Emerging Research Directions on 3D Video Quality Assessment*

Jaroslaw Bułat, Michal Grega, Lucjan Janowski, Dawid Juszka, Mikolaj Leszczuk, Zdzislaw Papir, and Piotr Romaniak

Department of Telecommunications,
AGH University of Science and Technology

Abstract. Motion picture producers, providers and equipment developers have to deeply consider end user perception of the application being often expressed in terms of a capacious Quality of Experience (QoE) concept. QoE is affected across the whole application delivery chain including content digitisation and compression, its network delivery and reproduction. During recent years enormous research effort and massive tests have been performed in order to identify factors affecting QoE and develop their mapping to scales like Mean Opinion Score for 2D content. Today, the digital video world is on the eve of 3D imaging which is far more complex and sophisticated not only because of the involved technology but also due to the multi–factor nature of the overall 3D experience. This paper discusses the current state of the research on the emerging problem of the user perceived quality of 3D content.

Keywords: Quality of Experience, QoE, 3D Video, subjective tests, quality metrics.

1 Introduction

Nowadays we are experiencing a revolution in the production and delivery of multimedia — the common introduction of the 3D content to cinemas and home television. In the age of strong competition when multiple applications and equipment deliver similar functionality, the key to success is to provide users with the highest possible experience.

The first approach towards assessment of the quality of multimedia applications was based on the Quality of Service (QoS) parameters. However, in recent years it has been realised that the quality of experience depends much on other measurable parameters which arise during acquisition and compression of multimedia content. What is more it has been emphasised that the whole application delivery chain can impact the QoE (Quality of Experience), therefore parameters of the screening environment have to be taken into account as well.

* This publication is an output from a research project "Future Internet Engineering" POIG.01.01.02-00-045/09, part of the "Innovative Economy 2007-2013" programme co–funded by the European Union.

F. Alvarez and C. Costa (Eds.): UCMEDIA 2010, LNICST 60, pp. 69–75, 2012.

The quality assessment approach driven by all the factors influencing what users get and how they perceive imaging content has contributed to countless research problems. These problems encompass objective (numerical) quality metrics, subjective tests (MOS - Mean Opinion Score), a dilemma of full-reference (FR) and no-reference (NR) measurements, correlation and mapping between objective metrics and MOS, statistical tools specific for test data analysis, and even testing credibility and homo/heterogeneity of an end user pool. All that efforts have been raised to a ITU-T recommendation level.

The 3D content QoE research inherits all research problems which had to be previously solved for 2D QoE measurements. Moreover, new 3D-specific factors have to be taken into account, such as depth perception, naturalness and comfort of viewing. All these makes the researchers return to the drawing boards and start development of 3D QoE measurement methodology from scratch.

The rest of the paper is structured as follows. Section 2 presents the current state of the art of the QoE measurements of 3D content. Emerging research directions are described in section 3. The paper is summarised in section 4.

2 Quality of Experience

Any service provided to a customer is evaluated by him/her. Such judgement decides if a customer uses the service ever again. As QoS parameters are not enough to predict a customer experience special metrics and analysis are provided to predict QoE. In order to predict QoE value we have to measure a customer's (called further subject) subjective judgement. The most commonly used way to do so are subjective tests where a group of subjects judges the provided service quality. The VQEG group tests showed that with this methodology a high accuracy can be obtained [1].

The history of creating such precise specification shows that many different test have to be run in order to specify which viewing conditions have strong impact on the observed quality.

In order to properly measure QoE for 3D displays a similar work has to be done. Nevertheless, in 3D content there are more variables which have to be considered. Although all those problems are important, the most significant one which has to be solved is creating methodologies making it possible to compare different displays. Now the differences between different technologies are so significant that only custom metrics (i.e. metrics build for a specific display) can be proposed.

2.1 State of the Art of the 2D QoE Measurement Methods

Image and 2D video quality was extensively analysed over last few decades. The dawn of the quality assessment was dominated by simple statistics–based metrics (e.g. MSE or PSNR) utilised in performance evaluation of image and video compression schemes. Low correlation with the user experience (recent discussion is presented by Wang and Bovik in [2]) enforced necessity of more sophisticated approaches. Several full reference metrics (e.g. SSIM and VQM) utilising the human visual system properties were developed and successfully verified upon subjective results. Certain limitation of full reference approach (i.e. availability of the reference) brought to the market reduced and no reference metrics, capable of the absolute quality assessment. No reference metrics aim at evaluation of a certain artifacts, like noise, blur or blockiness, known prior to the metric design.

The majority of the existing metrics are devoted to the source quality and compression artifacts. Another important aspect of the video delivery chain is addressed by so called quality of delivery approach, trying to assess the perceived quality based on quality of service parameters. Another approach is the bit stream analysis which is currently being investigated by the Joint Effort Group within the Video Quality Experts Group [1].

2.2 Challenges of 3D Imaging

Three dimensional imaging systems try to imitate the human visual system. That is why there are many new challenges in field of perceived quality of experience assessment with respect to those involved when 2D content is analysed. In order to obtain depth perception sensation, a scene is captured from slightly different positions by stereo cameras (either real or virtual) which actually take over the role of the eyes.

There are four main 3D content visualisation criteria that are analysed in literature: image quality, naturalness, viewing experience and depth perception. In terms of the depth preception criterion six important aspects must be considered: binocular disparity and stereopsis[3], accommodation and vergence[4], asymmetrical binocular combination and individual differences (stereoblindness, strabismus, interpupillary distance, age, display duration etc.). Human visual system to construct a perception of depth utilises monoculary available information (or cues) such as accommodation, occlusion, linear and aerial perspective, relative density, and motion parallax. The effectiveness of monocular cues can be easily proven by closing one eye and noticing a considerable appreciation of depth, the binocular cues, stereopsis and vergence, require both eyes to work together [5]. What is more, the accuracy of depth perception is strictly depending on *consistency* of specified cues.

2.3 3D Content Creation, Delivery and Presentation

Nowadays a lot of effort is being made to produce binocular stereopsis by means of delivering different images to each eye separately. Aided stereoscopic display is one of such methods. In general, separation of L and R views is achieved using specialised glasses. The simplest separation is done by colour filtration. This old method is called anaglyph. Unfortunately, the method produces strong color artifacts and is hardly acceptable by the audience. More advanced method is used in Dolby3D systems. It utilises complex multiband filters in order to separate video stream for each eye. This method produces better visual quality but requires expensive glasses and substantially decreases brightness. Light polarisation is another very popular method used in the IMAX cinemas, which is able to divide video streams by means of either linear or circular polarisation of the light. This method, applied in 3D IMAX cinema, requires dedicated projectors as well as non-depolarising screens. The last glasses–based solution employs shutters build from liquid–crystal. It is believed to be the most advanced separation technology. It requires double refresh rate capable screen, usually 120Hz, and is used in modern 3D LCD (Liquid Crystal Display) and Plasma displays as well as in some cinema systems.

Binocular head mounted displays (HMD) is another solution for 3D vision creation. It utilises active LCD or OLED (Organic Light Emitting Diode) displays mounted in front of each eye separately and thus, provides the most accurate channel separation. Since image depends on the head position, it has to be generated on–line. That is why HMD is practically limited to the virtual generated environment such as 3D games, simulations and medical applications.

Autostereoscopic displays are very attractive for the audience as glasses are not required for the 3D effect. This technology makes use of spatial multiplexing of left and right image by means of parallax barrier or microlens solution. Biggest shortcoming of spatial multiplexing is so called "sweet spot". Depending on the relative eye and screen position 3D image appears as depth-correct or incorrect (inverse depth). What is more, sweet spots tend to be narrow and thus, very confusing.

All the methods presented above may suffer from 3D distortions such as: optical cross talk, lack of motion parallax, wrong perspective, unbalanced colour, brightness reduction, insufficient refresh rate, etc. Some of them could be compensated by means of digital signal processing.

Since the most popular method of 3D content creation rely on delivery of two correlated video streams, it is easy to deal with compression by extensions of well known standards. For example, Multiview Video Coding (MVC) is an extension of H.264/AVC [6]. What is more, correlation between streams makes it possible to achieve better compression ratio than for two separate streams. It is also possible to encode 3D image as 2D image plus depth [7]. Such an approach requires reconstruction of 3D positional information using signal processing methods.

3D video content creation is usually done by simultaneous recording from multiple view points. Camera could be either real or virtual. 3D camera could mimic human visual system by means of two lenses separated by the distance of human's eyes (64 mm) or the 3D sensation could be achieved by processing of multiple video streams [8]. It is also possible to add depth to the 2D movies by digital video processing [9].

2.4 QoE of 3D Video

The quality assessment for 3D video applications is a much more complicated task than for 2D. The reason for this is that the overall 3D experience can be described as a combination of both 2D image quality and factors introduced by the 3rd dimension (as discussed in section 2.2). The only aspect covered up to date by the plethora of 2D quality metrics is the image quality. All other aspects of 3D quality are currently new research directions.

The relationship between three 2D video objective metrics (PSNR, SSIM and VQM) and the subjective results are analysed for 3D video content affected with network losses in [10]. Two video standards are considered, namely stereo (left and right images) and colour plus depth. Stereo sequences were screened using PC monitors with shutter glasses, while colour and depth were screened using autostereoscopic displays. Obtained results suggest strong correlation of the objective metrics with the subjective image quality and depth perception.

In [11] the authors present how the quality of 3D video encoded using H.264 as color plus depth is affected by the different bit budgets allocated for color and depth. The 3D quality was estimated using two quality metrics dedicated for 2D video material, namely PSNR and VSSIM (Video SSIM). Obtained results suggests that in order to optimize the overall 3D quality depth should be encoded with 15% - 20% of the available bit rate. The result obtained using objective metrics were verified in small subjective experiment.

A joint source channel coding scheme (JSCC) for color and depth 3D video was proposed in [12]. The proposed coding schema aims at optimization of the overall 3D experience for transmission over loss prone WiMAX network. The results show that the overall 3D experience is dominated by the color component. In consequence, higher bit rate and stronger protection should be allocated for this component. The overall image quality, the depth perception and the overall 3D experience were rated in the subjective experiment. It was also shown that the overall 3D experience can be fairly close approximated by a single 3D perceptual attribute or even by 2D objective quality metric.

In [13] the authors present an effect of the depth compression on the 3D perception, for sequences in colour plus depth standard. For the purpose of the experiment the depth component was compressed using H.264 and different bit rates. Prepared test set was screened using autostereoscopic displays. Subjective results show than the depth component may be significantly compressed while sustaining high 3D quality.

The above consideration contains a clear message that the overall quality metric suited for 3D does not exist. Current efforts tend to make use of existing 2D metrics even without any adaptation. The conclusion is simple — a design of the quality metric dedicated for 3D video content is a big challenge.

3 Emerging Trends

The 3D QoE is a fairly new research area. There are several research projects and organisations which put a lot of effort into development of the 3D quality metrics. This section provides a brief overview of the 3D QoE research initiatives.

3.1 VQEG

The VQEG (Video Quality Experts Group) [1] has recently started to point its attention to 3DTV video quality metrics and models. This activity is related to the ITU-R Question 128/6.

The group at first will try to investigate ways of measuring 3D quality, which will be followed be the development of standard metrics and models. These tasks differ significantly from previous VQEG efforts towards 2D models. It is not straightforward to transfer 2D expertise into the 3D area. For example, apart from classical image quality aspects, metrics for depth map quality, presentation room quality or viewing comfort quality (how long the user can watch 3D) have to be developed. Furthermore, some of the problems that have been already solved for 2D video technology, like blurriness, strike back in the 3D technology (such as crosstalk leading to ghosting images). The currently available quality metrics for stereoscopic images are not enough as they have been based on 2D ground-truth. Nevertheless, current creation of 3D video quality metrics is hampered by a lack of high quality and realistic reference content. In order to gather a reference test-set of 3D videos, a CDVL (The Consumer Digital Video Library) [14] library is currently being extended to accept and provide 3D content as well.

The currently investigated problems that may affect the perceived 3D video quality, include also screen luminance (being not equal to perceived luminance), monitor resolutions problems, viewing distances, depth rendering, depth of focus and naturalness. Other considered aspects take account of analysis of planes of stereoscopic voxels (3D pixels), depth resolution, depth rendering ability, number of planes within the focus range, comfortable viewing zone, angular depth plane interval, and field of view. More information about topics related to 3D quality metrics, being under investigation of VQEG member can be found in [15]. Complementary activities in this area have been initiated in the ICDM (International Committee for Display Metrology) group as well [16].

3.2 "Future Internet Engineering"

In January 2010 a Polish national project "Future Internet Engineering" was inaugurated with a total budget of approx. 10M EUR. The overall goal of the project is to develop and test the infrastructure and services for the future generation Internet. One of tasks of the project is to develop a measurement methodology for the 3D video and services. The research will cover not only the quality of 3D video but also the user perceived quality of 3D environments, such as virtual museums and the QoE of the user interfaces implemented in the 3D environments.

4 Summary

The presented position paper reminds the readership Quality of Experience issues that strongly affect design, development, and deployment of imaging services. As 2D imaging applications in most cases do involve end user perception little is known how QoE point of view will impact 3D imaging applications. The paper points to specific 3D challenges across a whole delivery chain and summarises some ongoing research projects.

References

1. VQEG, The Video Quality Experts Group, http://www.vqeg.org/
2. Wang, Z., Bovik, A.C.: Mean squared error: love it or leave it? - a new look at signal fidelity measures. IEEE Signal Processing Magazine 26(1), 98–117 (2009)
3. Woo, G., Sillanpaa, V.: Absolute stereoscopic threshold as measured by crossed and uncrossed disparities. American Journal of Optometry and Physiological Optics (56), 350–355 (1979)
4. Lambooij, M.T.M., Ijsselsteijn, W.A., Heynderickx, I.: Visual comfort of binocular and 3d displays. In: Visual Discomfort in Stereoscopic Displays: a Review, in Stereoscopic Displays and Virtual Reality Systems XIV. Proceedings of SPIE, vol. 6490, pp. 1–13 (January 2007)
5. Wijnand, A.I., Seuntins, P.J.H., Meesters, L.M.J.: Human factors of 3d displays. In: 3D Videocommunication - Algorithms, Concepts and Real-Time Systems in Human Centred Communication, pp. 219–233. John Wiley and Sons Ltd. (2005)
6. Merkle, P., Smolic, A., Muller, K., Wiegand, T.: Efficient prediction structures for multiview video coding. IEEE Transactions on Circuits and Systems for Video Technology 17(11), 1461–1473 (2007)
7. Smolic, A., Mueller, K., Merkle, P., Kauff, P., Wiegand, T.: An overview of available and emerging 3d video formats and depth enhanced stereo as efficient generic solution. In: Picture Coding Symposium, PCS 2009, pp. 1–4 (May 2009)
8. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision, 2nd edn. Cambridge University Press (2004) ISBN: 0521540518
9. Würmlin, S., Lamboray, E., Staadt, O.G., Gross, M.H.: 3d video recorder. In: PG 2002: Proceedings of the 10th Pacific Conference on Computer Graphics and Applications, p. 325. IEEE Computer Society, Washington, DC (2002)
10. Yasakethu, S.L.P., Hewage, C., Fernando, W., Kondoz, A.: Quality analysis for 3d video using 2d video quality models. IEEE Transactions on Consumer Electronics 54(4), 1969–1976 (2008)
11. Tikanmaki, A., Gotchev, A., Smolic, A., Miller, K.: Quality assessment of 3d video in rate allocation experiments. In: IEEE International Symposium on Consumer Electronics, ISCE 2008, April 14-16, pp. 1–4 (2008)
12. Yasakethu, S.L.P., Fernando, W.A.C., Kamolrat, B., Kondoz, A.: Analyzing perceptual attributes of 3d video. IEEE Transactions on Consumer Electronics 55(2), 864–872 (2009)
13. Leon, G., Kalva, H., Furht, B.: 3d video quality evaluation with depth quality variations. In: 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, May 28-30, pp. 301–304 (2008)
14. CDVL, The Consumer Digital Video Library, http://www.cdvl.org/
15. Chen, W., Fournier, J., Barkowsky, M., Callet, P.L.: New Requirements of Subjective Video Quality Assessment Methodologies for 3DTV. In: Fifth International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, Arizona, U.S.A. (January 2010), http://www.vpqm.org/
16. ICDM, International Committee for Display Metrology, http://icdm-sid.org/