-Woong, Jang Yeong-Min, "Link Recovery Scheme Using Cooperative Communication for VLC System, " The Journal of Korean Institute of Communications and Information Sciences, vol. 37 no. 5B, pp. 357-364, May 2012.
- [15] Nirzhar Saha, Ratan Kumar Mondal, Nam Tuan Le, Yeong Min Jang, "Mitigation of interference using OFDM in visible light communication," in Proceeding of 2012 International Conference on ICT Convergence, pp. 159-162, Oct. 2012.
- [16] Kapila Bandara, P.Niroopan, Yeon-Ho Chung, "PAPR reduced OFDM visible light communication using exponential nonlinear companding," in Proceeding of 2013 IEEE International Conference on Microwaves, Communications, Antennas and Electronics Systems, pp. 1-5, Oct. 2013.
- [17] Sunghwan Kim and Sung-Yoon Jung, "Modified Reed-Muller Coding Scheme Made From the Bent Function for Dimmable Visible Light Communications," IEEE Photonics Technology Letters, vol. 25 no. 1, pp.11-13, Jan. 2013.
- [18] Sridhar Rajagopal, Richard D. Roberts, Sang-Kyu Lim, "IEEE 802.15.7 Visible Light Communication: Modulation Schemes and Dimming Support," vol. 50 no. 3, pp.72-82, Mar. 2012.
- [19] Nam-Tuan Le, Trang Nguyen, Yeong Min Jang, "Frequency shift on-off keying for optical camera communication," in Proceeding of 2014 Sixth International Conf on Ubiquitous and Future Networks, pp. 22-25, July 2014.

- [20] Trang Nguyen, Nam Tuan Le, Yeong Min Jang, "Practical design of Screen-to-Camera based Optical Camera Communication," in Proceeding of 2015 International Conference on Information Networking, pp. 369-374, Jan. 2015.
- [21] Trang Nguyen, Nam Tuan Le, and Yeong Min Jang, "Asynchronous Scheme for Optical Camera Communication-Based Infrastructure-to-Vehicle Communication," International Journal of Distributed Sensor Networks, vol. 2015, Jan. 2015.
- [22] Nam-Tuan Le, Thithanhnhhan Le, Trang Nguyen, Yeong Min Jang, "Synchronization issue for optical camera communications," in Proceeding of 2015 Seventh International Conference on Ubiquitous and Future Networks, pp. 220-224, July 2015.
- [23] Trang Nguyen, Chang Hyun Hong, Nam Tuan Le, and Yeong Min Jang, "High-speed Asynchronous Optical Camera Communication using LED and Rolling Shutter Camera," 2015 Seventh International Conference on Ubiquitous and Future Networks, pp. 214-219, July 2015.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (no. 2013057922).



Nam-Tuan Le received diploma program from HaNoi University of Science and Technology, Viet Nam in 2005. He completed his MS and Ph.D in Electronics Engineering from Kookmin University, Korea in 2011 and 2014. His research interests are sensor networks, visible light communication. Since 2014, he has been with the School of Electrical Engineering, Kookmin University, Seoul, Korea.



Yeong Min Jang received the B.E. and M.E. degrees both in electronics engineering from Kyungpook National University, Korea, in 1985 and 1987, respectively. He received the doctoral degree in computer science from the University of Massachusetts, USA, in 1999. He worked for ETRI (Electronics and Telecommunications Research Institute) between 1987 and 2000. Since September 2002, he is with the School of Electrical Engineering, Kookmin University, Seoul, Korea. He had been the director of LED Convergence Research Center at Kookmin University since 2010. He has organized several conferences and workshops such as

International Conference on Ubiquitous and Future Networks (ICUFN), International Conference on ICT Convergence (ICTC), International Conference on Information Networking, and International Workshop on Optical Wireless LED Communication Networks. Currently he is the vice president of KICS (Korean Institute of Communications and Information Sciences). He serves as the Co-Editor-in-Chief for ICT Express which is published by Elsevier. He had been the steering committee chair of Multi Screen Service Forum since 2011 and the steering committee chair of Society Safety System Forum since 2015. Currently, he is serving as the Chairman of IEEE 802.15.7r1 Optical Wireless Communications Task Group. His research interests include 5G mobile communications, optical wireless communications, optical camera communication (OCC), public safety, and internet of things (IoT).

Let there be Light Again!

An Amendment to IEEE 802 Visible Light Standard is in Progress

*Tunçer Baykaş¹, Morteza Soltani¹, Mohamed Siala,² Trang Nguyen³, Yeong Min Jang³
¹Istanbul Medipol University, Turkey, ²SUP'Com,tunisia, ³Kookmin University, Korea
tbaykas@medipol.edu.tr*

1. Introduction

Visible Light Communications (VLC) use the visible spectrum (wavelengths between 390 and 750 nm or frequency band of 400-790 THz) and can provide wireless communications using illumination and display elements. From the ancient times to the 19th century, all VLC communication systems were relying on the human eye as the receiver. The invention of the Photophone by Alexander Graham Bell and Charles Sumner Tainter changed the nature of VLC communications. They used the fact that selenium resistance varies with respect to light intensity and used this property by connecting it to a phone receiver in order to send audio signals. Many improvements have been achieved on these systems until the 1950s, however most of the materials used for detection have higher sensitivity to infra-red radiations, hence precluding visible light to be used as a transmission medium.

The introduction of light-emitting diodes (LED) created a new interest for the use of visible light communications. More specifically, the introduction of GaN LEDs [1] and white light-emitting phosphors [2] provided visible light sources which can be modulated at higher speeds, without sacrificing their main illuminating role. In 2004, the first high-speed communication demonstrations with LEDs were made in Japan, using photodiodes. On the other hand, the proliferation of cellular phones with cameras, enabled them to be used as VLC receivers. Researchers started using LCD screens and other display elements as transmitters.

One of the first standardization bodies to work on a VLC standard was the Visible Light Communications Consortium (VLCC) of Japan. They expanded the IrDA standard for infrared communications to the visible light spectrum in 2008.

In the year of 2011, IEEE 802 LMSC published the IEEE Standard 802.15.7 for Short-Range Wireless Optical Communication Using Visible Light [3,4]. Due to the growing interest in visible light communications, in 2014, the IEEE standardization association approved another project authorization request from IEEE 802 to amend the previous standard for a faster, better and more application enabling standard. In this paper we summarize current activities in IEEE 802 for the new amendment.

2. IEEE 802.15.7r1 Amendment

In 2013, the IEEE 802.15 working group (WG) decided to investigate if an amendment to the 802.15.7 standard is necessary by establishing a study group. The discussions inside the group and suggestions from industry and academia led to a project authorization request [5], which states: "This standard defines a Physical (PHY) and Media Access Control (MAC) layer for short-range optical wireless communications in optically transparent media... The standard is capable of delivering data rates sufficient to support audio and video multimedia services and also considers mobility of the optical link, compatibility with various light infrastructures, impairments due to noise and interference from sources like ambient light and a MAC layer that accommodates the unique needs of visible links as well as the other targeted light wavelengths. It also accommodates optical communications for cameras where transmitting devices incorporate light emitting sources and receivers are digital cameras with a lens and image sensor. The standard adheres to applicable eye safety regulations...". The acceptance of the project authorization request by the IEEE Standards Association enabled the IEEE 802.15 WG to create a Task Group (802.15.7r1 TG) to work on an amendment, which is open to almost any type of VLC communications. According to the timeline of the group shown in Fig. 1, the 802.15.7r1 TG will collect

proposals in 2016 and finalize the amendment before 2018. In Fig. 1, CFA, TRD and CFP stand for call for applications, technical requirement documents and call for proposals, respectively.

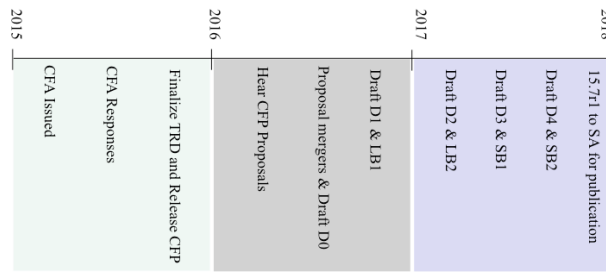


Figure 1. IEEE 802.15.7r1 Tentative Timeline.

A technical requirements document is prepared to guide prospective standard proposals [6]. The document uses the term “Optical Wireless Communication (OWC)” and classifies OWC into three modes:

- Image Sensor Communications (IMC).
- Low Rate Photodiode Receiver Communications (LR-PC).
- High Rate Photodiode Receiver Communications (HR-PC).

In regards to the definition of low speed and high speed, the throughput threshold data rate is 1 Mbps as measured at the Physical Layer output of the receiver. Throughputs less than 1 Mbps rate are considered low rate and higher than 1 Mbps are considered high rate. The WG determined the possible applications that can be served by each communication type. Regarding eye safety, the document states that the modulated light that can be seen by the human eye shall be safe in regards to the frequency and intensity of light (e.g., IEC 60825-1:2014) and the modulated light will not stimulate sickness, such as photosensitive epilepsy.

Image Sensor Communications (IMC)

IMC enable optical wireless communications using an image sensor as a receiver. The main applications of IMC are:

- Marketing/Public Information Systems.
- Internet of Things.
- Location-Based Services / Indoor Positioning.
- Vehicular Communications.
- Underwater Communications.
- Power Consumption Control.
- Vehicular Positioning.
- Seaside Communications.
- LED based tag applications.
- Point-to-(multi)point / relay/ communications.
- Digital signage.

The requirements to be observed by the IMC PHY layer are listed as: dimming control, power consumption control, coexistence with ambient light, coexistence with other lighting systems, simultaneous communication with multiple transmitters and multiple receivers (MIMO), nearly point image data source, identification of modulated light sources, low overhead repetitive transmission, image sensor compatibility and localization.

For MIMO communications, a MIMO MAC protocol may be incorporated into the standard so that the camera enabled receiving device knows how to process the received data. IMC will support at least one PHY mode that works when the light source appears as nearly a point source; i.e., the light source illuminates only a small number of image pixels. IMC will support at least one MAC mode that supports repetitive informational broadcast at very low data rates; that is, the frame format should present a very small overhead and be optimized for short payloads sent in a repetitive manner. Regarding image sensor compatibility, IMC will support a PHY mode that is compatible with a variety of cameras with different image sensing sampling rates (read-out time), resolutions and frame rates. Specifically, either constant frame rate or varying frame rate will be supported.

Low Speed Photodiode Communications (LS-PC).

Low Speed Photodiode Receiver communications require LEDs as transmitters and low speed photodiodes as receivers. The main applications are:

- Underwater/Seaside Communications.
- Point-to-(multi)point communications.
- Digital signage.
- Internet of Things.
- LOS Authentication.
- Identification based services.

LS-PC should support the LED Tags and the Smart Phone Flash lights as transmitters. The standard will support low speed photodiode receivers, which measure intensity of the visible light, IR and/or near UV. It may provide mechanisms to support handover between LED light sources, allowing the users to maintain a continuous network connection.

LS-PC may provide mechanisms that can be used to develop and deliver interference coordination techniques by higher layers and may support link recovery mechanisms to maintain connection in unreliable channels and reduce connectivity delays.

High Speed Photodiode Communications (HS-PC)

The use of high speed photodiode receivers will enable high-speed, bidirectional, networked and mobile wireless communications. The task group is planning to create a PHY mode within high speed photodiode receivers that supports peak data rates of 10 Gbps at the PHY Service Access Point (SAP). The main applications of this mode are:

- Indoor Office/Home Applications: (Conference Rooms, Shopping Centers, Museums, etc.)
- Data Centers / Industrial Establishments, Secure Wireless (Manufacturing Cells, Factories, etc.)
- Vehicular Communications.
- Wireless Backhauling (Small Cell Backhauling, Surveillance Backhauling, LAN Bridging).

In HS-PC, continuous data streaming for all applications will be supported with a bidirectional functionality as well as short packet transmissions where low latency is required. A MAC/PHY mechanism to support adaptive transmission as well as multiple users communicating with different data streams from the same light source (multiple access) will be included. HS-PC will provide mechanisms to support horizontal handover between light sources, allowing the users to maintain a continuous network connection for applications and providing efficient mechanisms that can be used to deliver interference coordination techniques by higher layers. HS-PC will co-exist with ambient lights. This may enable a receiver to communicate with a supported transmitter even in the presence of other modulated lights. In addition, coexistence shall be investigated with the existing IEEE802.15.7-2011 operating modes. HS-PC

may support MIMO communications and cooperative signal processing among multiple transmitters with negligible impact on latency. HS-PC will support efficient and reliable feedback and control channels. These may be used for adaptive transmission, multiple user support, MIMO support, cooperative signal processing or other features. HS-PC may provide internal metrics via an open interface. This information may be used to support cooperative signal processing, vertical handover and link aggregation with other wireless transmission techniques. For this purpose, HS-PC may report the following metrics with minimized overhead and low latency:

- Information about instantaneous metrics, such as SINR and detailed channel state information.
- Information about recent history of the metrics, such as temporal characteristics, signal blocking, frequency of signal losses.

VLC Channel Models

The TG received many contributions on the channel models for all mentioned communication modes. It has agreed that proposers do not need to evaluate their submissions against a channel model to hear Technical Proposals for the Image Sensor and LS-PC modes. On the other hand for HS-PC, detailed channel models are provided by TG for 4 scenarios:

- Open Office/Office with Cubicles (Fig. 2).
- Office with Secondary Light (Fig. 2).
- Home (Fig. 3).
- Manufacturing Cell (Fig. 3)

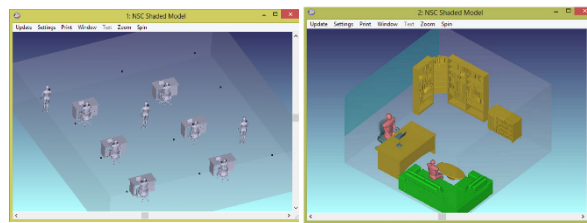


Figure 2. Office and Office with Secondary Light Scenarios

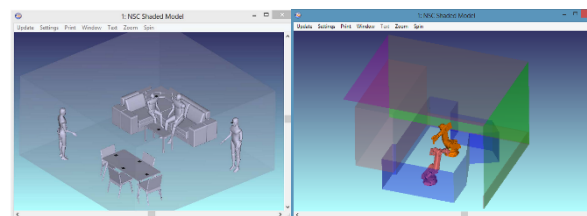


Figure 3. Home and Manufacturing Cell Scenarios

In an effort to develop more realistic channel models, a modeling approach based on ray tracing was used [7]. The proposed approach is based on Zemax®; a commercially available optical and illumination design software [8]. First, a three dimensional simulation environment is created where one can specify the geometry of the indoor environment, the reflection characteristics of the surface materials and the specifications of both light sources and detectors. The computer aided design (CAD) objects can further be imported in the simulation environment to model furniture and any other objects. The Zemax® non-sequential ray-tracing tool generates an output file, which includes all the data about rays, such as the detected power and path lengths for each ray. The data from the Zemax® output file is imported to Matlab® to deduce the resulting channel impulse response [8,9].

Observed channel impulse responses vary between an almost AWGN channel and channels with severe delay spreads as shown in Fig. 4.

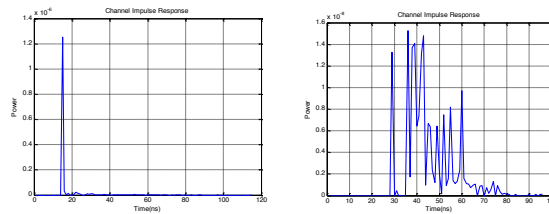


Figure 4. Channel Impulse Response Examples

3. Future of 802.15.7r1 Amendment

As of March 2016, the TG is hearing proposals for all communication modes. There is interest from both academia and industry. The group is planning to select and combine proposals to create the first draft for the amendment at the end of 2016.

Acknowledgements

The work of Tuncer Baykas is carried out as an activity of the “Centre of Excellence in Optical Wireless Communication Technologies (OKATEM)” funded by the Istanbul Development Agency under Innovative Istanbul Financial Support Program, 2015.

References

- [67] S. Nakamura, T. Mukai, and M. Senoh, “Candela Class High Brightness InGaN/AlGaN Double Heterostructure Blue Light Emitting Diodes,” *Applied Physics Letters*, vol. 64, no. 13, pp. 1687–1689, 1994.
- [68] J. S. Kim, *et al.*, “White-light Generation Through Ultraviolet-emitting Diode and White-emitting Phosphor,” *Applied Physics Letters*, vol. 85, no. 17, pp. 3696–3698, 2004.
- [69] IEEE Standard for Local and Metropolitan Area Networks--Part 15.7: “Short-Range Wireless Optical Communication Using Visible Light,” in *IEEE Std 802.15.7-2011*, vol., no., pp.1-309, Sept. 6, 2011.
- [70] R. D. Roberts, S. Rajagopal and S. K. Lim, “IEEE 802.15.7 physical layer summary,” *IEEE GLOBECOM Workshops*, pp. 772-776, Houston, TX, 2011.
- [71] The IEEE P802.15.7r1 Short-Range Optical Wireless Communications Task Group Project Authorization Request (PAR).
- [72] Jang Y. M., *et al.* “Technical Considerations Document,” IEEE 802 15-15-0492-03, 2015.
- [73] F. Miramirkhani, M. Uysal, E. Panayirci, “Novel channel models for visible light communications”. SPIE Photonics West, February 7-12, 2015.
- [74] Zemax® 13 Release 2, Radiant Zemax® LLC. www.radiantzemax.com/zemax
- [75] M. Uysal, *et al.* “TG7r1 Channel Model Document for High-rate PD Communications,”
- [76] M. Uysal, *et al.* “TG7r1 CIRs Channel Model Document for High-rate PD Communications,” Online: <https://mentor.ieee.org/802.15/dcn/15/15-15-0747-00-007a-tg7r1-cirs-channel-model-document-for-high-rate-pd-communications.zip>



Tuncer Baykas [SM] (tbaykas@ieee.org) works as an assistant professor at Istanbul Medipol University. He was the chair of IEEE 802.19.1 TG. He served as co-editor and secretary for 802.15 TG3c and contributed to many standardization projects, including 802.22, 802.11af and 1900.7. Currently he is the vice director of the “Centre of Excellence in Optical Wireless Communication Technologies (OKATEM)”. He contributed to the technical requirements document and the channel models of 802.15.7r1.



Morteza Soltani received his B.S. in electrical and computer engineering with highest honors from Ferdowsi University of Mashhad. He is now a research assistant and working towards his M.S. at Istanbul Medipol University, Istanbul, Turkey. His research interests broadly span the areas of wireless communications, signal processing and optimization.



Mohamed Siala is a Professor at Sup'Com, Tunis, Tunisia. His research interests are in the areas of wireless communications with special emphasis on OFDM with optimized pulse shaping, channel estimation, efficient synchronization algorithm and cognitive radio with efficient multiband sensing. He holds 230 journal and conference research papers and 6 patents on channel estimation for CDMA and OFDM.



Trang Nguyen received the B.E. degree from Hanoi University, Vietnam. Currently he works as a researcher of LED Convergence Research Center at Kookmin University. Also, he is a voting member of the IEEE 802.15.7r1 Task Group. His research areas include Biomedical Electronics, OCC, Localization, Digital Signage, and Augmented Reality services.



Yeong Min Jang received the doctoral degree in computer science from the University of Massachusetts, USA, in 1999. He worked for ETRI between 1987 and 2000. Since September 2002, he is with the School of Electrical Engineering, Kookmin University, Seoul, Korea. He had been the director of the LED Convergence Research Center at Kookmin University since 2010. Currently, he is serving as the Chairman of the IEEE 802.15.7r1 Task Group. His research interests include 5G mobile communications, optical wireless communications, OCC.

LiFi – Technological and Commercial Considerations

Nikola Serafimovski¹, Dobroslav Tsonev¹, Harald Burchardt¹, Stefan Videv¹, Harald Haas^{1,2}

¹pureLiFi Ltd, ²The Li-Fi R&D Center University of Edinburgh
Edinburgh, UK

{nikola.serafimovski, dobroslav.tsonev, harald.burchardt, harald.haas}@purelifi.com
{s.videv, h.haas} @ed.ac.uk

1. Introduction

One of the key drivers of the impending Spectrum Crunch is multimedia consumption.

Cisco has published their Global Mobile Data Traffic Forecast Update for 2015 – 2020 showing that global mobile data traffic grew 74% in 2015, experiencing a 4000 times growth in the last 10 years, where fourth-generation (4G) traffic exceeded third-generation (3G) traffic for the first time generating 53% of all network traffic [1]. Interestingly, mobile offload exceeded cellular traffic for the first time in 2015, with 51% of total mobile data traffic being offloaded onto the fixed network through Wi-Fi or femtocells in 2016 with a total of 3.9 Exabyte of mobile data.

More than half a billion mobile devices and connections were added in 2015 and mobile video traffic accounted for 55% of total mobile data traffic. Online streaming services such as YouTube, Netflix, Amazon Prime, etc. are the preferred choice for consumers aged 14 – 25 is streaming video with 72% citing it as one of the most valuable services [2]. Indeed, Netflix alone accounted for over 30% of all internet data in the US in 2015.

The estimates, according to Qualcomm, are that the next generation of wireless technologies need to achieve 1000 times greater energy efficiency and 1000 times greater wireless capacity.

In addition to the large amount of data going through the wireless networks to satisfy the growing demand for multimedia streaming, the Internet of Things is emerging. The expectation is that by 2020, there will be over 50 billion connected devices, each looking to use the unlicensed spectrum for connectivity.

This sustained growth in the demand for wireless capacity as well as the increasing density of wireless connections means that interference management is a key requirement for any system that hopes to address this issue.

At around 2000 prior to the introduction of 3G, the telecommunications industry pondered why base stations that deliver 2 Mbps are required, 'What is the killer app?' was the dominating question at that time.

Now, 15 years later, with a 400-million-fold increase in the mobile data traffic, WiGig is being introduced with maximum throughput of 7 Gbps [1]. The exponential Growth in demand for wireless internet has been identified by Cisco, Huawei, Qualcomm, the European Commission and others, with predictions of a sustained 60% compound annual growth rate (CAGR) in increasing demand.

If this trend continues, as all statistics indicate, then satisfying the demand for wireless data means that access points will be required to deliver 53 Gbps 2020. Indeed, improvements in technology will need to deliver access points that achieve 806 Gbps by 2025, unless a significant reduction in cell sizes is introduced.

Small cells are the key to enabling continued growth in wireless communications. Traditional Radio Frequency (RF) communications has limitations on the size of the small cells due to interference. Incorporating interference mitigation techniques is a key aspect of Qualcomm's strategy to achieve 1000 times greater efficiency (by 2020), however, this requires advanced algorithms that increase the cost and complexity of a small cell solution. High speed, networked and mobile wireless communications using light, or LiFi, enables both, high data rates from the access point and very small cells where interference is significantly more contained and predictable.

2. Optical wireless communications - LiFi

There is a clear difference between visible light communications (VLC) and LiFi. LiFi takes VLC further by using light emitting diodes (LEDs) to realize fully networked wireless systems. Synergies are harnessed as luminaries become LiFi attocells resulting in enhanced wireless capacity providing the necessary connectivity to realize the Internet-of-Things, and contributing to the key performance indicators for the fifth generation of cellular systems (5G) and beyond [3].

LiFi leverages the visible light spectrum that is 1000 times larger than the entire 300 GHz of RF spectrum. The visible light spectrum enables a huge, step-change growth in future data rates, future-proofing wireless networks for the continued exponential growth in demand for wireless data. Intensive system analyses show that LiFi attocellular networks will provide superior user data rates [4].

pureLiFi and the University of Edinburgh have shown that in the medium term LiFi light bulbs can achieve 100 Gbps, over 14 times higher than WiGig, reaching over 1 Tbps in the long term (1 Tbps) by exploiting the huge visible light spectrum resource [5]. This means LiFi technology will be able to deliver the predicted access point data rates. In addition, the atto-cell network of LiFi-enabled lights can deliver data density (Mbps / sqm) that is over 1000 times greater than state of the art wireless systems.

Combining the long term data rates of over 1 Tbps with the 1000 times greater data density of the overall system means that the problem of Spectrum Crunch will be resolved for the foreseeable future.

LiFi enables significantly more access points in a given space than RF/WiGig by utilising the lighting infrastructures. This results in a more uniform and easier accessible signal distribution (reduction of dead spots), and a massive scaling of user data rates, critically important for the IoT. The small cell concept has been the major contributor in cellular communications to achieving higher data rates and greater wireless capacity. Second generation systems (2G) typically exhibited 35 km cell radius in 2G (slightly reduced in the cities), 3G has approximately 5 km cell radius with around 1km in the cities, 4G has approximately 1 km to 500 m cell radius, and 5G is anticipated to have a 50 m cell radius. LiFi enables a 5 m cell radius.

Commercial LiFi system, such as the pureLiFi LiFi-X solution, offer data rates of over 40 Mbps for both the uplink and the downlink with an adjustable cell radius of less than 2 m [6].

Creating smaller cells with cell radius below 20 m is very difficult due to the propagation environment in RF [7]. This technical limitation, in combination with the increased cost of additional consumer devices and infrastructure, are some of the key reasons why femto-cellular networks have not experienced the massive uptake that was predicted.

The pursuit to limit the cell size has lead to systems operating in the mm spectrum (~300 GHz). However, mmWave communication requires antenna beamforming techniques to overcome the higher path loss. This means more complex antenna systems, potentially patchy coverage and compromised energy efficiency as multiple power amplifiers are required. Critically, the typical energy efficiency of a power

amplifier is below 50%. In addition, complex user tracking techniques are required to support user mobility, and the systems suffer significant communication degradations in non-line-of-sight conditions. To overcome this limitation, a large number of access points may be needed, however, there is a limitation to the number of access points that can be active in the same area due to interference and installation/backhaul cost.

In addition to increased wireless capacity, LiFi attocellular networks have the potential to be three orders of magnitude more energy efficient as they are much more efficient in serving multiple users in dense users/devices/things deployments.

With the coming of the IoT to prominence, as well as mounting environmental concerns, delivering an energy neutral communication system is crucial. LiFi has the potential to deliver on this front as well. Recent work [8]] has shown that solar panels can be used as signal detectors for LiFi, which enables simultaneous data communication and energy harvesting. With plentiful energy available for scavenging in the form of solar radiation, this is an enabler for the creation of simple and cheap IoT devices that can interface to LiFi networks without generating cost in terms of operational energy. Moreover, with tens of Mbps communication speeds possible [9], solar enabled LiFi is capable of delivering connectivity to remote and/or undeveloped areas of the world in a sustainable and infrastructure independent manner. This presents a solution that brings remote settlements into the connected world without the need for enormous capital expenditure. Finally, with the recent advances in flexible circuit technology and organic flexible solar cells, one can envision almost infinite applications of this technology. Virtually any surface be it solid or flexible could potentially be network connected via LiFi and serve as a sensor host or a smart agent.

3. Commercial Considerations

The technical considerations are only one aspect highlighting the advantages of LiFi in future deployments. There are also a number of key market trends that will facilitate the commercial adoption of LiFi. Enabling LiFi requires providing both power and data connectivity to each light. This connectivity can be delivered using power line communications (PLC) or Power over Ethernet (PoE).

The lighting industry, spearheaded by GE Lighting and Philips, with the support of network infrastructure providers is already delivering PoE enabled smart building solutions. A notable example is the Edge Building in Amsterdam, where Philips installed over 6500 PoE connected lights. The key here is that the lighting industry, independent of any requirements for LiFi, is already installing the required infrastructure.

The lighting industry is also experiencing a change in the underlying market metric. The transition from dollars per lightbulb is currently going through monetizing dollars per lux (or dollars per watt). However, this business model is also saturating and the focus is now on monetizing dollars per bit – putting lighting in the center for the developing IoT. Philips, ABB and Bosch have partnered with Cisco to enable connectivity in their Smart Building solutions. Sony and Toshiba are introducing a new all-inclusive, Wi-Fi connected smart-light similar to Apple that contains a host of sensors (occupancy, humidity, temperature, smoke detector, etc.), even a Bluetooth connected loudspeaker that would activate when it detects movement to a configuration relevant for the person near the device.

These developments point to a fundamental shift in the role that Lighting will play in the future. The integration of these sensors and the lifetime of LEDs means that the key value generation for the Lighting industry will be focused on data – learning from and predicting user behaviour, while providing access and localized services.

This creates a fundamentally new business model for the lighting industry that derives from the same market metric of dollars per bit. This is fundamentally the same market metric as Amazon, Google, Apple, and other technology giants in the Content creation industry as well as the Telecoms and network providers such as T-Mobile. LiFi provides a clear advantage for the lighting industry as it strives to ensure its place in the new market space.

4. Conclusion

In this paper, the increased consumption of multimedia is identified as the key driver for the emerging spectrum crunch. The potential for LiFi to access the light spectrum and provide a key enabler for future wireless solutions is highlighted. The role that new connectivity technologies can play in the IoT is identified and some of the drivers for the commercial adoption of LiFi are discussed.

References

- [1] Cisco Visual Networking Index, "Global Mobile Data Traffic Forecast Update, 2015-2020," CISCO, White Paper, Feb. 2016
- [2] <http://variety.com/2015/digital/news/streaming-overtakes-live-tv-among-consumer-viewing-preferences-study-1201477318/>
- [3] H. Haas, L. Yin, Y. Wang, and C. Chen, "What is LiFi?," in *Journal of Lightwave Technology*, vol. PP, no. 99,
- [4] C. Chen, D. A. Basnayaka, and H. Haas, "Downlink Performance of Optical Attocell Networks," in *IEEE Journal of Lightwave Technologies*, vol. 34, no. 1, pp. 137-156, Jan. 2016.
- [5] Dobroslav Tsonev, Stefan Videv, and Harald Haas, "Towards a 100 Gb/s visible light wireless access network," *Opt. Express* 23, 1627-1637 (2015)
- [6] C. Dandelski, B. Wenning, D. V. Perez, D. Pesch, J.P Linnartz, "Scalability of Sense Wireless Lighting Control Networks," in *IEEE Communications Magazine*, pp. 137-156, Jan. 2015.
- [7] <http://purelifi.com/lifi-products/lifi-x/>
- [8] Zixiong Wang, D. Tsonev, S. Videv and H. Haas, "Towards self-powered solar panel receiver for optical wireless communication," *Communications (ICC), 2014 IEEE International Conference on*, Sydney, NSW, 2014, pp. 3348-3353.
- [9] Z. Wang, D. Tsonev, S. Videv and H. Haas, "On the Design of a Solar-Panel Receiver for Optical Wireless Communications With Simultaneous Energy Harvesting," in *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 8, pp. 1612-1623, Aug. 2015.



Nikola Serafimovski in his role as Director of Business Strategy at pureLiFi Ltd. works as with major companies in the area of Li-Fi technology and commercialization, leading the creation and cultivation of the LiFi ecosystem, marketing, sales and standardization. His experience with T-Mobile and T-Home in Macedonia focused on mobile network deployment and analysis as well as database app development. Nikola worked for the UK-China Science Bridges project to successfully demonstrate the world's first practical implementation of the Spatial Modulation MIMO concept. He received a BSc in electrical engineering and computer science and an MSc in communications, system and electronics, both from Jacobs University Bremen, Germany. Nikola earned his PhD in digital communications and signal processing from the University of Edinburgh.



Dobroslav Tsonev received the B.Sc. degree in electrical engineering and computer science from Jacobs University Bremen in 2008, the M.Sc. degree in communications engineering with a specialization in communications electronics from the Munich Institute of Technology in 2010, and the Ph.D. degree in electrical engineering at the University of Edinburgh in 2015. His main research interests lie in the area of optical wireless communications with an emphasis on visible light communications. He is currently working on research and development of next generation Li-Fi networking equipment.



Harald Burchardt in his role as COO for pureLiFi works with major companies to assist them in integrating LiFi technology. Responsible for the operation and management of the company's activities as a whole, including business development and sales, marketing, engineering and development, administration, recruitment and financials. His experience includes DoCoMo Euro-Labs and Rohde&Schwarz GmbH in Munich, Germany, holds a patent and has filed three more in the field of communications engineering. He has a BSc in electrical engineering and computer science from Jacobs University Bremen, Germany, an MSc in communications engineering from the Technical University of Munich, Germany and a PhD in digital communications from the University of Edinburgh.



Stefan Videv received the B.Sc. in electrical engineering and computer science from Jacobs University Bremen, Germany, in 2007 and the M.Sc. in communications, systems and electronics in 2009 from the same university. He was awarded the Ph.D. degree for his thesis titled "Techniques for Green Radio Cellular Communications" from the University of Edinburgh. Stefan is currently employed as Experimental Officer at the Li-Fi R&D Center at the University of Edinburgh and working in the field of visible light communications (VLC). His research focus at the moment is on quick prototyping of communication systems, smart resource allocation, and energy efficient communications.



Professor Harald Haas received the PhD degree from the University of Edinburgh in 2001. He currently holds the Chair of Mobile Communications at the University of Edinburgh, and is co-founder and interim CEO of pureLiFi Ltd as well as the Director of the LiFi Research and Development Center at the University of Edinburgh. His main research interests are in optical wireless communications, hybrid optical wireless and RF communications, spatial modulation, and interference coordination in wireless networks. He first introduced and coined spatial modulation and LiFi. LiFi was listed among the 50 best inventions in TIME Magazine 2011. In 2012, he was the recipient of the prestigious Established Career Fellowship from the EPSRC (Engineering and Physical Sciences Research Council) within Information and Communications Technology in the UK. Prof. Haas is recipient of the Tam Dalyell Prize 2013 awarded by the University of Edinburgh for excellence in engaging the public with science. In 2014, he was selected by EPSRC as one of ten RISE (Recognising Inspirational Scientists and Engineers) Leaders in the UK.

***SPECIAL ISSUE ON
ULTRA-HIGH DEFINITION VIDEO COMMUNICATIONS***

Guest Editors: Pedro A. Assunção¹ and Rahul Vanam²

¹Instituto de Telecomunicações / IPLeiria, Portugal; ²InterDigital Communications, Inc., USA

amado@co.it.pt, rahul.vanam@interdigital.com

Ultra-high definition (UHD) video has been rapidly progressing into various fields of the consumer market, covering a whole technological chain from the acquisition and processing to storage, transmission and display. This evolution is sustained by research efforts worldwide with contributions from academia and industry to new developments and solutions capable of dealing with the huge amount of data associated with UHD video. Standardization activities have also been playing a crucial role in the technology and market development by providing the necessary elements to ensure interoperability among UHD video communications equipment from different manufacturers. As a result, an increasing number of diverse services and applications using UHD video content are expected to appear in near future, also bringing to light new challenging problems in the various elements of a communication chain. This Special Issue addresses some of these problems in six E-Letters from different authors, discussing relevant aspects of UHD video communications and sharing their latest results.

In the first article by Dragan *et al.*, from the University of Novi Sad and RT-RK, entitled “*Recent advances in UHD video coding technology: High Dynamic Range and wide color gamut*” the authors address the technical challenges associated with the evolution of current video coding standards to support high dynamic range, without losing backward compatibility, in conjunction with the need for a wider color gamut. The authors describe the relevant coding tools and metadata included in the HEVC standard for increasing the dynamic range and widening the color gamut of UHD format. The main standardization activities towards the specification of HExt extensions for HEVC are also presented, with reference to the core experiments and test conditions that support the standard development.

Bandoh *et al.* from NTT Corporation authored the second article entitled “*Analytical evaluation of high frame-rate video signal for encoder design.*” The current HEVC video coding standard supports higher frame rate in addition to higher resolution. This article studies the quantitative effect of increased frame-rate on encoded bit-rate. Two mathematical models are presented that provide relationship between frame-rate and motion-compensated prediction error. From experiments, the proposed model is demonstrated to be useful for determining a quantitative solution for the number of encoded bits required for motion-compensated prediction error when the frame-rate is increased.

The third article “*From HD to UHD video: implications for embedded systems implementations of software-base HEVC video encoders*”, by Monteiro *et al.*, from the Federal University of Rio Grande do Sul, address challenges posed on practical implementations of HEVC encoders for UHD formats in consumer devices. The article presents evidence to support the choice of cache memory configurations that ensure minimum energy consumption and reduced hardware costs. In the same context, the authors also present a method for reducing the computational cost of UHD HEVC encoding to allow encoder implementations based on embedded processors.

The fourth paper “*Contouring artefacts prevention in compressed UHD video sequences: tools and analysis of their performance*” by Seixas Dias *et al.* from the BBC, is concerned with contouring artefacts that appear in coded UHD image areas with low contrast and smooth variations of pixel intensities. The authors present two different methods to reduce the visibility of these artefacts in HEVC encoders in order to prevent deterioration of the increased quality of experience provided by UHD video formats.

Relevant solutions are presented based on two alternative approaches: post-processing at the decoder side or modified quantization at the encoder. It is shown that the efficiency of such methods in reducing visible artefacts is better for higher qualities, i.e., higher bit rates.

The fifth article is contributed by Asbun *et al.* from InterDigital Communications, and is titled “*DASH segment scheduling for scalable UHD video streaming*”. With increasing viewership of videos on mobile devices, there is growing need for multimedia clients that can stream videos while consistently maintaining high quality of user experience. To address this need, the authors present a DASH video segment scheduling algorithm for an SHVC streaming client that can yield higher video quality over traditional multi-rate streaming. The scheduling algorithm maximizes the video quality under a given bandwidth, and selects segments that minimize quality variations during bandwidth fluctuations, while maintaining the appropriate fullness of the download buffers.

Finally, the sixth article by Ferreira *et al.* from Instituto de Telecomunicações addresses the problem of video retargeting to enable seamless delivery of UHD video to mobile devices with small screens. Besides a brief review of related works, the article presents a method for retargeting UHD video based on aggregated saliency maps, which in turn are driven by visual attention models. To ensure temporal consistency of the resulting video sequence, jitter removal is implemented through filtering of the spatial trajectory of the cropping window. This has the effect of enhancing the visual experience, in addition to improving the coding efficiency in a consistent manner.

The extremely wide technological scope of UHD video communications is obviously impossible to be covered in the six E-Letters of this special issue. However, a good balance was achieved by addressing diverse issues, ranging from UHD content, HEVC encoding, streaming and adaptation to small displays. We would like to thank all authors for their great contributions and the E-Letter Board for making this special issue possible. Obviously we hope that the technical and scientific community can find these E-Letters interesting and inspirational for pursuing new ideas and making useful contributions for the field of UHD video communications, always bearing in mind the ultimate goal of improving the quality of user experience.



Pedro A. Amado Assunção received the Licenciado and M.Sc. degrees in Electrical Engineering from the University of Coimbra, Portugal, in 1988 and 1993, respectively, and the Ph.D. in Electronic Systems Engineering from the University of Essex, UK, in 1998. He is currently Professor of Electrical Engineering and Multimedia Communication Systems at the Polytechnic Institute of Leiria and researcher at the Instituto de Telecomunicações, Portugal. He is author/co-author of more than one hundred papers in conferences and journals, seven book chapters and four US patents. His current research interests include UHD and plenoptic video coding, complexity control and networking, multiple description coding, error concealment and quality evaluation.



Rahul Vanam received the B.E. degree in electronics and communication engineering from Bangalore University, Bangalore, India, in 2000, the M.S.E.E. degree from the New Mexico State University, Las Cruces, in 2005, and the Ph.D. degree in electrical engineering from the University of Washington, Seattle, in 2010. Since 2010, he has been working with InterDigital Communications, Inc., San Diego, USA, where he is currently a Staff Engineer. His current research interests include video and image compression and processing, and perceptual video coding. Dr. Vanam is an Associate Editor for the IEEE Transactions on Image Processing, and he also serves on the organizing committee and technical program committee of different conferences.

***Recent advances in UHD video coding technology:
High Dynamic Range and Wide Color Gamut***

Dragorad Milovanovic and Dragan Kukolj

Faculty of Technical Sciences, University of Novi Sad, Republic of Serbia

RT-RK Institute of Computer based Systems, Republic of Serbia

dragoam@gmail.com, dragan.kukolj@rt-rk.uns.ac.rs

1. Introduction

The paper presents of the most recent achievements in increasing the dynamic range and widening the color gamut of ultra high definition video (**UHD**) format and associated high efficiency video coding **HEVC/H.265** technologies in MPEG/ITU standardization framework. The main advantage of HEVC is better compression, and it is a better choice for increased spatial resolution in UHD application domains [1, 2, 3]. The HEVC compression standard has been approved in 2013, with the approved second version in 2014, including support for higher bit-depths and enhanced chroma formats. MPEG is currently studying whether HEVC is optimal to support higher dynamic range and wider color gamut, and launch a new standardization activity in 2015 to address the deficiencies.

UHD format would appear to be sufficient for video systems to display realistic images. ITU-R BT.2020 recommendation defines five parameters: pixel resolution (3840x2160 and 7680x4320), temporal properties (up to 120 Hz frame rates), wide-color gamut RGB primaries, signal format describing color encoding methods from RGB to luminance Y and chrominance $C_B C_R$ signals, and the way to digitize the RGB and $Y' C_B' C_R'$ signals. However, the human visual system has a significantly larger dynamic range than that supported by current display devices. Developed devices that are able to support a much larger dynamic range than conventional standard/high definition SD/HD systems are starting to appear on the market. A key technical challenge for video delivery is how to signal the higher dynamic range (**HDR**) while also supporting legacy standard dynamic range (**SDR**) displays. Next, current systems do not support the wide range (**WCG**) of colors that the human eye can perceive. The emissive color can be *both* bright and saturated. We emphasize the need for both a wider color gamut and a higher dynamic range, *simultaneously*.

In response to this trend, several standardization organizations have launched efforts to better enable these features in both the short- and mid-term. In this paper, we provide a timeline of MPEG standardization activities and a summary of the underlying HDR/WCG coding tools and metadata. Our emphasis is on both existing and potential extensions to the High Efficiency Video Coding (HEVC) standard.

2. HDR/WCG coding tools and metadata

The Moving Picture Experts Group (MPEG) as a working group of ISO/IEC has most recently developed the **HEVC** (*High Efficiency Video Coding*, officially known as ISO/IEC 23008-2 and ITU-T Recomm. H.265). HEVC enables the bit rate for video services to be approximately halved compared to the previous standard, H.264/AVC, while achieving the same video quality. The recommended bitrate for HEVC encoding of UHD format is 23.90 Mbps, for the HD format is 2.70 Mbps, while for the SD format is 0.90 Mbps.

The first version of this standard was approved in January 2013, where Main Profile and Main 10 Profile profiles were specified [2]. The Main Profile is used for mass market video services that historically require only 8 bits of precision in their processing, whereas Main 10 Profile, which supports up to 10 bits of processing precision, is expected to be used where additional accuracy is needed, notably for encoding video that has a high dynamic range.

The second version of this standard was approved in July 2014, where three amendments were added to the original specification [4]. The first, known as the *Range Extensions amendment*, added support for colour sampling formats beyond 4:2:0 and up to 16 bits of processing precision [5]. The second, known as *Multiview HEVC* (MV-HEVC), added support for providing efficient representation of video content with multiple camera views and optional depth map information, such as for 3D stereoscopic and auto stereoscopic video applications [6]. The third, known as *Scalable HEVC* (SHVC), added support for embedded bitstream scalability in which different levels of encoding quality are efficiently supported by adding or removing layered subsets of encoded data [7].

HEVC versions already comprise a set of coding tools and metadata relevant for increasing the dynamic range and widening the color gamut of UHD format: support of higher bit-depth signals (with the definition of a consumer profile, Main 10), BT.2020 color gamut for UHD format, PQ (*Perceptual Quantizer*, SMPTE ST 2084) and transfer functions (OETF) that can be used for HLG (*Hybrid Log-Gamma*, ARIB STD-B67) correction with RGB and YCbCr color components, one SEI message providing descriptive information on **color volume** (*the color primaries, white point, luminance range*) of the content display, three SEI messages embedding processing parameters for an efficient adaptation at rendering side (*tone mapping, knee function, color remapping*).

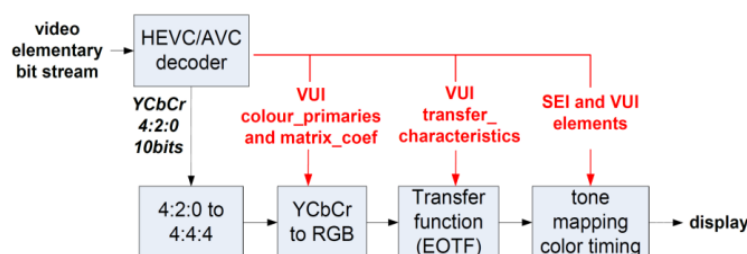


Figure 1. Usage of VUI (Video Usability Information) and SEI (Supplemental Enhancement Information) messaging at rendering side [8].

In the context of scalable coding, HEVC contains two tools, the color gamut scalability and bit-depth scalability, allowing the support of HDR and WCG with backward compatibility with standard SDR BT.709 HD format [9]. HEVC is therefore, already usable for HDR and WCG video distribution applications.

3. HEVC standardization process

Recently, MPEG has launched in June 2015 a fast-track standardization process to enhance the performance of the Main 10 profile for HDR and WCG video, that would lead to the HExt extension around mid-2016. To achieve this, immediately following its meeting in February 2015, MPEG issued a *Call for Evidence (CfE)* of new tools that may improve the performance of HEVC when used to encode high dynamic range and wide color gamut video [10, 11, 12].

In parallel to its rapid-track effort, MPEG plans to launch a second track to address long-term requirements that may include even further improved compression of HDR/WCG content. This second effort could be part of a larger future video compression effort within MPEG [13]. The current goal is to have the International Standard for this new video coding standard around 2020. A formal *Call for Proposals* for the second track is expected to be issued in 2017.

4. H2M test model for HExt video coding

After a successful call for evidence (*CfE*) for High Dynamic Range (HDR), the technical work starts in the video subgroup with the goal to develop an architecture (H2M) as well as core experiments (CE) [14, 15].

Document MPEG113 *Exploratory Test Model for HDR extension of HEVC* presents a first draft of the **H2M** test model. The test model focuses on possible normative process changes to the HEVC specification. The process is out of loop and *does not* involve changes in the normative decoding process. The only normative tool in the HDR reconstruction post-process in H2M is the inverse **reshaper**. H2M provides an example on how the syntax and semantics of the reshaper would be included in the SPS/PPS (Sequence/Picture Parameter Set) extension which contains general video parameters.

A functional diagram of the H2M test model is shown in Figure 2. The proposed system uses HEVC Main 10 profile for the bitstream generation and bitstream decoding, and uses *metadata* provided by the decoder to control decoder side processing used for HDR reconstruction and WCG representation.

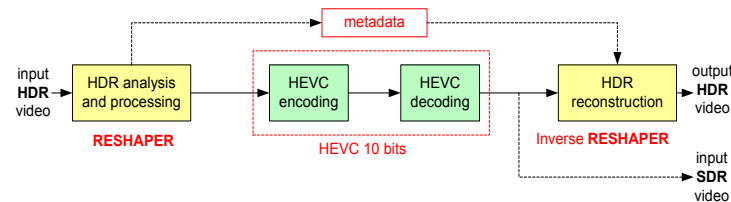


Figure 2. H2M functional diagram [14].

The input HDR signal is *pre-processed* to produce a modified HDR signal that is provided to the HEVC Main 10 encoder. The HEVC Main 10 decoder output is used to reconstruct the HDR signal. The pre-processing and *post-processing* steps primarily aim at improving the coding efficiency of HDR content, and at providing support for SDR backward compatibility.

The HDR analysis and processing is applied prior to the HEVC encoding. It maps the input HDR signal to a format adapted to the HEVC Main 10 profile. The normative encoding process utilizes the adaptive reshaper module after the HDR signal is converted from color sampling format 4:4:4 to 4:2:0. The purpose of the reshaper is to change the signal characteristics of YCbCr to improve the coding efficiency of the existing HEVC Main 10 codec, and also to potentially enable producing the SDR compatible reshaped version. The motivation of the reshaper has three aspects:

- ✓ PQ (Perceptual Quantizer) is designed to cover the full range of HDR signal from 0 to 10,000 nits (measure of luminance). However, due to the display limitation and director's intent, the video range might be smaller than the full range;
- ✓ For HDR and WCG signal has a much larger color volume than SDR which includes both color and intensity;
- ✓ In case of SDR-backward compatibility, the reshaper performs dynamic range reduction with control of color shift resulted from this dynamic range reduction.

Parametric reshaper model is designed to simultaneously improve texture sharpness and color performance. It can additionally enable SDR backward-compatible support. The reshaper is used to implement the following two features:

- ✓ Adaptive codeword re-distribution;
- ✓ Re-quantization of luma and chroma signal components.

Parameters of the luma and chroma reshaper models are signaled in the PPS (Picture Parameter Set) syntax and usually updated when there is a scene change or IRAP (Intra Random Access Point).

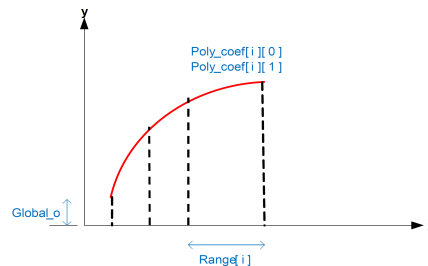
The luma reshaper is modeled using a piecewise 2nd order polynomial or piecewise linear model. The maximum number of pieces is 8. The use of 8 segments allows approximation of most reasonable curves, i.e., continuous and bounded derivative. The 2nd order polynomial model is used to approximate complex, non-linear smooth curves efficiently without the need to use a large number of segments. For less complex curves, the piecewise linear model with up to 8 pieces provides sufficient performance with

reduced complexity. For specific implementation, the piecewise polynomial and linear model can be implemented using a LUT or computation procedure. The segment locations are part of the reshaper parameters. The segment locations are derived at the decoder side from segment lengths, which are signaled for each segment. The freedom to select the segment locations allows adaptation to a particular signal being modeled.

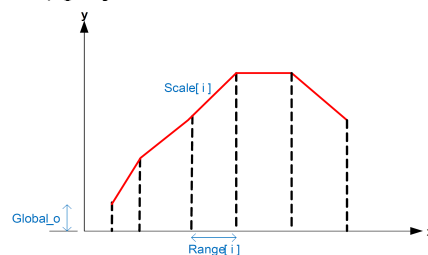
For chroma, the reshaper is based on piecewise linear model reshaper functions with up to 32 pieces. The piecewise linear model is constructed from parameters which include number of utilized segments, a length of each segment and a scale value applied for each segment, and a global offset value. Chroma reshaper can operate in two possible modes [14]:

- In Mode 0, reshaper is performed as *Dynamical Range Adjustment (DRA)*, which is applied independently to each chroma component using piecewise linear functions;
- In Mode 1, reshaper is performed as *Cross-Component Scaling (CCS)*, with a scaling factor applied to each chroma sample and being a function of the co-located luma sample. This scaling factor function is also modeled as a piecewise linear function.

Figure 3 provides an illustration of the reshaper functions for luma and chroma components.



a) polynomial function for luma



b) linear function for chroma sample value

Figure 3. Piecewise models of the inverse reshaper [14].

The further development of HExt extensions for HEVC are based on set of core experiments (CE) of specific tools under investigation and a time line of simulation and cross-check reports. Core experiment is the regular process for a tool to be adopted into the draft specification. Conceptually, a successful core experiment can be considered as the step before adopting a proposal into the draft specification. Document MPEG113 *Common test conditions for HDR/WCG video coding experiments* defines common test conditions (CTC) and software reference configurations [16] to be used in the context of core experiments conducted in the AhG for HDR/WCG video coding experiments [15]:

CE1. Optimization without HEVC specification change. The goal of this CE is to identify and investigate methods for optimization of the HEVC Main 10 coding of HDR/WCG content without any specification changes.

CE2. NCL fixed point for HDR video coding. This document describes the core experiment on normative 4:2:0 YCbCr non-constant luminance fixed point for HDR video coding.

CE3. Objective/subjective metrics. This document describes the core experiment on objective metrics for HDR/WCG video coding evaluation. This experiment focuses on the performance analysis of selected luminance and color metrics. Correlation with subjective testing of CfE responses will be conducted.

CE4. Consumer monitor testing. The aim is to investigate the visual performance of test reshaper technologies on consumer monitors and compare it against the performance on professional HDR monitors.

CE5. Colour transforms and sampling filters. The aim is to investigate the use of alternative colour transforms and sampling filters. The performance will be measured using the common test conditions, objective metrics and optional subjective assessments.

CE6. Non-normative post processing. This relates to HDR and WCG video coding technologies and focuses on the reconstruction of HDR video with the decoded output video from an HEVC Main 10 decoder. In this CE, only solutions that operate as pre-processing to the encoder and post-processing to the Main10 decoder are considered.

CE7. HLG investigation. The CE investigates how to generate compressed HDR/SDR video using HLG (Hybrid Log Gamma) transfer function and HEVC.

CE8. Viewable SDR testing. An identified use case and requirement of high-dynamic range (HDR) video is the potential to provide bit-stream compatibility with legacy HEVC decoders. This CE will examine how to evaluate (by means of a formal subjective test) and report the quality of the SDR images.

5. Conclusion

The enhanced resolution of ultra high definition video (UHD) format provides undoubtedly better picture quality. Further enhancements of higher dynamic range, wider color gamut, and higher frame rates, are becoming technologically feasible and are generating much interest within the industry. Also, we have observed the growing activities within the standardization community.

The HEVC compression standard is essential to make UHD format feasible. Its approved second version includes support for higher bit-depths and enhanced chroma formats. In June 2015, MPEG has launched a fast-track standardization process to enhance the performance of the Main 10 profile for HDR/WCG video. In parallel to its rapid-track effort, MPEG plans to launch a second track to address long-term requirements.

Recently, MPEG has published a first draft H2M test model for HDR/WCG video compression. The main pre/post-processing block, the parametric reshaper, is designed to improve texture sharpness and color performance. We emphasize the need for efficient processing, both for a wider color gamut and a higher dynamic range, simultaneously.

References

- [77] K.R. Rao, Z. Bojković, D. Milovanović, *Multimedia communication systems: techniques, standards and networks*, Prentice Hall PTR (Pearson Education) 2002.
- [78] G.J. Sullivan, J. Ohm, Han Woo-Jin, T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard", *IEEE Trans. on CSVT*, vol. 22, no. 12, pp. 1649-1668, Sept. 2012.
- [79] T. Tan, R. Weerakkody, M. Mrak, N. Ramzan, "Video quality evaluation methodology and verification testing of HEVC compression performance", *IEEE Trans. on CSVT*. (to appear)
- [80] G.J. Sullivan, J. Boyce, Y. Chen, J.-R. Ohm, "Standardized extensions of High Efficiency Video Coding (HEVC)", *IEEE Journal of Selected Topics in Signal Processing*, vol.7, pp.1001-1016, Dec. 2013.
- [81] D. Flynn, D. Marpe, M. Naccari, T. Nguyen, "Overview of the Range Extensions for the HEVC standard: tools, profiles and performance", *IEEE Trans. on CSVT*. (to appear)
- [82] G. Tech, Y. Chen, K. Muller, J.R. Ohm, A. Vetro, Y.K. Wang, "Overview of the Multiview and 3D extensions of high efficiency video coding", *IEEE Trans. on CSVT*. (to appear)
- [83] J.M. Boyce, Y. Ye, J. Chen, A.K. Ramasubramanian, "Overview of SHVC: Scalable extensions of the High Efficiency Video Coding (HEVC) standard", *IEEE Trans. on CSVT*. (to appear)

- [84] E. Francosi, C. Fogg, Y. He, X. Li, A. Luthra, C.A. Segall, "High Dynamic Range and Wide Color Gamut video coding in HEVC: Status and potential future enhancements", *IEEE Trans. on CSVT*. (to appear)
- [85] Y. Yan, P. Andrivon, "The scalable extensions of HEVC for ultra-high definition video delivery", *IEEE Multimedia*, Vol. 21, Issue 3, pp.58-64, July-Sept. 2014.
- [86] ISO/IEC JTC1/SC29/WG11 N15084, *Requirements and use cases for HDR and WCG content coding*, MPEG111 Feb. 2015.
- [87] ISO/IEC JTC1/SC29/WG11 N15083, *Call for Evidence (CfE) for HDR and WCG video coding*, MPEG111 Feb. 2015.
- [88] ISO/IEC JTC1/SC29/WG11 N15350, *Test results of Call for Evidence (CfE) for HDR and WCG video coding*, MPEG112, June 2015.
- [89] ISO/IEC JTC1/SC29/WG11 N15340, *Requirements for a Future Video Coding standard*, MPEG112, June 2015.
- [90] ISO/IEC JTC1/SC29/WG11 N15792, *Exploratory Test Model for HDR extension of HEVC*, MPEG113, Oct. 2015.
- [91] ISO/IEC JTC1/SC29/WG11 N15794, N15795, N15796, N15797, N15798, N15799, N15800, N15801, *HDR CE1-8*, MPEG113, Oct. 2015.
- [92] ISO/IEC JTC1/SC29/WG11 N15793, *Common Test Conditions for HDR/WCG video coding experiments*, MPEG113, Oct. 2015.



Dragorad Milovanovic received the Dipl. Electr. Eng. and M.Sc. degree from University of Belgrade, Serbia. From 1987 to 1991, he was working as research assistant at the Department of Electrical Engineering, where his research interests include analysis and design of digital communications systems. He has been working as R&D engineer for DSP software development in digital television industry. He is serving as ICT lecturer and consultant in implementation of standard-based solutions for multimedia communication. He has published more than 200 papers in international journals and conference proceedings.



Dragan Kukolj (M'97-SM'06) received his Diploma degree in control engineering in 1982, MSc degree in computer engineering in 1988, and PhD degree in control engineering in 1993, all from the University of Novi Sad, Serbia. He is currently a Professor of computer-based systems with Dept. of Computer Engineering, Faculty of Engineering, University of Novi Sad. His main research interests include digital signal processing, video processing, soft computing and techniques. He has published over 200 papers in referred journals and conference proceedings. Dr. Kukolj is the coordinator Intellectual Property Centre of University of Novi Sad.

Analytical evaluation of high frame-rate video signal for encoder design

Yukihiro Bandoh, Seishi Takamura and Atsushi Shimizu

Nippon Telegraph and Telephone Corporation, Japan

{bandou.yukihiro, takamura.seishi, shimizu.atsushi}@lab.ntt.co.jp

1. Introduction

Recent advances in imaging technologies and communication network technologies are expected to grow the market of high realistic communication using high quality videos. The visual representations for high realistic communication require improvements to the following four elements: spatial resolution, dynamic range, reproducing accurate color, and temporal resolution. For example, 4K UHDTV (3840×2160 pixels) and 8K UHDTV (7680×4320 pixels) [1] offer digital videos with high-resolution. The high-dynamic-range imaging technologies [2] are being developed for generating images with high contrast found in the real world scene. The efforts to reproduce accurate color are made with spectrum-based imaging technologies [3].

In our studies, we focus on video with high temporal resolution, that is, high frame-rate video. The current frame-rate (60 [frames/sec] or [fields/sec]) was selected from the viewpoint of suppressing flicker. Unfortunately, suppressing flicker is not directly connected to the representation of smooth movement needed for realistic communication. Relationship between frame-rate and human vision [4, 5] are evaluated from the viewpoints of human visual sensitivity to spatial and temporal frequencies. Studies on subjective quality evaluations related with frame-rate [6, 7] report that subjective quality of video can be improved through increasing frame-rate up to 240 fps or 300 fps. Additionally, cloud gaming applications [8] also demand high frame-rate to provide smooth experience to end users. The latest video coding standard H.265/HEVC [9, 10] supports high frame-rate video signal.

Since high frame-rate video generates far more encoded bits than low frame-rate video, it is important to understand the relationship between frame-rate and bit-rate for designing efficient encoder. It is easily predicted that increasing the frame-rate leads to a decrease in the amount of the encoded bits of inter-frame prediction error. However, the quantitative effect of frame-rate on bit-rate has not been clarified.

In this paper, we summarize studies on mathematical models [11, 12] of the relationship between frame-rate and bit-rate in anticipation of encoding high frame-rate video. The models quantify the impact of frame-rate on bit-rate of motion-compensated prediction error.

2. Modeling motion-compensated prediction error with frame-rate

In a study [11], we derive a mathematical model that describes the relationship between frame-rate and motion-compensated prediction error. In motion-compensated prediction scheme, a frame is divided into rectangular block, and each block is predicted from the previous frame by using estimated displacement. Our models approximate motion-compensated prediction error, by using the first order approximation of Taylor expansion and the assumption that the noise element is statistically independent of the video signal.

Furthermore, focusing on displacement estimation error between estimated displacement and the true displacement, we approximate displacement estimation error as the function of the frame-rate. Finally, using the approximation of motion-compensated prediction error, we have the following model for information of motion-compensated prediction error:

$$I(F) = \log_2(A_1 F^{-2} + A_2 F^{-1} + A_3) \quad (1)$$

where F is frame-rate. A_1, A_2, A_3 are constants that depend on predicted video signal. For details on

derivation of the above model and the above parameters, refer to the literature [11]. The derivation of equation (1) premises that frame-rate is set by temporal sub-sampling, that is, frame-skip. The model of equation (1) does not consider the effect of motion blur caused by the open interval of the shutter (shutter-open interval) in the image sensing device.

In modeling motion-compensated prediction error, it is important to consider the effect of shutter-open interval. Increasing the shutter-open interval increases the occurrence of motion blur in a frame. On the other hand, increase in the shutter-open interval reduces the jerkiness caused by temporal down-sampling. Therefore, shutter-open interval can control the trade-off between jerkiness and motion blur to suit the characteristics of the imaging system. By considering shutter-open interval, we extend the model of equation (1) to the following:

$$I(F, r) = \hat{A}_1 r^{-1} F^{-1} + \hat{A}_2 r^{-1/2} F^{-1/2} + \hat{A}_3 r^{-1} F + \hat{A}_4 \quad (2)$$

where r is called the shutter-open ratio, and is defined as the ratio of the shutter-open interval to the frame interval determined by frame-rate F . Parameters $\hat{A}_1, \hat{A}_2, \hat{A}_3, \hat{A}_4$ are constants that depend on predicted video signal. For details on derivation of the above model and the above parameters, refer to the literature [12].

3. Experiments

We performed evaluation to validate the proposed models using three sequences captured by high speed camera whose specification is shown in Table 1. Sequences taken with the high speed camera consisted 24 bits RGB format. We converted the color format from RGB data to YCbCr data. Y-data with 8 bits gray scale were used for these experiments. The frame rate for capturing the original sequences was 1000 Hz. We evaluated the models over sequences with frame rates set to 1000, 500, 250, 125, 62.5 and 31.25 Hz. Sequences with frame rates less than 1000 Hz were generated by applying the temporal averaging method on the original sequence captured at 1000 Hz. The motion-compensated prediction was configured as shown in Table 2. The information of motion-compensated prediction error was evaluated with entropy.

We evaluate our proposed models from the viewpoint of information criteria, and compare them with the conventional model [13] that is expressed as follows:

$$I(F) = ab(1 - \exp(-1/(bF)))$$

where a and b are constants that depend on the video signal. Table 3 and Table 4 show two kinds of information criteria Akaike's Information Criterion (AIC) [14] and Bayesian information criterion (BIC) [15] for each model, where the values of both AIC and BIC are the average values of three sequences with above-mentioned six kinds of frame rates. AIC and BIC evaluate a candidate model in terms of both the goodness of fit of the model and the complexity of the model. It is preferable to select the model with smaller value of AIC and BIC. Note that AIC and BIC can take negative value, because the goodness of fit in AIC and BIC is calculated as negative value. Table 3 and table 4 shows results in the case of shutter-open ratio $r = 50[\%]$ and $r = 100[\%]$, respectively. We confirmed that proposed models achieved smaller values of both AIC and BIC than the conventional one. Furthermore, by considering shutter-open interval, the proposed model of equation (2) reduced the values of both AIC and BIC compared with that of equation (1). These results verify that proposed model of equation (2) is better than the proposed one of equation (1) and the conventional one from the viewpoint of model selection based on AIC and BIC, even allowing for its increased degree of freedom.

The above evaluations were over VGA sequences, due to the upper limit of spatio/temporal pixel rate which can be captured by the high speed camera used in this experiment. However, since the proposed

models are not limited to low resolution sequences, they are applicable to higher resolution sequences such as 4K and 8K as well.

Table 1: Specifications of high speed camera

Item	Specifications
Frame-rate	1000 [Hz]
Shutter-speed	1/1000 [sec]
Resolution	640×480 [pixels]
Number of frames	480 [frames]
Color format	24bit RGB

Table 2: Configurations of MC prediction

Item	Configurations
Block size of MC prediction	16×16 [pixels]
Accuracy of MC prediction	Quarter pixel
Motion estimation scheme	Full search algorithm
Criterion of motion estimation	Sum of absolute difference (SAD)

Table 3: Information criteria
(shutter-open ratio $r = 50$ [%])

	AIC	BIC
Proposed model (eq.(2))	-58.4	-58.6
Proposed model (eq.(1))	-40.2	-40.5
Conventional model	-3.7	-3.8

Table 4: Information criteria
(shutter-open ratio $r = 100$ [%])

	AIC	BIC
Proposed model (eq.(2))	-78.1	-76.6
Proposed model (eq.(1))	-45.2	-44.0
Conventional model	-3.2	-2.3

4. Conclusion

In this paper we summarize mathematical models that explains the impact of frame-rate on motion-compensated prediction error. By using this model, we can describe the variation in the relationship between frame-rate and motion-compensated prediction error, and offer a quantitative solution to the question of how many encoded bits of motion-compensated prediction error are required when the frame-rate is increased.

References

- [1] ITU-R Recommendation BT.2020: Parameter values for ultra-high definition television systems for production and international programme exchange, Oct. 2015
- [2] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, "High Dynamic Range Imaging, Second Edition: Acquisition, Display, and Image-Based Lighting", Morgan Kaufmann, 2010.
- [3] M. Yamaguchi, H. Haneishi, and N. Ohyama, "Beyond Red-Green-Blue (RGB): Spectrum-Based Color Imaging Technology", Journal of Imaging Science and Technology, Vol. 52, No. 1, 2008.
- [4] A. B. Watson, "High frame rates and human vision: A view through the window of visibility", SMPTE Motion Imaging Journal vol. 122, no. 2, pp. 18-32, March 2013..
- [5] K. C. Noland, "The Application of Sampling Theory to Television Frame Rate Requirements", BBC R&D White Paper WHP 282, July, 2014.
- [6] Y. Kuroki, T. Nishi, S. Kobayashi, H. Oyaizu, and S. Yoshimura, "A psychophysical study of improvements in motion-image quality by using high frame rates", Journal of the Society for Information Display, Vol. 15, No. 1, pp.61-68, Jan., 2007.
- [7] R.A. Salmon, M. Armstrong, and S. Jolly, "Higher Frame rates for more Immersive Video and Television", BBC R&D White Paper WHP 209, Oct. 2011.
- [8] J. Wu, C. Yuen, N. Cheung, J. Chen, and C. W. Chen, "Enabling Adaptive High-Frame-Rate Video Streaming in Mobile Cloud Gaming Applications", IEEE Transactions on Circuits and Systems for Video Technology, vol. CSVT-25, No.12, pp.1988 - 2001, 2015.
- [9] ITU-T Recommendation H.265: High efficiency video coding, April, 2015.
- [10] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Trans. Circuits and Systems for Video Technology, Vol. 22, No. 12, pp. 1649-1668, Dec. 2012.
- [11] Y. Bando, K. Hayase, S. Takamura, K. Kamikura, and Y. Yashima, "Theoretical Modeling of Inter-Frame Prediction Error for High Frame-Rate Video Signal", IEICE trans. fundamentals, Vol.E91-A No.3 pp.730-739, 2008.
- [12] Y. Bando, S. Takamura, H. Jozawa, and Y. Yashima, "Generalized Theoretical Modeling of Inter-Frame Prediction Error for High Frame-Rate Video Signal Considering Integral Phenomenon", IEICE trans. fundamentals, Vol.E93-A No.8 pp.1442-1452, 2010.
- [13] N. Murayama, "The law of video data compression", PSCJ, No. 1-1, pp. 11-12, 1988. in Japanese.
- [14] H. Akaike, "A new look at the statistical model identification", IEEE Trans. on Automatic Control, Vol. AC-19, No. 6, pp. 716-723, 1974.
- [15] G. Schwarz, "Estimating the dimension of a model", Annals of Statistics, Vol. 6, No. 4, pp. 461-464, 1978.



Yukihiro Bandoh received the B.E., M.E., and Ph.D. degrees from Kyushu University, Japan, in 1996, 1998 and 2002, respectively. He granted JSPS Research Fellowship for Young Scientists from 2000 to 2002. In 2002, he joined Nippon Telegraph and Telephone (NTT) Corporation, where he has been engaged in research on efficient video coding for high realistic communication. He is currently a senior research engineer at visual media coding group of NTT Media intelligence Laboratories. Dr. Bandoh is a senior member of IEEE, IEICE, and IPSJ.



Seishi Takamura received the B.E., M.E. and Ph.D. degrees from Department of Electronic Engineering, Faculty of Engineering, the University of Tokyo in 1991, 1993 and 1996 respectively. In 1996 he joined Nippon Telegraph and Telephone (NTT) Corporation, where he has been engaged in research on efficient video coding and ultra-high quality video coding. He has fulfilled various duties in the research and academic community in current and prior roles including Associate Editor of IEEE Transactions on Circuits and Systems for Video Technology (2006–2014), Executive Committee Member of the Institute of Electrical and Electronics Engineers (IEEE) Tokyo Section, the IEEE Japan Council. He has also served as Chair of ISO/IEC JTC 1/SC 29 Japan National Body, Japan Head of Delegates of ISO/IEC JTC 1/SC 29, and as an International Steering Committee Member of the Picture Coding Symposium. From 2005 to 2006, he was a visiting scientist at Stanford University, California, USA. He is currently a senior research engineer, supervisor, distinguished technical member at visual media coding group of NTT Media Intelligence Laboratories. Dr. Takamura is a senior member of IEEE, IEICE, and IPSJ and a member of MENSAs, IEEEJ, and ITE.



Atsushi Shimizu received the B.E. and M.E. degrees in electronic engineering from Nihon University in 1990 and 1992 respectively. He joined NTT in 1992 and has been engaged in video compression algorithm and software development. He is currently a Senior Research Engineer, Supervisor, and the head of visual media coding group in NTT Media intelligence Laboratories.

***From HD to UHD video: implications for embedded systems
implementations of software-based HEVC video encoders***

Eduarda Monteiro¹, Guilherme Corrêa², Luis A. da Silva Cruz³, and Sergio Bampi¹

¹Universidade Federal do Rio Grande do Sul, ²Universidade Federal de Pelotas, Brazil

³Instituto de Telecomunicações/Universidade de Coimbra-DEEC, Portugal

ermonteiro@inf.ufrgs.br, gcorrea@inf.ufpel.edu.br, lcruz@deec.uc.pt, bampi@in.ufrgs.br

1. Introduction

The HEVC standard [1] achieved the goal set earlier of halving the bitrate of encoded video with respect to that attainable with its predecessor H.264/AVC. However this efficiency gain was accompanied by an increase in both computational costs [2] and memory requirements [3]. The coming of age of UHD capture and display technology with large amounts of video data to be transmitted and stored fits nicely with the high coding efficiency of HEVC.

Unfortunately the implementation of HEVC encoders poses serious challenges both in hardware and software based platforms. As shown in the following sections, the move from HD to UHD formats enabled by the inclusion of UHD capable cameras in consumer devices such as tablets and smartphones further complicates the design of software-based encoders to run in embedded platforms with CPUs such as the line of Cortex processors. The work presented will show that a judicious choice of cache memory configurations is necessary to ensure minimum memory energy expenditure and reduce the hardware costs. It will also be shown through controlled experiments that HEVC encoding of UHD video is much more computationally complex than encoding HD video, to the point that heuristic-based optimizations of the encoding process are absolutely necessary. A computation reduction method will be presented which reduces UHD HEVC encoding computational costs to levels that allow encoding using embedded processors.

TABLE I. CACHE MEMORY SETTINGS

Device	Hierarchy	Capacity	Associativity	Line Size
Cortex A8	L1C	16 KB	4-way	64 bytes
	LLC	512 KB	8-way	
Cortex A53	L1C	64 KB	4-way	
	LLC	2 MB	16-way	

2. Experimental Setup

The experimental setup used in this work is described in the following three subsections, *video coding setup*, *cache memory analysis setup*, and *computational complexity analysis setup*.

Video Coding Setup

The encoding experiments reported were performed using HEVC test Model (HM), version 16.2 [4], with Quantization Parameters (QP) set to 22, 27, 32 and 37. The configuration used was Main, Low Delay following the HEVC Common Test Conditions (CTC) [5]. Four video sequences from the Ultra Video Group repository [6] (Beauty, Jockey, HoneyBee, and YachtRide), both at HD (1920x1080) and UHD (3840x2160) resolutions were used as test content.

Cache Memory Analysis Setup

The move from HD to UHD formats, complicates the design of software-based encoders that run in embedded platforms with application processors such as those of the Cortex line. To study the effect of the video resolution on the cache memory behavior and energy consumption two Cortex processors were chosen, A8 and A53. Table 1 lists the specifications of their memory systems.

Computational Complexity Analysis Setup

The computational complexity analysis was performed using the HM software, compiled with the Microsoft Visual Studio C++ compiler and run on a dedicated core of an Intel® Xeon® E5520 (2.27 GHz). Computational complexity was measured in terms of processing time. Additionally, changes in compression efficiency resulting from the use of the computational complexity reduction strategy used in this work were quantified in terms of bit rate and image quality variations, measured by the Bjontegard Delta (BD)-rate and BD-PSNR objective metrics [7].

3. Cache Memory Energy Consumption Analysis and Optimization

This section presents a comparative analysis of the behavior of a cache memory hierarchy when encoding video with HD and UHD resolutions and also a study of cache behavior of two Cortex processors when encoding UHD video. Firstly a comparison will be made of the cache memory access statistics, hits and misses, and energy consumption for both HD and UHD resolutions. Afterwards a comparison will be made of the cache memory behavior, accesses and energy consumption, when two different processors, Cortex A8 and 53, are used to encode UHD video. The cache memory usage parameters recorded and analyzed are the miss rates, number of accesses to last level memory and the energy consumption.

Methodology

The cache memory evaluations described in the following paragraphs were done using the Memory Energy Estimation Framework (MEEF) composed of three main tools and estimation models.

The MEEF is a set of Python scripts that handle the communication interfaces between the tools and the models of the framework: **(i)** HEVC Encoder: the HEVC reference software (HM) used to encode the video sequences; **(ii)** Cachegrind: part of the instrumentation framework Valgrind [8], used to collect and summarize information on cache behavior while running HM; **(iii)** Cacti [9]: a tool from the HP Labs used to estimate the energy consumption costs of each access to the cache memory. MEEF includes analytical models to compute the energy consumed by each read and write operation to the level 1 cache (L1C) and lower level cache (LLC). The model equations use as input parameters the number of cache hits and misses that occur at each level. When a hit occurs, energy is consumed only to read data from the corresponding level. On other hand, when a misses happens, it is necessary to write in the cache the block fetched from lower levels before reading it. The energy equations used are listed below.

$$\begin{aligned} Read_{En_{L1}} = & (n_{Read_Hits_{L1}} * C_{Read_{L1}}) \\ & + (n_{Read_Misses_{L1}} \\ & * (C_{Read_{L1}} + C_{Write_{L1}})) \end{aligned} \quad (1)$$

$$\begin{aligned} Read_{En_{LL}} = & (n_{Read_Hits_{LL}} * C_{Read_{LL}}) \\ & + (n_{Read_Misses_{LL}} \\ & * (C_{Read_{LL}} + C_{Write_{LL}} \\ & + C_{Read_{RAM}})) \end{aligned} \quad (2)$$

$$\begin{aligned} Write_{En_{L1}} = & (n_{Write_Hits_{L1}} * C_{Write_{L1}}) \\ & + (n_{Write_Misses_{L1}} \\ & * (C_{Read_{L1}} + C_{Write_{L1}})) \end{aligned} \quad (3)$$

$$\begin{aligned} Write_{En_{LL}} = & (n_{Write_Hits_{LL}} * C_{Write_{LL}}) \\ & + (n_{Write_Misses_{LL}} \\ & * (C_{Read_{LL}} + C_{Write_{LL}} \\ & + C_{Read_{RAM}})) \end{aligned} \quad (4)$$

$$\begin{aligned} Total_{Energy} = & Read_{En_{L1}} + Read_{En_{LL}} \\ & + Write_{En_{L1}} + Write_{En_{LL}} \end{aligned} \quad (5)$$

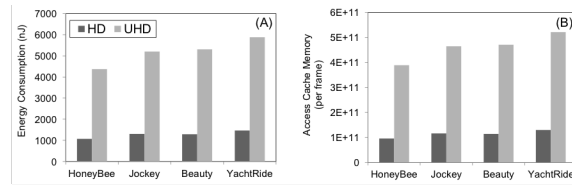


Figure 1. From HD to UHD: Energy Consumption (A) and Number of Memory Accesses (B)

The variables $n_{ReadHits*}$ and $n_{ReadMisses*}$ are respectively the hit and miss counts output by the Cachegrind profiler. C_{Read*} and C_{Write*} stand for the energy costs of read and write operations for each level, also estimated by Cacti. The LL cache read energy equation in (2) is analogous to (1), except that an extra main memory (RAM) read operation is considered in case of a miss (at LL level). The write energy models of eqns. (3) and (4), were formulated analogously to the read energy models. The costs of read and write operations at all levels (L1, LL, and RAM) are considered. Lastly, (5) expresses the total energy consumed by the memory hierarchy computed by adding the read and write energy costs.

Comparative Analysis of Cache Memory Behavior for HD and UHD

This section presents a comparison of cache memory behavior when encoding UHD and HD resolution video. Fig. 1 presents data for the energy consumption and the number of cache memory accesses according to video resolution and test sequence used. From the plots in Fig. 1 (A) one realizes without surprise that the energy consumption incurred during UHD video encoding is higher than that observed for HD resolution video encoding. Furthermore the energy costs grow, also as expected, proportionally to the increase in the number of pixels, with UHD encoding requiring 4.05 times the energy needed for HD resolution video encoding. Fig. 1 (B) presents data about the number of cache memory accesses during the HD and UHD video encoding process. These results show that encoding of UHD resolution video generates 4.05 times more accesses (per frame) to the cache memory hierarchy than encoding HD video.

In conclusion, the results show that both energy consumption and cache memory accesses grow in direct proportion to the increase in the number of pixels of the video frame. Other results not presented reveal a similar trend although with small differences in the growth factor.

Cortex Cache Memory Analysis of UHD Video Coding

The previous section showed that going from HD to UHD resolutions increases the cache memory accesses and energy consumption. A second important question which is now analyzed is whether cache memory architectural parameters can be chosen which result in less misses and less energy consumption during encoding of UHD video. This question is partially answered through a comparative study of the behavior of two distinct cache memory configurations employed in two current Cortex processors. The main goal of this evaluation is to show that cache configurations are indeed important and should be chosen carefully to ensure minimum memory energy consumption and to reduce the hardware costs.

This investigation considers two conventional Cortex processors, listed in Tab. I, A8 and A53. Fig. 2 shows a comparison of the cache miss rates for both levels of the cache (A) and the absolute number cache memory accesses to the LLC (B) observed with these two cache memory configurations. Observing Fig. 2 one notices that these cache memory configurations exhibit different behaviors, reflecting the simpler cache memory hierarchy of the A8 Cortex device when compared to that of the A53 model. The results in Fig. 2 (A) show that the A8 cache memory organization suffers from higher miss rates than that of the A53. On average the A8 cache suffers 3.68 and 4.54 times more misses in the L1C and LLC, respectively, when compared with the A53 cache memory configuration. In addition, the plot presented in Fig. 2 (B) shows that the LLC of the A8 processor has, on average about 270% more accesses than that of the A53.

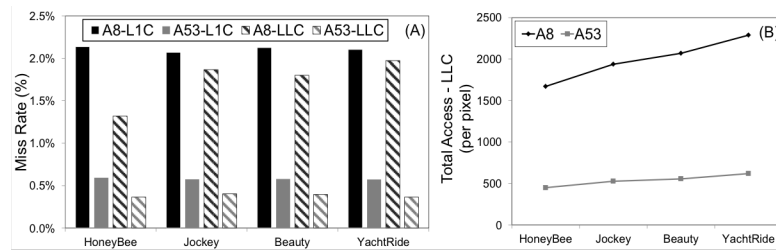


Figure 2. A8 vs. A53: Miss Rate (A) and Number of Memory Access (B)

Therefore we conclude that the higher capacity cache memories provide higher cache memory performance both at L1 and LL levels. The other important variable, energy consumption, is compared in Fig. 3. The plot in this Figure shows that the A53 processor cache requires about 20% more energy than that of the A8. In addition, the A53 processor cache is more expensive compared with the latter. These energy costs are significantly higher, since read and write operations on the A53 cache consume 61.2 and 82.2 times more power than the A8 counterparts. It should be noted that the more economical cache memory variants were selected from those offered by the processor manufacturer. In conclusion, the study indicates that the A8 processor cache is a better choice for energy and cost constrained platforms to be used in software-based UHD video encoding.

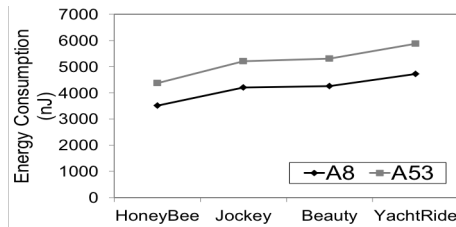


Figure 3. A8 vs. A53: Energy Consumption.

TABLE II. COMPLEXITY OF HD VERSUS UHD VIDEOS (HM)

Video Sequence	Enc. Time (s)		Increase Factor
	HD	UHD	
Beauty	2,610	12,737	4.9
Jockey	2,531	10,770	4.3
HoneyBee	1,825	8,383	4.6
YachtRide	2,871	11,695	4.1
Average	2,459	10,896	4.5

This brief study shows that controlled HEVC video encoding experiments can be a useful tool to choose cache architectures with low hardware costs and energy consumption, both in general and in the specific case of software based UHD video encoding. The next section will analyze the related problem of UHD video encoding computational cost.

4. Computational Complexity Analysis and Reduction

This section presents a discussion on the issue of HEVC UHD video computational complexity costs. Initially, a comparative analysis between the computational complexity of encoding HD and UHD videos is presented. Then, the application of a state-of-the-art complexity reduction strategy for HEVC encoders is evaluated for the encoding of UHD sequences in terms of computational complexity and coding efficiency.

Complexity Analysis of UHD Video Coding

In order to evaluate the computational complexity demanded by the process of encoding UHD sequences according to HEVC, the video sequences mentioned in section II.A have been encoded with the HM software under the experimental setup described in section II.C. For each sequence, both HD and UHD resolution versions were encoded and their respective encoding times were recorded for comparison. Table II presents the results obtained in terms of average absolute encoding time (in seconds) and relative

encoding time increase (in times) of the UHD version over the HD version of each video sequence.

Average results were calculated considering the four QPs mentioned in section II.A. On average, the change from HD to UHD represents an increase of 4.5 times in computational time. Notice that this increase is roughly the same as the increase in the number of pixels of UHD over HD videos (4 times).

Complexity Reduction applied to UHD Video Coding

Several works have been published in the last few years proposing techniques for computational complexity reduction of HEVC encoders. However, they are usually tested in accordance with the sequences of the CTC [5], which does not include any UHD video.

In this section, we analyze the performance of a data mining-based strategy proposed by Correa et al. [10], which is considered a state-of-the-art work, since it decreases the computational complexity of HEVC encoders by 2.9 times (or 65%) with BD-rate increases of only 1.3%, on average. When considering only HD video sequences, the method proposed in [10] is capable of decreasing the computational complexity by 3.3 times (or 70%), on average.

Table III presents the average results in terms of absolute encoding time (in seconds) for UHD videos encoded with the original HM encoder and with the fast encoder proposed in [10]. The table also presents in the rightmost column the encoding time reduction factor obtained when the fast encoder is used instead of the original HM. Notice in the last row of Table III that the encoding time can be decreased by a factor of 4.1 (about 75%), on average, which is very close to the increase caused by changing from HD to UHD representations (see Table II). This large reduction means that the undesired drawback of increasing the computational complexity when encoding UHD instead of HD videos can be nearly overcome if fast encoding strategies are applied. Also, notice that these experiments revealed that the data mining-based strategy proposed in performs still better with UHD videos than with HD and other lower resolutions. Finally, the compression efficiency results have shown a small average increase of 1.7% in bit rate (BD-rate) and a decrease of only 0.03 dB in image quality (BD-PSNR), when compared to the original HM encoder. These values are close to those obtained for HD videos and smaller resolutions, which proves that the strategy is suitable for UHD video encoding.

5. Conclusion

This letter showed that software-base encoding of UHD resolution video does indeed require more resources, both computational and memory related, when compared to the case of HD video encoding. The increase factors were quantified with recourse to simulations based on the HEVC reference software together with accurate memory usage monitoring tools and time-based complexity estimation procedures. The results presented show that the increases in both memory related costs (hardware cost and energy expenditure) and computational complexity incurred when moving from HD to UHD resolution video can be minimized with the use of a careful cache memory size and architecture choice and the use of heuristic-based HEVC computational complexity reduction methods.

Acknowledgements

This work was supported by R&D Unit UID/EEA/50008/2013, FCT grant SFRH BSASB 113682 2015 and the Parallel and Distributed Group (GPPD) at UFRGS, Brazil.

References

- [1] G. Sullivan, J. Ohm, Woo-Jin Han; T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," in IEEE TCSVT , vol.22, no.12, pp.1649-1668, Dec. 2012.
- [2] G. Correa, P. Assuncao, L. Agostini, L. A. da Silva Cruz, "Performance and Computational Complexity Assessment of High Efficiency Video Encoders", IEEE TCSVT, vol.22, no.12, pp. 1899-1909, 2012..
- [3] A. Iranpour and K. Kuchcinski. Memory Architecture Evaluation for Video Encoding on Enhanced Embedded Processors. 4017, 2006.
- [4] ISO/IEC-JCT1/SC29/WG11, "High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Improved Encoder Description," doc. N14970, 2014.
- [5] ISO/IEC-JCT1/SC29/WG11, "Common test conditions and software reference configurations", doc. JCTVC-J1100, 2012.
- [6] Ultra Video Group. Available at: <http://ultravideo.cs.tut.fi/>.
- [7] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," Austin, TX, 2001.
- [8] N. Nethercote and J. Seward. Valgrind: a framework for heavyweight dynamic binary instrumentation. In PLDI '07, New York, NY, USA, 2007.
- [9] N. Muralimanohar, R. Balasubramonian, N. P. Jouppi, CACTI 6.0, tech. report HPL-2007-167, HP Lab., 2007.
- [10] G. Correa, P. Assuncao, L. Agostini, and L. A. da Silva Cruz, "Fast HEVC Encoding Decisions Using Data Mining," IEEE TSCVT, vol.25, no.4, pp. 660-673, 2015.

Eduarda Monteiro received the B.Sc in Computer Science from the Catholic University of Pelotas (UCPel, 2009), Brazil, and the M.Sc. degree in Computer Science from the Federal University of Rio Grande do Sul (UFRGS, 2012), Brazil. She is a PhD student at UFRGS, where she works on memory and energy-aware video encoder optimizations.

Guilherme Correa has a Ph.D. in Electrical and Computer Engineering from the University of Coimbra (UC, 2015), Portugal. He is a postdoctoral researcher with the Group of Architectures and Integrated Circuits (GACI) at UFPel, where he works on video coding algorithms, complexity-aware video coding.

Luis A. da Silva Cruz has a Ph.D. degree in Electrical Computer and Systems Engineering from Rensselaer Polytechnic Institute (RPI), Troy, NY, US in 2000. He is with the Department of Electrical and Computer Engineering of the University of Coimbra and a researcher at the Instituto de Telecomunicacoes, where he works on video processing and coding.

Sergio Bampi Sergio Bampi received the M.Sc. and Ph.D. degrees in electrical engineering from Stanford University (USA) in 1986. He is full professor of the UFRGS since 1986 where he has been project leader of the Microelectronics and Computer Science Programs since 1988.

***Contouring artefacts prevention in compressed UHD video sequences:
tools and analysis of their performance***

André Seixas Dias, Matteo Naccari and Marta Mrak

British Broadcasting Corporation - R&D Department, London, UK

Andre.SeixasDias@bbc.co.uk, Matteo.Naccari@bbc.co.uk, marta.mrak@bbc.co.uk

1. Introduction

With the rapidly increasing availability and affordability of Ultra High Definition (UHD) consumer displays in the market, UHD video content and services are expected to become more popular in the near future. With its increased spatial resolution, UHD video allows a more detailed representation of the visual information and therefore is expected to provide end users with a more immersive and compelling quality of experience. Because of increased spatial resolution, UHD content is associated with significantly higher volumes of data, which challenges video compression technologies needed for its distribution.

The High Efficiency Video Coding (H.265/HEVC, [1]) standard significantly improves bitrate reductions with respect to its ancestors (e.g. H.264/AVC) for the same video quality [2]. Nevertheless, in certain scenarios, e.g. when the bitrate is low, hybrid block-based video codecs such as the ones specified by HEVC can introduce visual artefacts in the decoded output video, with the most prominent being blocking, ringing and blurring. Moreover, in image areas with low contrast and smooth variation of pixel intensities, contouring artefacts may also appear with false edges located along pixel variations, contributing to an undesirable degradation of the perceived video quality. For some UHD video sequences, this degradation occurs even when the objective quality measured in terms of PSNR is very high (e.g. around 45 dB). It is therefore important to reduce the visibility of these artefacts, since they can significantly deteriorate the superior quality of experience that UHD formats can deliver.

During compression, quantisation of residuals in the frequency domain causes the suppression of some small magnitude transform coefficients, leading to the loss of fine variation of pixel values in smooth image areas, thus introducing contouring. It should be noted that contouring mainly appears in standard dynamic range content (i.e. 8 bits per pixel) while is not present in high dynamic range content (e.g. 10 bits per pixel). This letter presents two main contouring prevention techniques specifically designed for hybrid block-based video codecs such as the one specified by HEVC. The first technique varies the Quantisation Parameter (QP) in contour prone areas in intra coded frames and propagates the reconstructed pixels in inter coded frames so that fine pixel value variations are preserved. To increase the granularity at which the quantisation step is varied, the second approach uses an adaptive quantisation method whereby the quantiser dead-zone is adjusted for some coefficients whose rounding to zero would lead to contouring. This second technique was originally proposed and tested only for Intra coding. In this letter, a further extension is proposed for Inter coding.

2. Related Work

Several solutions have been proposed in the literature to decrease the visibility of contouring artefacts. These solutions can typically be grouped into two main classes. The first class comprises post-processing techniques performed at the decoder side, without modifying the decoding process. As an example, Wang et al. [3] proposed a block-based dithering method to recover gradient and boundary smoothness after compression. This technique was applied only in image blocks prone to contouring, identified in the original frames based on the average pixel value level difference between a given block and its 8 neighbouring blocks.

The second class comprises methods which modify the encoding/decoding process. Yoo et al. [4] proposed a method to reduce the visibility of contouring by injecting pseudo-random noise after the in-loop deblocking filter of an H.264/AVC encoder. Visual quality improvements are reported with low RD losses. Contouring artefacts were also addressed during the development of HEVC [5]. The problem was considered to stem from a discontinuity in the reference samples used to perform Intra prediction in 32×32 blocks. This discontinuity was then propagated to the remaining frames. The solution consisted of modifying the reference samples by applying a bi-linear interpolation of the samples in the corners only in 32×32 blocks.

Given the increased computational resources needed at the receiver side to decode a UHD video, post-processing techniques may not be suitable for contouring removal. Therefore, the two techniques addressed in this letter ([6], [7]) fall into the second class.

3. Adaptive quantisation tools for contouring artefacts prevention

In this section, two proposed tools for contouring prevention in UHD video content are described. Figure 1 illustrates the integration of these tools in the workflow of a typical HEVC video encoder. As depicted in Figure 1, both techniques aim to adapt the quantisation process in areas that might suffer from contouring. The first technique [6] relies on the reduction of the QP and modified RD costs while the second [7] relies on an adaptive variation of the quantisation dead-zone.

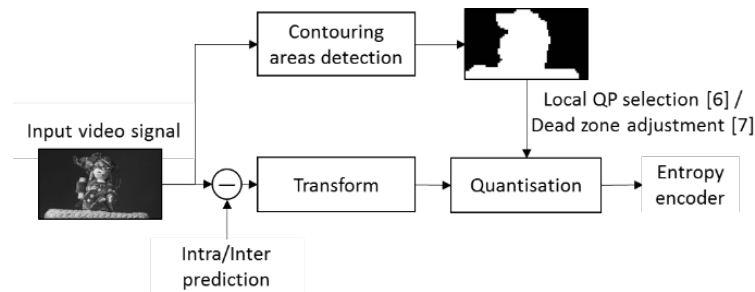


Figure 1: Proposed tools integrated in a HEVC encoder workflow

QP reduction and modified RD costs.

This technique relies on reducing the QP used by a video encoder in Intra frames, to enhance the perceived quality of contour prone areas [6]. As depicted in Figure 1, the contouring areas detection process takes as input the original video signal and determines which areas of each frame are prone to contouring after compression (contouring areas) and which areas are not (non-contouring areas). This classification was achieved by the block-based method proposed in [3], using a granularity of 64×64 blocks. The generated contouring map is then used to command the QP used in the quantisation process, only in Intra frames. This way, for areas where the compression process is likely to generate contouring artefacts, the QP used is lower in order to preserve the details in the prediction residual. For the remaining areas, the base QP is set according to the desired level of compression. It is important to note that decreasing the QP in contouring areas comes with the drawback of increasing the bitrate of the compressed bitstream. This bitrate increase also depends on the reduced QP used in contouring areas (the contouring QP), as in some cases, very low QPs need to be used to avoid contouring. In this letter, a QP = 20 for contouring areas was used in all sequences.

For Inter-coded frames, the QP reduction technique applied in Intra frames is replaced by a modification of the RD costs. This method was selected to prevent bitrate increases and possible persistence of

contouring artefacts in the decoded video [6]. The RD cost is modified to enable copying of previously encoded blocks as predictions in Inter frames, avoiding completely flat blocks in the output video signal. In order to do this, the computation of the cost for each merge mode in Inter frames is slightly modified according to

$$J = p \cdot D + \lambda \cdot R, \quad (1)$$

where D denotes the distortion introduced by the merge mode being evaluated, λ denotes the Lagrangian multiplier, R denotes the associated bitrate of the merge mode being tested and p is a parameter introduced to intentionally reduce the weight of the MSE-based distortion. The *Ningyo* sequence from the test set was used to assess the trade-off between coding penalty and contouring prevention at different coding points. From these experiments, the parameter p was set to

$$p = \begin{cases} 0.1 & \text{if } CBF = 0 \\ 0.7 & \text{otherwise} \end{cases}, \quad (2)$$

where CBF represents the Coded Block Flag for the merge mode being tested. A null CBF indicates that the residual after quantisation for the particular merge mode being tested is null, meaning that the final reconstructed block will be provided only from the selected predictor samples. The parameter p was introduced based on the assumption that the computed MSE-based distortion is not well correlated with the visual quality observed in contouring areas and that copying areas of previously encoded frames results in a better subjective quality for these cases. By introducing p , the weight of the distortion in the cost computation is reduced, since MSE is not a reliable metric to assess the perceptual quality of the decoded video in contouring areas.

Adaptive variation of the quantiser dead-zone

The technique presented in the previous section selects one QP for the whole coding block. When a finer level of granularity to adjust quantisation is required, the technique in [7] prevents contouring by adaptively varying the quantiser dead-zone so that the frequency coefficients associated with smooth transitions are retained. As shown in Figure 1, the luma component of each coded frame is used to detect contour-prone image areas by means of the same detector described in the previous section [3]. The map computed by the detector is then used to perform adaptive quantisation. More precisely, a dead-zone adjustment process is performed by tuning the rounding offset θ . In order to minimise the bitrate increment associated with the coding of more coefficients in contour prone image areas, the dead-zone adjustment is only performed over some particular coefficients. These coefficients represent the frequency distribution of image areas with smooth variations of pixel intensities. The coefficients where the adaptive quantiser dead-zone is performed are shown in Figure 2 for Transform Block (TB) sizes of 32×32 , 16×16 and 8×8 . The adjustment is not performed for 4×4 blocks as it was concluded that contour-prone areas are coded with larger block sizes. For more details about the frequency analysis performed to derive the coefficient positions in Figure 2, the reader is referred to [7]. For all the remaining coefficients, i.e. the non-shaded blocks in Figure 2, the value for θ is set to the default one used by the HEVC reference software, that is $1/3$ or $1/6$ for Intra- and Inter-coded slices, respectively.

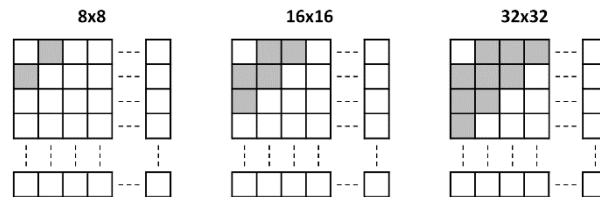


Figure 2. Selected coefficients for dead-zone adjustment, [7].

Given the critical coefficients highlighted in Figure 2, the value for θ is selected from a set of predefined values in the range $\Theta = [0.5, 0.8]$ spaced by a step $s = 0.05$. The optimal offset is selected on a TB basis

by counting the number of non-zero levels after quantizing the critical coefficients in each TB for a given offset $\theta' \in \Theta$. When the number of non-zero levels is greater or equal than a given threshold τ , offset θ' is selected for the current TB and quantisation takes places as specified in the HEVC standard. In this letter, $\tau = 1$ for 8×8 blocks while $\tau = 2$ for 16×16 and 32×32 blocks. These values were determined after visual inspection over some sequences in the test set where contouring artefacts were more noticeable.

As mentioned in the Introduction, this adaptive quantisation has been originally proposed for Intra coding [7]. In order to adapt this technique to Inter coding, the Random Access Structure Of Picture (SOP) is considered. This SOP is employed in broadcasting video coding applications given the efficient exploitation of temporal redundancy and the possibility to minimise the delay incurred when a channel switch occurs (which requires the insertion of an Intra-coded or random access point frame). Typical Random Access SOP has a length of eight frames and a hierarchy for temporal prediction is defined with four temporal layers. Each temporal layer in the hierarchy depends on layers in the higher levels with level 1 being the highest. Given this dependency, the adaptive dead-zone adjustment is applied as described above to frames in temporal layers 1 and 2. This choice is introduced to control the bitrate increments when more coefficients are retained. Finally, over TBs affected by contouring, the Rate-Distortion Optimised Quantisation (RDOQ, [8]) is disabled since its cost function does not take into account contouring artefacts.

4. Experimental Results

This section reports and analyses the performance of both techniques described in Section 3. The test set comprises five sequences with 8 bits per component, 4:2:0 chroma format, 3840×2160 spatial resolution and frame rate of 50 and 60 fps. Figure 3 depicts the thumbnails for the first frame for each video. The test set is composed of content representative of broadcasting (e.g. Badminton and ShowDrummer) with all sequences containing large contour prone areas. The HEVC reference software (HM Version 14.0) was used as the base software for the implementation of both methods using the Random Access SOP and the JCT-VC common test conditions [9]. For the sake of simplicity, the QP reduction with modified costs technique is denoted as QPMC while the adaptive dead-zone adjustment is denoted as ADZA.



Figure 3. Original UHD test sequences.

Performance indicators

The performance of both contouring prevention methods is assessed in terms of coding efficiency penalty and image quality improvement. The coding efficiency penalty is measured by using the Bjøntegaard Delta-rate (BD-rate, [10]) assuming HM as anchor and considering the luma component only since in the YCbCr colour space contouring mainly affects this component. For picture quality, the Pixel Variation Preservation (PVP) score [7] is used. PVP quantifies the preservation of pixel variations over contour prone blocks. PVP assumes values in the range $[0, +\infty)$ wherein the higher the score the less visible is contouring.

Results and analysis

Table 1 reports the BD-rate values for both the QPMC and ADZA techniques with respect to the HM reference software. All coded sequences have been also visually inspected to ensure that no noticeable contouring was present in the content coded with both methods. From the results in Table 1, it can be seen

that the impact in terms of BD-rate penalty associated with the ADZA is smaller than QPMC. This can be explained by the fact that in QPMC, a very low QP needs to be selected in contouring areas to completely remove contouring artefacts, even for lower target qualities (e.g. base QP = 37). This means that the QPMC technique is more suitable for high target qualities, where the difference between the base QP and the contouring QP is not significantly high.

Table 1. BD-rate results for QPMC and ADZA.

Sequence	Y BD-rate [%]	
	QPMC	ADZA
<i>Badminton</i>	10.0	2.3
<i>CandleSmoke</i>	12.4	2.0
<i>Ningyo</i>	13.9	7.9
<i>ShowDrummer</i>	15.7	2.0
<i>YoungDancers</i>	12.7	5.2
Average	12.9	3.9

The results of the PVP score are depicted in Figure 4. Both techniques described in this letter achieve higher PVP scores than the anchor, which means that both succeed in preserving the small details in contouring areas that are discarded after compression. Due to the QP reduction used in contouring areas, the QPMC method is able to preserve these details consistently better than the ADZA technique.

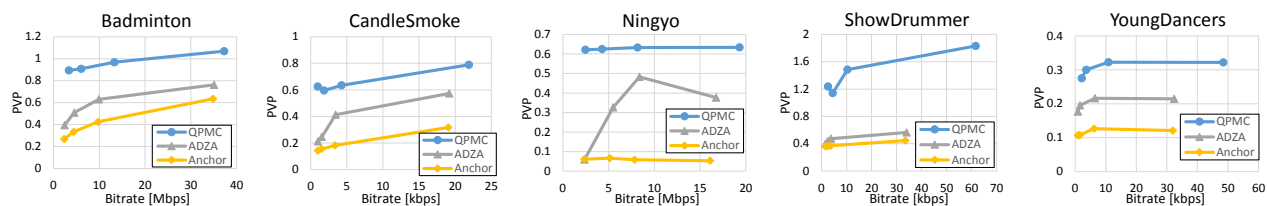


Figure 4. PVP Scores for QPMC and ADZA.

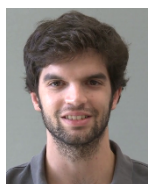
4. Final Remarks

This letter described and compared the performance of two methods proposed to prevent contouring artefacts in UHD video content. The first technique (QPMC) is based on the reduction of the QP and modified RD costs while the second (ADZA) relies on an adaptive variation of the quantisation dead-zone. The results show that both methods are able to avoid contouring in HEVC encoded video and therefore improve the perceptual quality of the output video with a moderate BD-rate penalty. Although the QPMC method is able to preserve better the pixel variations in contouring areas according to the PVP score, higher bitrate increases are introduced for lower qualities, which makes this method more suitable for higher target qualities, where the associated bitrate increase is lower.

References

- [1] G. Sullivan, J. Ohm, W.-J. Han and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE TCSVT.*, vol. 22, no. 12, pp. 1649 - 1668, Sep. 2012.
- [2] T. K. Tan, R. Weerakkody, M. Mrak, N. Ramzan, V. Baroncini, J. Ohm and G. Sullivan, "Video Quality Evaluation Methodology and Verification Testing of HEVC Compression Performance," *IEEE TCSVT*, (to appear).

- [3] Y. Wang, C. Abhayaratne, R. Weerakkody and M. Mrak, “Multi-scale dithering for contouring artefacts removal in compressed UHD video sequences,” in *IEEE Global Conf. Signal Inform. Process.*, Atlanta, GA, USA, Dec. 2014.
- [4] K. Yoo, H. Song and K. Sohn, “In-loop selective processing for contour artefact reduction in video coding,” *Electronics Letters*, vol. 45, no. 20, p. 1020–1022, Sep. 2009.
- [5] T. K. Tan and Y. Suzuki, “Contouring artefact and solution - Technical Report JCTVC-K0139,” in *JCTVC 11th meeting*, Shanghai, China, Oct. 2012.
- [6] A. S. Dias and M. Mrak, “Region-Adaptive Quantisation for Prevention of Contouring in Coded Video,” in *IEEE Multimedia Sig. Proc. Conf.*, Xiamen, China, Oct. 2015.
- [7] N. Casali, M. Naccari, M. Mrak and R. Leonardi, “Adaptive Quantisation in HEVC for Contouring Artefacts Removal in UHD Content,” in *IEEE Int. Conf. on Image Processing*, Quebec, Canada, Sep. 2015.
- [8] M. Karczewicz, Y. Ye and I. Chong, “Rate distortion optimized quantization,” in *ITU-T SG16/Q6 VCEG, Doc. VCEG-AH21*, Antalya, Turkey, Jan. 2008.
- [9] F. Bossen, “Common HM test conditions and software reference configurations,” in *JCTVC-L1100, JCTVC 12th meeting*, Geneva, Switzerland, Jan. 2013.
- [10] G. Bjontegaard, “Calculation of average PSNR differences between RD-curves,” in *VCEG-M33, VCEG 13th meeting*, Austin, Texas, USA, Apr. 2001.



André Seixas Dias received his M.Sc. in electrical and computers engineering from Instituto Superior Técnico, Lisbon, in 2013. In 2014, he joined BBC R&D as an early stage researcher to work on perceptually optimised video compression. His main research interests are video compression and perceptual video analysis.



Matteo Naccari received his Laurea and Ph.D. degrees from the Technical University of Milan in 2005 and 2009, respectively. In 2011 he joined BBC R&D as Senior Research Engineer working on the definition of the HEVC standard and its extensions. His main research interests are compression of UHD content with HEVC, perceptual video coding and high dynamic range imaging.



Dr Marta Mrak is Lead R&D Engineer at British Broadcasting Corporation. Her main research activities are related to video coding, standardisation and signal processing, and on related subjects she co-authored over 100 publications and co-edited two books. Dr. Mrak is a member of the IEEE Multimedia Signal Processing Technical Committee, an Area Editor of *Signal Processing: Image Communication* (Elsevier), and a Guest Editor of several special issues in relevant journals.

DASH Segment Scheduling for Scalable UHD Video Streaming*Eduardo Asbun¹, Ahmed Hamza², Srinivas Gudumasu³, and Yong He¹**¹InterDigital Communications, Inc., U.S.A., ²Simon Fraser University, Canada,**³Aricent Technologies Ltd., U.S.A.**{ed.asbun, yong.he}@interdigital.com, srinivas.gudumasu@aricent.com, aah10@sfu.ca***1. Introduction**

Streaming of multimedia content over the HTTP is ubiquitous, using technologies based on international standards (e.g., MPEG DASH [1]) and on proprietary solutions (e.g., Apple's HTTP Live Streaming (HLS) and Microsoft's Smooth Streaming). In HTTP-based streaming, content coded at different bit rates is partitioned into segments, typically between 2 and 10 seconds in length. Segments are stored in streaming servers and distributed via content delivery networks (CDN) and/or edge server networks. When users request multimedia content, the HTTP origin server sends back to the client a media presentation description (MPD) manifest file, describing all available bit rates of the requested content as well as available URLs from which video may be downloaded. At the beginning of the session, clients typically request segments at low rates, allowing them to assess network conditions. Bit rate of the following segments is chosen by optimizing a combination of factors, including quality requirements, available network bandwidth and device capabilities, allowing clients to receive better quality video with lower latency and shorter start-up and buffering time.

As device capabilities increase, streaming of immersive audio and of video content of higher spatial resolution (e.g., ultra-high-definition, UHD – 3840×2160 and above), frame rate (e.g., 60-120 Hz), bit depth (e.g., 10-12 bits), high dynamic range (HDR) and wide color gamut (WCG), is becoming increasingly common. This poses new challenges for streaming systems because new video coding techniques must be used to efficiently carry a much larger amount of information. The HEVC video coding standard and its scalable extensions, referred to as SHVC [2], are the latest technologies used to address this problem.

2. Ultra-High-Definition SHVC Streaming System

In this paper, we describe an algorithm used by a proprietary multimedia client to stream ultra-high-definition video using SHVC. The goal of the algorithm is to maintain consistent video quality during playback; that is, it aims to avoid abrupt video quality fluctuation, which may degrade user's quality of experience. In our system we use layered video coding, a core technology of SHVC, along with DASH, to efficiently deliver content to devices of different capabilities. For example, in a simple SHVC configuration for spatial scalability, a base layer (BL) may carry an HD video stream, while an enhancement layer (EL) may be combined with the base layer to obtain a UHD video stream. When DASH is used to stream SHVC content, video may be segmented by layers and stored separately: one set of segments contains the BL bitstream, while other set holds the EL bitstream. (Other sets may exist if more than one enhancement layer is used.) In this example, some clients, due to limited device capabilities, may request BL segments only and decode them using a HEVC decoder to obtain a HD stream. Other clients may request both BL and EL segments and pass them to a SHVC decoder to obtain a UHD stream, as shown in Figure 1. Other more elaborate configurations for streaming video systems are described in [3].

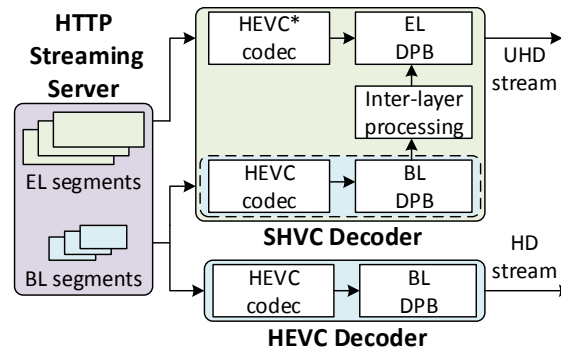


Figure 1. Example of a SHVC system with spatial scalability. (DPB: Decoded Picture Buffer)

In DASH, dependencies across layers are signaled in the MPD. Let $S(i, j)$ be the j -th segment of the i -th layer, with the $(i+1)$ -th layer being dependent on the i -th layer. A client can decode $S(i+1, j)$ only if $S(i, j)$ has already been decoded. That is, a layer can only be decoded if all its dependent layers have already been decoded. This dependency across layers is shown in Figure 2. SHVC is flexible in that a layer may be dependent on only a subset of layers below it. However, to simplify the presentation in this paper, this use case is not discussed

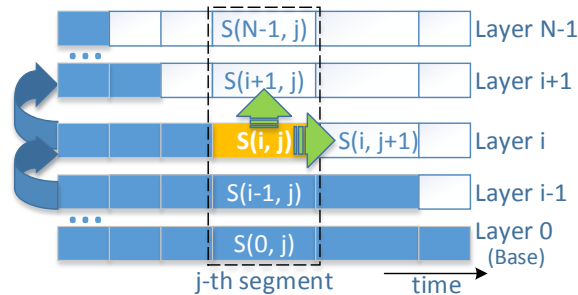


Figure 2. Dependencies across layers in SHVC.

3. Video Segment Scheduling

In this section, we describe the segment download scheduling algorithm of our SHVC streaming client. After downloading segment $S(i, j)$, the client must decide whether to download a higher quality segment $S(i+1, j)$ or the next segment in temporal order, $S(i, j+1)$, as shown by green arrows in Figure 2. Our algorithm uses the following constraints to make the selection:

- The selected segment must maximize the video quality under current bandwidth constraints.
- After downloading the selected segment, the buffer fullness of each layer must be between low and high thresholds to avoid underflow or overflow.
- The selected segment must minimize quality variation despite network bandwidth fluctuation.

An algorithm for scheduling SHVC segment downloads is different from multi-rate HEVC simulcast, because of inter-layer dependence. That is, segments from multiple layers must be downloaded and passed together as a set to the SHVC decoder. The conventional downloading approach solely based on the network bandwidth may lead to frequent changes in video quality as the client adapts to varying network conditions. We address this challenge with the algorithm described next.

Algorithm description

To satisfy the constraints listed above, we keep separate download buffers for each layer and set a “backfilling threshold” for each buffer. The client fetches segments of layer i until the backfilling threshold is reached. Then, it fetches segments from layer $i+1$, as shown in Figure 3.

The backfilling threshold, $BF[i]$, represents the buffer fullness level at which sufficient segments are considered to be available in the i -th layer’s buffer, such that the client may fetch earlier segments of the $(i+1)$ -th or higher layers. Denote the last segment in the i -th layer buffer as $S(i, j)$. When $BF[i]$ is reached, the client will start fetching segments $S(l, k)$, for which $l \geq i$, and $k \leq j$. In other words, the backfilling threshold is a threshold that determines when a condition (for example, available bandwidth) is sufficient for the client to backfill the buffer with higher quality (i.e., higher layer) segments from a time earlier than the latest downloaded segment. Figure 3 illustrates buffer thresholds for a three-layer buffer model where layer-2 depends on layer-1 and layer-1 depends on layer-0. Segment $S(1,3)$ (#7, in yellow) is the latest downloaded segment and the backfilling threshold for this layer ($BF[1]$) has been reached. Therefore, the client determines that $S(2,1)$ (#8) should be downloaded next, assuming bandwidth is sufficient. Buffer fullness ($F[i]$) is determined by the latest segment downloaded in each layer.

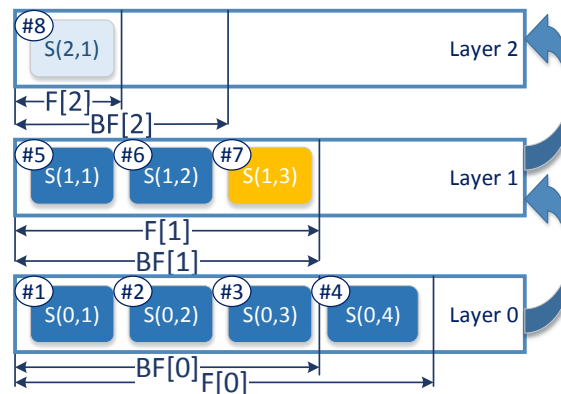


Figure 3. Backfilling threshold (BF) and buffer fullness (F) of download buffers of a three-layer SHVC system.

The backfilling thresholds are chosen to balance between consistent video quality vs faster adaptation to bandwidth changes. For example, a high backfilling threshold value would maintain more consistent video playback quality when bandwidth changes, but may not switch up to higher quality video quickly when more bandwidth becomes available. A low backfilling threshold value would increase the video quality faster when bandwidth increases, but may introduce quality oscillation when network bandwidth fluctuates. Due to the layer dependency, the backfilling threshold of a higher layer must be less than or equal to the backfilling threshold of its dependent layers ($BF[i+1] \leq BF[i]$).

Segment selection process

As segments are removed from the download buffer to be passed to the SHVC decoder (e.g., #1, #5 and #8 in Figure 3), it is possible that download of a higher layer segment may not be completed in time (e.g., #8), forcing the client to discard the partially downloaded segment, thus increasing network utilization without improving visual quality. Therefore, the client must determine whether download will complete before a segment’s decode time. Otherwise, a more effective use of resources would be to download a later segment of the same layer. Thus, for layer i , the client has to calculate the download time of segment using the current bandwidth estimate, and compare it with the decode time of the segment. This is done as

follows for layer i : Let $S[j]$ be the segment size in bits, $D[j]$ be the decoding time of the segment and BW_{est} be the estimated network bandwidth. Further, assume that the duration of each segment (T), indicated in the MPD, is the same. (While this assumption may not be accurate, the client only learns about the actual size in bytes and duration in seconds when the segment is received.) Figure 4 shows the candidate segments for layer i .

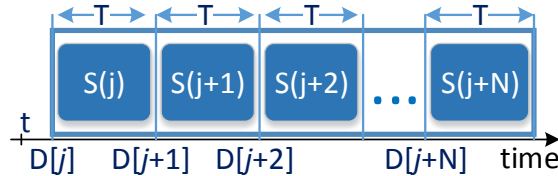


Figure 4. Candidate segments for layer i .

Given the current time t , the minimum bandwidth $BW_{min}[j]$ required to download $S(j)$ can be obtained by:
 $BW_{min}[j] = S[j]/(D[j]-t)$ The maximum download time for $S(j)$ is:
 $Dw_{max}[j] = D[j] - t$

The maximum download time for $S(j+1)$ is:

$$Dw_{max}[j+1] = D[j+1] - t = D[j] + T - t$$

Therefore, the minimum bandwidth required to download $S(j+1)$ is:

$$BW_{min}[j+1] = S[j+1]/Dw_{max}[j+1] = S[j+1]/(D[j]+T-t)$$

BW_{min} for all candidate segments is calculated. The segment $S(k)$ to be downloaded is the first segment for which the following condition is true:

$$BW_{est} \geq BW_{min}[k]$$

4. Performance evaluation

We implemented a SHVC streaming client, using the open source OpenHEVC as decoder [4]. In our testbed, shown in Figure 5, we varied the bandwidth available to the client with Linux traffic control tools that use a token bucket filter (TBF) as the queueing discipline. In our tests we evaluated performance by simulating varying network conditions using artificial test patterns and a bandwidth trace from an LTE network.

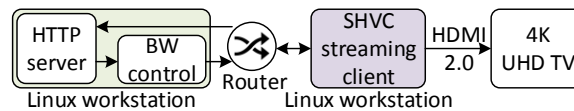


Figure 5. Testbed for SHVC streaming client.

In Figure 6, we show a plot of bandwidth utilization during a portion of a UHD streaming session. In the plot, the current time ("now") is when $t=0$, and events for $t>0$ occurred in the past. In this test, we controlled the available bandwidth using a binary pattern (20 and 45 Mbps) for a two-layer SHVC stream encoded at 16 and 43 Mbps (BL: 1536×832 / EL: 3072×1664 , 24 fps, 2 seconds per segment). The plot

shows a streaming session that lasted approximately 40 seconds. The client initially fetches base layer segments only (shorter blue bars) to fill up its buffer and because the available bandwidth (20 Mbps) is not enough to download EL. When available bandwidth increases to 45 Mbps, backfilling takes place, that is, consecutive EL segments (longer blue bars) are downloaded. This improves the visual quality of decoded video because these EL segments correspond to BL segments that have already been downloaded. Thereafter, the client fetches BL and EL segments because enough bandwidth is available to fetch both. When available bandwidth decreases, the client continues to fetch BL and EL segments but at a slower pace. If available bandwidth remains the same, the client would eventually fetch BL segments only (not shown in Figure 6).

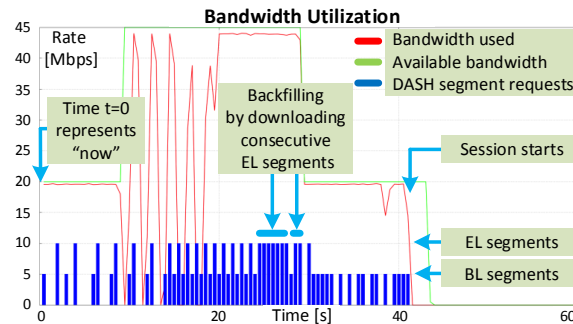


Figure 6. Bandwidth utilization of the proposed segment scheduling algorithm.

Figure 6 also shows one of the advantages of scalable video, namely improved visual quality when available bandwidth increases. The client can fetch EL segments that correspond to BL segments already downloaded and combines them to show higher quality video, that is, it can improve quality incrementally. In the same scenario but using traditional multi-rate streaming, lower rate segments already downloaded would have to be discarded and replaced by higher rate segments, effectively wasting data; or there would be a delay in transitioning to higher rate, decreasing user experience.

5. Conclusion

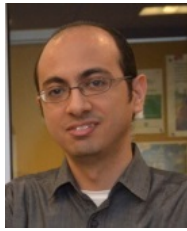
In this paper, we described an algorithm used for scheduling download of DASH segments by a SHVC streaming client. Goals of the algorithm are to minimize variation of visual quality, while using network bandwidth effectively. Development of segment scheduling algorithms for scalable video has challenges different from the ones found in conventional multi-rate streaming, in particular, inter-layer dependence. Thus, development of custom algorithms that address these challenges is important given the increasing importance of scalable video streaming.

References

- [1] I. Sodagar, "The MPEG-DASH Standard for Multimedia Streaming Over the Internet", IEEE Multimedia, vol. 18, no. 4, pp. 62-67, 2011.
- [2] Y. Ye and P. Andrivon, "The Scalable Extensions of HEVC for Ultra-High-Definition Video Delivery," IEEE Multimedia, vol. 21, no. 3, pp. 58-64, 2014.
- [3] Y. Ye, Y. He, and X. Xiu, "Manipulating Ultra-High Definition Video Traffic," IEEE Multimedia, vol. 22, no. 3, pp. 73-81, 2015.
- [4] Wassim Hamidouche, Mickael Raulet and Livier Deforges, "Multi-core Software Architecture for the Scalable HEVC Decoder", IEEE ICASSP, 2014



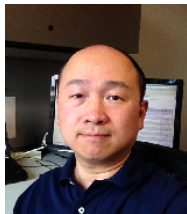
Eduardo Asbun received a B.S.E.E. from Metropolitan University, Mexico, and M.S.E.E. and Ph.D. degrees from Purdue University, Indiana, USA. He is a Member of Technical Staff in InterDigital Communications, Inc., San Diego, CA, USA.



Ahmed Hamza is a Ph.D. candidate at the School of Computing Science, Simon Fraser University, Canada. He worked as a research intern with InterDigital Communications, Inc., San Diego, CA, USA. His research interests are in the areas of multimedia networking, distributed systems, and future video technologies.



Srinivas Gudumasu is a technical leader in Aricent Technologies (Holdings) Ltd, USA and working as consultant to InterDigital Communications, Inc., San Diego, CA, USA. He received a master's degree from Indian Institute of Technology Madras and bachelor's degree from Acharya Nagarjuna University.



Yong He is a Member of Technical Staff in InterDigital Communications, Inc., San Diego, CA, USA. He received a Ph.D. degree from Hong Kong University of Science and Technology, and M.S. and B.S. degrees from Southeast University, China.

UHD Video Retargeting based on Visual Attention Models and Temporal Consistency*Lino Ferreira^{1,2,3} and Luís A. da Silva Cruz^{1,2} and Pedro Assuncao^{1,3}**¹Instituto de Telecomunicações, Portugal**²Universidade de Coimbra/DEEC, ³Intituto Politecnico de Leiria/ESTG, Portugal**lino.ferreira@ipleiria.pt, {lcruz, amado}@co.it.pt***1. Introduction**

The widespread use of mobile devices with increasing computational and communications resources has been pushing users to access and subscribe a great variety of multimedia content and services. However, the current trend towards Ultra High Definition (UHD) video still imposes constraints in small screen devices, which may limit the access to such type of content. For this reason, video and image retargeting methods have been receiving increasing attention from the research community due to its importance for content and service providers to guarantee flexible and seamless access to all types of content [1][2].

In general, video retargeting may include image resize, crop and even frame rate adaptation, but common to all solutions is the fact that the most relevant information of each frame should be maintained, while less relevant regions might be discarded according to the service or application requirements. In addition, temporal consistency is required in retargeted video sequences to ensure smooth transitions between consecutive cropped images such that the equivalent camera motion appears realistic, i.e., without noticeable visual artifacts [3]. Among the most common visual artifacts one may identify object truncation, jitter and/or flickering in temporal domain. To address these problems, several content-aware retargeting methods have been proposed. Some of them are focused on keeping the visual salient regions (e.g. [4]) while others are designed to reduce visual artifacts and ensure the temporal consistency [5][6].

This e-Letter deals with UHD video retargeting based on visual saliency regions and temporal consistency. A new method is presented to adapt UHD video to lower resolution formats based on aggregated saliency maps, which identify the salient regions to be cropped in each image followed by temporal filtering of their spatial location in the UHD resolution to improve temporal consistency and fluidity. The visual saliency maps were computed by fusion of three intermediate saliency maps and center-bias weighting to better match the observer behavior. The results show that enforcing temporal consistency leads to better visual quality of the retargeted video as well as higher coding efficiency.

2. Related work

In general, video retargeting is a content-aware adaptation process achieved through cropping and scaling of high-resolution images [7]. As mentioned before, its application to mobile communications has been driving research and development of efficient methods for seamless content adaptation. Although fixed approaches may be used, such as cropping the center of all images in a sequence, useful retargeting methods should take into consideration the type and location of objects in the visual scene [8].

Useful methods of video retargeting include constraints to enforce various types of visual consistency in order to achieve improved performance [9,10]. Such methods can be classified into: content-aware, seam carving, warping and hybrid methods. The content-aware methods are mostly based on the principle that salient regions must be preserved while the content in the borders of the frame can be dropped [8]. Seam carving changes the size of a video according to its content but the saliency regions are not considered, which may cause artifacts in large objects. Kolf [11] and Grundmann [12] used image seam carving methods with adaptations to keep temporal coherence in video. Yan *et al.*, [13] present a motion-aware seam carving solution to guarantee temporal smoothness of seams. More recently, Wang *et al.*, [14] proposed a shape matching solution to protect the shapes of the salient curves from deformation. In

addition to the above methods, some combinations may be used. The most advanced methods use visual attention models to identify the important regions of high-resolution images [15]. However, adequate temporal consistency is not guaranteed with visual attention models, because even those that include motion in saliency computation cannot provide adequate smooth transition between consecutive cropped windows.

3. UHD video retargeting with temporal consistency

The UHD video retargeting method presented in this e-Letter is comprised of two main steps (Figure 1): (i) the saliency map computation based on attention modeling and (ii) image cropping based followed by jitter removal through temporal filtering of the spatial location of the cropping windows. Although this e-Letter includes examples of retargeting from UHD 4K to HD 720p, application to any lower resolution is straightforward.

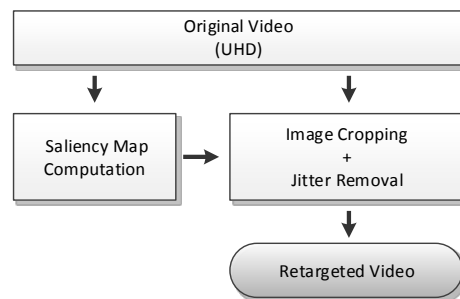


Figure 1. UHD video retargeting method

Saliency map computation

The saliency map computation is based on the method described in [16], which aggregates three saliency feature maps, computed from the spatial (texture), temporal (motion) and face presence information. These saliency feature maps are combined into a single saliency map per frame. Figure 2b shows the saliency map obtained for the *Jockey* sequence, where the white regions are the most relevant and the white square represents the face-region of the horseman.

Image cropping

A cropping window identifies the most important region of each UHD frame, based on the saliency map. The size of the cropping window is the same as that of retargeted images and it is given as a user-defined parameter. A search is performed within a search window (SW) of the saliency map to find the best sub-image with the pre-defined resolution, i.e., the cropped image. This is the one that maximizes the energy of the saliency map signal located inside the cropping window. For each UHD frame, a single cropping window is determined to define the cropped image, as shown in Figure 2

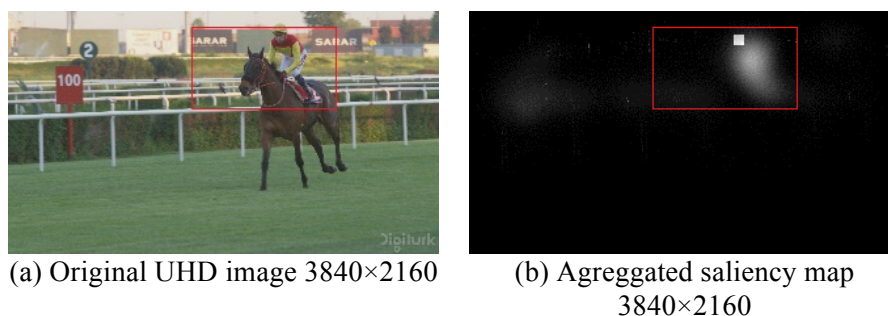


Figure 2. Red box: cropping window of size 1280×720 *Jockey* sequence

Filtering for temporal consistency (jitter removal)

Temporal consistency is achieved by filtering abrupt changes in the spatial location of cropped images across consecutive frames, i.e., jitter removal. Figure 3 shows the temporal evolution of the up-left corner of the cropping window (horizontal and vertical position). The blue and red lines represent the x and y spatial positions before and after median filtering, respectively. As shown in Figure 3, after jitter removal, the up/down displacements of the cropping window follows a much smoother trajectory in comparison with the initial one, i.e., the raw data obtained directly from the maximum-energy criterion applied to the saliency map.

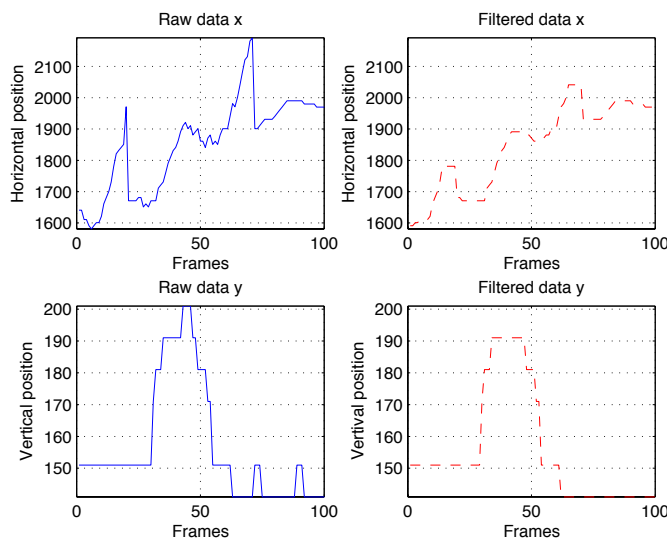


Figure 3. Temporal evolution of the spatial location of the cropping window (*Jockey* sequence).

4. Performance evaluation

The performance of the previous UHD retargeting method was evaluated through comparison with a downsizing method (using the MPEG-4 downsampling filter) and centered cropping. A visual comparison was performed as a user-driven approach while the impact of enforcing temporal consistency in the coding efficiency of retargeted video was also evaluated. Rate-distortion (R-D) efficiency was used for the latter. The *Bosphorus* and *Jockey* UHD test sequences were used in the evaluation. The same HD 720p resolution (1280×720) was used for all cases of retargeting from the UHD 4K resolution (3840×2160). The HEVC reference software HM-16.6 was used to encode the retargeted video sequences. The encoder configuration was set to *Random Access*. The R-D operational points used to obtain the R-D function were obtained from the set of $QP=\{22,27,29,32\}$ for 100 frames.

Visual comparison

Figures 4b, 4c and 4d show the visual results of three different retargeting methods: (i) downsizing, (ii) centered cropping and (iii) the method presented in Figure 1, for sequence *Jockey*. Downsizing preserves the context of the scene but some important objects may not be recognizable due severe loss of detail. For example in the *Jockey* scene, the horse number and rider face are not discernible. In centered cropping it is assumed that the region of interest is always located in the center of the image, but such assumption is not true in many cases, particularly in UHD because the number of regions of interest tend to be highly spread over larger resolution images. For instance, in Figure 4c, one may clearly observe that center cropping is not an acceptable solution. Regarding the method described in this e-Letter, the retargeted images contain the essential information of the scene, as shown in Figure 4d. Since downsizing is not used, it is possible to keep some high-resolution details, such as the bird on the fence and the face of the horseman.



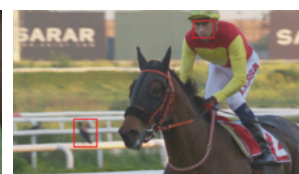
(a) Original 3840×2160



(b) Downsize



(c) Centered cropping

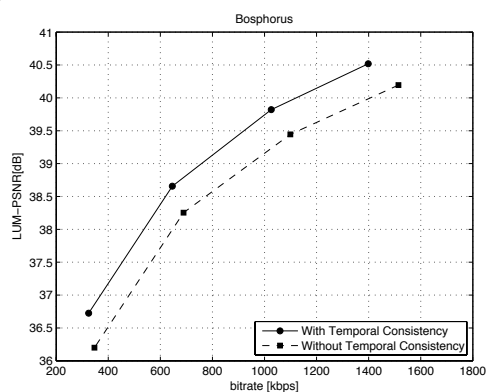
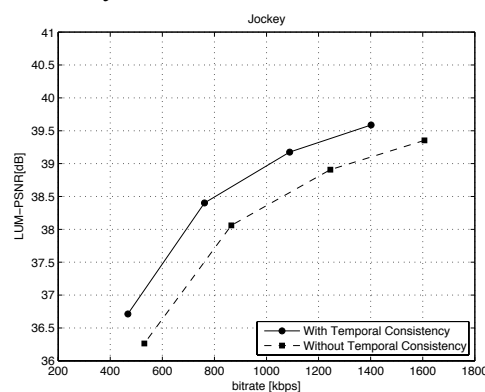


(d) Cropping from visual saliency

Figure 4. Visual comparison of *Jockey* sequence.

The influence of temporal consistency on coding efficiency

To evaluate the impact of enforcing temporal consistency in the coding efficiency of retargeted video, a set of simulations were carried out specifically for this purpose. In these tests, the R-D efficiency of the retargeted video with and without temporal consistency is compared for two test sequences, *Jockey* and *Bosphorus*. Figures 5 and 6 show that better efficiency is always obtained for retargeting with temporal consistency. The difference between the two methods lies in range 0.5~1.0dB for both test sequences. This is because temporal consistency ensures greater correlation between adjacent frames, which improves motion compensated prediction and benefits coding efficiency. When jitter removal is not used and just the maximum-energy cropping window is used to produce the retargeted video, there are sharp discontinuities between the spatial locations of adjacent frames. This leads to non-matching regions in the borders, thus more failures in motion estimation and lower coding efficiency. Implicitly, a smooth trajectory of the cropping window over the time dimension benefits the R-D efficiency. From additional experiments carried out with different sequences, a PSNR gain of approximately the same magnitude was obtained for various types of content, which seems to indicate that more important than the content itself is the jitter removal filter used to enforce the temporal consistency.

Figure 5. R-D of *Bosphorus* sequenceFigure 6. R-D of *Jockey* sequence

4. Conclusion

In this paper UHD video retargeting was discussed and a method based on visual saliencies driven by visual attention models was presented. A visual comparison of the outputs generated by such method shows its ability to preserve the relevant content of the UHD scene. Furthermore, when HEVC RA is used to encode retargeted UHD video, consistently better PSNR is obtained when temporal consistency is enforced through filtering for jitter removal, which shows a double benefit, both for the visual quality and for the coding efficiency.

Acknowledgements

This work was supported by R&D Unit UID/EEA/50008/2013 and PhD Grant SFRH/BD/37510/2007, co-funded by FEDER-PT2020, FCT/MEC, Portugal.

References

- [1] Fule, R.Y.; Theng, D.P., "Issues and Challenges in Content-Based Media Retargeting," 4th International Conference on Communication Systems and Network Technologies, pp.881-885, April 2014.
- [2] Tan, W.; Yan, B.; Li, K.; Tian, Q., "Image Retargeting for Preserving Robust Local Feature: Application to Mobile Visual Search," IEEE Transactions on Multimedia, vol.18, no.1, pp.128-137, Jan. 2016.
- [3] Bo Yan; Binhang Yuan; Bo Yang, "Effective Video Retargeting With Jittery Assessment," IEEE Transactions on Multimedia, vol.16, no.1, pp.272-277, Jan. 2014.
- [4] L. Wolf, M. Guttman, and D. Cohen-Or, "Non-homogeneous Content-driven Video-retargeting," IEEE 11th International Conference on Computer Vision, 2007. ICCV 2007, pp. 1–6, Oct. 2007.
- [5] W.-L. Chao, H.-H. Su, S.-Y. Chien, W. Hsu, and J.-J. Ding, "Coarse-to-fine temporal optimization for video retargeting based on seam carving," IEEE International Conference on Multimedia and Expo, pp. 1–6, July 2011.
- [6] Zheng Yuan; Taoran Lu; Huang, T.S.; Dapeng Wu; Yu, H., "Addressing Visual Consistency in Video Retargeting: A Refined Homogeneous Approach," IEEE Transactions on Circuits and Systems for Video Technology, vol.22, no.6, pp. 890-903, June 2012.
- [7] Wen Gao; Ling-Yu Duan; Jun Sun; Junsong Yuan; Yonggang Wen; Yap-Peng Tan; Jianfei Cai; Kot, A.C., "Mobile media communication, processing, and analysis: A review of recent advances," IEEE International Symposium on Circuits and Systems (ISCAS), , vol., no., pp.869-872, May 2013.
- [8] Engelke, U.; Junle Wang; Marendy, P., "Perceptual Relevance Based Image Retargeting", IEEE Signal Processing Letters, vol.22, no.6, pp.705-708, June 2015.
- [9] Zheng Yuan; Taoran Lu; Huang, T.S.; Dapeng Wu; Yu, H., "Addressing Visual Consistency in Video Retargeting: A Refined Homogeneous Approach," IEEE Transactions on Circuits and Systems for Video Technology, vol.22, no.6, pp.890-903, June 2012.
- [10] Bing Li; Ling-Yu Duan; Jinqiao Wang; Rongrong Ji; Chia-Wen Lin; Wen Gao, "Spatiotemporal Grid Flow for Video Retargeting," IEEE Transactions on Image Processing, vol.23, no.4, pp.1615-1628, April 2014.
- [11] S. Kopf, J. Kiess, H. Lemelson, and W. Effelsberg, "FSCAV: Fast Seam Carving for Size Adaptation of Videos," Proceedings of the 17th ACM International Conference on Multimedia, 2009, pp. 321–330.
- [12] M. Grundmann, V. Kwatra, M. Han, and I. Essa, "Discontinuous seam-carving for video retargeting," IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 569–576.

- [13] B. Yan, K. Sun, and L. Liu, "Matching-Area-Based Seam Carving for Video Retargeting," IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 2, pp. 302–310, Feb. 2013.
- [14] B. Wang, H. Xiong, Z. Ren, and C. W. Chen, "Deformable Shape Preserving Video Retargeting With Salient Curve Matching," Emerg. Sel. Top.
- [15] Mateescu, V.; Bajic, I., "Visual Attention Retargeting," IEEE MultiMedia , vol.PP, no.99, pp.1-1, July 2015.
- [16] L. Ferreira, P. Assuncao, and L. Cruz, "A method to compute saliency regions in 3D video based on fusion of feature maps," IEEE International Conference on Multimedia and Expo, July 2015.



Lino Ferreira received the electrical engineering degree in 2001 and he is currently pursuing the Ph.D. degree in Electrical and Computer Engineering at the University of Coimbra, Portugal. Since 2003, he has also been with the Department of Electrical Engineering, Polytechnic Institute of Leiria, and Instituto de Telecomunicações, Leiria, Portugal. His research interests include digital signal and image/video processing, 2D and 3D video coding.



Luis A. da Silva Cruz has a Ph.D. degree in Electrical Computer and Systems Engineering from Rensselaer Polytechnic Institute (RPI), Troy, NY, US in 2000. He is with the Department of Electrical and Computer Engineering of the University of Coimbra in Portugal as an Assistant Professor. He is a researcher of the Instituto de Telecomunicações in Coimbra, where he works on video processing and coding.



Pedro Assuncao received the PhD in Electronic Systems Engineering from the University of Essex, UK, in 1998. He is currently professor of Electrical Engineering and Multimedia Communications at the Polytechnic Institute of Leiria and a senior researcher at Instituto de Telecomunicações, Portugal. His current research interests include video coding, video stream adaptation to diverse networking and user environments, error concealment and quality evaluation.

MMTC OFFICERS (Term 2014 — 2016)

CHAIR

Yonggang Wen
Nanyang Technological University
Singapore

STEERING COMMITTEE CHAIR

Luigi Atzori
University of Cagliari
Italy

VICE CHAIRS

Khaled El-Maleh (North America)
Qualcomm
USA

Liang Zhou (Asia)
Nanjing University of Posts & Telecommunications
China

Maria G. Martini (Europe)
Kingston University,
UK

Shiwen Mao (Letters & Member Communications)
Auburn University
USA

SECRETARY

Fen Hou
University of Macau, Macao
China

E-LETTER BOARD MEMBERS (Term 2014—2016)

Periklis Chatzimisios	Director	Alexander TEI of Thessaloniki	Greece
Guosen Yue	Co-Director	Broadcom	USA
Honggang Wang	Co-Director	UMass Dartmouth	USA
Tuncer Baykas	Editor	Medipol University	Turkey
Tasos Dagiuklas	Editor	Hellenic Open University	Greece
Chuan Heng Foh	Editor	University of Surrey	UK
Melike Erol-Kantarci	Editor	Clarkson University	USA
Adlen Ksentini	Editor	University of Rennes 1	France
Kejie Lu	Editor	University of Puerto Rico at Mayagüez	Puerto Rico
Muriel Medard	Editor	Massachusetts Institute of Technology	USA
Nathalie Mitton	Editor	Inria Lille-Nord Europe	France
Zhengang Pan	Editor	China Mobile	China
David Soldani	Editor	Huawei	Germany
Shaoen Wu	Editor	Ball State University	USA
Kan Zheng	Editor	Beijing University of Posts & Telecommunications	China