Guest Editorial Special Issue on Quality of Experience for Advanced Broadcast Services

During the last decade, the evolution of the TV market has been remarkable. Broadcasters have been facing fresh challenges to cope with an increasing user demand for new services. With second-screen adoption and the increase of real-time news consumption via social channels, the broadcast landscape underwent major transformation in recent years: viewers have begun to demand highly customized experiences that meet their individual needs.

In addition to traditional Terrestrial/SAT/Cable broadcast, global service providers have begun to offer fixed/mobile advanced media delivery on the customer premises, enabling consumers to enjoy the emergence of new services offered by IPTV, 3DTV, UHDTV advancements in cloud services, and over-the-top (OTT) content providers. Within this framework, 5G is gaining more and more attention as a possible vehicle for these new features, whereas artificial intelligence (AI) through machine learning (ML) is starting to be adopted to predict user/network behaviour. Moreover, new technologies such as multi-sensorial media, virtual/augmented reality (AR/VR), holographic screens and the proliferation of connected devices through the Internet of Things (IoT), could create an immersive environment that will enrich a rapidly growing array of customer experiences and become the next frontier of advanced broadcast services. To this end, there is a need to evaluate the level of enhancement of these experiences and to compare their functionalities and requirements so operators can properly design their networks and regulators can assess the services offered to the audience.

The present special issue comes at an opportune time to provide an overview of advances in research and in the state-of-the-art technologies of fundamental areas that are critical to the Quality of Experience (QoE) for advanced broadcast services. To cover the wide scope of topics, this special issue consists of 23 formally submitted papers that are briefly introduced below.

In “No Reference Quality Assessment of Stereo Video Based on Saliency and Sparsity” [item 1] in the Appendix, J. Yang, C. Ji, B. Jiang, W. Lu and Q. Meng consider various video attributes in concert to construct a simple model to combine and analyze the diverse features based on saliency and sparsity. Firstly, they utilize the 3D saliency map or sum map, which remains the basic information of stereoscopic video, as a valid tool to evaluate the video’s quality. Secondly, they use the sparse representation to decompose the sum map of 3D saliency into coefficients, and then calculate the features based on sparse coefficients to obtain the effective expression of video’s message. Within that process, they take advantage of saliency and sparsity to extract and simplify features.

J. Korhonen in “Study of the Subjective Visibility of Packet Loss Artifacts in Decoded Video Sequences” [item 2] in the Appendix presents the results of a subjective study, using a methodology where a video sequence is displayed on a touchscreen and the users tap it in the positions where they observe artifacts. He also analyzes the objective features derived from those artifacts and proposes different models for combining those features into an objective metric for assessing the noticeability of the artifacts. The proposed method can be applied for developing packetization and error recovery schemes to minimize the subjective distortion in video broadcasting over error-prone packet-switched networks.

Q. Wu, H. Li, F. Meng and K. N. Ngan in “Towards A Blind Quality Metric for Temporally Distorted Streaming Video” [item 3] in the Appendix propose an efficient quality metric to blindly evaluate the user experience for stalled streaming video without using its original sequence. With the rapid progress of mobile Internet, the streaming video service has flourished over wireless networks in recent years. A smooth playback experience is crucial for the popularization of these services. However, limited by fluctuating bandwidth and various network impairments, streaming video inevitably suffers all kinds of stalling events, which significantly distorts its temporal structures and results in annoying jerky playback. Instead of requiring buffer or manifest information like existing methods, authors only access the decoded video and extract two complementary image features, i.e., global intensity and local texture, to estimate the stall number and duration. Then, by means of a straightforward and easy-to-use linear combination model, they map the normalized stall number and duration information to a quantitative quality score. Experimental results on the publicly available LIVE-Avvasi mobile video database show that the predicted video quality is highly consistent with the user experience and outperforms many state-of-art QoE models.

In “Quality of Experience for 3-D Immersive Media Streaming” [item 4] in the Appendix, A. Doumanoglou, D. Griffin, J. Serrano, N. Zioulis, T. Phan, D. Jimenez, D. Zarpalas, F. Alvarez, M. Rio, and P. Daras undertake one of the first studies into the QoE of real-time 3D media content streamed to VR headsets for entertainment purposes, in the context of game spectating. They focus on tele-immersive media that embed real users within virtual environments of interactive games. In this study, they investigate varying network conditions for a set of tele-immersive media sessions produced in a range of visual quality levels. Further, they investigate user navigation issues that inhibit free viewpoint
VR spectating of live 3D media. After reporting on a study with multiple users, they analyze the results and assess the overall QoE with respect to a range of visual quality and latency parameters, and propose a neural network QoE prediction model for 3D media, constructed from a combination of visual and network parameters.

P. Paudyal, F. Battisti, M. Carli and P. Le Callet, in “Toward the Assessment of Quality of Experience for Asymmetric Encoding in Immersive Media” [item 6) in the Appendix] study the assessment of QoE for stereoscopic 3D video. The focus of this contribution is the development of a QoE assessment framework for understanding the visual effect of asymmetric and symmetric encoding for immersive media. Asymmetric stereoscopic video coding exploits the binocular suppression of the human vision system by representing one of the two views with a lower quality. Although this processing, has limited effect on image quality, it may influence the overall QoE. Many studies show that the QoE of immersive media such as 3DTV can be thought as the combination of perceived visual quality, perceived depth quality, visual fatigue, and visual discomfort. In this work, authors exploit the concept of Preference of Experience and protocols recently standardized for characterizing QoE. They conduct a case study using these standardized protocols to investigate the factors involving visual discomfort in stereoscopic video sequences with a focus on binocular rivalry, present the results of subjective experiments performed, by using the perceptual quality and preference of experience assessment protocols, for evaluating the impact of symmetrical, asymmetrical, and alternate coding schemes.

“Evaluation of the Concept of Dynamic Adaptive Streaming of Light Field Video” [item 6) in the Appendix] by P. Kara, A. Cserkaszky, M. Martini, A. Barsi, L. Bokor and T. Balogh presents a research on the dynamic adaptive streaming of light field video. Light field visualization has progressed and developed significantly in the past years. At the time of this special issue, light field displays are utilized in the industry and they became commercially available as well. Although their appearance on the consumer’s market is approaching, many potential applications of light field technology have not yet been addressed, such as video streaming, which may have a potential impact in future broadcast services. A series of subjective tests have been carried out to evaluate the presented concept of quality switching. Test participants were shown light field videos containing stallings and switches in spatial and angular resolution.

A.-F. Perrin, M. Rerabek, W. Husak and T. Ebrahimi in “Quality assessment of an HDR dual-layer backward-compatible codec compared to uncompromised SDR and HDR solutions” [item 7) in the Appendix] investigate the performance of a dual-layer backward-compatible compression codec, when compared to state-of-the-art high dynamic range (HDR) compression strategies, in terms of perceived quality. The evaluated system is a dual-layer compression scheme enabling the transmission of a backward-compatible standard dynamic range (SDR) stream along with an HDR stream, reconstructed from the residual-based enhancement layer and SDR mapping (i.e., prediction). In their conclusion, authors provide guidance regarding the compression strategy to use as well as bandwidth allocation for HDR delivery, ensuring both SDR and HDR contents with perceptually acceptable quality.

M. T. Vega, C. Perra, F. De Turck and A. Liotta, in “A Review of Predictive Quality of Experience Management in Video Streaming Services” [item 8) in the Appendix], review the most significant ‘predictive’ QoE management methods for video streaming services, showing how different ML approaches may be used to perform proactive control. Authors pinpoint a selection of the best-suited ML methods, highlighting advantages and limitations in specific service conditions. The review leads to lessons learned and guidelines to address QoE requirements in complex video services.

M. Narwaria in “Towards Better Statistical Validation of Machine Learning Based Multimedia Quality Estimators” [item 9) in the Appendix] sheds light on the limitations of the current approach to objective assessment of multimedia quality using ML which has been gaining popularity especially in the context of both traditional (e.g., terrestrial and satellite broadcast) and advanced broadcast services (such as Over-the-top media services, IPTV). In fact, such methods rely on training to find the optimal model parameters and suffer from a few limitations related to the qualitative aspects of training and testing data, the use of improper sample size for statistical testing, and the lack of focus on quantifying the learning ability of the ML based objective quality predictor. Therefore, the author proposes an alternate approach to overcome some of them. As a major advantage, the proposed guidelines not only offer a theoretically more grounded statistical comparison but also provide useful insights into how well the ML based objective quality predictors exploit data structure for learning.

“Subjective Panoramic Video Quality Assessment Database for Coding Applications” [item 10) in the Appendix] by Y. Zhang, Y. Wang, F. Liu, Z. Liu, Y. Li, D. Yang and Z. Chen considers the geometric transformation in projection and the limited resolution of Head-Mounted Device (HMD). This study proposes a modified display protocol of high resolution sequences for subjective rating test in which an optimal display resolution is determined based on the geometry constraints between the screen and human eyes. By sampling the videos to the optimal resolution before coding, the proposed method significantly alleviates the interference of HMD sampling while displaying, thus ensuring the reliability of subjective quality opinion in terms of video coding. Using the proposed display protocol, a subjective quality database for panoramic videos is established for video coding applications. Distortions are introduced with the High Efficiency Video Coding (HEVC) compression. Each sequence is evaluated by 30 subjects for video quality, following the Absolute Category Rating with Hidden Reference (ACR-HR) method. The rating scores and Differential Mean Opinion Scores (DMOS) are recorded and included in the database.

databases constructed with hand-crafted test sequences, the SQoE-III database consists of a total of 450 streaming videos created from diverse source content and diverse distortion patterns, with 6 adaptation algorithms of diverse characteristics under 13 representative network conditions. All streaming videos are assessed by 34 subjects, and a comprehensive evaluation is conducted on the performance of 15 objective QoE models from four categories with regards to their efficacy in predicting subjective QoE. Detailed correlation analysis and statistical hypothesis testing are carried out. The results of this study shed light on the future development of adaptive bitrate streaming algorithm and video QoE monitoring system.

S. Yasuhito and T. Morita in “Estimating Depth Range Required for 3D Displays to Show Depth-Compressed Scenes without Inducing Sense of Unnaturalness” [item 12] in the Appendix] deal with the problem of reproducing deep 3D scenes with sufficient quality on 3D displays with light-ray reconstruction, such as light field and integral 3D displays, due to their depth-reconstruction range restriction. To reveal the size of the depth range required for showing deep 3D scenes with perceptually appealing quality, they conducted evaluation experiments under practical viewing conditions using a 3D-display simulator that provides binocular and motion disparities. Even with deep 3D scenes (originally with a depth of 250 m), they found that a depth range of at least 1 m was necessary to show these scenes with sufficient quality in terms of naturalness using a nonlinear depth compression method. These results provide a design goal for future 3D-television development, suggesting that 3D displays can naturally show a variety of scenes that originally had substantial depths if they can reproduce a depth of 1 m.

Y. Shishikui and S. Yasuhito in “Effects of Viewing Ultra-high-resolution Images with Practical Viewing Distances on Familiar Impressions” [item 13] in the Appendix] investigate the psychological effects of viewing ultra-high-resolution images. This study conducted subjective evaluation experiments, in which images with different resolutions using familiar subjects were presented to viewers with practical viewing distances, and their ratings of impressions were obtained. They found an enhancement of the impressions of “beautiful” or “delicious” with an increase in the resolutions of the presented images. Furthermore, the tendency of this impression enhancement was observed even when viewing it as far as four times the design viewing distance. The results of multiple regression analyses provide insight on the production and processing of ultra-high-resolution images for impression enhancement.

In “Blind Image Quality Estimation via Distortion Aggravation” [item 14] in the Appendix], X. Min, G. Zhai, K. Gu, Y. Liu, and X. Yang introduce multiple pseudo reference images (MPRIs) by further degrading the distorted image in several ways and to certain degrees, and then compare the similarities between the distorted image and the MPRIs. Via such distortion aggravation, they obtain some references to compare with, i.e., the MPRIs, and utilize the full-reference (FR) image quality assessment framework to compute the quality. Validation is conducted on four mainstream natural scene image and screen content image quality assessment databases, and the proposed method is comparable to or outperforms the state-of-the-art blind image quality assessment measures.

H. Z. Nafchi, M. Cheriet, in “Efficient No-Reference Quality Assessment and Classification Model for Contrast Distorted Images” [item 15] in the Appendix] propose an efficient Minkowski Distance based Metric (MDM) for no-reference (NR) quality assessment of contrast distorted images which can occur during media acquisition when poor and varying illumination conditions and poor camera’s quality can drastically change image contrast and visibility. In this study, authors show that higher orders of Minkowski distance and entropy provide accurate quality prediction for the contrast distorted images. The proposed metric performs predictions by extracting only three features from the distorted images followed by a regression analysis. Furthermore, the proposed features are able to classify type of the contrast distorted images with a high accuracy. Experimental results on four datasets CSIQ, TID2013, CCID2014, and SIQAD show that the proposed metric with a very low complexity provides better quality predictions than the state-of-the-art NR metrics.

M. Torcoli, J. Herre, H. Fuchs, J. Paulus and C. Uhle, in “The Adjustment/Satisfaction Test (A/ST) for the Evaluation of Personalization in Broadcast Services and its Application to Dialogue Enhancement” [item 16] in the Appendix] work on content personalization in media delivery for advanced broadcast services with focus on audio. To enable the personalization of the relative level of the dialog and the background sounds, they recently proposed the Adjustment/Satisfaction Test (A/ST), a perceptual test where subjects interact with a user-adjustable system and their adjustments and the resulting satisfaction levels are studied. In this study, two configurations of this test paradigm are implemented and compared for the evaluation of Dialogue Enhancement (DE). The test configuration closer to the final application is found to provide less noisy data and to be more conclusive about the QoE. For this configuration, DE is tested both in the case in which the original audio objects are readily available and in the case in which they are estimated by blind source separation. The results show that personalization is extensively used, resulting in increased user satisfaction, in both cases.

In “Modeling User Quality of Experience of Olfaction-enhanced Multimedia” [item 17] in the Appendix] N. Murray, G. Muntean, Y. Qiao, S. Brenann and B. Lee, conduct a research on multisensory experiences which have recently gained significant traction in the research community as a novel method to enhance QoE beyond what is possible with traditional media. This paper presents a model developed based on empirical data as to the estimation of user QoE of olfaction-enhanced multimedia. The proposed model considers the influence of system, user and content factors on perceived QoE. The achieved results suggest that human, content, and system factors play a key role in perceptual multimedia quality of olfaction enhanced multimedia.

L. Jalal, M. Anedda, V. Popescu and M. Murroni, in “QoE Assessment for IoT based Multi Sensorial Media Broadcasting” [item 18] in the Appendix] propose the
use of the IoT paradigm to enable multi sensorial media delivery in smart home. In this study, the requirements in terms of synchronization between media and devices is analyzed and the architecture of the system is defined accordingly. Furthermore, a prototype is implemented in a real smart home scenario with real consumer devices, which allowed a subjective test measurement campaign to assess the QoE of the users and the feasibility of the proposed multi sensorial media TV service.

A. Martin, J. Egaña, J. Florez, J. Montalban, I. Olaizola, M. Quartulli, R. Viola, M. Zorrilla, in “Network Resource Allocation system for QoE-aware delivery of media services in 5G Network” [item 19] in the Appendix] provide a Network Resource Allocator system as the main contribution which enables autonomous network management aware of QoE in a 5G network scenario. Software Defined Networking (SDN) and Network Function Virtualization (NFV) techniques, as well as HTTP-based Adaptive Streaming (HAS) and Machine Learning technology are used to reach this 5G vision. This system predicts demand to foresee the amount of network resources to be allocated and the topology setup required to cope with the traffic demand. Furthermore, the system dynamically provisions the network topology in a proactive way, while keeping the network operation within QoS ranges. To this end, the system processes signals from multiple network nodes and end-to-end QoS and QoE metrics. This paper evaluates the system for live and on-demand Dynamic Adaptive Streaming over HTTP (DASH) and HEVC services. From the experimental results, it is concluded that the system is able to scale the network topology to address the level of resource efficiency required by media streaming services.

A. Bentaleb, A. C. Begen and R. Zimmermann, in “QoE-Aware Bandwidth Broker for HTTP Adaptive Streaming Flows in an SDN-enabled HFC network” [item 20] in the Appendix], propose a SDN based bandwidth broker solution for improving viewer experience for any type of content delivered to any type of consumer device using HAS in a hybrid fiber coax (HFC) network. This solution is designed to meet per session and per-group QoE objectives, to avoid common HAS distortion culprits such as video instability, unfair and unequal quality distribution and network resource undersutilization, and to scale to a large number of concurrent HAS sessions without introducing excessive overhead. Results confirm the effectiveness of the proposed solution over the state-of-the-art bitrate adaptation and bandwidth allocation schemes.

In “Buffer State is Enough: Simplifying the Design of QoE-Aware HTTP Adaptive Video Streaming” [item 21] in the Appendix], W. Huang, Y. Zhou, X. Xie, D. Wu, M. Chen and E. Ngai, research on DASH to make streaming smooth under limited bandwidth by only exploiting client side buffer state information and propose a pure buffer-based DASH scheme to optimize user QoE. They explicitly define an integrated user QoE model, which takes playback freezing, bitrate switch and video quality into account, and then formulate the problem into a non-linear stochastic optimal control problem. Next, they utilize control theory to design a dynamic buffer-based controller for DASH, which determines video bitrate of each chunk to be requested while stabilizing the buffer level. Extensive experiments have been conducted to validate the advantages of the proposed approach, and the results show that it can achieve the best performance compared with other alternative approaches.

In “Quality of Experience Driven Rate Adaptation for Adaptive HTTP Streaming” [item 22] in the Appendix] by M. V. P. Kumar and S. Mahapatra a QoE driven rate adaptation approach is proposed for adaptive HTTP streaming which jointly considers both bandwidth savings and video quality adaptation for rate adjustment, beneficial to both service providers and subscribers. At each decision epoch, the adaptation algorithm accumulates the time varying QoE of a viewer by accounting for all the impairments, namely, initial delay, quality transitions and playback interruptions that commonly occur during the playback of a video in adaptive streaming. Based on the possible bandwidth savings and resulting QoE variations, the algorithm decides how to adapt the bitrate dynamically and accordingly maximizes viewers’ QoE. The experimental evaluation carried over real-world wireless network environments demonstrate that the proposed adaptation approach can maximize viewers’ QoE in adaptive streaming, especially, under highly variable throughput conditions.

Finally, to test the generalizability and to facilitate the wide usage of QoE measurement techniques in real-world HTTP DASH applications, “5G-QoE: QoE Modelling for Ultra-HD Video Streaming in 5G Networks” [item 21] in the Appendix] by J. Nightingale, P. Salva Garcia, J. M. Alcaraz Calero and Q. Wang, proposes a 5G-QoE framework to address the QoE modelling for UHD video flows in 5G networks. Particularly, it focuses on providing a QoE prediction model that is both sufficiently accurate and of low enough complexity to be employed as a continuous real-time indicator of the ‘health’ of video application flows at the scale required in future 5G networks. The model has been developed and implemented as part of the EU 5G PPP SELFNET autonomic management framework, where it provides a primary indicator of the likely perceptual quality of UHD video application flows traversing a realistic multi-tenanted 5G mobile edge network testbed. The proposed 5G-QoE framework has been implemented in the 5G testbed, and the high accuracy of QoE prediction has been validated through comparing the predicted QoE values with not only subjective testing results but also empirical measurements in the testbed. As such, 5G-QoE would enable a holistic video flow self-optimization system employing the cutting-edge scalable H.265 video encoding to transmit UHD video applications in a QoE-aware manner.

The Guest Editors of this Special Issue would like to thank all the authors who have made this special issue possible and all the reviewers for their dedication and timely review of the submissions that are important for ensuring the quality of the publication. We would also like to express our deepest appreciation to Dr. Yiyan Wu (Editor-in-Chief) for his encouragement and support. Lastly, but not the least, we extend our thanks to Ms. Jennifer Barbato (IEEE BTS publication office) for her capable and tireless support in making our task easy and smooth.
Related Work


Maurizio Murroni (S’99–M’02–SM’13) received the M.Sc. degree in electronic engineering and the Ph.D. degree in electronic engineering and computers from the University of Cagliari, in 1998 and 2001, respectively. He was Graduate Visiting Scholar with the School of Electronic Engineering, Information Technology and Mathematics, University of Surrey, Guildford, U.K., in 1998 and a Visiting Ph.D. Scholar with the Image Processing Group, Polytechnic University, Brooklyn, NY, USA, in 2000. In 2006, he was Visiting Lecturer with the Department of Electronics and Computers, Transilvania University of Brașov, Romania, and a Visiting Professor with the Department of Electronics and Telecommunications, Bilbao Faculty of Engineering, University of the Basque Country (UPV/EHU), Spain, in 2011. He is currently with the Department of Electrical and Electronic Engineering, University of Cagliari, where he is a Coordinator of the research unit of the Italian University Consortium for Telecommunications. Since 2016, he has been a Chair of the IEEE Broadcast Technology Society Italy chapter. He has co-authored of an extensive list of journal articles and peer-reviewed conference papers and received several best paper awards. He served as a chair for various international conferences and workshops. He was co-authored of the 1900.6–2011—IEEE Standard for Spectrum Sensing Interfaces and Data Structures for Dynamic Spectrum Access and other Advanced Radio Communication Systems. His research currently focuses on broadcasting, cognitive radio, signal processing for radio communications, and multimedia data transmission and processing.
Reza Rassool received the B.Sc. degree in physics from King’s College, London, in 1984. He is a CTO with RealNetworks, where in addition to his oversight of corporate-wide technology initiatives, he personally drives both machine learning and codec developments. He evangelizes strategic innovations through technical papers, public speaking, and customer engagements. His entrepreneurial career in digital media technology include many industry firsts. He pioneered “non-linear” editing with OSCAR/EMMY winning Lightworks and brought the world’s first disk-based VOD system to NAB in 1994 while a Chief Engineer of Micropolis. He worked on ground-breaking cochlear implant and bionic eye developments. He was the Founder of Widevine Technologies transforming digital rights management for online television. He has 24 U.S. patents granted, he continues to drive new innovation in next generation video coding, computer vision, and machine learning.

Li Song (M’07–SM’17) received the B.E. and M.S. degrees in engineering from the Nanjing University of Science and Technology, in 1997 and 2000, respectively, and the Ph.D. degree in electrical engineering from Shanghai Jiao Tong University (SJTU) in 2005. He joined SJTU as a Faculty Member, where he is currently a Full Professor with the Department of Electrical Engineering, Santa Clara University, CA, USA, from 2011 to 2012. He has over 100 publications, over 30 issued patents, and 18 standards technical proposals in field of video coding and image processing. His current research interests include video processing, computer vision, and multimedia system. He was a recipient of the National Science and Technology Progress Award in 2015, the Okawa Foundation Research Grant in 2012, the Second Place Award from International Conference on Image Processing Grand Challenge jointed with Twitch in 2017, and several best paper awards from IEEE. He has been serving as an Associate Editor for *Multidimensional Systems and Signal Processing* since 2012. He is a Technical Committee Member of Visual Signal Processing and Communications in IEEE Circuits and Systems Society and has served as the area or session chair for various international conferences and workshops.

Rafael Sotelo (S’90–M’05–SM’13) is member of the AdCom of BTS, the Past Chair of IEEE Cono Sur Council, IEEE Uruguay Section, the Chair of IEEE BTS Uruguay Chapter, and a Full Member of the National Academy of Engineering of Uruguay.

He received the Electrical Engineer degree in 1994, the M.B.A. degree in 1999, and the Ph.D. degree in telematics engineering from the Universidad de Vigo, Spain, in 2010. He is the Director of the ICT Department, Universidad de Montevideo, Uruguay, an Associate Professor with the Universidad de la República, Uruguay, and an Advisor of the Ministry of Industry/National Telecommunications Agency. He has been involved with the broadcasting and cable TV industry in Uruguay since 1991.

Dr. Sotelo’s main research interests are in the area of quality of experience in multimedia applications, recommender systems, IoT, DTV, and IBB.