CROSS-LAB STUDY ON PREFERENCE OF EXPERIENCE IN 3DTV: INFLUENCE FROM DISPLAY TECHNOLOGY AND TEST ENVIRONMENT

Jing Li[a], *Ondřej Kaller*[b], *Francesca De Simone*[c], *Jussi Hakala*[d], *Dawid Juszka*[e], *Patrick Le Callet*[a]

[a]LUNAM Université, Université de Nantes, IRCCyN UMR CNRS 6597, Polytech Nantes.
rue Christian Pauc BP 50609 44306 Nantes Cedex 3, France
[b]Dept. of Radio Electronics, Brno University of Technology, CZ-61200 Brno, Czech Republic
[c]Telecom ParisTech, Paris, France. [d]Aalto University School of Science, Finland
[e]AGH-University of Science and Technology, Cracow, Poland

ABSTRACT

Quality of Experience (QoE) in 3DTV may be influenced by many factors, such as the viewing conditions, the characteristics of the panel of viewers, etc. In this study, the impact of two factors, namely the test environment and the display technology are analyzed. Particularly, two paired comparison experiments have been carried out to assess the overall preference of experience (PoE) for a set of 3D video sequences. The first experiment, performed at University of Nantes, allows comparing the results obtained in a standard laboratory versus living-room-like environment. No significance difference has been found. The second experiment, performed at Brno University, allows comparing the results obtained using stereoscopic display with shutter glasses versus polarized glasses. Significance difference could be identified depending on the video content. Since the same test material has been used in both locations, inter-laboratory correlation is also analyzed.

Index Terms— Pair comparison, 3DTV, Preference of Experience, display technology, test environment

1. INTRODUCTION

How dependent are the results of a subjective quality assessment test on the particular viewing conditions used? Many studies have tried to answer this question when considering quality assessment of 2D multimedia content. For example, in [1], the results obtained for 2D audio-visual quality assessment in standard laboratory versus public environment have been compared. The impact of the environment resulted to be minimal when a wide range of quality is considered. When considering 3D content, instead, few studies on this topic are available in literature. In [2], the effect of the display technology on subjective quality has been studied. The results show that different technologies lead to differences in terms of image quality, perceived depth, and visual discomfort. Recently, a cross-lab study has been performed which analyses, among others, the effect of display technology (passive polarized versus active shutter glasses) and viewing conditions (viewing distance and number of parallel viewers). The results show that slight differences occur depending on the video degradation and the viewing distance.

In this paper, the impact of two factors, namely the test environment (standard laboratory versus living-room-like environment) and the display technology (passive polarized versus active shutter glasses), are analysed for 3DTV. As a major difference with respect to [2] and [3], the Paired Comparison (PC) methodology has been used. The outcome of subjective assessment by PC is referred to as Preference of Experience (PoE). The choice of the PC method is motivated by the fact that Quality of Experience (QoE) in 3DTV can be considered as a "multidimensional" subjective impression, resulting from "monodimensional" factors, such as image quality, depth quantity and visual comfort. While traditional 2D subjective quality assessment methods, such as Absolute Category Rating (ACR), are recommended by ITU-R BT.2021 to quantify the "monodimensional" terms of 3D QoE, it is not clear whether they are applicable to quantify "multidimensional" quality. Furthermore, the use of PC avoids the problems of language dependency and cross-lab score alignment that occur for example with ACR [3].

2. EXPERIMENTAL DESIGN

2.1. Test video sequences

The test sequences are from NAMA3DS1-COSPAD1 database [3]. Four stereoscopic Full HD (1920×1080) video sequences, i.e., *Soccer*, *Basket*, *Lab*, *Tree branches* were used as SRCs. Three degradation types are selected: Still image coding (JPEK2000) with bitrate 8Mb/s; Video coding (H264) with QP 32; Down sampling by a factor of 4 and then sharpening through edge enhancement. They covered different degradation types in this database and their perceptual image qualities are distributed in the middle range of MOS (around 3 or less) and not significantly different. In this way, the PC method is capable of evaluating the degradations where the artefacts are difficult to identify and their influences on QoE.

2.2. Experimental setup

The experiments in this study are conducted in two labs. One is the Image and Video Communication (IVC) lab in University of

This work has been financially supported by Qualinet COST IC 1003, the PERSEE project which is financed by ANR (project reference: ANR-09-BLAN-0170), and also supported by Project number P102/10/1320, national project no. LD12005 (Czech) and the SIX project; the registration number CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation.

Contact: jing.li2@univ-nantes.fr, xkalle00@stud.feec.vutbr.cz

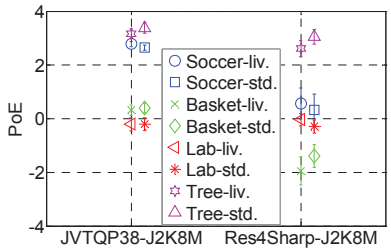
Table I. Experimental Setup

Laboratory	IRCCyN/IVC, Nantes		UREL, Brno	
Experiment	Experiment 1	Experiment 2	Experiment 1	Experiment 2
Display	Philips 46PFL9705H (1920×1080)		LG 32LW570S (1920×1080)	Panasonic TX-P42GTT20E (1920×1080)
Technology	Shutter glasses		Circular polarization	Shutter glasses
Viewing environment	Living room	Standard room	Standard room	
Viewing distance	3H (1.72m)		3H (1.2m)	3H (1.57m)
Gender (m/f); Age (mean)	15/17; 19-29 (22.9)		42/0; 22-25(23.7)	40/0; 23-27(23.7)

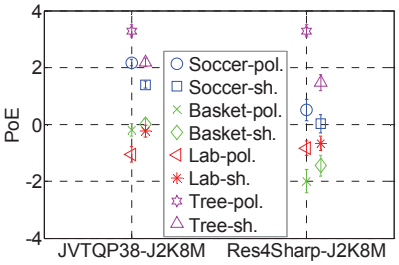
Nantes. The other is the UREL lab in Brno University of Technology. The details are shown in Table I. The pair comparison method is used and the question for each observer is "Which one do you prefer?" Comparisons are conducted only on the same SRC, both presentation orders are considered in the test (A-B, B-A). Thus, there are in total 4×3×2 = 24 pairs for each observer. All the pairs are randomly displayed for each observer.

3. RESULTS

The Bradley-Terry model is used to convert the PC data to a scale value (PoE) for each stimulus. HRC J2K8M is set as reference. The PoE of each processed video sequence (PVS) compared to J2K8M is shown in Fig. 1. The Pearson Correlation Coefficients (PLCC) of the PoEs between the labs are shown in Table II.



(a) IVC results: comparison of different test environments (liv: living room; std: standard room)



(b) UREL results: comparison of different display technology (pol: polarize display; sh: shutter glasses)

Fig. 1. PoE value of each PVS to J2K8M with 95% confidence intervals.

Table II.
Table III. PLCC matrix of PoEs of four experimental results

Exp.	IVC-1	IVC-2	UREL-1	UREL-2
IVC-1	-	0.9821	0.9367	0.9403
IVC-2	0.9821	-	0.9508	0.9262
UREL-1	0.9367	0.9508	-	0.9354
UREL-2	0.9403	0.9262	0.9354	-

The results of IVC1 and IVC2 correlate very well. There is no significant difference of PoE in different test environments as shown in Fig. 1(a). However, the performance of display technology is different due to the video contents (Fig.1 (b)). In *Tree* and *Soccer* sequences, the JVTQP38 and Res4Sharp showed significantly higher PoEs in polarized display than shutter glasses. However, in *Basket* and *Lab* sequences, the differences between the two display technologies are less significant. *Tree* and *Soccer* contain large area of high spatial frequency components, e.g., tree leaves and football net, the ringing and blocking artifacts are quite annoying through shutter glasses, however, these artifacts are less visible in polarized display due to the halved resolution, and thus, the discrimination of video quality is much more dominant than in shutter glasses.

The N-way ANOVA test is applied on the PoE scores of IVC and UREL, respectively. Both labs' results show that SRC, HRC and their interactions are significant influence factors on PoE (p-value <0.05). In UREL, the display technology doesn't show any significant influence (p-value = 0.12), but the interactions between SRC and display technology show significant influence with p-value = 0.04. Another very important finding is that in IVC, test environment is not a significant factor where p-value = 0.79 under the condition that the living room's lighting is changing with the sunlight. When comparing the results between IVC Experiment 2 and UREL Experiment 2, the PoEs in IVC are significantly higher than in UREL (p-value = 0.04) which might be explained by the small screen size, number of observers and biased gender distribution of observers in UREL.

4. CONCLUSIONS

In this study, the paired comparison method is used for reliable measurement of QoE. The results indicate that test environment is not a significant influence factor in 3D PoE which is an extension of the conclusions from [1]. The performance of display is affected by video content. Screen size, observers are possible influence factors on 3D PoE.

5. REFERENCES

[1] M. Pinson, L. Janowski, et al. "The Influence of Subjects and Environment on Audiovisual Subjective Tests: An International Study." *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 6, pp.640-651, 2012.

[2] M. Slanina, T. Kratochvil, et al. "Testing QoE in Different 3D HDTV Technologies." *Radioengineering*, vol. 21, no. 1, pp.445-454, 2012.

[3] M. Barkowsky J. Li, et al. "Towards standardized 3DTV QoE assessment: Cross-lab study on display technology and viewing environment parameters," *Proc. SPIE 8648, Stereoscopic Displays and Applications XXIV*, 2013.

[4] M. Urvoy, M. Barkowsky, et al. " NAMA3DS1-COSPAD1: Subjective video quality assessment database on coding conditions introducing freely available high quality 3D stereoscopic sequences." *Fourth International Workshop on Quality of Multimedia Experience (QoMEX),* IEEE. 2012.