An overview of developments and standardization activities in immersive media

Dragorad Milovanovic and Dragan Kukolj

Introduction

This article provides the recent and current activities related to the MPEG development and research activities of the emerging immersive media technologies. Especially we outline the need for perceptual tools in MPEG AhG *Immersive Media Quality Evaluation*[1] towards specification the new standard ISO/IEC 23090 *Coded Representation of Immersive Media*. Next, we provide ideas regarding quality assessment of immersive media (VR/AR, LF, 360-video) in QUALINET Task Force *Immersive Media Experiences* (IMEx) and VQEG *Immersive Media Group* (IMG).

Most of these standards activities are currently in early phase. An important aspect, not yet fully addressed is the *Quality of Experience* (QoE) of immersive applications and services.

Up-to-date standardization activities

In June 2016, MPEG started working on MPEG-VR initiative (currently MPEG-I *Collection of standards to digitally represent immersive media*) to develop a roadmap and coordinate the various activities related to VR within MPEG and to liaison also with other SDOs. Other consortia working on innovative products and services in this domain are 3GPP collaboration group of telecommunications associations (TR 26.918 *Virtual*

---

[1] https://lists.aau.at/mailman/listinfo/immersive-quality

Reality media services over 3GPP), DVB promoted a study mission to official group CM-VR (*Commercial Module on Virtual Reality*), and QUALINET (European network on quality of experience in multimedia systems and services) established Task force IMEx (*Immersive Media Experiences*).

The need for perceptual tools and assessment

Recently, MPEG established an *ad hoc* group *Immersive Media Quality Evaluation* with the goal to document requirements for VR QoE, collect test material, study existing methods for QoE assessment, study VR experience metrics and their measurability in VR services, and develop a test methodology.

In order to optimize the standard ISO/IEC 23090 (Part 2 *Immersive video*, Part 5 *Point cloud compression*) for the intended applications, MPEG-I[2] is calling for video test material to assess algorithm performance for different setups where information is combined from different cameras to generate virtual views scene (*Call for immersive visual test material*, April 2017). Different levels of experience are achieved by the user who may freely move his head around three rotational axes 3DoF (yaw, pitch, roll), and along three translational directions 6DoF (left/right, forward/backward, up/down). Test material should comply to the attributes as follow:

- General considerations. Still image and video sequences from both indoor and outdoor scenes can be submitted, with sufficient complexity to test the limits of the algorithms under study - natural content is highly preferred over computer-generated content. Color components, depth, and metadata are provided separately (particular for the camera parameters). Types of cameras and camera array arrangements (highly dense array of images along a

[2]http://mpeg.chiariglione.org/standards/mpeg-i

predefined track - 2D linear with parallel cameras, 2D linear with convergent cameras, 2D cylindrical surface, 2D spherical surface). Accurate temporal synchronization of multiple cameras is preferred.

- Omnidirectional video with depth data. The content should be captured with an arrangement of cameras that records divergent views, preferably in an arrangement that supports the capture of a full 360-degree field of view. Both the texture and depth data must be provided at the same resolution with an input greater than or equal to 4K, and the same projection - preferably in the equirectangular projection.

- Video material recorded by divergent/convergent camera arrangement with significant overlap preferably in an arrangement that supports the capture of a full 360-degree field of view / volume of visual data. Both the intrinsic and extrinsic camera parameters must also be provided.

- 2D camera array arrangement following a planar, cylindrical or spherical surface. Dense video sequences are particularly sought with a baseline distance between cameras not more than 20cm, and the distance from one end of the array to the other end as wide as possible.

- Plenoptic cameras with density of micro-lenses supposed to be large enough to ensure a good angular sampling of the light field. Resolution of the plenoptic image should be no less than 15 mega-rays.

- Systems of simultaneous multiple acquisitions shall simultaneously acquire the same scene following the specifications defined above.

Currently, the ITU-T started a new work program G.QoE-VR on parametric bitstream-based quality assessment (P.NATS promoted to P.1203). In this context Video Quality Experts Group (VQEG) has an *Immersive Media Group* (IMG) with the mission on quality assessment of immersive media,

including virtual reality, augmented reality, stereoscopic 3DTV, and multiview. The initial goals for new established VQEG and QUALINET joint team on immersive media (JQVIM[3]) are collecting and producing open source immersive media content and data set, establishing and recommend best practices and guidelines, collecting and producing open source immersive media tools, and survey of standardization activities.

Dragorad Milovanovic *received the Dipl. Electr. Eng. and M.Sc. degree from Faculty of Electrical Engineering, University of Belgrade, Serbia. He has working as research assistant in DSP, R&D engineer in multimedia communications and ICT lecturer.*

Dr. Dragan Kukolj *is a Professor of computer-based systems with Dept. of Computer Engineering, Faculty of Engineering, University of Novi Sad, Serbia.*

---

[3]https://www3.informatik.uni-wuerzburg.de/qoewiki/qualinet:imex:jqvim