Quality of Experience in 3D Multimedia

Antoine Lavignotte, Christophe Gravier, Julien Subercaze, Jacques Fayolle

Quality of Experience in 3D multimedia

Antoine Lavignotte, Christophe Gravier, Julien Subercaze and Jacques Fayolle
Université de Lyon, F-42023, Saint-Étienne, France;
Université de Saint-Étienne, Jean Monnet, F-42000, Saint-Étienne, France;
Télécom Saint-Étienne, école associée de l'Institut Télécom, F-42000, Saint-Étienne, France;
Laboratoire Télécom Claude Chappe (LT2C), F-42000, Saint-Étienne, France.
{*firstname.lastname*}*@telecom-st-etienne.fr*

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

Abstract: The Quality of Experience (QoE) is a key element in the search for rapid growth of three-dimensional video technology. Research concerning the study of QoE in a 3D environment is on the move. This paper proposes a review of 3D QoE specificities. We will also present the actual findings in this area as well as the projects in process around the world. Thus, any researcher wishing to know more about this domain will be able to start with the foundations necessary to understand the issues related to 3D.

1 INTRODUCTION

A lot of technologies and services which contain HCI (Human Computer Interface) can be deployed due to their ready-to-use technology. Most of them struggle to emerge because they fall short of users' expectations.

The users' perception (QoE) is a key element to take into account during the development of a new technology or service, especially in Multimedia (Jain, 2005).

ITU-T (International Telecommunication Union) defines QoE by: -*"The overall acceptability of an application or service, as perceived subjectively by the end-user*" (12, 2009). That short definition contains two additional notes:
- *Quality of Experience includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc).*
- *Overall acceptability may be influenced by users' expectations and context.*

A more general notion of QoE is proposed by *Alben* to define all the different interaction aspects between a product and a person (Alben, 1996). Unlike Quality of services (QoS), QoE has subjectivity, expressed or bound to a given context.

A technology is currently trying to be adopted by users: the third dimension (3D). More and more TV manufacturer offers this feature in their product range. A rapid adoption will strongly depend on its impact on the user. 3D will have to therefore significantly enhance the user experience in a multimedia broadcast. Taking into account QoE notion during 3D service development will be a crucial step to its durability.

The rest of this paper is organized as follows. Section 2 describes the different approaches to improve the QoE measured by a user during a movie diffusion. In section 3, a review of the existing works on QoE associated with 3D is provided. Section 4 presents the various national or international projects whose work concerns the study of the QoE associated with 3D. Finally, the paper is summarized in section 5.

2 QOE APPROACHES

Several studies have been carried out on QoE. We can actually identify two principal approaches to evaluate QoE: The objective methods and the subjective methods.

2.1 The objective evaluation

The objective evaluation is based on objective metrics associated with the final service. There are two main families of studies: Model-based or Feature-based

evaluations.
The model-based family recreates the human perception and study the impact of artifacts and degradations on a subject.
The Feature-based method consist of studying the impact of video's degradation by signal processing methods.
In this case, the most often used metrics are MSE (Mean Squared Error) and PSNR (Peak Signal to Noise Ratio) (Zepernick, 2007; **?**; **?**). These metrics are widely used due to their ease of computation. However, (Sheikh, 2006; **?**) demonstrate a moderate correlation between PSNR and human perception.
Another type of model-based method consists in focusing on altering effects produced during the content transport. This approach traditionally focuses on QoS parameters (Varga et al., 2006; **?**; **?**). This view ignores the most important part of QoE: the end user.

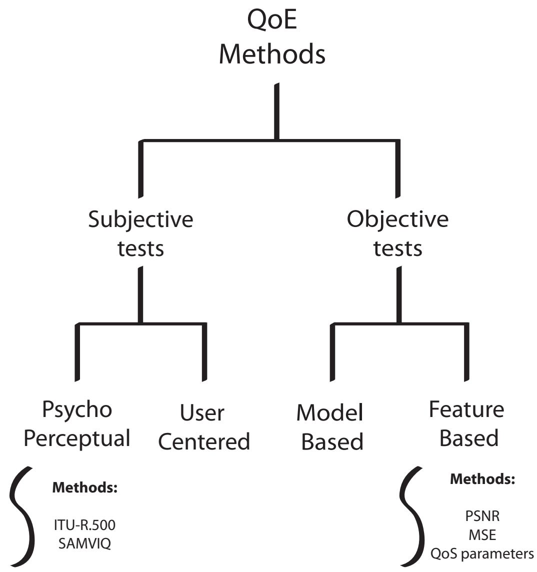
2.2 Subjective evaluation

Subjective evaluation is usually based on tests with a group of users who need to grade different configurations to which they are exposed (Chen et al., 2009; **?**; **?**). In scientific and industrial fields, subjective evaluation is the most direct way to evaluate human perception of a service and allows one to optimize 3DTV systems. Several tests using this method have been established during the last few decades.
This kind of evaluation can be divided into two groups: psycho-perceptual and user-centered approaches.

The psycho-perceptual approach needs a controlled laboratory environment and study the relation between physical stimuli and sensorial experience. Concerning 2D subjective video quality assessment methodologies, the ITU-R BT.500 (ITU-R, 2002) recommendation is widely used for assessing television pictures quality. Another very popular method in 2D is subjective assessment methodology for video quality (SAMVIQ) (Brotherton et al., 2006).

The user-centered approach is quite different due to the perspective chosen to conduct this kind of studies. We need to put ourselves in the user's place rather than that of the system. This is achieved by taking into account typical users, required system characteristics, context of potential usage scenario and goal-related evaluation tasks.

The figure 1 depicts these QoE assessment methods.



QoE methods

A more complete description of the different tests can be found in (Goldmann and Ebrahimi, 2010)

In the following part, we will explain the difference between 2D and 3D insofar as QoE is concerned.

3 FINDINGS IN 3D QOE

One may assume that 3D is just a sum of a 2D technology and the depth added to the pictures. If this were the case, one could simply extends existing 2D QoE assessment methods to take depth into account. Unfortunately, As Goldman states in (Goldmann and Ebrahimi, 2010), it is not so simple.

3.1 The 3D content visualization criteria

The third dimension tries to imitate the human visual system. Unfortunately, each person's 3D perception system is unique. Unlike other technological leaps, such as the transition from black and white to color where everybody was in agreement on the improvement provided by this new technology, for 3D, it is completely different. Some viewers complain about visual discomfort, fatigue, eye strain, problems of visual quality and distortions in 3DTV broadcasting, ...
In order to understand the cause of these problems, we must define the different criteria that play a role in the perceived quality by a user.

As stated in (Bulat et al., 2010), we can find four main criteria in the literature: image quality, naturalness, viewing experience, and depth perception. The first three criteria are traditional: the same than in 2D QoE. However, Depth perception is the new criteria introduced by 3D. It includes several important characteristics that are described in the following paragraph.

3.1.1 Binocular disparity and stereopsis (Woo and Sillanpaa, 1979; ?):

This refers to the difference seen by the left and the right eye for the same object. The brain uses the binocular disparity to extract the depth information from images perceived by the two eyes.

3.1.2 Accomodation and vergeance (Ukai and Howarth, 2008; ?):

The accommodation is the adaptive optical modification that permits a clear image for different vision distances. The vergence allows for the measurement of the focusing properties of the users.

3.1.3 Asymmetrical binocular combination and individual differences:

One must take into account the specificities of 3D perception of each user. The last aspect has a large research filing which contain:

Stereoblindness: *Richards* (Richards, 1970) shows that a part of the population was unable to see in 3D using stereo vision due to the inability to perceive depth information by combining and comparing images from their two eyes. A study by Eyecaretrust, an English health organism, showed that 12% of the english population has a visual impairment and can not enjoy 3D films in all their glory [1].

Strabismus: A part of the population is inflicted with a non alignment of the optical axes of the two eyes. The gaze of each eye is not perfectly aligned on the point in the space and affects the depth perception.

Interpupillary distance: Dodgson pointed out in (Dodgson, 2004) and (BAHN, 2002), the mean interpupillary distance (IPD) is an important and oft-quoted measure in stereoscopic works. He proved that the mean IPD is around 63mm, but a range 40-80mm has to be taken into account to cover the whole population (adults and children).

Age: Age is also an important variable. Norman (Norman et al., 2000) & Hayslip (Hayslip and Panek, 1989) demonstrated that older adults can extract depth and shape from optical patterns containing differential motion or binocular disparities but these abilities are often manifested with lower levels of performance.

Display duration: The test display duration has to be defined. (Tam and Stelmach, 1998) and (Patterson and Fox, 1984) investigated the role of display duration in stereoscopic depth perception. Results confirmed large individual differences in the display duration required for a stereoscopic depth perception.

3.2 Specificity of 3D assessments

As explained in 2.0.2, ITU-R BT.500 (ITU-R, 2002) is a recommendation widely used to assess the quality of 2D videos. The previously exposed singularities of 3D over 2D make that recommendation inappropriate for 3D assessments. *Chen et al* propose an extend of the ITU recommendation greater completeness and to take into account the specificity of 3D which is next presented. In this part, we will present the particularities of 3D assessment exposed by (Chen et al., 2010).

3.2.1 General viewing conditions

We can list several new important conditions for a good experience in 3D:

Problems of luminance and contrast ratio: They appears du to the use of additional optical instruments for 3D viewing (glasses and filters).

Greater importance of the display position: Has to be taken into account compared to 2D tests. The perception of real background and the perceived display depth depend on it.

Room illumination: It has to be specified precisely to avoid conflicts and perturbations with the media broadcasted to the user.

[1] `http://www.eyecaretrust.org.uk/view.php?item_id=566`

Monitor resolution: The monitor resolution is also an important parameter.

Recommendations: Minimum values for spatial and temporal view resolution and stereoscopic resolution must be addressed (Holliman, 2010).

Preferred Viewed Distance: The PVD recommended in BT-500 does not take into account the depth perception. It should certainly be added in the recommendation (Patterson, 2007).

Viewing position: is an important parameter which can cause luminance reduction depending of the viewing angle or motion parallax on multiview autostereoscopic displays.

Depth rendering: Finally, the depth rendering has a significant influence on the QoE for autostereoscopic displays (Barkowsky et al., 2009)

3.2.2 Source signals

Actually, several video formats are available for 3D videos as ”video plus depth” (Fehn, 2003), ”Multi-View video plus depth representation and Coding(MVC)” (Merkle et al., 2007) and ”Layer Depth Video (LDV)” (S. Jolly, 2009), etc... In 2011, the MPEG 3D committee issued a Call for Proposal (CfP) on 3D video coding technology with the objective to “define a data format and associated compression technology to enable the high-quality reconstruction of synthesized views for 3D displays”. Unlike many studies that have focused on the stereoscopic technology, this CfP also contained the auto-stereocopic multi-view display technology. The results are not actually known but a decision will be taken during the year 2012.

3.2.3 Selection of test materials

Main elements of 2D video complexity defined by ITU-T P910 (ITU,) are the Spatial perceptual Information (SI) and the Temporal perceptual Information (TI). In 3D, new elements as Depth perceptual information (DI) should also be added (Chen et al., 2010).

3.2.4 Test methods

A new element is revealed during 3D assessments: the visual discomfort. It is an important point. It can be measured by optometric tests, ERP (event-related potential) (Li et al., 2008), eye tracking or questionnaires.

3.2.5 Observers

The number of 15 participants defined in ITU-BT.500 can be discussed du to the inter-individual differences (Ukai and Howarth, 2008). Moreover, about 10-15% of the population cannot perceive the depth in 3D due to eyes malfunctions. We are therefore not capable of addressing all the population with only one test.

3.2.6 Test duration

During a 3D diffusion, an accommodation time is needed to the viewer. The 10 seconds proposed by the recommendation BT.500 seems too short for good results.

3.3 Collaboration for future research

The domain is currently structured. The different actors work more and more in tandem to advance their research. In this part, we will present important work to propose a base for future research.

3.3.1 A standard for evaluation tests

An European COST Action named QUALINET (European Network on Quality of Experience in Multimedia Systems and Services) has been invited by MPEG for the CfP to participate during the evaluation part of the different 3D video coding algorithms.
The main scientific objective of QUALINET is the development of methodologies for subjective and objective quality metrics by taking into account current and new trends in multimedia communication systems as witnessed by the appearance of new types of contents and interactions.

The ability to compare the different video quality assessment methods is crucial, therefore the development of a standard certification mechanism is critical. To demonstrate the importance of a certification procedure, a group of four institutions (e.g. NTNU, EPFL, UBC and Acreo) conducted a cross-laboratory analysis to estimate the correlation of quality scores obtained by each laboratory for an identical test (Perkis et al., 2012). Results demonstrated that different laboratories employing different subjects can still produce highly correlated results, as they follow similar guidelines to carry out assessments.
That is an interesting basis for future QoE tests that

will be able to be carried out by many laboratories in the same conditions for greater impartiality.

3.3.2 3D Databases

To unify the future research in the domain, EPFL (Goldmann et al., 2010) proposed a comprehensive stereoscopic video database that contains a large variety of scenes and different captured parameters. The database also provides subjective quality scores. It can be used to evaluate the performance of visual quality metrics as well as for the comparison and design of new metrics. More information can be found at `http://mmspg.epfl.ch/3dvqa`.

4 Current projects & reflexions for QoE 3D

Since 2010, national and international projects on 3D have been launched around the world. The QoE is a key element in the production chain of a 3D movie. So, it is often an important part of these projects. In this chapter, we will scan all the past and actual projects over the world that address the QoE notion in 3D.

4.1 3D Live

This french project was composed of 9 academic and industrial partners namely Orange, Institut Telecom, Technicolor, INRIA, AMP, Binocle 3D, Thales Angenieux, Thomson and Grass Valley. The project covered all the 3D diffusion chain from coding to the quality perceived by the user. Their main objective was to improve the user experience by working on all the 3D creation chain. Their research covered: - specific equipments as cameras or optics
- software development for disparity correction, content adaptation (upscaling, downscaling) and data protection
- metadata management (depth, camera type)
- text and graphical insertions without conflicts in the scene depth (subtitles, logos)
The program finished on march 2012. Informations and publications can be found at `www.3dlive-project.com`.

4.2 Jedi 3D

The project Jedi 3D is a consortium of 21 partners in 3 countries (Belgium, France and Spain). It aims to study and develop a high quality end to end 3DTV for consumer, with a user centric approach, with stereoscopic as basis, toward multi-view. Two work-packages are interesting for the QoE community: WP3 ”3D User experience” and WP4: ”Demonstrations & User assessment” which are led by the societies NXP and Barco. More information can be found at `http://www.jedi-itea2.org/`

4.3 Future Internet Engineering

The ”Future internet Engineering” is a Polish national project financed by the European Union. The project covers the development and testing of infrastructure and services for future Internet generation. One part of the project consist of developing a measurement methodology for 3D video and services. Several interesting publications has been produced on QoE (Stankiewicz and Jajszczyk, 2011; **?**). The project just ended on december 2012.

4.4 3D4YOU

3D4YOU is an European project funded by ICT (Information and Communication Technologies), a thematic priority for research and development under the specific program ”Cooperation” of the Seventh Framework Program 2007-2013. Industrial and academic partners that work on multimedia joined together on this project (Philips NL, BBC UK, Fraunhofer Heinrich Hertz Institut DE, Thomson R&D FR, Orange France Telecom FR, Christian Albrechts Universitat DE and KUK Filmproduktion DE). Their objective is to develop the key elements of a practical 3D television system. A first approach of the content requirements for subjective testing is presented in (3D4YOU, 2008) A second document (3D4YOU, 2011), much more complete, has been produced in march 2011 to list the consortium proposals to achieve subjectives quality tests for 3D multimedia.

4.5 VQEG & 3DTV

VQEG (Video Quality Experts Group) is an expert group born in 1997 to deal with the field of video quality assessment by investigating new and advanced subjective and objective techniques (subjective assessment methods and objective quality metrics). Several projects are actually in progress. One of them is called 3DTV. It’s goal is to investigate how to assess 3DTV subjective video quality. This activity is linked to the ITU-R Question 128/6. A first draft is actually available [1]. The ITU’s schedule provides a project completion in 2015.

[1] `http://www.its.bldrdoc.gov/vqeg/project-pages/3dtv/3dtv.aspx`

4.6 QUALINET

The European COST named Qualinet is an expert group which want to create a network for multidisciplinary QoE research in Europe: Their objectives are:
- the development of methodologies to assess the perceived QoE for multimedia presentations,
- the identification of measurable features which are relevant for the subjectively perceived QoE,
- the development of metrics to measure the QoE perceived by a human observer,
- the development of tools to optimize existing systems,
- and the design of new interaction models between users and content multimedia.

It is actually composed of 24 COST countries and 6 non-european. Every year, they organize workshops on QoE as QoMEX (Quality of Multimedia Experience). A white paper written on QoE's definition is available at `http://www.qualinet.eu/images/stories/whitepaper_v1.1_dagstuhl_output_corrected.pdf`. More details can be found at `http://www.qualinet.eu/`

5 Conclusion

The presented paper permits to discover the 3D QoE basics. It explains the new characteristics of 3D assessments compared to 2D. We also introduce national, european and international projects. We hope these elements will enable each researcher to better understand the concept and the importance of QoE in 3D multimedia.

REFERENCES

12, I.-T. S. G. (2009). Itu-t rec. p.10/g.100 amendment 2 (07/2008) vocabulary for performance and quality of service amendment 2: New definitions for inclusion in recommendation itu-t p.10/g.100. pages 1–10.

3D4YOU (2008). Requirements on post-production and formats conversion. *ICT- 215075, Deliverable D2.1.2*.

3D4YOU (2011). 3D video formats and conversion. *ICT-215075, Deliverable 4.2.2*.

Alben, L. (1996). Defining the criteria for effective interaction design. *Interactions*.

BAHN, J. K. (2002). Effects of Interpupillary Distance and AC/A Ratio on Binocular Fusion and Depth Perception. *Proc Int Disp Workshops*.

Barkowsky, M., Cousseau, R., and Le Callet, P. (2009). Influence of depth rendering on the quality of experience for an autostereoscopic display. In *2009 International Workshop on Quality of Multimedia Experience*, pages 192–197. IEEE.

Brotherton, M. D., Huynh-thu, Q., and Hands, D. S. (2006). Subjective Multimedia Quality Assessment. *IEEE Signal Processing Magazine*, E89-A(November):2920–2932.

Bulat, J., Grega, M., Janowski, L., Juszka, D., Leszczuk, M., Papir, Z., and Romaniak, P. (2010). Emerging Research Directions on 3D Video Quality Assessment. *UCMedia 2010 2nd International ICST Conference on User Centric Media Palma de Mallorca September 13 2010*.

Chen, K., Wu, C., Chang, Y., and Lei, C. (2009). A crowdsourceable qoe evaluation framework for multimedia content. *Proceedings of the seventeen ACM international conference on Multimedia*, pages 491–500.

Chen, W., Fournier, J., Barkowsky, M., and Le Callet, P. (2010). New requirements of subjective video quality assessment methodologies for 3DTV. In *Proceedings of VPQM*, volume 2010.

Dodgson, N. A. (2004). Variation and extrema of human interpupillary distance. *Proceedings of SPIE*, 5291(January):36–46.

Fehn, C. (2003). A 3D-TV system based on video plus depth information. *The ThritySeventh Asilomar Conference on Signals Systems Computers 2003*, 2:1529–1533.

Goldmann, L., De Simone, F., and Ebrahimi, T. (2010). A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video. *Database*, 7526(1):75260S–75260S–11.

Goldmann, L. and Ebrahimi, T. (2010). 3d quality is more than just the sum of 2d and depth. *Signal Processing*, pages 2–3.

Hayslip, B. and Panek, P. (1989). Adult Development and Aging. *Harper and Row*.

Holliman, N. S. (2010). 3D display systems. In *Science*, volume 38, pages 31–36. IOP Press.

ITU. "subjective video quality assessment methods for multimedia applications" in recommendation p.910.

ITU-R (2002). Methodology for the Subjective Assessment of the Quality of Television Pictures.

Jain, R. (2005). Quality of experience. *Multimedia, IEEE*, 11(1):96.

Li, H.-c. O., Seo, J., Kham, K., and Lee, S. (2008). Measurement of 3D visual fatigue using event-related potential (ERP): 3D oddball paradigm, Kwangwoon University , 447-1 Nowon-Gu , Wolgae-Dong , Seoul , Korea Kangwon National University , 192-1 Hoja-Dong , Chuncheon , Korea. *Measurement*, pages 213–216.

Merkle, P., Smolic, A., and Muller, K. (2007). Efficient Prediction Structures for Multiview Video Coding. *IEEE Transactions on Circuits and Systems*, 17(11):1461–1473.

Norman, J. F., Dawson, T. E., and Butler, A. K. (2000). The effects of age upon the perception of depth and 3-D shape from differential motion and binocular disparity. *Perception*, 29(11):1335–1359.

Patterson, R. (2007). Human factors of 3-D displays. *Journal of the Society for Information Display*, 15(11):861.

Patterson, R. and Fox, R. (1984). The effect of testing method on stereoanomaly. *Vision Research*, 24(5):403–408.

Perkis, A., You, J., Xing, L., Ebrahimi, T., De Simone, F., Rerabek, M., Nasiopoulos, P., Mai, Z., Pourazad, M., Brunnstrom, K., Wang, K., and Andren, B. (2012). Towards certification of 3D video quality assessment. In *Proceedings of the 6th International Workshop on Video Processing and Quality Metrics for Consumer Electronics - VPQM 2012*.

Richards, W. (1970). Stereopsis and stereoblindness. *Experimental Brain Research*, 10(4):380–388.

S. Jolly, e. a. (2009). 3D Content Requirements and Initial Acquisition Work. *3D4YOU project*.

Sheikh, H. (2006). Image information and visual quality. *Image Processing*.

Stankiewicz, R. and Jajszczyk, A. (2011). A survey of QoE assurance in converged networks. *Computer Networks*, 55(7):1459–1473.

Tam, W. J. and Stelmach, L. B. (1998). Display duration and stereoscopic depth discrimination. Technical Report 1, Communications Research Center, Ottawa, Ont., Canada.

Ukai, K. and Howarth, P. (2008). Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations. *Displays*, 29(2):106–116.

Varga, P., Kn, G., Sey, G., Moldovn, I., and Gelencsr, P. (2006). Correlating user perception and measurable network properties: Experimenting with qoe. 4268:218–221.

Woo, G. C. and Sillanpaa, V. (1979). Absolute stereoscopic thresholds as measured by crossed and uncrossed disparities. *American journal of optometry and physiological optics*, 56(6):350–355.

Zepernick, H. (2007). Perceptual-based quality metrics for image and video services: A survey. *Next Generation Internet Networks*.