MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
http://www.comsoc.org/~mmc

MMTC Communications - Frontiers

Vol. 13, No. 3, May 2018

CONTENTS

Message from MMTC Chair

Dear MMTC colleagues:

Time flies! The current MMTC leadership team was elected at IEEE ICC 2016, May 23-27, 2016 in Kuala Lumpur, Malaysia. Our two-year term has ended by end of May 2018. In the Spring of 2018, a major task for us is to elect the new leadership team of MMTC for the next term, which is critical for the continued success of this vibrant technical committee.

According to the MMTC bylaw, a Selection Committee was formed with past MMTC chairs as members, including Dr. Yonggang Wen as Chair, and Drs. Jianwei Huang, Haohong Wang, and Luigi Atzori as members. A call for nominations/volunteers was sent to all MMTC members on Feb. 20, 2018, with many nomininations/self-nominations received till mid April. The Selection Committee carefully reviewed the nominations, had many online discussions, communicated with the candadidates, and eventually voted to select qualified candidates. The voting took place at the MMTC meeting at IEEE ICC 2018, Kansas City, MO, USA, May 23, 2018. About 40+ members attended in person and voted to select the next MMTC leadershiop team as follows.

- MMTC Chair: Honggang Wang, University of Massachusetts, USA
- MMTC Steering Committee Chair: Sanjeev Mehrotra, Microsoft Research, USA
- MMTC Vice Chair (America): Pradeep K Atrey, University of Albany, USA
- MMTC Vice Chair (Asia): Wanqing Li, University of Wollongong, Australia
- MMTC Vice Chair (Europe): Lingfen Sun, Plymouth University, UK
- MMTC Vice Chair (Letters and Member Communications): Jun Wu, Tongji University, China
- MMTC Secretary: Shaoen Wu, Ball State University, USA
- MMTC Standard Liaison: Guosen Yue, Huawei, USA

Congratulations! I am glad that the MMTC will be in good hands. And I am sure MMTC will continue to be a high active TC and to be successful in the next term!

The outgoing MMTC leadership team was elected at ICC'16. It was a great honor and highly rewarding experience for me to work with the team in the past two years. I would like to thank all these friends and colleagues, our newsletter editor Dr. Mugen Peng, our webmaster Dr. Haixia Zhang, all the board directors and members, and all the IG chairs and members for your kind support and collaboration! I would also like to thank our past chairs, Drs. Yonggang Wen, Jianwei Huang, Haohong Wang, Qian Zhang, Nelson Fonseca, Heather Yu, Stan Moyer, Tak Kamae, Alex Gelman, Jeff Derby, Sid Ahuja, Charlie Judice, Fritz Froelich, Birendra Prasada, who laid out the foundation of MMTC and made it grow to such a vibrant TC we have today.

Although stepped down, I will certainly continue to volunteer for MMTC activities. If you are interested in getting involved in MMTC, please do not hesitate to contact the new chairs. Wish all the best for the 2018-2020 team!

Sincerely,

Shiwen Mao
Outgoing Chair, Multimedia Communications TC of IEEE ComSoc

SPECIAL ISSUE ON QUALITY OF EXPERIENCE FOR MULTIMEDIA COMMUNICATIONS: APPLICATIONS AND CHALLENGES

Guest Editor: Hantao Liu, Cardiff University, United Kingdom

hantao.liu@cs.cardiff.ac.uk

Nowadays, multimedia systems have become an integral part of human activity, including entertainment, education, security and medicine. In many real-world applications, humans rely upon multimedia content to communicate information or to accomplish a task. To improve human experience and task performance, it is critical to understand how humans perceive the quality of multimedia content and use what is learned to develop algorithms and tools for improved multimedia quality. With the growing amount of media information and the emerging multimedia technologies, many challenges remain in developing methodologies that can reliably assess quality of multimedia content.

This special issue on “Quality of Experience for Multimedia Communications: Applications and Challenges” brings a collection of perspectives of leading researchers in the field and provides cutting-edge results from their research groups.

In “Exploring the challenges of Higher Frame Rates: from Quality Assessment to Frame Rate Selection”, researchers Angeliki V. Katsenou, Alex Mackin, Di Ma, Fan Zhang, and David R. Bull from the University of Bristol present their study on high frame rate (HFR) video content. This paper has summarised their recent work on quality assessment and frame rate selection.

The paper entitled “Objective Image Retargeting Quality Assessment”, from Nanyang Technological University, discusses the objective perceptual measures for image retargeting quality assessment (IRQA). The researchers Yabin Zhang, Qiaohong Li, Qiuping Jiang, and Weisi Lin present recent advances in modelling of image retargeting and the corresponding IRQA measures.

In “State of the Art on Medical Image and Video Quality Assessment”, the researchers Lucie Lévêque and Hantao Liu from Cardiff University provide an overview on medical image and video quality assessment. The paper has summarised methodologies for the assessment of the quality of medical content and investigated the impact of specialty settings on the perceived quality.

Hantao Liu received the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands, in 2011. He is currently an Assistant Professor with the School of Computer Science and Informatics, Cardiff University, Cardiff, U.K. He is currently serving for IEEE MMTC as the Chair of the Interest Group on Quality of Experience for Multimedia Communications, and he is an Associate Editor of the IEEE Transactions on Human-Machine Systems and the IEEE Transactions on Multimedia. His research interests include visual media quality assessment, visual attention modelling and applications, visual scene understanding, and medical image perception.

Exploring the challenges of Higher Frame Rates: from Quality Assessment to Frame Rate Selection

Angeliki V. Katsenou, Alex Mackin, Di Ma, Fan Zhang, and David R. Bull
Visual Information Lab (VI-Lab)
Dept. of Electrical & Electronic Engineering, University of Bristol
{angeliki.katsenou, a.mackin, di.ma, fan.zhang, dave.bull}@bristol.ac.uk

1. Introduction

Visual experiences are key drivers, not just for the entertainment sector but also for business, security and communication technologies [1]. Immersion is the ultimate goal that signals the success of the visual formats and towards this are all latest attempts in visual technologies aiming at [1], [2], [3]. Various efforts to enhance experiences have led to technologies such as 3D, 360°, virtual reality (VR)/ augmented reality (AR), high dynamic range (HDR) and more. The success of 3D content in consumer applications was very limited. However, with lessons learned, the rest of the aforementioned technologies have started to gain interest. New efforts in this respect have focused on extending the video parameter space with greater dynamic range, wider colour gamut, higher spatial resolution and increased frame rates [1], [2], [3].

The frame rates used for television and cinema have remained fairly constant since the 1930's and have rarely exceeded 60 frames per second (fps). Frame rates this low can cause problems for motion portrayal and this has become evident with the state-of-the-art display resolutions, HD and UHD [3], [4]. High frame rate (HFR) video content can supply more realistic portrayal of a scene by reducing motion blur, minimizing temporal aliasing and visual artifacts. Using HFR video content, directors also have greater flexibility over the 'look' of the content (using the shutter angle to control the levels of motion blur in the scene) and it makes it easier to convert to existing formats [1], [2], [3]. Moving beyond 60 fps (high frame rates) has been one of scenario that the industry and researchers have been discussing for video broadcast, cinema, gaming, virtual reality, and more [2], [3]. Recently, frame rates up to 120 fps have been recommended by the ITU in the latest UHDTV video standard (Rec. 2020) [5].

2. Coping with the arising HFR challenges

Shifting to HFR is not straightforward as it raises many challenging questions [4], [6], [7]. A first issue is that the perceptual aspects of the new video parameters needs to be investigated and studied, especially considering that it is content dependent. Another important issue for the video technology stakeholders is that HFR leads to an unavoidable increase of data rates [8], which in combination with the enhanced resolutions, dramatically increases the flow of information through the complete video pipeline (from acquisition to delivery).

By investigating the role that frame rates play in both the acquisition of video content, and with respect to human perception, we can attempt to make appropriate frame rate predictions across a range of different video content types and genres, addressing both perceptual quality and compression requirements. Focusing on the above problems, Ou *et al*. [9], [10], and Emoto *et al*. [11] studied the relationship between frame rate and perceptual quality, the former considering frame rates up to 30 fps, and that the latter up to 240 fps. Sugawara *et al*. [12] explored the relations between the motion blur and perceptual quality but did not present how frame rate influences the perceived quality. In a recent work [13], Huang et al. also researched how perceptual quality is affected by frame rate, however, the maximum frame rate employed was only 60 fps. A major issue that influences these studies is the lack of representative high frame rate content - very few HFR databases are currently publicly available [14].

Addressing the HFR challenges requires extensive research in a number of key areas, namely the perceptual properties of HFR, the design of suitable objective quality metrics, the evaluation of current encoding technologies on HFR, and more. In VI-Lab, we have identified these challenges and have designed projects around these topics. In this paper though, only a few examples of recent work from VI-Lab on HFR video are briefly summarized. The work that has been selected to be presented in this paper includes the description of a unique high frame rate (up to 120 fps) video quality database (BVI-HFR) [15], which can be used as a basis for research in the HFR area (see

Section 3). Next, an objective quality metric tailored for HFR video sequences and with better correlation with perceived quality is reviewed and compared to state-of-the-art objective quality metrics [16] (see Section 4). Last, motivated by the fact that not all video content benefits from HFR, we outline our frame rate selection method [17] that is solely based on spatio-temporal features extracted from the original sequences is discussed (see Section 5).

3. The BVI-HFR video quality database

VI-Lab captured and created a new data set, the Bristol Vision Institute HFR (BVI-HFR) [15]. BVI-HFR contains 22 source sequences at 1920×1080p (HD) resolution and was down-sampled by averaging frames to 60, 30 and 15 fps, thus allowing comparisons across frame rates to be made. An example of a frame at different frame rates from the cyclist sequence is depicted in Figure 1, in which the amount of perceptible motion blur decreases with frame rate increase.

BVI-HFR was validated in terms of its range of low level characteristics by employing three low-level descriptors: spatial information (an estimate of spatial complexity), temporal information (the motion energy between adjacent frames), and colorfulness (the variety and intensity of colors), were used to characterize the content distribution of the BVI-HFR video database [14]. Based on these metrics, it was evident that BVI-HFR spans a variety of scenes, motions and colors, and hence that it can provide a basis for fair and sufficient scrutiny across a range of applications.

(a) 120 fps (b) 60 fps (c) 30 fps

Figure 1: Example frames of the cyclist sequence from the BVI-HFR at different frame rates.

The first study on this dataset was the evaluation of the relationship between frame rate and perceptual quality taking into account the content dependence of frame rates. Therefore, a subjective study with 51 participants was performed at the specially designed studios in VI-Lab. The studios comply with ITU recommendations ITU-R BT.500-13 and a single stimulus methodology with a continuous quality scale evaluation method was followed. After the mean opinion scores were collected and outliers rejected, a one-way repeated measures ANOVA on the participants opinion scores comparing 60 fps and 120 fps for every sequence in the BVI-HFR video database was employed. The results revealed the anticipated content dependence of the frame rate to the perceived quality and suggested that higher frame rates are beneficial when fast global motions i.e. camera motions, are present.

The BVI-HFR video database, alongside the subjective scores, is publicly available in order to enable other researchers to replicate the study, investigate the role that frame rates play from capture to delivery in their own defined use cases as for example in assessing objective video quality metrics or video compression in HFR content. So far, as no other similar data set is currently publicly available, it has played a significant role in all HFR related research in University of Bristol and in the whole research community. As it is a relatively new area of research it is expected to be widely used in the near future.

4. Objective Quality Metrics for HFR videos

It is well known that subjective tests do provide reliable and effective quality assessment, but they are also expensive, time consuming and are not suitable for real-time applications. Therefore, objective quality metrics, which attempt to correlate the statistics of video content with subjective opinions, and generally used for video quality assessment are developed and offer an alternative way to realise automatic quality assessment. For video content with various frame rates, existing generic full-reference image and video quality metric can provide objective video quality assessment by comparing the high frame version with its temporally up-sampled lower frame rate version. Moreover, several bespoke quality metrics, such as TCFQ, and MNQT have also been developed for assessing the perceptual quality difference between videos with various frame rates. However, none of these models work well on HFR databases (more than 60 fps), e.g. BVI-HFR.

In this context, we have proposed a Frame Rate dependent Quality Metric (FRQM) in [16], which predicts visual quality difference between video content with different frame rates through temporal wavelet decomposition, subband combination and spatiotemporal pooling, as shown in Figure 2. This bespoke quality metric is with relatively low complexity but offers superior correlation performance with subjective results when it is compared with popular generic quality metrics and other frame rate dependent models. The performance of FRQM alongside six generic and bespoke quality metrics, Peak Signal-to Noise Ratio(PSNR) [18], Structural Similarity Index (SSIM) [19], Video Quality Metric (VQM) [20], Spatio-temporal Most Apparent Distortion Model (ST-MAD) [21], TCFQ [22] and MNQT [9] on the BVI-HFR video quality database are summarized in Figure 3. Here three commonly used correlation statistics, Pearson Linear Correlation Coefficient (LCC), Spearman Rank Correlation Coefficient (SROCC), and Root Mean Squared Error (RMSE) are employed.

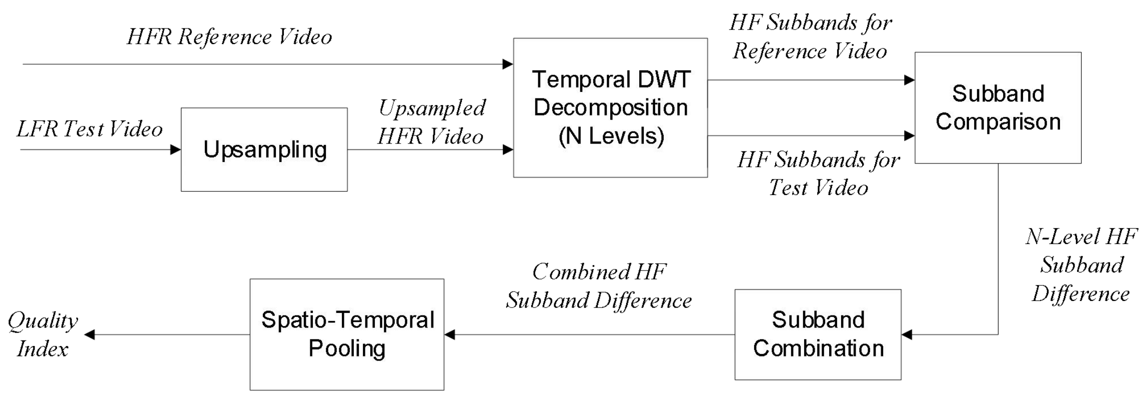


Figure 2: The architecture of the FRQM model. Here DWT stands for Discrete Wavelet Transform. More algorithmic detail of FRQM can be found in **[16]**.

FRQM has been further employed as a key component in our resolution adaptation based video coding framework, ViSTRA [23], [24], which won first prize in the IEEE ICIP 2017 Video Compression Grand Challenge [25].

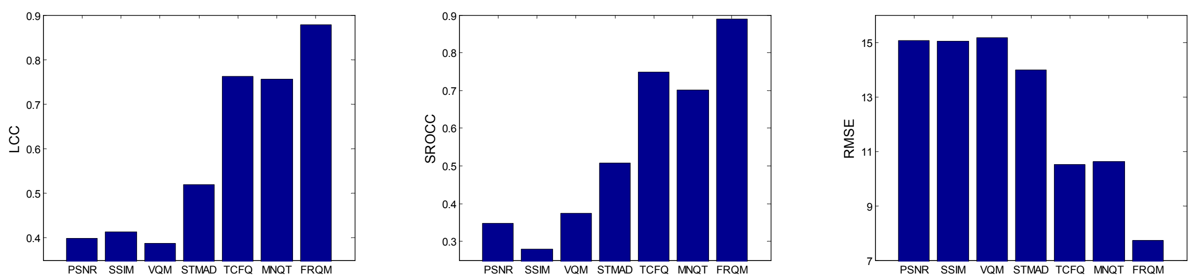


Figure 3: The performance of the tested quality metrics on the BVI-HFR database.

5. Optimal frame rate selection using spatio-temporal video features

As the impact of acquiring video sequences at a high frame rate on the perceived quality is content dependent, we have tried to identify the cases for which using the highest available acquisition rate would provide significantly better video quality. Therefore, in [17], we proposed a method that recommends the lowest frame rate at which the perceived quality is not significantly different than the original video. The method comprises an offline and an online phase, as illustrated in Figure 4. During the offline phase, low-level features and their high order statistics are extracted from the original video sequences (namely at the acquisition frame rate). Following that, a feature selection process based on Random Forests, is employed. After selecting the best performing subset of features with the lowest cardinality, the selected features and ground truth (extracted from the ANOVA analysis in [15]) are fed into a supervised machine learning method to build a frame rate selection model. The recommended frame rate is the minimum frame rate for which the perceived video quality is equivalent to that of the acquisition frame rate.

Compared to the recent literature [9], the proposed method has the advantage of employing only a small number of spatio-temporal features extracted from the original video sequence (only at the acquisition frame rate) resulting in reduced complexity. Furthermore, it has been tested with BVI-HFR which as mentioned above includes sequences captured at 120 fps and it achieves very high accuracy (95.1%).

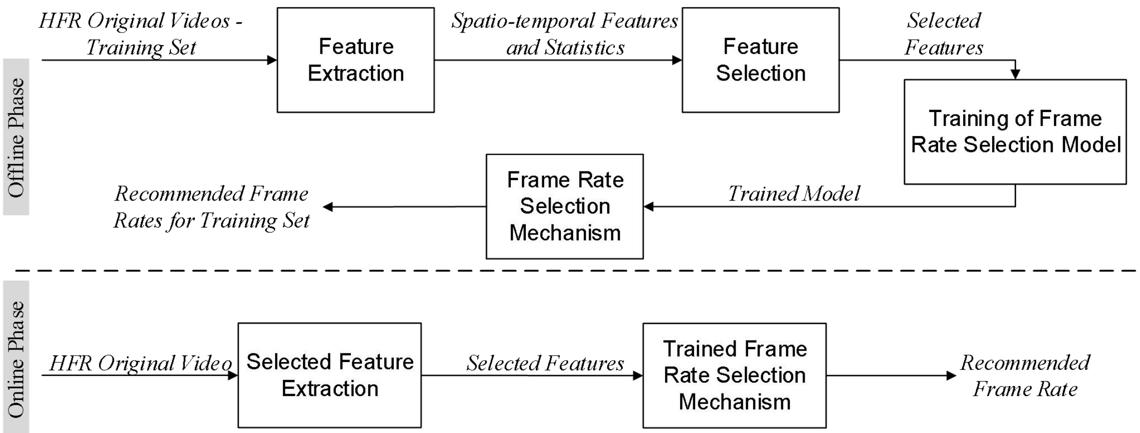


Figure 4: A diagrammatic illustration of the frame rate selection method.

We would like to note that the proposed method has been trained and tested only on a small data set, BVI-HFR, as no other high frame data sets are currently publicly available. Further to this, the proposed method has been tested for sequences shot with a full shutter angle. This means that all machine-learning-based models for feature selection and for frame rate selection have been trained on data with these specific acquisition parameters. However, it is important to point out that the method can be generalized for different acquisition parameters if the respective data become available.

The proposed frame rate selection method can be used in future video broadcast systems that will allow the adaptation/switching of the frame rate based on the content characteristics. Also, the method will be extended to decide on the lowest spatial resolution for which the upscaling of the sequence would not degrade the perceived quality. This will provide a complete framework for the adaptation of both spatial and temporal parameters without compromising the perceptual video quality.

5. Summary and Future Work

This paper has summarised recent work from VI-Lab on high frame rate video research and, in particular, on quality assessment and frame rate selection. The reviewed works have been using the BVI-HFR database as it is the only publicly available HFR dataset up to 120 fps that also includes subjective quality data (mean opinion scores). This dataset was used to assess existing objective quality metrics on HFR content and to devise a more appropriate quality metric that has the advantage of being independent of frame rate. Furthermore, a frame rate selection approach was presented which can recommend the frame rate to achieve video quality perceptually equivalent to the original video sequence.

Future work in this area in VI-Lab will focus on the investigation of other immersive video parameters, e.g. spatial resolution and dynamic range, and their influence on visual quality. All these parameters have a common important factor, which is the content dependence. Therefore, all these parameters should be ideally studied in a more systematic manner. This could lead in a unified framework from acquisition to compression, next to communication and finally to display in order to provide an optimally engineered perceptual experience.

Acknowledgement

The authors acknowledge funding and support from UK EPSRC platform grant EP/M000885/1, BBC Research and Development, and UK EPSRC EP/L016656/1.

References

[1] D. R. Bull, Communicating pictures: a course in image and Video Coding, Academic Press, 2014.

[2] D. R. Bull, E. J. Delp, S. Takamura, T. Wiegand and F. Wu, "Introduction to the issue on emerging technologies for video compression," *IEEE Journal of Selected Topics in Signal Processing,* vol. 5, pp. 1366-1377, 2011.

[3] K. Noland, "The application of sampling theory to television frame rate requirements," BBC Research & Development White Paper 282, 2014.

[4] M. Armstrong, D. Flynn, M. Hammond, S. Jolly and R. Salmon, "High frame-rate television," *BBC Research White Paper 169,* 2008.

[5] Recommendation ITU-R BT.2020, "Parameter values for ultra-high definition television systems for production and international programme exchange," 2012.

[6] R. A. Salmon, M. G. Armstrong and S. J. E. Jolly, "Higher Frame Rates for More Immersive Video and Television," BBC Research and Development White Paper 209, 2011.

[7] A. Watson, "High Frame Rates and Human Vision: A View through the Window of Visibility," *SMPTE Motion Imaging,* vol. 122, no. 2, pp. 18-32, 2013.

[8] A. Mackin, F. Zhang, M. A. Papadopoulos and D. R. Bull, "Investigating the Impact of High Frame Rates on Video Compression," in *IEEE International Conference on Image Processing (ICIP)*, Beijing, 2017.

[9] Y.-F. Ou, T. Liu, Z. Zhao, Z. Ma and Y. Wang, "Modeling the Impact of Frame Rate on Perceptual Quality of Video," in *Proc. IEEE Int Conf. on Image Processing*, San, 2008.

[10] Y.-F. Ou, Y. Xue and Y. Wang, "Q-star: a perceptual video quality model considering impact of spatial, temporal, and amplitude resolutions," *IEEE Trans. on Image Processing,* vol. 23, pp. 2473-2486, 2014.

[11] M. Emoto, Y. Kusakabe and M. Sugawara, "High-frame-rate motion picture quality and its independence of viewing distance," *Journal of Display Technology,* vol. 10, pp. 635-641, 2014.

[12] M. Sugawara, K. Omura and M. Emoto, "Temporal Sampling parameters and Motion Protrayal of television," in *SID Symposium Digest of Technical Papers*, 2009.

[13] Q. Huang, S. Y. Jeong, S. Yang, D. Zhang, S. Hu, H. Y. Kim, J. S. Choi and C.-C. J. Kuo, "Perceptual Quality Driven Frame-Rate Selection (PQD-FRS) for High-Frame-Rate Video," *IEEE Trans. on Broadcasting,* vol. 62, pp. 640-653, 2016.

[14] S. Winkler, "Analysis of Public Image and Video Database for Quality Assessment," *IEEE Journal of Selected Topics in Signal Processing,* vol. 6, pp. 1-10, 2012.

[15] A. Mackin, F. Zhang and D. R. Bull, "A study of subjective video quality at various frame rates," in *Proc. IEEE Int Conf. on Image Processing*, 2015.

[16] F. Zhang, A. Mackin and D. R. Bull, "A New Frame Rate Dependant Quality Model Based on Temporal Wavelet Decomposition and Spatiotemporal Pooling," in *Proc. IEEE Int Conf. on Image Processing*, 2017.

[17] A. V. Katsenou, D. Ma and D. R. Bull, "Perceptually-aligned Frame Rate Selection using Spatio-temporal Features," in *Picture Coding Symposium*, 2018.

[18] S. Winkler and P. Mohandas, "The evolution of video quality measurement: From PSNR to hybrid metrics," *IEEE Transactions on Broadcasting,* vol. 54, no. 3, pp. 660-668, 2008.

[19] Z. Wang, A. Bovik, H. Sheikh and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Processing,* vol. 13, pp. 600-612, 2004.

[20] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans. on Broadcasting,* vol. 50, no. 3, pp. 312-322, 2004.

[21] P. V. Vu, C. T. Vu and D. M. Chandler, "A spatiotemporal most-apparent-distortion model for video quality assessment," in *Proc. IEEE Int Conf. on Image Processing*, 2011.

[22] Z. Ma, M. Xu, Y.-F. Ou and Y. Wang, "Modeling of rate and perceptual quality of compressed video as functions of frame rate and quantization stepsize and its applications," *IEEE Trans. on Circuits and Systems for Video Technology,* vol. 22, pp. 671-682, 2012.

[23] D. R. Bull, F. F. Zhang and M. Afonso, "Video processing method". Patent GB1714791.9.

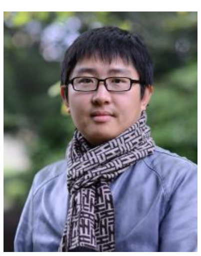
[24] D. R. Bull, F. Zhang and M. Afonso, "Description of SDR video coding technology proposal by University of Bristol," JVET-J0031, 2018.

[25] "Grand Challenge at ICIP 2017: Video Compression Technology," 2017. [Online]. Available: http://www.provision-itn.eu/grand-challenge-videocompression-icip2017.htm.

Angeliki V. Katsenou is a Leverhulme Early Career Fellow and is with the Visual Information Lab at the University of Bristol since 2015. She obtained her Ph.D. degree from the Dept. of Computer Science and Engineering, University of Ioannina, Greece. She received her Diploma in Electrical and Computer Engineering and the M.Sc. degree in Signal and Image Processing from the University of Patras, Greece. She has experience in several EC-funded and EPSRC projects, such as MSCA PROVISION ITN and EPSRC Platform Grant EP/M000885/1. Her research interests include perceptual video analysis, video compression, and transmission.

Alex Mackin obtained his MEng in Engineering Mathematics (2013) and PhD in Electrical and Electronic Engineering (2017) from the University of Bristol. He is currently a research associate in the Visual Information Laboratory Department of Electrical and Electronic Engineering, University of Bristol. His research interests include high frame rates, human visual system modelling, adaptive acquisition and immersive video formats.

Di Ma received the BEng degree in Electronic Information Engineering from the Civil Aviation University of China, Tianjin, P. R. China, in 2013, and the M.Eng. degree in Information and Communications Engineering from the Tianjin Key Lab for Advanced Signal Processing, Tianjin, P. R. China. He is currently working toward the PhD degree in the EPSRC Centre for Doctoral Training in Communications, University of Bristol. His research interests include image and video processing, video quality assessment, adaptive acquisition and immersive video formats.



Fan Zhang received the B.Sc. (Hons) and M.Sc. degrees from Shanghai Jiao Tong University (2005 and 2008 respectively), and his Ph.D. from the University of Bristol (2012). He is currently working as a Senior Research Associate in the Visual Information Laboratory, Department of Electrical and Electronic Engineering, University of Bristol, on projects related to perceptual video compression and immersive video processing. His research interests include perceptual video compression, video quality assessment and immersive video formats.

David R. Bull received the B.Sc. degree from the University of Exeter, Exeter, U.K., in 1980; the M.Sc. degree from University of Manchester, Manchester, U.K., in 1983; and the Ph.D. degree from the University of Cardiff, Cardiff, U.K., in 1988. Dr Bull has previously been a Systems Engineer with Rolls Royce, Bristol, U.K. and a Lecturer at the University of Wales, Cardiff, U.K. He joined the University of Bristol in 1993 and is currently its Chair of Signal Processing and Director of its Bristol Vision Institute. In 2001, he co-founded a university spin-off company, ProVision Communication Technologies Ltd., specializing in wireless video technology. He has authored over 450 papers on the topics of image and video communications and analysis for wireless, Internet and broadcast applications, together with numerous patents, several of which have been exploited commercially. He has received two IEE Premium awards for his work. He is the author of three books and has delivered numerous invited/keynote lectures and tutorials. Dr. Bull is a fellow of the Institution of Engineering and Technology.

Objective Image Retargeting Quality Assessment

Yabin Zhang, Qiaohong Li, Qiuping Jiang, and Weisi Lin
School of Computer Science and Engineering, Nanyang Technological University
zhan0398@e.ntu.edu.sg; qli013@e.ntu.edu.sg; n1609861g@e.ntu.edu.sg; wslin@ntu.edu.sg

1. Introduction

Over the past decade the increasing diversity of display devices such as mobile devices has imposed an urgent demand for image adaptation in terms of the aspect ratio. Meanwhile, graphic designers call for the automatic image adaption techniques to efficiently edit images into different sizes. The traditional image retargeting methods such as cropping (CR) and uniform scaling (SCL) are usually criticized for severe salient content loss and squeezed/stretched visual distortion, respectively. To alleviate such problems, many content-aware methods have been developed to attain perceptually better image retargeting results. The results of some influential retargeting methods [1] are illustrated in Fig. 1, including seam-carving (SC) [12], shift-map (SM) [13], non-homogeneous warping (WARP) [14], streaming video (SV) [15] and scale-and-stretch (SNS) [16]. Generally, existing retargeting methods can be categorized into discrete (e.g., SC, SM) and continuous (e.g., WARP, SV, and SNS) approaches, which dedicate to eliminating pixels judiciously or deforming image into the desired size based on the visual importance analysis of image content. To date, there is no particular image retargeting method that works consistently well for all the images. One important reason is the lack of effective quality metric as the guidance towards the satisfied subjective experience of retargeted images. To further advance image retargeting techniques, it is necessary to design effective objective perceptual measures for image retargeting quality assessment (IRQA).

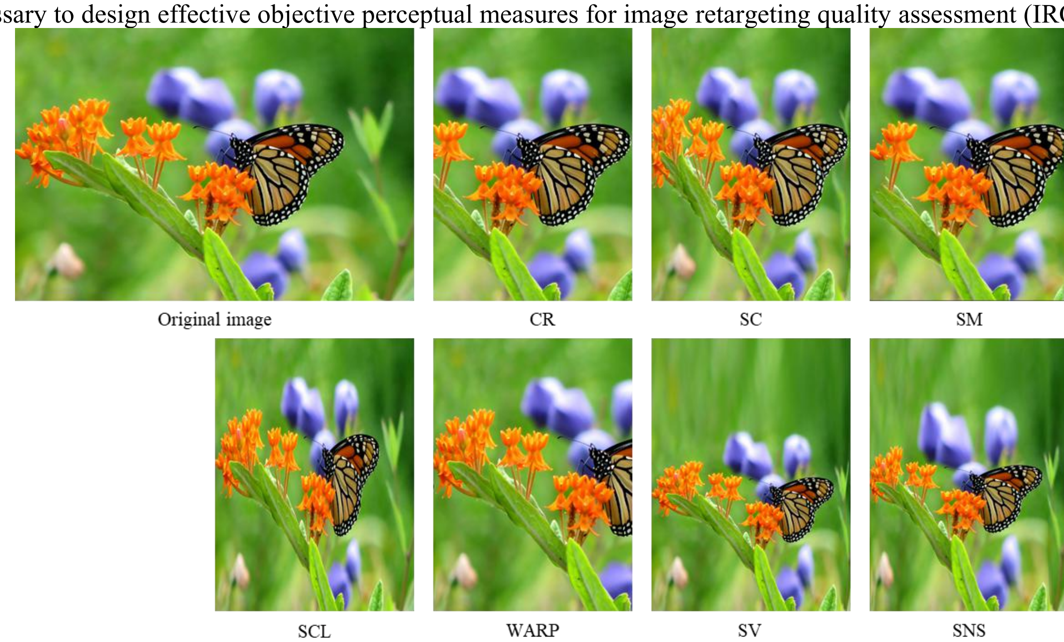


Fig. 1: Retargeted *butterfly* images with half width reduction using different image retargeting methods.

The salient information loss or visual distortion varies greatly across different image contents and retargeting methods, and it is a high-level semantic task to estimate the perceptual quality of retargeted images. The most reliable way for IRQA is the subjective test, but it is commonly cumbersome, non-automatic and expensive. Objective IRQA methods are more favorable to be an integrated part of the image retargeting process. As the performance of most retargeting methods is principally determined by the adopted objective quality, the lack of effective IRQA methods limits the further development of image retargeting. To facilitate the fair comparison of different objective IRQA methods, Rubinstein et al. first conducted a comparative study of existing influential retargeting methods and evaluated six simple objective quality metrics in accordance with the subjective ground-truth. The results suggest that the Bidirectional Similarity (BDS) [3], which has been utilized in existing retargeting methods, are not significantly better than those randomly sampled from the τ distribution. The study also shown that the SIFT flow [6], one recent image alignment algorithm, and the Earth mover's distance (EMD) [7], which measures the minimal cost to transform one distribution into another, show statistically better agreement with the ground-truth subjective rating. The possible reason may be attributed into their ability to capture the retargeting modification in the constrained matching between images. In the later dedicated IRQA works, Liu et al. [8] estimated global structure and local correspondence from coarse scales to fine scales in a top-down manner and designed a local and global correspondence based quality metric. Fang et al. [9] first utilized the SIFT flow to explore the dense correspondence between the images and employ the famous Structural Similarity (SSIM) metric to measure the quality of preserved structural information in retargeted images. Both bottom-up and top-down image saliency estimation are considered as the visual importance, and importance-weighted SSIM score is calculated as the overall quality index for the retargeted images. One of widely agreed principles for IRQA is that the quality degradation factors are mainly the information loss and visual distortion. The complicated relationship between the original and retargeted images makes it challenging to measure these two major quality factors effectively. Thus, to address the IRQA problem, it is important for us to seek an effective correspondence estimation between the original and retargeted images, which provide the guidance to measure the occurred information loss and visual distortion reliably. Here we introduce the recent advance about the matching algorithm with effective modelling of image retargeting and the corresponding IRQA measures.

2. Unified Interpretation for Image Retargeting

When we retarget one image, the major objective is to avoid the perceptually important image content loss and suppress the introduced annoying visual distortion or artifacts. Therefore, IRQA methods should be able to measure both the visual distortions and information loss in the retargeted image effectively. Because the image modification during retargeting process can be extremely complex, it is hard to directly measure these two quality-related aspects. One practical way is to first reveal the artificial retargeting modification without restricting to a specific retargeting method, and then infer the content changes with the estimated dense correspondence. In our recent work [2], to effective estimate inter-image dense correspondence, we have first provided an interpretation about different image retargeting methods in a unified framework consisting of resampling grid generation and forward resampling.

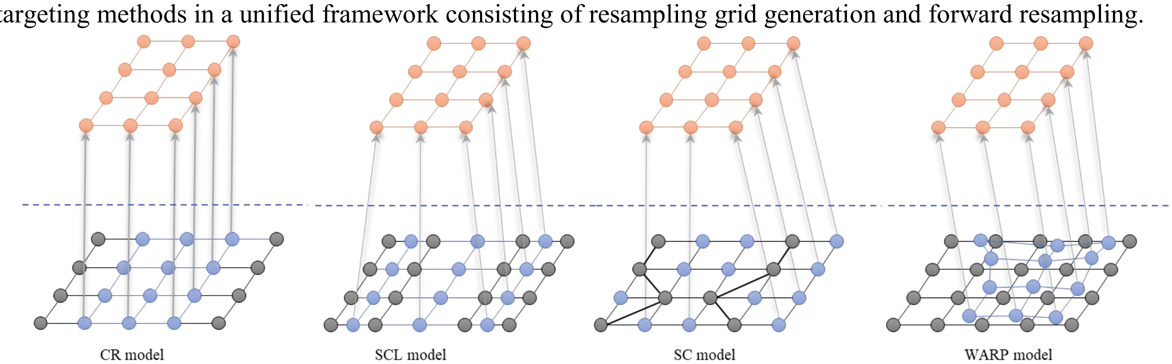


Fig. 2: Grid models to interpret image retargeting process

To introduce the modelling of image retargeting process, we take four typical image retargeting methods to explain the grid-based modelling for image retargeting as shown in Fig. 2. The gray grids correspond to the original image and the pink grids correspond to the image retargeted by CR, SCL, SC and WARP, respectively. The blue grids overlapped with gray grids are the resampling grids, and their patterns vary for different retargeting methods. The CR method is to crop out an optimal window from the original image, where the resampling grid is just part of the original grid, while the SCL method scales the image along one direction, where the resampling grid is uniformly distributed across the original grid. The content-aware seam-carving method [1] works by removing the seams one by one. The resampling grid is generated with the rest grid nodes after the seam removal at each iteration. The WARP method employs the mesh to manipulate the image by optimizing a warping function, where the resampling grid can be more flexible compared with the other three models. Next, the forward resampling based on the resampling grid will generate the retargeted image from the original image. During forward resampling, the direct pixel copy happens when the resampling grid node is overlapped with the original one, otherwise the resampling grid at the sub-pixel location need the interpolation such as bi-cubic interpolation. Therefore, the image retargeting can be interpreted as the resampling grid generation and forward sampling.

Obviously, we can find out that the resampling grid has a significant influence on the final retargeting results. For IRQA, the resampling grid will provide us the strong evidence about how the original image is retargeted into different resolutions in detail and facilitate the quality measurement of image content. In our work [2], the estimation problem of the resampling grid is modelled using Markov random field. We adopt the dual-layer loopy belief propagation algorithm for this problem and utilize a coarse-to-fine scheme for speed-up.

3. Quality Evaluation Methodology

The information loss and visual distortion are considered as the two major factors that degrade the quality of retargeted image. Since the resampling grid indeed provide the detailed information on how images are retargeted, e.g. content removal, squeeze and stretch, it is feasible to measure the local quality change regarding the information loss and visual distortion based on the resampling grid. Our recent works [2, 3] develop different kinds of effective features for IRQA based on this principle.

Here we first introduce the simple but effective aspect ratio similarity measure. The underlying principle is to estimate whether the local blocks are well-preserved, removed or distorted as shown in Fig. 3. With the regular partitioned $N \times N$ blocks in the original image, the estimated resampling grid reveals the corresponding deformed blocks in the retargeted image, so that the local deformation can be measured with the linked block pairs. According to the bounding box of the retargeted block, we have the width w_{ret} and height h_{ret}, so the two-dimensional change ratios can be calculated by $r_w = w_{ret}/N$ and $r_h = h_{ret}/N$. In addition, the mean ratio $\mu_r = (r_w + r_h)/2$ is introduced to complement the absolute block size change information. The similarity between the linked block pair is calculated by

$$S = \left[\frac{2 \cdot r_w \cdot r_h + C}{r_w^2 + r_h^2 + C}\right] \cdot \left[e^{\alpha - (\mu_r - 1)^2}\right],$$

where C is a constant to avoid the division by zero. The first aspect ratio term measures the aspect ratio similarity between the linked block pairs, and it always prefer the block with aspect ratio close to 1:1. As we notice that the aspect ratio cannot quantify the changes on block size, the absolute size term is utilized to take into account the influence of block size change. If the similarity score is close to 1, it means that the local content is preserved in high quality, and when the score decreases, the retargeted block may suffer from certain information loss or visual distortions. As addressed previously, IRQA task is highly subjective. Like most related works, we need the importance map to identify the visually important regions. The content loss or distortion in the important regions has a greater influence on the final quality of retargeted image. The quality prediction score for the retargeted image is obtained by pooling aspect ratio similarity score with visual importance weight as

$$ARS = \sum_i S_i \cdot V_i / \sum_i V_i ,$$

where S_i represents similarity score for block i, V_i denotes the visual importance weight. The final ARS index can thus predict the perceptual quality of retargeted image in a content-aware manner.

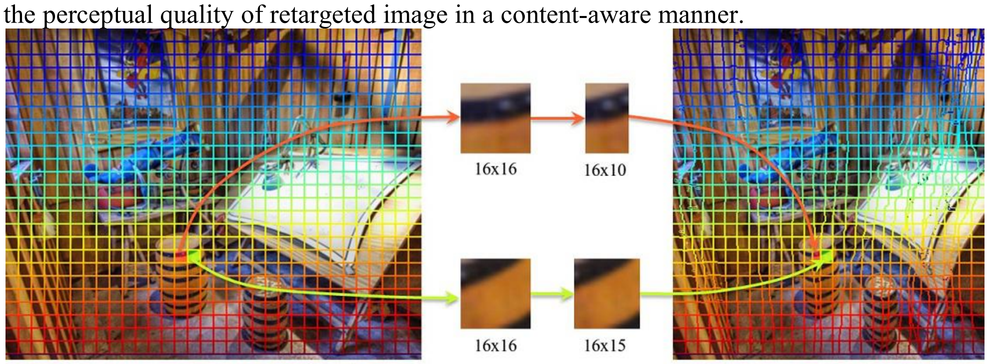


Fig. 3: Illustration of aspect ratio similarity for IRQA.

In the following, we notice that most existing works rely on the importance map for content-aware purpose. However, accurate importance is too challenging to obtain. Moreover, distorted regions in retargeted image always attract human attention, which cannot be captured by importance map based on content analysis. Therefore, we further develop the new features such as the edge group similarity and facial block similarity, to measure the structure/shaped and facial region related distortions, directly and explicitly. The IRQA model by fusing the different effective features can predict the visual quality more accurately and reliably.

4. Experimental Study

To demonstrate the quality prediction accuracy, we report the performance of different objective quality metrics on the widely-used CUHK image retargeting subjective quality database, according to the correlation with their subjective scores. The database contains 57 source samples and 171 retargeted samples, and the size change options are 25% or 50% reduction along either dimension. Four evaluation metrics commonly used in IQA, including Pearson Linear Correlation Coefficient (PLCC), Spearman Rank-Order Correlation Coefficient (SROCC), Root Mean Squared Error (RMSE) and Outlier Ratio (OR), are utilized to evaluate the correlation between the objective scores and the provided MOSs. In Table 1, we show the performance of Bi-Directional Similarity [4], Edge Histogram [10], SIFT flow [6], Earth Mover's Distance [7], CSim [8], GLS [5], PGDIL [11], ARS [2] and MLF [3]. We can observe that the developed methods have shown great improvement on the retargeting quality prediction accuracy compared with related methods.

Methods	PLCC	SROCC	RMSE	OR
Bi-Directional Similarity [4]	0.290	0.289	12.92	0.216
Edge Histogram [10]	0.342	0.329	12.69	0.205
SIFT flow [6]	0.314	0.290	12.82	0.146
Earth Mover's Distance [7]	0.276	0.290	12.98	0.170
CSim [8]	0.437	0.466	12.14	0.152
GLS [5]	0.462	0.476	10.93	0.135
PGDIL [11]	0.540	0.541	11.36	0.152
ARS [2]	0.684	0.669	9.86	0.070
MLF [3]	0.758	0.738	8.53	0.030

Table 1 Performance of different IRQA methods.

5. Conclusion

Retargeting modifications by different image retargeting operators are usually unknown and diverse, and the complicated image relationship impedes the effective quality evaluation. To understand the image retargeting process better, in our recent work we model the image retargeting explicitly with a unified grid based resampling process and provide a highly accurate backward registration method, which reconstructs out the dense correspondence relationship between the original and retargeted images. Accordingly, the new developed features directly based on the estimated retargeting modification have advanced the image retargeting quality prediction performance dramatically. However, the existing IRQA measures are still far from being perfect. As the real-world image contents vary greatly, it is still quite challenging to develop effective quality measures to help retargeting techniques generate perceptually satisfied retargeted images. Due to the recent large-scale database and deep learning techniques, there is a substantial progress on the studies on image aesthetic assessment, and it is also a promising research direction to transfer the learned knowledge from image aesthetic to the evaluation of image retargeting.

References

[1] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir. "A comparative study of image retargeting", ACM Trans. Graph., 29(6):160, 2010.
[2] Y. Zhang, Y. Fang, W. Lin, X. Zhang, and L. Li. "Backward registration-based aspect ratio similarity for image retargeting quality assessment", IEEE Trans. Image Processing, 25(9):4286–4297, 2016.
[3] Y. Zhang, W. Lin, Q. Li, W. Cheng, X. Zhang, "Multiple-Level Feature-Based Measure for Retargeted Image Quality", IEEE Trans. Image Processing 27(1): 451-463 (2018).
[4] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani. "Summarizing visual data using bidirectional similarity", In IEEE Computer Society Conference on CVPR, 24-26 June 2008, Anchorage, Alaska, USA.
[5] J. Zhang and C.C. Kuo. "An objective quality of experience (QoE) assessment index for retargeted images", In Proceedings of the ACM International Conference on Multimedia, Orlando, FL, USA, November 03 - 07, 2014, pages 257–266, 2014.
[6] C. Liu, J. Yuen, and A. Torralba. "SIFT flow: Dense correspondence across scenes and its applications", IEEE Trans. Pattern Anal. Mach. Intell., 33(5):978–994, 2011.
[7] O. Pele and M. Werman. "Fast and robust earth mover's distances", In IEEE 12th International Conference on Computer Vision, ICCV, Kyoto, Japan, September 27 – October 4, 2009, pages 460–467.
[8] Y.J. Liu, X. Luo, Y. Xuan, W. Chen, and X. Fu. "Image retargeting quality assessment", Comput. Graph. Forum, 30(2):583–592, 2011.
[9] Y. Fang, K. Zeng, Z. Wang, W. Lin, Z. Fang, and C. W. Lin. "Objective quality assessment for image retargeting based on structural similarity", IEEE J. Emerg. Sel. Topics Circuits Syst., 4(1):95–105, 2014.
[10] B. S. Manjunath, J. Ohm, V. Vasudevan, and A. Yamada. "Color and texture descriptors", IEEE Trans. Circuits Syst. Video

Techn., 11(6):703–715, 2001.
[11] C. C. Hsu, C. W. Lin, Y. Fang, and W. Lin. Objective quality assessment for image retargeting based on perceptual geometric distortion and information loss", J. Sel. Topics Signal Processing, 8(3):377–389, 2014.
[12] M. Rubinstein, A. Shamir, and S. Avidan. "Improved seam carving for video retargeting", ACM Trans. Graph., 27(3), 2008.
[13] Y. Pritch, E. Kav-Venaki, and S. Peleg. "Shift-map image editing", In IEEE 12th International Conference on Computer Vision, ICCV, Kyoto, Japan, September 27 -October 4, 2009, pages 151–158.
[14] L. Wolf, M. Guttmann, and D. Cohen-Or. "Non-homogeneous content-driven video-retargeting", In IEEE 11th International Conference on Computer Vision, ICCV, Rio de Janeiro, Brazil, October 14-20, 2007, pages 1–6.
[15] P. Krähenbühl, M. Lang, A. Hornung, and M. H. Gross. "A system for retargeting of streaming video", ACM Trans. Graph., 28(5), 2009.
[16] Y. S. Wang, C. L. Tai, O. Sorkine, and T. Y. Lee. "Optimized scale-and-stretch for image resizing", ACM Trans. Graph., 27(5):118, 2008.

Yabin Zhang received the B.E. degree in Electronic Information Engineering in the Honors School, Harbin Institute of Technology in 2013. He is currently pursuing the Ph.D. degree from the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests include video coding, image/video processing, image quality assessment, image retargeting quality assessment and computer vision. He is now a Project Officer with the School of Computer Science and Engineering, Nanyang Technological University, Singapore.

Qiaohong Li received the B.E. and M.E. degree in School of Information and Communication Engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2009 and 2012. She received the Ph.D. degree from the School of Computer Science and Engineering, Nanyang Technological University, Singapore, in 2017. She is now a Research Fellow with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. Her research interests include image quality assessment, speech quality assessment, computer vision, and visual perceptual modelling.

Qiuping Jiang is currently working towards the doctoral degree in Signal and Information Processing at Ningbo University, Ningbo, China. From January 2017 to May 2018, he was a visitng Ph.D. student with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. He has published over 20 academic papers in refereed journals and conferences in the areas of image/video processing, visual quality assessment, and computer vision. He is the receipt of JVCI2017 Best Paper Award Hornerable Mention and a reviewer for IEEE-TIP, IEEE-TCSVT, and IEEE-TMM.

Weisi Lin received the B.Sc. degree in electronics and the M.Sc. degree in digital signal processing from Sun Yat-Sen University, Guangzhou, China, in 1982 and 1985, respectively, and the Ph.D. degree in computer vision from King's College, London University, London, U.K., in 1993. He was involved in teaching and research with Sun Yat-Sen University, Shantou University, Shantou, China, Bath University, Bath, U.K., the National University of Singapore, the Institute of Microelectronics, Singapore, and the Institute for Infocomm Research, Singapore. He has been the project leader of 13 major successful projects in digital multimedia technology development. He was the Laboratory Head of Visual Processing and the Acting Department Manager of Media Processing, Institute for Infocomm Research. He is currently a Full Professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His current research interests include image processing, perceptual modeling, video compression, multimedia communication, and computer vision.

State of the Art on Medical Image and Video Quality Assessment

Lucie Lévêque and Hantao Liu
School of Computer Science and Informatics, Cardiff University
LevequeL@cardiff.ac.uk

1. Introduction

Medical imaging provides clinical information that is either unavailable by other means or captured with reduced invasiveness, playing a key role in diagnosis and treatment planning and monitoring. In radiology, for example, there are approximately a billion imaging examinations conducted worldwide every year [1]. The technologies used to acquire medical images include X-ray, ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), positron-emission tomography (PET), single-photon emission computed tomography (SPECT), etc. Both 2D and 3D content may be generated from these modalities. Furthermore, with the advances in telemedicine, and particularly in tele-surgery, medical images and videos are also more and more utilised in real-time communication frameworks. Due to the widespread use of imaging technology in medicine, a large amount of visual media is continuously being created and viewed by medical professionals in their routine practice. The methodologies used to acquire, process, transmit, store and display medical images and videos vary and, consequently, the ultimate visual information received by clinicians or other health professionals differs significantly in perceived quality. Visual signal distortions, such as various types of noise and artifacts arising in medical imaging, affect the perceptual quality of images and potentially impact the interpretation of medical images and the rendering of diagnoses [2]-[4].

To minimise diagnostic errors or to achieve satisfactory quality for image-guided surgery for instance, it is critical to understand how the medical professionals perceive the quality of medical visual media. Psychovisual experiments have been conducted to learn how visual distortions affect the perceptual quality of medical imaging, and with the view to investigate how visual quality can fulfil the medical professionals' expectations and help with their routine practice. In recent years, medical image perception research has been growing, on the one hand to evaluate the quality of diagnostic imaging, and on the other hand to consider imaging modalities beyond the traditional radiology, such as pathology and surgical video. In both areas, subjective quality assessment significantly contributes to the definition of default conditions and parameters, in terms of image acquisition, image/video encoding and content display, assuring adequate quality in various medical image applications.

In this paper, a review of the methodologies for the assessment of the perceived quality of medical images/videos is conducted, illustrated by representative example studies. The studies chosen as examples are relatively recent (i.e., not older than 10 years old) and represent diverse modalities and applications. We discuss the presented studies and make recommendations for future research in medical image quality assessment. We also present our recent study on the impact of specialty settings on the perceived quality of ultrasound video.

2. Review of the Methodologies for the Subjective Quality Assessment

Various methodologies prescribed by the International Telecommunication Union (ITU) have been used to assess the perceived quality of medical content. In this section, we present an overview of the methods and studies in the literature. The existing methods can be divided into two categories: single stimulus (SS) and multi stimulus (MS). The stimulus – a 2D image or a short video consisting of medical content – is presented to an observer, who then scores the perceived quality of the stimulus of interest using a discrete or continuous scale, typically containing five descriptors: *Bad*, *Poor*, *Fair*, *Good*, and *Excellent* quality. With the SS methods, the test stimuli are individually presented and scored. With the MS methods, both the references and test stimuli are presented but only the test stimuli are scored.

Kara et al. [5] used the Absolute Category Rating (ACR) scale [6] to assess the effects of angular resolution and light field reconstruction on 3D heart images. The ACR is a SS method where the stimulus is evaluated on a discrete quality scale. Kara et al. chose a 10-point scale for their tests and recruited twenty observers, including eight medical experts and twelve non-experts. Based on the regression analysis, the results identified a breakpoint of excellence at 75 views and showed that observers were more sensitive to degradations in texture than the number of views. Platisa et al. [7] investigated the effects of blurring, colour, gamma parameters, noise, and image compression on animal digital pathology images (dog gastric fundic glands and foal liver). They conducted an image assessment study with six veterinary pathologists, seven veterinary students, and eleven imaging experts using the Single Stimulus Hidden Reference Removal (SS-HRR) method [8] with a 6-point ACR scale. Based on the median opinion scores and

Kruskal-Wallis non-parametric one-way Analysis of Variance (ANOVA), they observed disagreement between the quality ratings made by different expertise groups, warning against the use of psycho-visual responses of non-expert subjects to guide the development of any pathology-specific image algorithms or imaging systems. Tulu et al. [9] studied the effects of delay, jitter, and packet loss ratio on ophthalmology videos in the context of telemedicine, using the Single Stimulus Continuous Scale (SSCQS) method [10]. The scoring scale was similar to the ACR scale, but a continuous scale was used in this case (i.e., a 100-point scale). Using ANOVA, they found that the perceived quality depended not only on technical parameters such as jitter and delay, but also on the transmission success of critical frames, i.e., the frames which are used to make a diagnosis, as they play a decisive role for the viewer.

Suad et al. [11] and Chaabouni et al. [12], respectively, used the Double-Stimulus Impairment Scale (DSIS) [6] and the Double-Stimulus Continuous Quality Scale (DSCQS) [10] for their experiments. These two DS methods employ a similar presentation protocol. With the DSIS method, the reference stimulus is presented first, followed by the corresponding test stimulus. For the DSCQS method, the reference and test stimuli are presented successively in random order and this pair presentation is repeated a second time. Suad et al. used the DSIS method to analyse the impact of different common types of distortion (i.e., additive Gaussian noise, blurring, JPEG compression, salt and pepper and sharpness) on brain MR images. A group of fifteen doctors participated in the study, where they were asked to evaluate the quality of the images. The results showed that the perceived quality is strongly affected by the distortions, with the highest quality for the sharpness and the poorest quality resulted from Gaussian noise. Chaabouni et al. made use of the DSCQS method to evaluate the impact of H.264 compression on laparoscopic surgery videos and analysed the resulting scores with some correlation metrics and a regression analysis. They found that compression artifacts could be noticeable for compression ratios of 100:1 up to 270:1. Another study on H.264 encoded laparoscopic videos was conducted by Münzer et al. [13]. A group of 37 medical experts participated in a test, using the DSCQS method to evaluate the impact of resolution and the constant rate factor (CRF) changes on overall image and semantic quality. The results suggested that an acceptable quality level may be achieved even reducing resolution down to 640×360 and with CRF = 26. With this setting, storage requirements would drop to 12.5% of current practice. Two studies on compressed videos were carried out by Razaak et al. [14] and by Usman et al. [15]. Both assessed the impact of the quantization parameter (QP) on perceptual quality of HEVC compressed videos, using the DSCQS method. The first study addressed ultrasound video, the second study investigated the video from wireless capsule endoscopy. A total of twenty observers participated in the first study (sixteen non-experts and four medical specialists), and 25 observers in the second study (nineteen non-experts and six medical experts). Experimental results, analysed with correlation metrics, recommended QP threshold values below 35 for acceptable diagnostic quality in the first study, and between 35 and 37 in the second study for satisfactory diagnostic and visual quality, respectively. Nouri et al. [16] also made use of the DSCQS method to study four videos representing different stages of a laparoscopic surgery (a type of minimally invasive surgery). Using a regression analysis, the authors found a quality threshold at bit rate = 3.2 megabits per second (or compression ratio 90:1 for MPEG2 compression), below which the surgeons considered the perceived image quality was too poor to perform the tasks. Chow et al. [17] carried out subjective experiments to assess the quality of MR images of the human brain, spine, knee and abdomen distorted with six types of distortion (Rician noise, Gaussian white noise, Gaussian blur, discrete cosine transform (DCT), JPEG compression and JPEG 2000 compression) at five different levels. They made use of the Simultaneous Double Stimulus for Continuous Evaluation (SDSCE) methodology [8], where the reference image and its distorted version are displayed side by side on a monitor. The observers rated the distorted image by judging the differences with respect to the original image. According to the results obtained by t-test, correlation and regression analysis, they claimed that Rician and Gaussian noise strongly affected the quality of MR images. The impact of a set of common distortions on the perceived quality of brain, liver, breast, foetus, hip, knee, and spine MR images was also studied by Liu et al. [18]. Ghosting, edge ghosting, white noise and coloured noise were simulated on MR scans of different content and acquisition protocols. The first experiment was divided in two parts: the first one included ghosting and white noise, while the second one included edge ghosting and coloured noise. Five different energy levels were defined for each type of artifacts. In each part, a total of 30 stimuli were shown to fifteen and seventeen expert participants, respectively, using the SDSCE method with a scale from 0 to 100. For the second experiment, a similar procedure was followed, and the study was conducted with eighteen expert subjects, using the two higher energy levels of all types of artifacts plus three new variations of the coloured noise (i.e., 112 stimuli). The results obtained from a one-way ANOVA indicated that artifacts with a flat spectral power density were nearly twice as annoying as artifacts with a spectral power density similar to the original image, at the same energy level. The study also concluded that differences in content affect the visibility of artifacts.

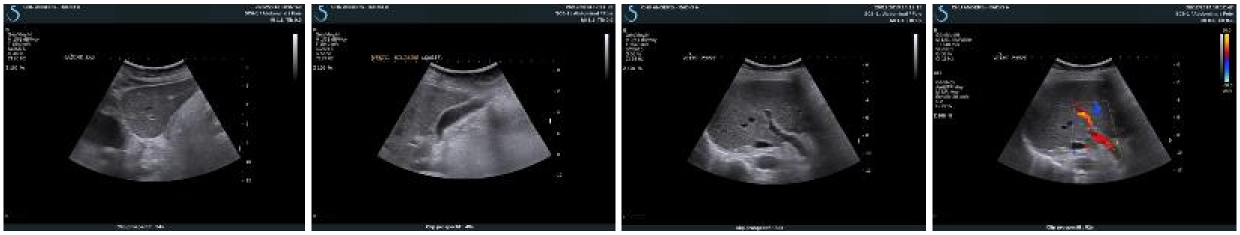
3. Discussion on the Existing Approaches and Suggestions

As noticed from the studies presented in the previous section, both single- and multi- stimulus methods can be used for the subjective assessment of perceptual quality of medical images and videos that are produced from different acquisition modalities, such as endoscopic, ultrasound, pathology, ophthalmology, etc. Each methodology has its own advantages. Experiments conducted using the single stimulus (SS) method are usually quicker to conduct than with the multi stimulus (MS) method, and they avoid potential vote inversions as only one stimulus is rated at a time [19]. However, SS-based experiments may lead to score drift over the course of a session [20]. New methods have therefore been proposed to combine the advantages of both SS and MS approaches, such as the Subjective Assessment Methodology for Video Quality (SAMVIQ) [21]. SAMVIQ offers the visualisation of a short video through a graphic interface, where the observers can navigate among the reference and the distorted versions of the reference content. Depending on the content type, i.e., images or videos, some methodologies may be more appropriate than others, regarding the length of the experiments for instance. A major problem in medical image and video quality assessment may be finding medical professionals to carry out the experiments.

Furthermore, some observers may generate dubious scores during experiments, due to the misunderstanding of the instructions or the lack of engagement in the task [22]. It is therefore recommended to use an outlier detection and subject exclusion procedure, as suggested in ITU-R recommendation BT.500-11 [8]. Four of the eleven studies mentioned above explicitly mentioned this 2-step pre-processing method: Chow et al. [17], Münzer et al. [13], Liu et al. [18], and Usman et al. [15]. No subject was rejected in the first three studies, while the last two studies each discarded one observer. Two studies used other methods: a graphical technique by Platisa et al. [7] and removal of extreme scores by Kara et al. [5]. Finally, the other five studies did not mention any outlier removal procedure. Subjective experiments in medical imaging can have different requirements compared to natural scene experiments. In a medical context, it is particularly important to test whether participants are consistent in their quality scoring, as their years of experience in medicine may affect their perception of visual quality [23]-[24]. Therefore, it may be necessary to divide the observers into groups depending on their experience and/or specialty. The criterion used to categorise the groups should be fixed prior to recruitment of participants. In addition, a common way to analyse the impact of the participants on scoring is to conduct an Analysis of Variance (ANOVA) on the scores. ANOVA is used to compare the means of two or more independent samples when assuming normality and homogeneity of the variance. Two studies used an ANOVA for statistical analysis: Tulu et al. [9] found no significant difference between the experts, whereas Liu et al. [18] found a significant difference and thus chose to normalise the raw scores using z-scores [25].

4. The Impact of Specialty Settings on the Perceived Quality of Medical Ultrasound Video

We conducted two studies with diverse medical content: surgery video [26] and ultrasound video [27]. In this section, we describe our second study. Unlike other related visual quality assessment studies in the literature which are either limited to a specific compression scheme or a small degree of stimulus variability, we aim to study a more comprehensive set of stimuli of a larger diversity in visual content and distortion. By this, we mean the dataset would include alternative popular compression schemes and various source stimuli and degradation levels. In the meantime, we seek to limit the total number of stimuli in order to make the subjective testing realistic so that the results are reliable. The source videos used in our experiment were extracted from four distinctive hepatic ultrasound scans by a senior radiologist from Angers University Hospital, France. It should be noted that the videos were purposely selected so that there was no apparent pathology. The participants would not be informed of the indications for the scans. The reason behind above choices is to encourage the participants to consider all plausible clinical uses of the stimuli rather than focusing on a specific pathology. All source videos last twelve seconds each and have a resolution of 1920×1080 pixels at a frame rate of 25 frames per second (fps). Fig. 1 illustrates one representative frame of each source video. The source videos were compressed using two popular compression schemes: H.264 [28] and HEVC [29]. H.264 is the most widely used video codec in current digital imaging systems. HEVC is the successor of H.264 and is meant to provide a better perceptual quality than H.264 at the same bit rate [30]. Both compression schemes could be potentially applied to the compression of clinical ultrasound video. To vary the perceptual video quality, for each source video, seven compressed sequences were created using the following bit rates: 512, 1000 and 1500 kbps (kilobits per second) for H.264 and 384, 512, 768 and 1000 kbps for HEVC. This resulted in a database of 32 video stimuli including the originals (i.e., 4 source videos + 4×7 compressed videos). It is well known that bit rate is not equal to quality for natural scenes, and that using the same bit rate to encode different natural contents could result in dramatically different visual quality. However, studies on how compression can affect the quality perception of medical content, and to what extent that perception is dependent on the specific user group are largely unexplored.



(a) (b) (c) (d)

Fig. 1: Illustration of one frame from each of the four source videos used in our experiment: (a) Content 1, (b) Content 2, (c) Content 3, and (d) Content 4 (Content 4 includes a Doppler ultrasound used to follow the blood flows).

Since standardised methodology for the subjective assessment of the quality of medical images and videos does not exist, we seek to make use of the experimental methodologies established for assessing natural images and videos. These methodologies are already described in Section 2, and in detail in [8] and [10]. To make our experiment feasible for radiologists, we conducted a user study where a few medical experts were surveyed for their preference in scoring quality of ultrasound videos. Based on the results of the survey, we decided to adopt a similar concept proposed by SAMVIQ (Subjective Assessment Methodology for Video Quality) [21], as described in Section 3.

Fig. 2 illustrates the final scoring interface developed in our study. In the experiment, the subjects are asked to assess the overall quality of each video by inserting a slider mark on a vertical scale. The grading scale is continuous (with the score range [0, 100]) and is divided into three semantic portions to help clinical experts in placing their opinions on the numerical scale. The associated terms categorising the different portions are: “Not annoying” (i.e., [75, 100]) corresponding to “the quality of the video enables you to conduct clinical practice without perceiving any visual artifacts”; “Annoying but acceptable” (i.e., [25, 75]) referring to “the visual artifacts are noticeable but the quality of the video suffices for the conduct of clinical practice”; and “Not acceptable” (i.e., [0, 25]) meaning “the visual artifacts are very noticeable and interfere with the clinical practice”. Fig. 2 also shows an example of the test organisation for each source scene, where an explicit reference (i.e., noted to the subjects), a hidden reference (i.e., a freestanding stimulus among other test stimuli) and seven compressed versions (placed in a different random order to each participant) are included. For each participant, the experiment is carried out scene after scene; and the order of scenes is randomised. Within a test (per scene), as shown in Fig. 2, subjects are allowed to view and grade any stimulus in any order; and each stimulus can be viewed and assessed as many times as the subject wishes (note the last score remains recorded). Note the entire methodology was developed in consultation with clinical experts to make sure the scoring experiment is more relevant and realistic to the reading environments in real clinical practice.

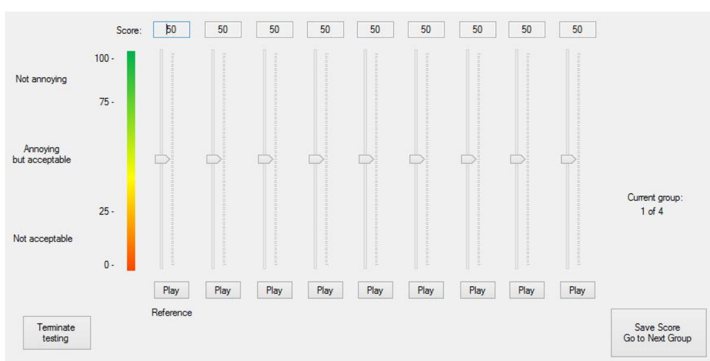


Fig. 2: Illustration of the rating interface used in our experiment.

The experiment was conducted in a typical radiology reading room environment. The venue represented a controlled viewing environment to ensure consistent experimental conditions: low surface reflectance and constant ambient light. The stimuli were displayed on a Dell UltraSharp 27-inch wide-screen liquid-crystal display with a native resolution of 2560×1440 pixels, which was calibrated to the Digital Imaging and Communications in Medicine (DICOM): Grayscale Standard Display Function (GSDF) [31]-[33]. The viewing distance was approximately 60 cm. No video adjustment was allowed. Before the start of the actual experiment, each participant was provided with instructions on the procedure of the experiment, and a training session was conducted. Since the goal of the study is to investigate visual quality perception across different specialities, the experiment was conducted with eight radiologists from Angers University Hospital, France, and nine sonographers from Castle Hill Hospital and Hull Royal Infirmary, United Kingdom.

The two sets of raw data, one collected from radiologists and one from sonographers, were individually processed in the same way. First, a simple outlier detection and subject exclusion procedure was applied to the raw scores within a subject group [8]. As a result of the outlier removal and subject exclusion procedure, none of the scores was detected as an outlier in both datasets and thus no radiologist or sonographer was excluded from further analysis.

Fig. 3 illustrates the mean opinion score (MOS), averaged over all subjects (within a subject group), for each compressed video in our experiment. It can be seen clearly from Fig. 3 that sonographers appear to be more annoyed by the low-quality videos than radiologists, as sonographers scored the highly compressed videos (i.e., H.264: 512 kbps and HEVC: 384 kbps) lower in quality than radiologists. However, the difference is less obvious for the higher quality videos. The observed tendencies are further statistically analysed. In the case of the low-quality videos, i.e., H.264: 512 kbps and HEVC: 384 kbps, a statistical significance test is performed with the quality as the dependent variable and the specialty, i.e., radiologist vs. sonographer, as the dependent variable. As the test for the assumption of normality is not satisfied, a nonparametric version (i.e., the Mann-Whitney *u*-test) analogue to an independent samples *t*-test is conducted. The test results (i.e., statistic=2591, p-value=0.004) indicate that there is a statistically significant difference between radiologists and sonographers in rating low-quality videos. Similarly, in the case of higher quality videos, i.e., H.264: 1000 and 1500 kbps and HEVC: 512, 768 and 1000 kbps, preceded by a test for the assumption of normality, a Mann-Whitney *u*-test is performed and the results (i.e., statistic=13420, p-value=0.207) reveals that there is no statistically significant difference between radiologists and sonographers in rating higher quality videos. Fig. 3 also shows that compression settings – both variables of compression scheme and compression ratio – affect the video quality. Also, the effect tends to depend on video content, for example, in both cases of radiologists and sonographers, the quality of "Content 1" is consistently scored higher than the quality of "Content 2", independent of the compression scheme or compression ratio. To further understand the impact of compression and content on video quality, we performed a statistical analysis, i.e., ANOVA (Analysis of Variance). In each case, the perceived quality is selected as the dependent variable, the video content and compression as fixed independent variables and the participant as random independent variable. The 2-way interactions of the fixed variables are included in the analysis. The results are summarised in Table I.

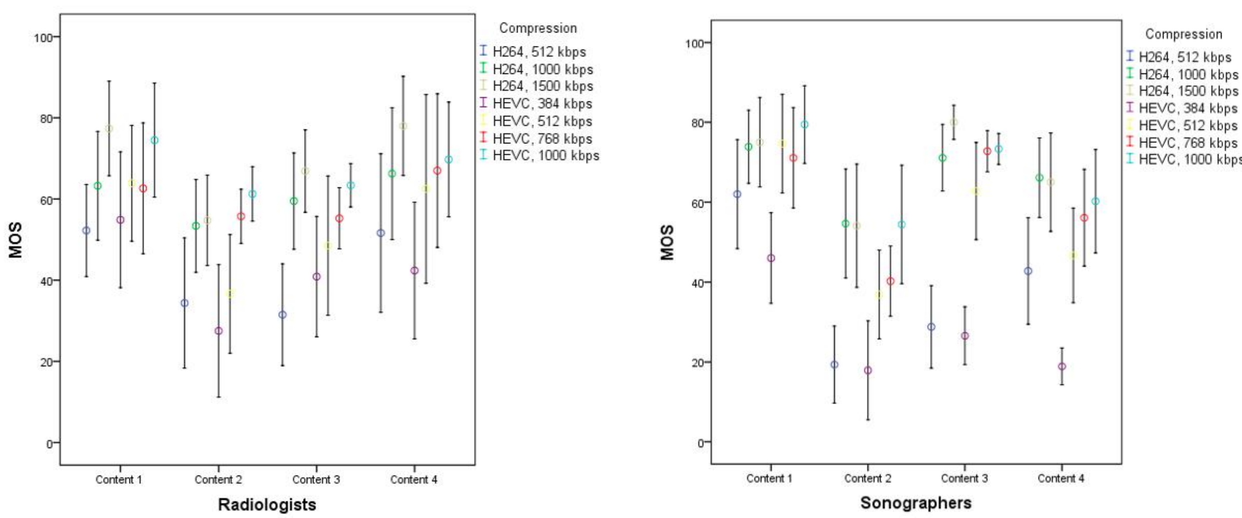


Fig. 3: Illustration of the mean opinion score (MOS) averaged over all subjects (within a subject group, i.e., radiologists or sonographers) for each compressed video. "Content" refers to a source video. Error bars indicate a 95% confidence interval.

Table I: Results of the ANOVA to evaluate the effect of "Participant", "Content" and "Compression" on the quality.

	Radiologists			**Sonographers**		
	df	F	p-value	df	F	p-value
Participant	7	1.365	0.266	8	1.264	0.306
Content	3	3.645	0.029	3	14.324	<0.001
Compression	6	21.520	<0.01	6	68.959	<0.001

Content*Compression	18	1.495	0.102	18	3.862	<0.001

6. Conclusions

We have presented diverse methodologies used in the literature for the subjective assessment of medical images and videos and detailed their assets and drawbacks depending on the application context. We suggest that future studies consider the following points of view: some methodologies may be faster and therefore better suited for busy medical specialists; the expertise level or speciality of the observers should be carefully selected; outlier detection methods should be applied to the subjective scores and properly documented. Furthermore, we have presented our study on the impact of specialty settings on the perceived quality via a dedicated subjective experiment. We designed and conducted a perception experiment where videos of different ultrasound exams distorted with various compression schemes and ratios were assessed by radiologists and sonographers. For both specialty groups, the impact of visual content and compression configuration on the perceived quality is found to be significant. Statistical analyses showed that the way the video quality changes with the content and compression configuration tends to be consistent for radiologists and sonographers. However, the results demonstrated that for the highly compressed stimuli, sonographers are more annoyed by the distortions than the radiologists; and that for the moderately compressed stimuli, radiologists and sonographers behave similarly in terms of scoring quality of visual experience.

References

[17] E. Krupinski, "Current perspectives in medical image perception", *Attention, Perception & Psychophysics*, vol. 72, 2010.

[18] P. Le Callet, S. Möller, and A. Perkis, "Qualinet white paper on definitions of quality of experience". *European Network on Quality of Experience in Multimedia Systems and Services,* 2012.

[19] E. Krupinski, and Y. Jiang, "Anniversary paper: evaluation of medical imaging systems", *Medical Physics*, vol. 35, pp. 645-659, 2008.

[20] E. Krupinski, "Improving Patient Care Through Medical Image Perception Research", *Policy Insights from the Behavioral and Brain Sciences*, vol. 2, pp. 74-80, 2015.

[21] P. Kara, P. Kovacs, S. Vagharshakyan, M. Martini, S. Imre, A. Barsi, and T. Balogh, "Perceptual Quality of Reconstructed Medical Images on Projection-Based Light Field Displays", *eHealth*, pp. 476-483, 2017.

[22] Recommendation ITU-T P.910, *Subjective video quality assessment for multimedia applications*, 2008.

[23] L. Platisa, L. Van Brantegem, Y. Vander Haeghen, C. Marchessoux, E. Vansteenkiste, and W. Philips, "Psycho-visual evaluation of image quality attributes in digital pathology slides viewed on a medical color LCD display", *Proceedings of SPIE Medical Imaging,* vol. 8676, 2013.

[24] Recommendation ITU-R BT.500-13, *Methodology for the subjective assessment of the quality of television pictures*, 2012.

[25] B. Tulu, and S. Chatterjee, "Internet-based telemedicine: An empirical investigation of objective and subjective video quality", *Decisions Support Systems*, vol. 45, pp. 681-696, 2008.

[26] Recommendation ITU-R BT.500-11, *Methodology for the subjective assessment of the quality of television pictures*, 2002.

[27] J. Suad, and W. Jbara, "Subjective quality assessment of new medical image database", *International Journal of Computer Engineering and Technology*, vol. 4, pp. 155-164, 2013.

[28] A. Chaabouni, Y. Gaudeau, J. Lambert, J.-M. Moureaux, and P. Gallet, "Subjective and objective quality assessment for H.264 compressed medical video sequences", *Proceedings of the 4th International Conference in Image Processing,* pp. 1-5, 2014.

[29] B. Münzer, K. Schoeffmann, L. Böszörmenyi, J. Smulders, and J. Jakimowicz, "Investigation of the Impact of Compression on the Perceptual Quality of Laparoscopic Videos," *IEEE 27th International Symposium on Computer-Based Medical Systems*, pp. 153-158, 2014.

[30] M. Razaak, M. Martini, and K. Savino, "A study on quality assessment for medical ultrasound video compressed via HEVC", *IEEE Journal of biomedical and health informatics*, vol. 18, pp. 1552-1559, 2014.

[31] M. A. Usman, M. R. Usman, and S. Shin, "Quality assessment for wireless capsule endoscopy videos compressed via HEVC: From diagnostic quality to visual perception", *Computers in Biology and Medicine*, vol. 91, pp. 112-134, 2017.

[32] N. Nouri, D. Abraham, J. Moureaux, M. Dufaut, J. Hubert, and M. Perez, "Subjective MPEG2 compressed video quality assessment: Application to tele-surgery", *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp.

764-767, 2010.

[33] L. Chow, H. Rajagopal, R. Paramesran, and Alzheimer's Disease Neuroimaging Initiative, "Correlation between subjective and objective assessment of magnetic resonance (MR) images", *Magnetic Resonance Imaging*, vol. 34, pp. 820-831, 2016.

[34] H. Liu, J. Koonen, M. Fuderer, and I. Heynderickx, "The relative impact of ghosting and noise on the perceived quality of MR images", *IEEE Transactions on Image Processing*, vol. 25, pp. 3087-3098, 2016.

[35] A. Kumcu, K. Bombeke, L. Platisa, L. Jovanov, J. Van Looy, and W. Philips, “Performance of Four Subjective Video Quality Assessment Protocols and Impact of Different Rating Preprocessing and Analysis Methods”, *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, pp. 48-63, 2017.

[36] M. Pinson, and S. Wolf, “Comparing subjective video quality testing methodologies”, *Visual Communications and Image Processing*, vol. 5150, pp. 573-582, 2003.

[37] Recommendation ITU-R BT.1788, *Methodology for the subjective assessment of video quality in multimedia applications*, 2007.

[38] R. K. Mantiuk, A. Tomaszewska, and R. Mantiuk, “Comparison of four subjective methods for image quality assessment”, *Computer Graphics forum,* vol. 31, pp. 2478-2491, 2012.

[39] C. Nodine, H. Kundel, S. Lauver, and L. Toto, “Nature of expertise in searching mammograms for breast masses”, *Academic Radiology*, vol. 3, pp. 1000-10006, 1996.

[40] E. Krupinski, and R. Weinstein, “Changes in visual search patterns of pathology residents as they gain experience”, *Proceedings of SPIE*, vol. 7966, 2011.

[41] Z. Wang, H. Sheikh, and A. Bovik, “Objective Video Quality Assessment”, *The Handbook of Video Databases: Design and Applications*, pp. 1041-1078, 2003.

[42] L. Lévêque, W. Zhang, C. Cavaro-Ménard, P. Le Callet, and H. Liu, “Study of Video Quality Assessment for Telesurgery”, *IEEE Access*, vol. 5, pp. 9990-9999, 2017.

[43] L. Lévêque, W. Zhang, P. Parker, and H. Liu, “The Impact of Specialty Settings on the Perceived Quality of Medical Ultrasound Video”, *IEEE Access*, vol. 5, pp. 16998-17005, 2017.

[44] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard”, *IEEE Transactions Circuits and Systems for Video Technology,* vol. 13, pp. 560-576, 2003.

[45] Recommendation ITU-T, H.265, *High efficiency video coding*, 2013.

[46] G. Sullivan, J. Ohm, H. Woo-Jin, and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard”, *IEEE Circuits and Systems Society*, vol. 22, pp. 1649-1668, 2013.

[47] E. Samei, “Assessment of display performance for medical imaging systems”, *Report of the American Association of Physicists in Medicine, Task group 18*, Medical Physics Publishing, 2005.

[48] B. Hemminger, R. Johnston, J. Rolland, and K. Muller, “Introduction to perceptual linearization of video display systems for medical image presentation”, *Journal of Digital Imaging*, vol. 8, pp.21-34, 1995.

[49] National Electrical Manufacturers Association, Digital Imaging and Communications in Medicine (DICOM) Part 14: Greyscale Standard Display Function, Rosslyn, USA, 2004. [Online]. Accessed on June 2017. Available: http://medical.nema.org/dicom/2004/04_14PU.pdf

Lucie Lévêque received the B.Eng. and M.Eng. degrees in cognitics from the École Nationale Supérieure de Cognitique, Bordeaux, France, in 2013 and 2015, respectively, and the M.Sc. degree in biomedical imaging from the University of Angers, Angers, France, in 2015. She is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering at Cardiff University, Cardiff, U.K. Her research interests include visual media quality assessment, and medical image perception.

Hantao Liu received the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands, in 2011. He is currently an Assistant Professor with the School of Computer Science and Informatics, Cardiff University, Cardiff, U.K. He is currently serving for IEEE MMTC as the Chair of the Interest Group on Quality of Experience for Multimedia Communications, and he is an Associate Editor of the IEEE Transactions on Human-Machine Systems and the IEEE Transactions on Multimedia. His research interests include visual media quality assessment, visual attention modelling and applications, visual scene understanding, and medical image perception.

SPECIAL ISSUE on OPTICAL WIRELESS COMMUNICATION

Guest Editor: Tuncer Baykas
Istanbul Medipol Univeristy, Turkey
tbaykas@medipol.edu.tr

This special issue of Frontiers focuses on the recent progresses of an emerging technology in wireless communication, which is the optical wireless communication (OWC). Various research groups all around the globe are currently working on OWC, which have recently also attracted the interest of the industry including the launch of three dedicated standardization activities.

In the first paper, IEEE 802.15.7m standardization activity is explained. The main focus of this group is using camera sensor as OWC receivers.

The second paper focuses on a specific technology within the IEEE 802.15.7m called LED-ID.

The third paper summarizes activities of 802.15.13 group, which prepares OWC specifications in optically transparent media enabling high data rate transfer among end points at rates up to 10 Gb/s and ranges up to 200 m unrestricted line of site.

The fourth paper is about a specific modulation scheme in OWC, which is called Variable Transparent Amplitude-Shape-Color (VTASC) Modulation.

The last but not least, the paper entitled "IEEE 802.11 Light Communications Study Group", summarizes OWC related activities in 802.11, where a new task group start preparing specifications.

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

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.

IEEE802.15.7m OWC PHY Specification Overview

Vinayagam Mariappan[1], Jaesang Cha[1†]
[1]*Seoul National University of Science and Technology, Seoul, Korea*
vinayagam_m@hotmail.com , chajs@seoultech.ac.kr†

1. Introduction

The OWC is a secured and harmless promising wireless communication technique alternative to frequency saturated and highly interference additive RF based wireless communication. The technology development in LED enables high speed illumination switching control on LED's and baseband processing power with imaging devices enabled the new research to use computer vision based camera applications on smart devices. The technology trend changes in high speed illumination control on LED lighting devices and increasing use of smart devices with cameras enabled the new research to use camera as a visible light communication (VLC receivers). The VLC researchers started using lighting devices, LCD screens and other all type of display elements as transmitters.

The IEEE standardization association approved new project authorization request from IEEE 802 for VLC to amend the IEEE 802.15.7-2011 standard [1] for better and more application enabling standard in 2014 due to the growing interest on VLC technology is IEEE.802.15.7m. The IEEE802.15.7m OWC standardization activities started with three main topics consideration as:

- LED-ID (Low Rate PD Communications) : LED-ID is defined as wireless light ID (Identification) system using various LED Lights
- OCC (Image Sensor Communications) : OCC based Positioning Technology connected with Smart Device Image Sensor
- Li-Fi (High Rate PD Communications)

Then High Rate PD communication moved as a separate Task Group (TG13) named with Multi-Gigabit per Second Optical Wireless Communications (OWC) with Ranges up to 200 meters (IEEE. 802.15.13). The IEEE.802.15.7m continues with LED-ID and OCC amendment as standardization activities for IEEE.802.15.7 OWC amendment. The IEEE.802.15.7m proposed different PHY modes for LED-ID and OCC with revised MAC control methodology. This letter summarize the PHY modes used for IEEE.802.15.7m OWC amendment.

2. IEEE802.15.7m PHY Modes

The IEEE802.15.7m PHY layer supports multiple PHY types as follows [6],

a) PHY I: This PHY type is intended for outdoor usage with low data rate applications. This mode uses on-off keying (OOK) and variable pulse position modulation (VPPM) with data rates in the tens to hundreds of kbps.
b) PHY II: This PHY type is intended for indoor usage with moderate data rate applications. This mode uses OOK and VPPM with data rates in the tens of Mbps.
c) PHY III: This PHY type is intended for applications using color-shift keying (CSK) that have multiple light sources and detectors. This mode uses CSK with data rates in the tens of Mbps.
d) PHY IV: This PHY type is intended for use with discrete light sources with data rates up to 22 kbps using various modulations.
e) PHY V: This PHY type is intended for use with diffused surface light sources with data rates up to in the kbps.
f) PHY VI: This PHY type is intended for use with video displays with data rates in the kbps using various modulations.

The PHY I, II, III use of over-the-air PHY frame configuration as stated in IEE802.15.7-2011 [1]. The use of over-the-air PHY frame configuration is forbidden for PHY types IV, V and VI [2]. The PHY frame configuration is mandatory that PHY frame configuration be done via the PHY PIB. This is due to the fact that unlike traditional wireless LAN/PAN, the data rates associated with OCC are such that the configuration overhead cannot be tolerated. This means that there is no "base default" transmission mode. In addition, it is anticipated that configuration will be with application layer "APPS" that are specifically loaded to support a particular OCC PHY mode. The PHY PIB is not transmitted; rather, it is written by the Device Management Entity and is read by the PHY layer.

Table 1 shows the supported PHY IV, V, and VI operating modes on amendment process of the IEEE802.15.7m TG.

Table 1. PHY IV, V and VI Operating Modes with supported modulations [3][4][5].

OCC MCS ID	Modulation	RLL	Optical clock rate	FEC	Data Rate
PHY IV operating modes					
0	UFSOOK	NA	multiple of frame rate	MIMO path dependent	10 bps (60 fps camera with 1/3 rate FEC)
1	Twinkle VPPM	NA	4x bit rate	RS(15,11)	4 kbps
2	S2-PSK	Half-rate code	10 Hz	Temporal error correction	5 bps
3	HS-PSK	Half-rate code for S2-PSK	10 kHz	RS(15,7)	22 kbps
4	Offset VPPM	None	25 Hz	RS(15,2)/RS(15,4)/None	18 bps
PHY V operating modes					
5	RS-FSK	None	30 Hz	XOR FEC	120 bps (16-FSK without FEC)
6	C-OOK	Manchester/4B6B	2.2 kHz / 4.4 kHz	Inner Hamming code, Optional outer RS(15,11)	400 bps (4.4KHz, 4B6B, no FEC)
7	CM-FSK	None	10 Hz	Optional outer RS(15,11)	60 bps (64-FSK, no FEC)
8	MPM	None	25 kHz	Temporal error correction	12.5 kHz
PHY VI operating modes					
9	A-QL	None	10 Hz	Outer RS(15,11), Inner CC(1/4)	5.54 kbps (32x32 cells TX, CC(1/4) RS(15,11))
10	HA-QL	Half-rate code	10 Hz	Outer RS(15,7), Inner CC(1/4)	140 bps
11	VTASC	None	30 Hz	RS(64,32)/RS(160,128)/None	512 kbps (FEC None)
12	SS2DC	None	30 Hz	RS(64,32)/RS(160,128)/None	368 kbps (FEC None)
13	IDE-MPFSK-BLEND	None	30 Hz	RS(64,32)/RS(160,128)/None	32 kbps (FEC None)
14	IDE-WM	None	30 Hz	RS(64,32)/RS(160,128)/None	256 kbps (FEC None)

3. Conclusion

As the process of the IEEE 802.15.7m to work on an amendment, which adds three more PHY addition to IEEE802.15.7-2011 PHY's to support LED-ID and OCC technology to support almost any type of VLC communications services. OWC merges lighting and data communications in applications such as area lighting, signboards, streetlights, vehicles, traffic signals, status indicators, displays, LED Panel, and digital signage by IEEE 802.15.7m amendment work.

Acknowledgement
This research was partially supported by the MSIT (Ministry of Science and ICT Korea, under the ITRC (Information Technology Research Center) support program (IITP-ㅏ 2018-2016-0-00311) supervised by the IITP (Institute for Information & Communications Technology Promotion).

References

[1] 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, September 2011.

[2] Rick Roberts, "Result of two email reflector discussions: MPDU text reorg and PIB configuration," online: https://mentor.ieee.org/802.15/dcn/17/15-17-0144-00-007a-mpdu-text-reorg-and-pib-configuration-email-discussion.docx, March 2017.

[3] Trang Nguyen, Yeong Min Jang, "OCC operating modes - table," online: https://mentor.ieee.org/802.15/dcn/17/15-17-0488-00-007a-table-81-phy-4-5-6-operating-modes.docx, September 2017.

[4] Jaesang Cha, Vinayagam Mariappan, et al., "OCC Operating Modes – Table Revision," online: https://mentor.ieee.org/802.15/dcn/17/15-17-0306-03-007a-occ-operating-modes-table.docx, August 2017.

[5] Hideki Aoyama, "OCC Operating Modes – Table Revision," online: https://mentor.ieee.org/802.15/dcn/17/15-17-0306-03-007a-occ-operating-modes-table.docx, June 2017.

[6] Jaesang Cha, Vinayagam Mariappan, et al., "LB-D2-Comment-Resolution-Based-Changes-On-PHY-Layer-General-Description," online: https://mentor.ieee.org/802.15/dcn/18/15-18-0245-00-007a-lb-d2-comment-resolution-based-changes-on-phy-layer-general-description.docx, May 2018.

Vinayagam Mariappan is a Ph.D. scholar majoring in Broadcasting and Communication Fusion Program from Graduate School of Nano IT Design Fusion, Seoul National University of Science and Technology, Seoul, Korea. He received his B.Tech degree in Electronics from Anna University, MIT Campus, Chennai in 1998, India and M.S degree in Media IT Engineering from Seoul National University of Science and Technology, Seoul, Korea in 2018. Also he is a voting member of the IEEE 802.15.7r1 Task Group. His research interests are OWC, OCC, LiFi, IoT/IoL, LED ID, Smart Digital Signage, Video Analytics.

Jaesang Cha received the Ph.D. degree from the Department of Electronic Engineering, Tohoku University in 2000. He worked for ETRI between 2002 and 2005. He is currently Professor at the Department of Electronics and IT Media Engineering, Seoul National University of Science & Technology, Seoul, Korea. He is also serving Head of IoT Research Center and IoT Convergence Research

Institute, Seoul National University of Science and Technology. He is serving as the Technical Editor of the IEEE 802.15.7r1 TG. His research interests are LED-ID, OCC, LiFi, IoT/IoL, LBS, ITS, Wireless Home Network.

LED-ID Technology for IEEE802.15.7m OWC

Jaesang Cha[1], Vinayagam Mariappan[1]
[1]*Seoul National University of Science and Technology, Seoul, Korea*
chajs@seoultech.ac.kr, vinayagam_m@hotmail.com

1. Introduction

The Optical Wireless Communication (OWC) technology in optically transparent media using light wavelengths 10,000 nm to 190 nm through which baseband signals are modulated on the Light Emitting Diode (LED) emitting lights in addition to light illumination. The OWC is a secured and harmless promising wireless communication technique alternative to frequency saturated and highly interference additive RF based wireless communication. The IEEE Standard 802.15.7 for Short-Range Wireless Optical Communication Using Visible Light in 2011 [1] is the first IEEE visible light communication (VLC) standard from IEEE 802 LMSC. The IEEE 802.15.7 OWC standardization was carried over use of photodiodes receivers with limited application usage.

The technology trend changes in high speed illumination control on LED lighting devices and increasing use of smart devices with cameras enabled the new research to use camera as a VLC receivers. The VLC researchers started using LCD screens and other all type of display elements as transmitters. The IEEE standardization association approved another project authorization request from IEEE 802 for VLC to amend the IEEE 82.15.7-2011 standard for better and more application enabling standard in 2014 due to the growing interest on VLC technology is IEEE.802.15.7m.

Light Emitting Diode Identification (LED-ID) is one of the key technology considered for IEEE.802.15.7m standardization which is used for identification, data transmission and illumination simultaneously in different environments including shops and supermarkets, museums, plenum spaces, etc. This paper summarize the LED-ID Technology and standardization activities for IEEE.802.15.7m OWC amendment.

2. LED-ID Technology

An LED-ID technology works same as like another electronic Tag technologies which transmits the short digital broadcast message. The LED-ID technology transmits voice/video/data via light waves through Tag designed using LED. The LED-ID system consist of two main components as shown in Figure 1, i.e. Reader and Tag. The LED-ID Tag have at least one LED and LED-ID Reader can be a PD / Image Sensor based photo detection device. The LED-ID Tags are classified into active Tags designed using PD and Semi-active Tags designed using Image Sensor.

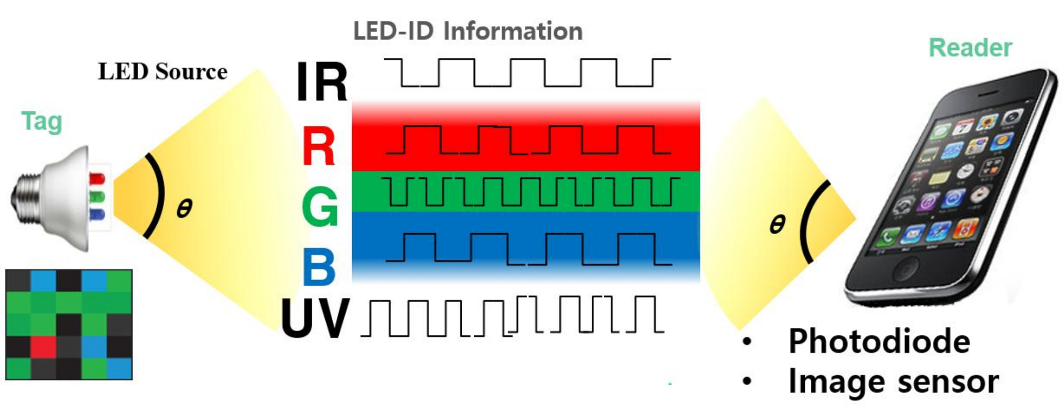


Figure 1. LED-ID System Model [2]

LED-ID technology with their low complexity on design and usage, considered to be a replacements of popular RFID, Bluetooth, and Wi-Fi Tags. The LED-ID technology advantages with RF based ID technologies are low power consumption, harmless to human, green environment, high security, high speed read and write capability, no interference with RF, supports QoS. The comparison between LED-ID and RF-ID is shown in Table 1.

Table 1. Comparison between LED-ID & RF-ID System

Tech item		LED-ID	RF-ID
Difference	Source	LED lights	RF
	Security	High	Low
	Harmful	X	O (Due to Electromagnetic)
	Promotional in market	High	Low
	Dedicated Reader	Unnecessary (using smart devices with Camera & PD)	Necessary
Common	Configuration	LED Tag & Smart device / PD	RF-ID Tag and Reader
	Directivity	LED-ID Tag to Smart device / PD	RF-ID Tag to RF-ID Reader

The LED-ID technology can be used in the different application service scenarios like supermarket/department store, museum, intelligent menu system (airplane, restaurant, etc.), advertisement, LBS applications, technology product expo, underwater / seaside Communication [9], car parking, entertainment / movie/ amusement park, and guiding system for blind person, etc. [7].

For An example of LED-ID systems in use are "smart" supermarket carts (shown in Figure 2 (a)), which via illumination infrastructure record a shoppers' path for sub-sequent analysis. LED-ID systems may also be used to "tag" particular shop shelves and areas to enable fast product localization.

(a) Indoor Store Service (b) Exhibition Service

Figure 2. Applications Service Scenario of LED-ID [3]

The digital signage systems used in museums, exhibitions, etc. are another example of LED-ID technology (shown in Figure 2 (b)). These signage systems may be used with specialized applications for mobile platforms to provide information about objects in proximity.

3. LED-ID on IEEE802.15.7m OWC Standardization

The starting of IEEE802.15.7m OWC standardization activities consist of three main topics including LED-ID as:

- LED-ID (Low Rate PD Communications) : LED-ID is defined as wireless light ID (Identification) system using various LED Lights
- OCC (Image Sensor Communications) : OCC based Positioning Technology connected with Smart Device Image Sensor
- Li-Fi (High Rate PD Communications)

The LED-ID technology considered the transmitter to support the LED Tags & display/image patterns (QR/Color code like Bar code), smart phone flash lights, lighting source, etc. for various applications and receiver to support PD / Camera [4] to measure intensity of visible light, IR and/or near UV, as receiver [4] [8]. The LED-ID technology for IEEE 802.15.7m technical consideration in the standardization history [5] is shown in Figure 3.

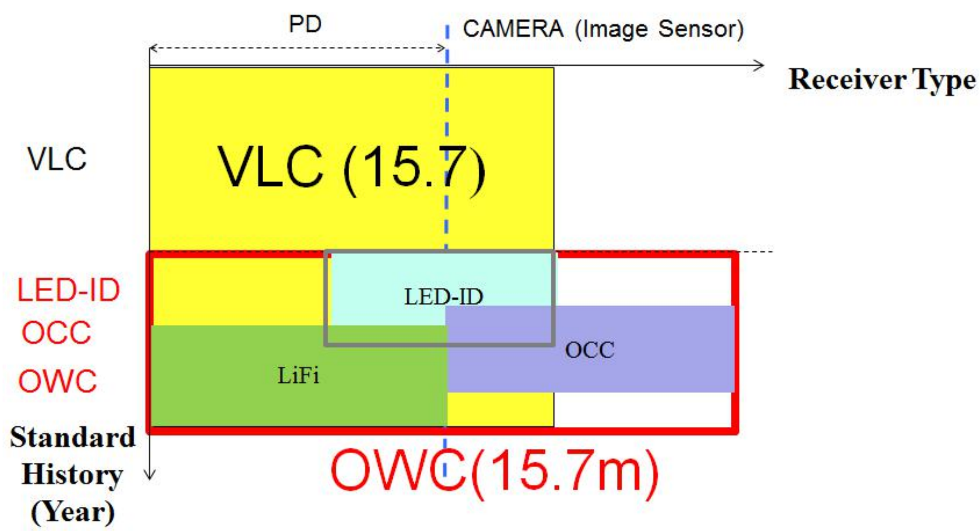


Figure 3. VLC and OWC Technologies History [5]

In the IEEE802.15.7m standardization activity, LED-ID technology channel models commonly classified according to two criteria, namely, degree of directionality of the transmitter and receiver and whether the link relies upon the existence of a line-of-sight (LOS) path between them and Non-LOS [6] as shown in Figure 4.

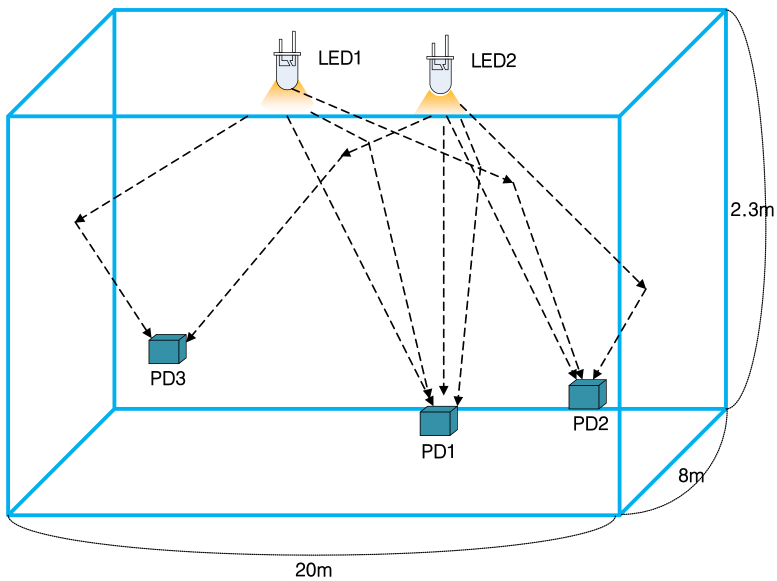


Figure 4. LED-ID Channel Model [6]

4. Conclusion

From the starting of IEEE802.15.7m standardization Task Group (TG), LED-ID is the one of the main technology considered to be standardized in the amendment of IEEE 802.15.7-2011. The LED-ID technology considered the LED Tags, Display / Signage, smart phone flash lights, lighting source, etc. as a transmitting device and PD / Camera as receiver to support various applications services. The present IEEE802.15.7m Letter ballot (LB) recirculation 2 draft released with LED-ID as a one of core technology as specified in IEEE802.15.7m PAR and CSD.

Acknowledgement
This research was partially supported by the MSIT (Ministry of Science and ICT Korea, under the ITRC (Information Technology Research Center) support program (IITP- 2018-2016-0-00311) supervised by the IITP (Institute for Information & Communications Technology Promotion).

References

[7] 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.

[8] Jaesang Cha, "Introduction of LED-ID and Smart Device Camera based Applications: For Short-Range Optical Wireless Communications Tutorial," online: https://mentor.ieee.org/802.15/dcn/15/15-15-0196-00-007a-introduction-of-led-id-and-smart-device-camera-based-applications.pptx, Mar 2015.

[9] Jaesang Cha, et al., "Home-page document for LED-ID: LED-ID Related Technical Features and Applications," online: https://mentor.ieee.org/802.15/dcn/15/15-15-0410-00-007a-home-page-document-for-led-id-led-id-related-technical-features-and-applications.pptx, May 2015.

[10] Jaesang Cha, et al., "Draft of LED-ID Part of TG7r1 Technical Considerations Document," online: https://mentor.ieee.org/802.15/dcn/15/15-15-0411-05-007a-led-id-part-of-tg7r1-technical-considerations-document.docx, May 2015.

[11] Jaesang Cha, Vinayagam Mariappan, et al., "IEEE 802.15.7m Technical Consideration History," online: https://mentor.ieee.org/802.15/dcn/16/15-16-0808-00-007a-ieee-802-15-7m-technical-

consideration-history.ppt, Nov 2016.

[12] Jaesang Cha, et al., "Necessity Review of Channel Models for LED-ID LOS Applications," online: https://mentor.ieee.org/802.15/dcn/15/15-15-0716-00-007a-necessity-review-of-channel-models-for-led-id-los-applications.pptx, Sep 2015.

[13] Rick Roberts, "15.7 Revision: Short-Range Optical Wireless Communications Tutorial," online: https://mentor.ieee.org/802.15/dcn/15/15-15-0112-02-007a-short-range-optical-wireless-communications-tutorial.pdf, Jan 2015.

[14] Yeong Min Jang, Jaesang Cha, et al., "Technical Considerations Document," online: https://mentor.ieee.org/802.15/dcn/15/15-15-0492-00-007a-technical-considerations-document.docx.

[15] Yeong Min Jang, et al., "Response to 15.7r1 CFA: OCC Application in Light house-to-Ship Communication," online: https://mentor.ieee.org/802.15/dcn/15/15-15-0404-00-007a-kookmin-university-response-to-15-7r1-cfa-occ-application-in-light-house-to-ship-communication.ppt.

Jaesang Cha received the Ph.D. degree from the Department of Electronic Engineering, Tohoku University in Japan in 2000. He worked for ETRI between 2002 and 2005. He is currently Professor at the Department of Electronics and IT Media Engineering, Seoul National University of Science & Technology, Seoul, Korea. He is also serving Head of IoT Research Center (IoT/IoL ITRC) and IoT Convergence Research Institute, Seoul National University of Science and Technology. And he is serving as the Technical Editor of the IEEE 802.15.7r1 Task Group. His research interests are LED-ID, OCC, LiFi, IoT/IoL, LBS, ITS, Wireless Home Network.

Vinayagam Mariappan is a Ph.D. candidate majoring in Broadcasting and Communication Fusion Program from Graduate School of Nano IT Design Fusion, Seoul National University of Science and Technology, Seoul, Korea. He received his B.Tech degree in Electronics from Anna University, MIT Campus, Chennai in 1998, India and M.S degree in Media IT Engineering from Seoul National University of Science and Technology, Seoul, Korea in 2018. Also he is a voting member of the IEEE 802.15.7r1 Task Group. His research interests are OWC, OCC, LiFi, IoT/IoL, LED ID, Smart Digital Signage, Video Analytics, and Deep Learning.

IEEE 802.15.13 Multi Giga bit/sec Optical Wireless Communications Project

Qiang Li[1] Tunçer Baykaş[2], Volker Jungnickel[3]
[1] *Huawei Technologies, China ,* [2]*İstanbul Medipol University, Turkey ,*[3]*Fraunhofer Heinrich Hertz Institute, Germany*
tbaykas@medipol.edu.tr

1. Introduction

IEEE 802.15 working group started a standardization project to define a Physical (PHY) and Media Access Control (MAC) layer using light wavelengths from 10 000 nm to 190 nm in optically transparent media for optical wireless communications in 2017. The standard will be capable of delivering data rates up to 10 Gb/s at distances in the range of 200 m unrestricted line of sight. It will be designed for point to point and point to multi point communications in both non-coordinated and coordinated topologies. For coordinated topologies with more than one peer coordinator there will be a master coordinator. The standard will include adaptation to varying channel conditions and maintaining connectivity while moving within the range of a single coordinator or moving between coordinators. In this paper we will summarize the latest status of the project P802.15.13. [1]

2. IEEE P802.15.13 Purpose

The purpose of the standardization project is to define optical wireless communication (OWC) specifications in optically transparent media enabling high data rate transfer among end points at rates up to 10 Gb/s and ranges up to 200 m unrestricted line of site and which are capable of meeting the needs of industrial and similar classes of applications requiring, secure, high performance, high data rate communications which are non-interfering with existing Radio Frequency (RF) systems. [1]

1. OWC Channel Models

4 scenarios are considered as channel models in P802.15.13

- Open Office/Office with Cubicles (Fig. 1).
- Office with Secondary Light (Fig. 1).
- Home (Fig. 2).
- Manufacturing Cell (Fig. 2)

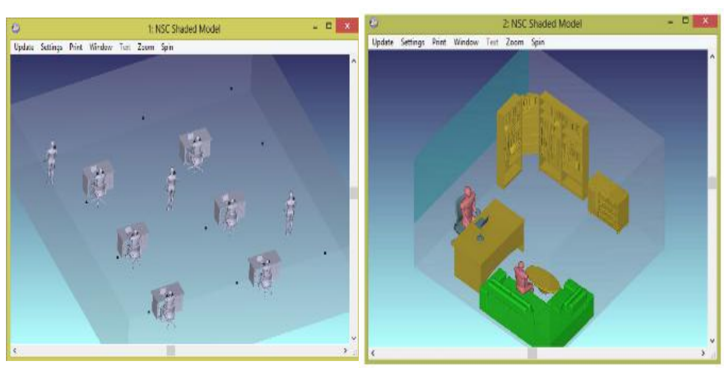


Figure 1. Office and Office with Secondary Light Scenarios

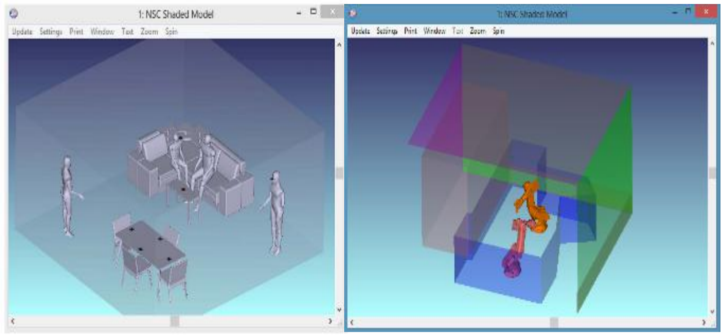


Figure 2. Home and Manufacturing Cell Scenarios

In an effort to develop more realistic channel models, a modeling approach based on ray tracing was used [2]. The proposed approach is based on Zemax®; a commercially available optical and illumination design software [3]. 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 [4,5]. Observed channel impulse responses vary between an almost AWGN channel and channels with sever delay spreads as shown in Fig. 3.

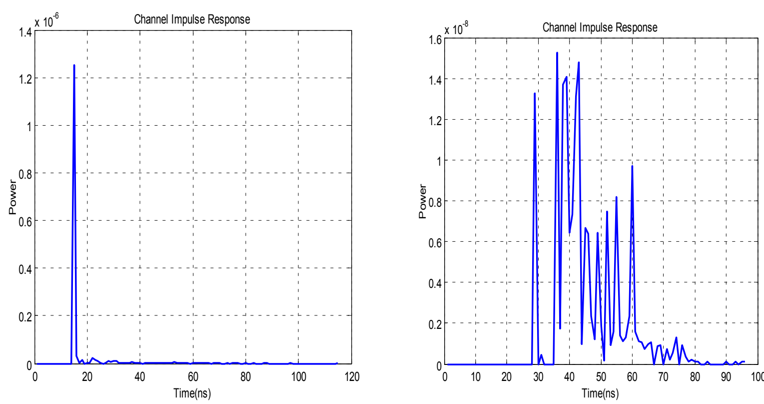


Figure 3. Channel Impulse Response Examples

2. OWC Network Topologies and advanced network functionalities

The IEEE 802.15.13 optical wireless personal area network (OWPAN) standard maps the intended applications to four topologies: peer-to-peer, star, broadcast and coordinated, as shown in Figure 4. Moreover, two advanced network functionalities are supported, relaying and heterogeneous networking of OWC and RF, as shown in Figure 5. In the star topology, the communication is established between devices and a single central controller, called the coordinator. In the peer-to-peer topology, one of the two devices in an association takes on the role of the coordinator. In the coordinated topology, multiple devices communicate with multiple coordinators, supervised by a master coordinator. The master coordinator connects to each coordinator via a backhaul link. Note that the functionality of the master coordinator is not part of this standard. In addition, two advanced network functionalities may be enabled: relaying and heterogeneous RF-OWC. With the relaying functionality, an intermediate relay node is introduced between the coordinator and the device. With the heterogeneous RF-OWC functionality, data transmission over the optical wireless link can combine with a parallel radio-based wireless link. The coordinator, the relay node and the master coordinator might often be mains powered, while the devices will often be battery powered.

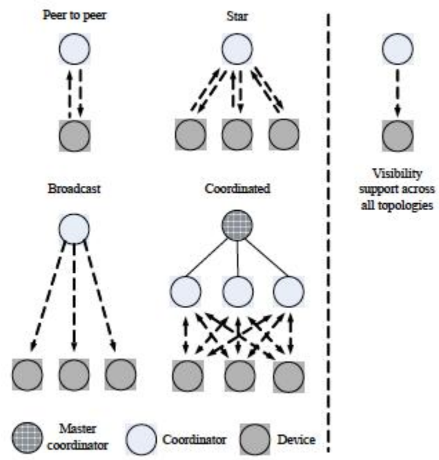


Figure 4. Different Network Topologies [6]

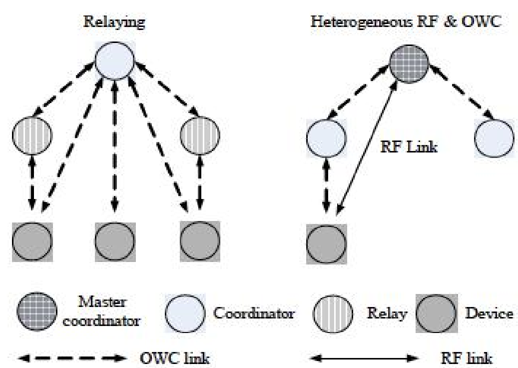


Figure 4. Advanced Network functionalities [6]

3. OWC Physical Layer

The MAC layer intends to support three PHY types.

a) Pulsed Modulation PHY (PM-PHY): This PHY is intended for indoor usage with moderate data rate applications and it allows operation at very low power (uplink, IoT). This mode uses 2-PAM modulation together with 8B10 line code or M-PAM (M=2, 4, 8, 16) together with Hadamard-coded modulation (HCM) at variable data rates from 1 to 150 Mb/s.

b) Low bandwidth OFDM PHY (LB-PHY): This PHY is intended for moderate data rate optical wireless transmissions using OFDM based modulation with variable data rates ranging from 1 to tens of Mb/s.

c) High bandwidth OFDM PHY (HB-PHY): This PHY is intended for high rate optical wireless transmissions using OFDM modulation and bit loading at data rates between 10 Mb/s up to several Gb/s.

4. Future of 802.15.13 OWC Group

As of May 2018, the TG is at the comment resolution stage and collects revised submissions in the agreed-upon writing style to generate a stable draft. The group is planning to finalize the project in March 2019.

Acknowledgements
The work of Tuncer Baykas is carried out as an activity of the Ardeb project 215E311 “"5g Ve Sonrası Için Yenilikçi Optik Kablosuz Haberleşme Teknolojileri"”funded by the Tubitak.

References

[16] The IEEE 802.15.13 Multi Giga bit/sec Optical Wireless Communications Task Group Project Authorization Request (PAR).

[17] F. Miramirkhani, M. Uysal, E. Panayirci, “Novel channel models for visible light communications”. SPIE Photonics West, February 7-12, 2015.

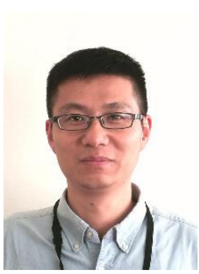
[18] Zemax® 13 Release 2, Radiant Zemax® LLC. www.radiantzemax.com/zemax

[19] M. Uysal, *et al.* “TG7r1 Channel Model Document for High-rate PD Communications,”

[20] 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

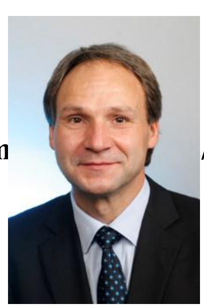
[6] John Li, “On unifying PPDU formats”, Online: https://mentor.ieee.org/802.15/dcn/17/15-17-0372-03-0013-on-unifying-ppdu-formats.pptx

[7] V. Jungnickel *et al.*, "A European view on the next generation optical wireless communication standard," *2015 IEEE Conference on Standards for Communications and Networking (CSCN)*, Tokyo, 2015, pp. 106-111.



Qiang Li received his B.S. from Beijing University of Posts and Telecommunications in 2004 and his M.S. from Peking University in 2007. He served as technical editor and secretary for 802.15 TG13 and contributed to many standardization projects, including 802.11. Currently he works as a senior research engineer in Huawei technologies.

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 802.15.13, 802.11af and 1900.7. Currently he is the vice director of the “Centre of Excellence in Optical Wireless Communication Technologies (OKATEM)”.



Volker Jungnickel (M) received doctoral and habilitation degrees in Physics and Communications Engineering from Humboldt University and Technical University in Berlin in 1995 and 2015, respectively.

He joined Fraunhofer HHI in 1997 working on optical wireless communication, multiple antenna techniques in mobile networks and new fixed access network infrastructures. Besides, he serves as Privatdozent at Technical University in Berlin with lectures and supervises Bachelor, Masters and Ph.D. thesis. Volker served as Chair for 802.15 TG13 and contributed to 802.11 TIG/SG on Light Communication.

VTASC – Light based Flexible Multi-Dimensional Modulation Technique for OWC

Jaesang Cha[1], Minwoo Lee[1], Vinayagam Mariappan[1]
[1]Seoul National University of Science and Technology, Seoul, Korea
chajs@seoultech.ac.kr, alsdnya@gmail.com, vinayagam_m@hotmail.com

1. Introduction

The Optical Wireless Communications (OWC) can enable secure and interference free wireless links alternative to the frequency saturated and inference additive radio frequency (RF) technology due to the high directionality of narrow light beams between the light emitting device and photo detection device. The photo detection device can be PD or Camera (Image Sensor). The he high directionality of narrow light beams and RF interference free characteristics make OWC an attractive solution for high-speed, short-range wireless communications in many scenarios such as indoor and outdoor service. The Optical Camera Communication (OCC) / Image Sensor Communication (ISC) is the one of the targeted OWC technology for the IEEE 802.15.7m OWC Task Group (TG7r1), also known as the amendment of the IEEE 802.15.7-2011 Visible Light Communication (VLC) standard [1].

As a part of technical consideration for standardization activities, SNUST proposed Light based Flexible Multi-Dimensional Modulation Technique for OWC called VTASC. The VTASC proposal considered light sources like a LED Pulbs, LED Tags, Display / Signage, LED Patch, etc. as a transmitting device and photodetection device as receiver to support various applications services as shown in Figure 1.

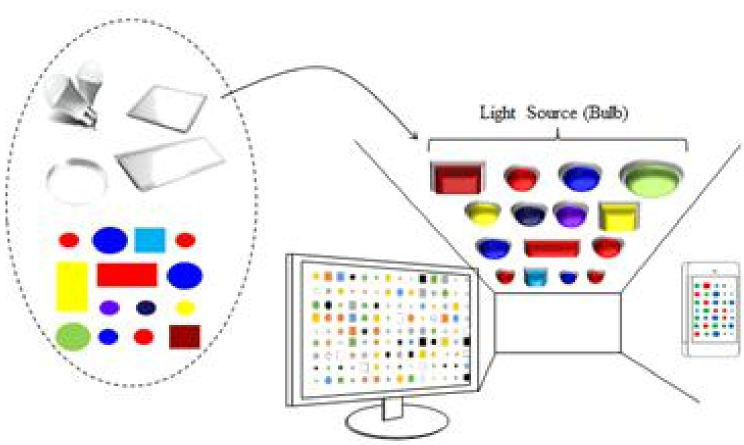


Figure 1. VTASC Transmitter Example [2]

The VTASC modulation technique is designed with specific key features in consideration to have error free and effective display to camera communication in the real-time usage of end system [3]. The design goals are,

- Angle and Distance Free Communication
- Rx Distance Adaptive Communication by Screen with interactive Camera
- Asynchronous Communication
- Rx Frame Rate independent Transmission
- Multi-Display Model for Transmission

To achieve the above described design goal, the VTASC uses Spread Spectrum as part of modulation process. The detailed VTASC modulation description is discussed in the following section.

2. VTASC Modulation Technique

The Color Shift Keying (CSK) is one of the promising modulation formats specifically for display based OWC system [1]. In order to improve the distance and angle free with higher bitrate, the new proposed color based modulation scheme called **V**ariable **T**ransparent **A**mplitude-**S**hape-**C**olor (**VTASC**) Modulation is proposed [4]. VTASC is a modulation scheme for visible-light communication involving single or multiple display (Panel, LED etc.) or light bulbs with variable transparency levels, sizes, shape models, and colors. VTASC enhances the OWC system performance with improved OWC throughput by increasing the bit per symbol rate, and avoiding the single color interference [6].

The VTASC is encoded by T (Transparency level) / A (Amplitude nothing but block size) / S (Shapes) / C (Colors) State as described in the Figure 2.

Figure 2. VTASC Code States [6]

The following illustration example shows the number of code levels in the VTASC modulation is (m x n x p x q) with two transparency levels, four block sizes, four shape models, and eight colors is 256 = 28 and this makes place to code 8 bit symbol with two levels of transparency, four size of blocks, four models of shape, and eight colors. The shape model design illustration shown in Figure 3.

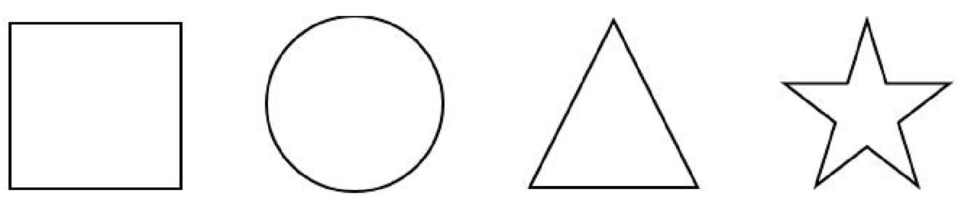
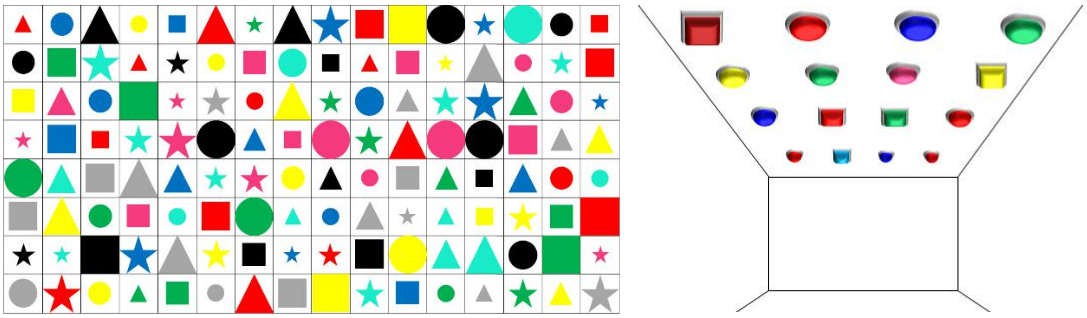


Figure 3. VTASC Shape Models Design Illustration with Four Shapes

The shapes inside the symbols pixel region is equally spaced in the VTASC symbols coding region. The coded symbols are ordered sequentially row by row same order as English text order and the coded region background color used shall be white. The zero padded VTASC coded symbols generated if the available number of data bits is less than the symbol mapping in the defined coding region. The VTASC code illustration is given Figure 4.



(a) Display (Panel, LED, etc.) Source (b) Light Bulb Source

Figure 4. VTASC Code Illustration with Four Shapes

The VTASC modulation works by overlaying the data mapped color code on visual area of lighting source. The overlaying visual symbols means that updating coded region pixel value according to VTASC coded symbol on display frame buffer to refresh on display. The overlaying coded symbol on frame buffer and data rate achievement vary based on the kind of display used to design the transmitter and the distance between transmitter and receiver. Table 1 describes the example data rate supported by VTASC code design with symbol size of 32x32 pixels on 42 inches full HD display with 16:9 aspect radio.

Table 1. VTASC Data Rate Example

Modulation (mxnxpxq)	RLL Code	Optical Clock Rate	FEC	Data Rate (Kbps)
2 Color VTASC Code (m = 2,n=4,p=4,q=2)	None	30Hz	RS(64,32)/ RS(160,128)/ None	384 Kbps (FEC None)
4 Color VTASC Code (m = 2,n = 4,p=4, q=4)	None	30Hz	RS(64,32)/ RS(160,128)/ None	448 Kbps (FEC None)
8 Color VTASC Code (m = 2,n=4,p=4,q=8)	None	30Hz	RS(64,32)/ RS(160,128)/ None	512 Kbps (FEC None)
2 Color SS VTASC Code (m = 2,n=4,p=4,q=2)	None	30Hz	None	192 Kbps[1]
4 Color SS VTASC Code (m = 2,n=4,p=4,q=4)	None	30Hz	None	224 Kbps[1]
8 Color SS VTASC Code (m = 2,n=4,p=4,q=8)	None	30Hz	None	256 Kbps[1]

[1] where spreading factor is 2

3. VTASC Coding Features

3.1 Spreading Sequence

The spread spectrum code used as a synchronization sequence. The spread spectrum used with VTASC based display / Light Pulbs to camera OWC to have effective asynchronous, distance adaptive scalable data rate controlled communication. The lighting device to camera communication adopted the binary zero-correlation duration (ZCD) code sequences as an optical spread code with the spreading code length [6].

3.2 Asynchronous Communication

The Asynchronous communication achieved when transmitting data, different spreading code is used per data frame. Each code sets repeated for spreading data according to spreading factor and each spreading code set 1, 2, 3, and 4 are assigned for successive 4 frames. The receiver side knows the spreading code of the transmitter synchronize the receiver application automatically. If Rx received the same frame, for example #1 frame receive twice, then receiver will de-spread frames using spreading code #1, spreading code #2. When processing using spreading code #2, dominant value will not appear so the frame will be discarded.

3.3 Scalable Bitrate Controller

To achieve robust communication, the scalable data transmission mode is proposed in VTASC model design. The frame is divided into multiple sub regions and each region has different data rate controlled data transmission is enabled. This approach adds robustness on system performance for receiver decoding rate adaptive communication based on the receiver performance.

4. Conclusion

Professor. Jaesang Cha (SNUST) proposed the VTASC modulation with differential encoding feature to achieve robustness and high data rate OCC/ISC as part of IEEE802.15.7m standardization activity of Task Group (TG7r1). The present IEEE802.15.7m Letter ballot (LB) recirculation 2 draft released with VATSC as a one of PHY mode in PHY VI.

Acknowledgement

This research was partially supported by the MSIT (Ministry of Science and ICT Korea, under the ITRC (Information Technology Research Center) support program (IITP- 2018-2016-0-00311) supervised by the IITP (Institute for Information & Communications Technology Promotion).

References

[21] 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, September 2011.

[22] Jaesang Cha, Vinayagam Mariappan, et al., “Display Light Pattern based Tx with SCAM,” online: https://mentor.ieee.org/802.15/dcn/16/15-16-0024-03-007a-display-light-pattern-based-tx-with-scam.ppt, January 2016.

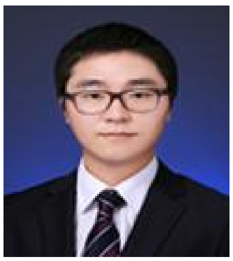
[23] Jaesang Cha, Vinayagam Mariappan, et al., “PHY/MAC for Variable Transparent Amplitude-Shape-Color (VTASC) modulation,” online: https://mentor.ieee.org/802.15/dcn/16/15-16-0278-00-007a-phy-mac-for-variable-transparent-amplitude-shape-color-vtasc-modulation.docx, March 2016.

[24] Jaesang Cha, Vinayagam Mariappan, et al., “PHY/MAC Draft D0 Comments Update for VTASC Modulation,” online: https://mentor.ieee.org/802.15/dcn/16/15-16-0528-01-007a-phy-mac-draft-d0-comments-update-for-vtasc-modulation.docx, July 2016.

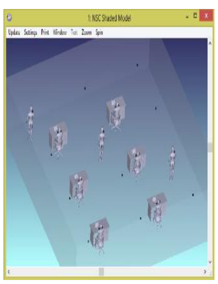
[25] Jaesang Cha, Vinayagam Mariappan, et al., “SNUST – VTASC PHY Layer Operating Modes and Specifications Revision,” online: https://mentor.ieee.org/802.15/dcn/17/15-17-0093-00-007a-snust-vtasc-phy-layer-operating-modes-and-specifications-revision.docx, January 2017.

[26] Jaesang Cha, Vinayagam Mariappan, et al., “SNUST Text for PHY VI Revision,” online: https://mentor.ieee.org/802.15/dcn/17/15-17-0571-01-007a-snust-text-for-phy-vi.docx, October 2017.

Jaesang Cha received the Ph.D. degree from the Department of Electronic Engineering, Tohoku University in Japan in 2000. He worked for ETRI between 2002 and 2005. He is currently Professor at the Department of Electronics and IT Media Engineering, Seoul National University of Science & Technology, Seoul, Korea. He is also serving Head of IoT Research Center (IoT/IoL ITRC) and IoT Convergence Research Institute, Seoul National University of Science and Technology. And he is serving as the Technical Editor of the IEEE 802.15.7r1 Task Group. His research interests are LED-ID, OCC, LiFi, IoT/IoL, LBS, ITS, Wireless Home Network.



Minwoo Lee received his Ph.D. degree in Graduate School of Nano IT Design Fusion, Seoul National University of Science and Technology, Seoul, Korea in 2018. Now he is working for IoT Convergence Research Institute, Seoul National University of Science and Technology. His research interests are LED-ID, OCC, LiFi, IoT/IoL, LBS, and Image Processing.



Vinayagam Mariappan is a Ph.D. candidate majoring in Broadcasting and Communication Fusion Program from Graduate School of Nano IT Design Fusion, Seoul National University of Science and Technology, Seoul, Korea. He received his B.Tech degree in Electronics from Anna University, MIT Campus, Chennai in 1998, India and M.S degree in Media IT Engineering from Seoul National University of Science and Technology, Seoul, Korea in 2018. Also he is a voting member of the IEEE 802.15.7r1 Task Group. His research interests are OWC, OCC, LiFi, IoT/IoL, LED ID, Smart Digital Signage, Video Analytics, and Deep Learning.

IEEE 802.11 Light Communications Study Group

Nikola Seramovski[1] , Tunçer Baykas[2]
[1]pureLiFi Ltd, *Scotland UK [1]İstanbul Medipol University, Turkey*
nikola.serafimovski@purelifi.com

1. Introduction

IEEE 802.11 Working Group (WG) is one of the most successful standardization groups among 802 LAN/MAN Standardization Committees. The group has published standards from TV bands to mmWave bands. In 2017, a topic interest group was formed as part of the IEEE 802.11 WG to study the feasibility of light communications (LC). The interest group evolved to a study group (SG), which prepares the project authorization request (PAR) and criteria for standards development (CSD) documents. At the time of the preparation of this document the PAR and CSD had been approved by the IEEE 802.11 WG as well as received unanimous consent from the IEEE 802 Executive Committee. The PAR was in the final stages of review before the official formation of the Task Group (TG), which would write the standard for LC as part of the IEEE 802.11 family. This paper summarizes the activities of the study group.

2. Need for the project

The study group explained the need for the project as: "We live in an increasingly connected world. The demand for wireless communications is increasing at nearly 50% per year according to the Cisco Visual Networking Index [1]. Three numbers capture the global ever-accelerating need for wireless bandwidth: by 2021 over 11 billion connected devices will be mobile, 70% of the IP traffic will be from mobile devices, 78% of the internet traffic will be video requiring high speed wireless. This enormous utilization results in a need for a continued increase of the connection speed and the capacity of wireless networks. This capacity could be satisfied by introducing additional unlicensed spectrum.

There are multiple solutions that can provide an increase in the available spectrum and increase the spectrum reuse efficiency in a given area, as well as increased speed. 60 GHz radio solutions, originally defined in IEEE Std 802.11ad (now part of IEEE Std 802.11-2016) and being extended in IEEE Std 802.11aj and IEEE P802.11ay are such examples. However, the continuous deployment and growth of IEEE 802.11 technology relies on exploiting further unlicensed spectrum based on the expected growth in the future. Additionally, non-radio frequency (RF) based wireless solutions may be preferred for multiple complementary use cases, like environments where traditional RF solutions are not allowed due to safety and/or security considerations, underwater communications.

The light spectrum, for the most part, has been underutilized for free space communication. Both the visible light spectrum and the infrared (IR) spectrum are unlicensed and could be used primarily in short-range wireless scenarios. In addition, the use of light for communications also supports the increasingly dense deployment of smaller and smaller cells.

The deployment of high-power solid state light sources together with large-area photodiodes and advanced electronics are key for the success of light communications (LC). In addition, physical (PHY) layer and medium access control (MAC) technologies have evolved significantly and are able to address existing use-cases for LC with enhanced performance as well as additional use-cases. Among those use-case is the complimentary deployment in traditional markets for IEEE 802.11, such as industrial wireless, home and enterprise networks, backhauling scenarios, underwater communication and wireless access in medical environments.

LC is a powerful complement to RF in environments where communications should be more secure (banks, R&D centres, defence, etc.) and where radio waves may be restricted (hospitals, electromagnetic interference (EMI) sensitive industrial facilities such as natural gas compression stations, nuclear power plants, etc.). The selection of use cases is driven by the facts that communications using the light spectrum do not interfere with any radio communications, the light spectrum is unlicensed for communications and the communications occur primarily inside the cone of the light.

With people in industrialized nations spending more than 85% of their time indoors (https://www.nature.com/articles/7500165.pdf), lighting has the opportunity to become an important communications infrastructure in the future."[1]

2. Scope of the project

This amendment specifies a new PHY layer and modifications to the IEEE 802.11 MAC that enable operation of wireless light communications (LC).

This amendment specifies a PHY layer that provides:
1) Uplink and downlink operations in 380 nm to 5,000 nm band,

2) All modes of operation achieve minimum single-link throughput of 10 Mb/s and at least one mode of operation that achieves single-link throughput of at least 5 Gb/s, as measured at the MAC data service access point (SAP),

3) Interoperability among solid state light sources with different modulation bandwidths.

This amendment specifies changes to the IEEE 802.11 MAC that are limited to the following clauses:

1) Hybrid coordination function (HCF) channel access,

2) Overlapping basic service set (OBSS) detection and coexistence,

3) Existing power management modes of operation (excluding new modes), and modifications to other clauses necessary to support these changes. [2]

3. Possible Use Cases

The group considered many use cases, as shown in Figure 1.

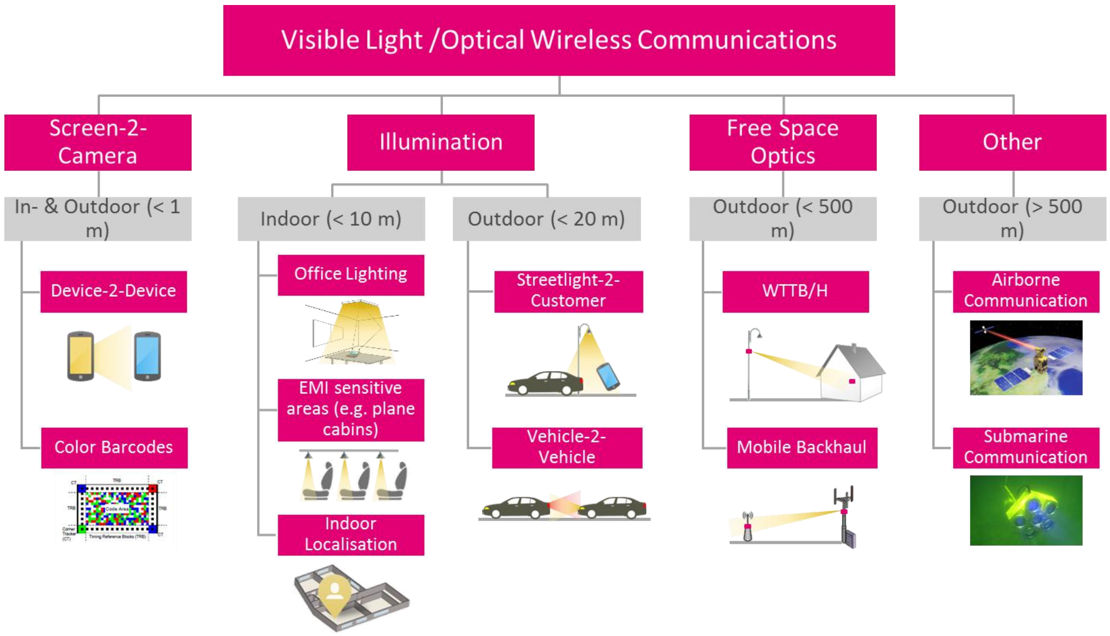


Figure 1. Possible Use Cases for LC communication. [3]

Among them illumination and free space optics use cases are considered to be used with 802.11 systems.

3. Considerations about the Physical Layer

Current 802.11 systems are based on mostly OFDM modulation. The Study Group showed in Figure 2 how OFDM can be easily modified and deployed in LC systems.

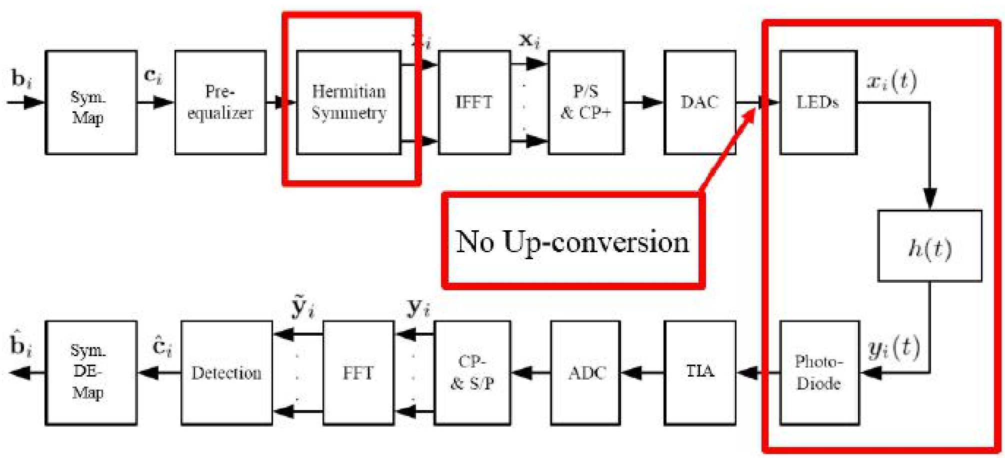


Figure 2. Use of OFDM in LC communication. [3]

3. The Future of the LC in 802.11

In Figure 3 and Figure 4, the historic and proposed timeline of the potential 802.11 LC TG standardization activity is provided

Topic Interest Group	2016	Nov	Technical and Economic considerations complete, including link budget and broad ecosystem support
	2017	Jan	
		Mar	
		May	Proposal to transition into a SG
Study Group		Jul	Use case considerations, collection of key performance indicators and first draft of Usage Model.
		Sep	
		Nov	
	2018	Jan	Limitations on the scope of the TG to focus on critical parts required for LC in 802.11
		Mar	
		May	Usage models and possible timelines
Task Group		Jul	TG 1st meeting – call for proposals TG Leadership selection Start on TG documents - approve work from TIG/SG (usage models, use-cases, KPIs, timelines, functional req. (FR), simulation models, channel) Technical Presentations
		Sep	Complete TG documents Start Specification Framework Doc. (SFD) Technical Presentations
		Nov	Continue with SFD

Figure 2. 802.11 LC activities 2016-2018. [4]

Task Group	2019	Jan	Continue with SFD
		Mar	Draft 0.1 ready for comment collection (CC)
		May	
		Jul	Complete CC, comment resolution and prepare Draft 1.0 Start WG LB
		Sep	WG LB comment collection and resolultion
		Nov	
	2020	Jan	
		Mar	Draft 2.0 and Recirculation
		May	
		Jul	
		Sep	Draft 3.0 and Recirculation
		Nov	**Sponsor Ballot (Target from PAR)**
	2021	Jan	
		Mar	Draft 4.0 and Recirculation
		May	Draft 5.0 and Recirculation
		Jul	**Submission to RevCom (Target from PAR)**
		Sep	

Figure 3. Planned 802.11 LC activities. [4]

The group is planning release a stable draft text by Nov. 2020 and to finalize its activities in line with the timeline outlined and agreed in the PAR in July 2021 by submitting the standard document to the revision committee.

Acknowledgements

Nikola Serafimovski and pureLiFi would like to acknowledge to acknowledge the H2020 WORTECS project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761329, as well as the H2020 ALC project, which has received funding from the Clean Sky 2 Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No 737645, for contributing to this work.

The work of Tuncer Baykas is carried out as an activity of the Ardeb project 215E311 “"5g Ve Sonrası Için Yenilikçi Optik Kablosuz Haberleşme Teknolojileri"”funded by the Tubitak.

References

[1] Nikola Serafimovski *et. al.* “a-csd-proposal-for-light-communications” Online: https://mentor.ieee.org/802.11/dcn/17/11-17-1603-09-00lc-a-csd-proposal-for-light-communications.docx

[2] The IEEE 802.15.11bb Amendmend Light Communications Project Authorization Request (PAR).

[3] Nikola Serafimovski, Tuncer Baykas *et al.* “ LC-for-802.11.pdf” Online: https://mentor.ieee.org/802.11/dcn/17/11-17-1048-04-00lc-lc-for-802-11-pdf.ppt

[4] Nikola Serafimovski, Ryan Mennecke *et. al.* “proposed-timeline-for-LG-TG” Online: https://mentor.ieee.org/802.11/dcn/18/11-18-0908-01-00lc-proposed-timeline-for-lg-tg.ppt

Nikola Serafimovski in his role as Vice-President of Standardization and Business Strategy at pureLiFi Ltd. works as with major companies in the area of LiFi technology and commercialization, leading the creation and cultivation of the LiFi ecosystem, marketing, sales and standardization. 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.

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.