JUXTAPOSITION BETWEEN COMPRESSION AND DEPTH FOR STEREOSCOPIC IMAGE QUALITY ON PORTABLE AUTO-STEREOSCOPIC DISPLAY

Satu Jumisko-Pyykkö[1], *Tomi Haustola*[1], *Atanas Boev*[2], *Atanas Gotchev*[2]

[1]Human-Centered Technology, [2]Department of Signal Processing,
Tampere University of Technology, Finland

ABSTRACT

Presence of depth and compression artifacts have been identified as major factors that influence subjective quality of experience and satisfaction of 3D imagery visualized on a portable auto-stereoscopic display. The juxtaposition between these two factors has been investigated through psychometric subjective experiment, where participants have evaluated the overall quality of stereoscopic images with varying both presences of depth and compression levels. The results reveal that compression artifacts dominate the quality of experience compared to varying depth range.

Index Terms— Mobile 3D, image quality, subjective quality

1. INTRODUCTION

Quality is a key issue for the whole delivery chain of mobile 3D television. From content creation through compression, transmission to display, quality has to be properly maintained in order to ensure a pleasant visual experience for the users. The user preference toward 3D (stereoscopic) presentation mode against 2D (monoscopic) presentation mode has been confirmed in a series of recent studies [1], [2]. However, it has been put also in strong relation with the level of visible artifacts. Namely, the added value of depth has been favored only if the level of artifacts was low [1], [2]. Source of such artifacts are the 3D video processing stages, which have to cope with restricted computational power of the terminal mobile device and the limited bandwidth of error-prone wireless transmission channels. Thus, they introduce artifacts affecting low-level visual characteristics such as quality of depth, space and motion [3]. Furthermore, auto-stereoscopic displays, the displays of choice for mobile 3DTV systems [4], [5], offer small spatial resolution and limited depth comfort zone, determined by their size and optical characteristics [6]. All these factors can influence not only low-level perceptual experiences such as visibility of depth and error-freeness but also high-level aspects of viewing experience such as ease of viewing, visual discomfort and level of immersion [7].

The recent studies have examined relation between depth and compression artifacts for video quality using multidimensional measures. The issue of visual discomfort [8] for the case of stereoscopic video on portable 3D displays has been studied by conducting five extensive experiments with different display technologies, variable depth levels and system parameters [9]. The study has demonstrated the superiority of a lenticular display technology, causing no negative symptoms for short viewing time of less than 30 min. The visual discomfort for combinations of high motion and stereo has been also studied in mobile gaming context [10]. More recently, the influence of depth and compression artifacts on quality of experience of mobile stereoscopic video has been addressed [11]. Along with psychometric tests, the study also employed a descriptive model of 3D quality of experience to draw deeper understanding on 3D video quality characteristics [11]. In a continuation of this study, here we juxtapose the compression artifacts and depth levels perceivable on a mobile auto-stereoscopic display, however in the case of stereo still images. Thus, we aim at quantifying the effect of varying stereo and blocking artifacts on the perceived quality in the case of no motion. This requires new tests independent from the previous ones on videos. The study should be also instructive for quantifying the influence of compression artifacts in different depth levels toward designing perceptually-driven objective metrics for the mobile 3D video quality.

The paper is organized as follows: The research method in terms of test procedure, participants, test stimuli and viewing conditions is described in Section 2. Section 3 summarizes the main results, while Section 4 underlines the discussion points and conclusions rising from the study.

2. RESEARCH METHOD

Participants - The test included a total of 32 participants equally stratified by gender and age group (18-45 years). The majority of the participants (80%) were categorized as naïve evaluators (defined as having little or no prior experience of quality evaluation experiments, they were not experts in technical implementation and they were not studying, working or otherwise engaged in information technology or multimedia processing [12]. All participants had normal or corrected-to-normal vision.

Test Procedure - The test procedure was divided into three phases. In the pre-test session, sensorial tests (Landolt

This work was supported by EC within FP7 (Grant 216503 with the acronym MOBILE3DTV).

chart 20/40; Ishihara test, Randot stereo test .5 arcsec), an pre-immersive measurement of Simulator Sickness Questionnaire (SSQ), and a combined training and anchoring took a place. SSQ is composed of 16 physical symptoms rated on a categorical labeled scale (none, slight, moderate, severe). The symptoms are contributing to groups of 1) nausea (e.g. stomach awareness), 2) oculomotor (e.g. eyestrain), and 3) disorientation (e.g. dizziness) [13].

At the beginning of actual test part, a combined anchoring and training took place. The goal was to familiarize the participants to the contents used in the study, the extremes of quality range of stimuli, and evaluation task. During the experiment, the stimuli were presented one by one, rated independently and retrospectively [12]. After each stimulus still-image, the overall quality was rated using 11-point unlabelled scale and acceptance of quality on a nominal (yes/no) scale [12], [14]. The duration of the stimuli was seven seconds followed by answer time of five second when mid-grey still image was presented on a screen on a zero disparity level.

In the post-test session, the post-immersive SSQ survey was collected after viewing of approximately 30 minutes excluding the answer time.

Stimuli –content, parameters, and preparations - The independent variables of the experiment were: Depth levels (3) x Quantization parameters (5) x Contents (4). Four test contents with variable visual spatial and depth characteristics were used in the experiment (Figure 1). Three depth levels and five quantization parameters were varied. The depth levels contained mono presentation, stereoscopic short and wide baselines. The values of varied quantization parameters, QP, were: 25, 30, 35, 40 and 45 representing a range from low to high quality. The goal of the selection of these parameters was to tackle the juxtaposition between the positive influence of depth and negative influence of artifacts on experienced quality [1].

Figure 1 Four still-image stimuli contents were used in the experiment

The source content was selected from the MPEG multi-view video datasets [15]. For each video content, the second pair of left and right frames was extracted as still image stereo pair. Varying depth levels of the same 3D scenes were achieved by selecting stereo pairs from cameras with different baselines. The following cases of varying depth were manipulated: 1) *Monoscopic* (2D), image, where left and right views were the same; 2) *Short baseline* – a camera baseline producing 3D scene with a limited disparity range creating mild yet visible 3D effect; 3) *Wide baseline* – camera baseline, corresponding to the optimal disparity range for the stereoscopic display in use. With the above notations we addressed two scenarios: the case of short baseline corresponds to stereo images repurposed for mobile use from higher-resolution sources. The case of wide baseline corresponds to content specifically optimized to fit the comfort zone of the portable display [16].

Each stereoscopic sequence was converted from its original resolution to the resolution of the target display by using a four-step procedure of disparity range analysis, image cropping in high-resolution domain, down-scaling and again cropping in low-resolution domain [11]. The procedure ensured an equalization of the absolute positive and negative disparity values, avoidance of frame violation and eventually fitting within the target display comfort zone and its aspect ratio with no geometrical distortions. Cropping was implemented by cubic spline interpolation while resizing was implemented by least-squares cubic spline projection [17].

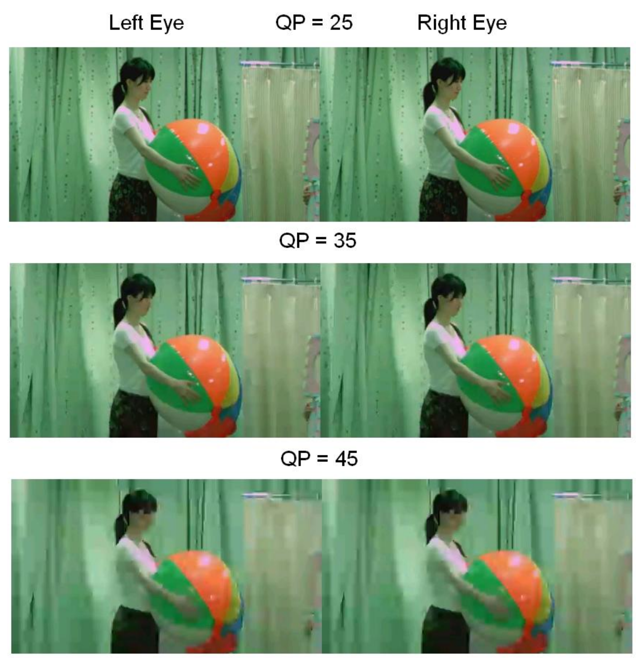

Figure 2 Example of the content Akkokayo with three quantization parameter (QP) values.

After downscaling, each test stereo image was compressed by H.264 reference encoder in intra-frame mode, applied independently to the left and right channels. Five values of the quantization parameter (QP), namely QP=[25, 30, 35, 40, 45] were selected leading to the presence of varying levels of blocky compression artifacts (Figure 2).

Viewing conditions and presentation of the stimuli - The experiments were conducted in the controlled laboratory conditions [14] using a portable autostereoscopic LCD display. The display, produced by NEC LCD has physical size of 3,5'' and resolution of 427x240px at 155DPI and utilizes horizontal double density pixel (HDDP) arrangement. Due to the HDDP arrangement, the display has the same resolution in 2D and 3D mode and has a low crosstalk in 3D mode as well [6]. The screen was connected to an external laptop (Asus G51J) used to store and playback the stimuli with player (MS Media Player 12). During the experiment the display used was located on a stable stand and the viewing distance to the viewer was 40 cm. All stimuli were presented twice in pseudo random-order during the quality evaluation task 0 .

3. RESULTS

Acceptance – In overall, quality was experienced as acceptable on the two lowest QP levels (25, 30) in all depth levels (Cochran's Q(14)=1516, p<.001; pair-wise comparisons McNemar p<.001 Figure 3). When QP was 35 or higher, quality reached the 50% threshold between acceptable and unacceptable quality or was experienced as unacceptable.

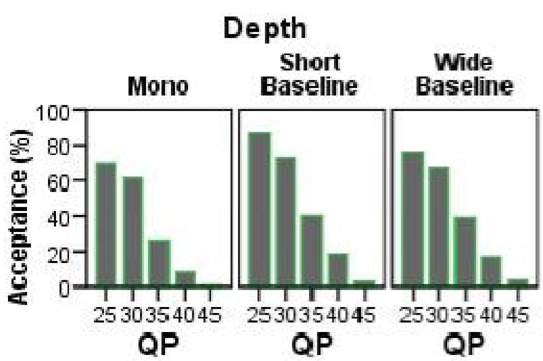

Figure 3 Influence of quantization parameter and depth levels on acceptance of quality across the contents.

Quality satisfaction – Both QP and depth showed the main effects but no interaction on satisfaction of quality when averaged across the contents (Figure 4). Quantization has the most significant influence on quality satisfaction (Repeated measures ANOVA - QP: $F_{2,107}$ = 605.208, p<.001, ŋ=.908). Perceived quality decreases when level of quantization increases and the differences are significant between all studied quantization levels (post-hoc comparisons p<.001). Quality satisfaction is also influenced by the depth level where the most satisfying quality is presented using stereoscopic presentation mode ($F_{1,78}$ = 15.759, p<.001, ŋ=.205; pairwise-comparisons: p<.05). Finally, the results do not show interaction between these two variables ($F_{5,308}$ = 1.594, p>.05, ŋ=.025, *ns*). Similar trend is shown in the content by content analysis.

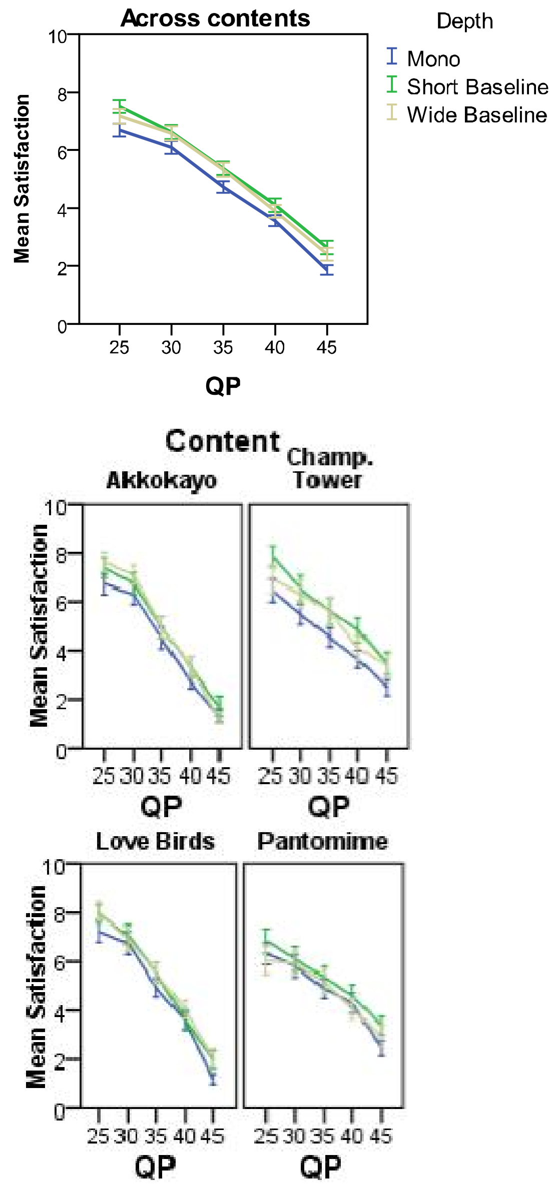

Figure 4 Influence of quantization parameter (QP) and level of depth quality satisfaction across the contents and content by content. The bars show 95% CI of mean.

Simulator sickness - In the analysis, the total SSQ scores of the four groups are calculated by summing the ratings of related symptoms in each group [13]. Each sum is further multiplied by a weighting score (nausea = 9.54, oculomotor = 7.58, disorientation = 13.92 and total score = 3.74) defined by varimax factor weights of SSQ model [13]. The analysis is based on absolute values. The comparisons between pre- and post-immersive simulator sickness symptoms show small increase in three of these symptoms (Figure 5).

The increase is significant in the categories called total (Wilcoxon: Z=-3.072, p<.01), oculomotor (Z=-3.109, p<.01), and disorientation (Z=-3.063, p<.01) while for nausea the difference is non-significant (Z=-0.206, p=-84, ns). Although these results show small increase in post-immersive symptoms, the maximum intensity of these symptoms (mean below 20) is in line with our previous studies with the same display conducted with variable visual video stimuli [9].

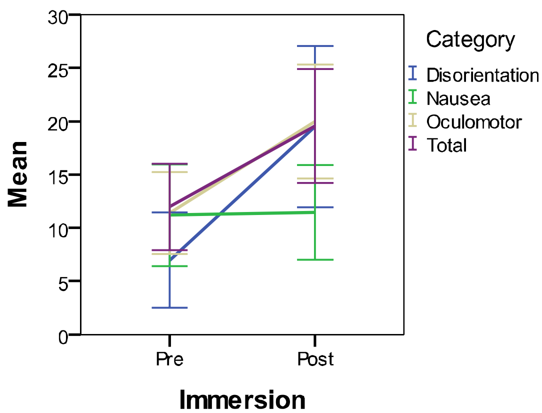

Figure 5 Simulator sickness – pre- and post-immersive influence of quality evaluation task. The bars show 95% CI of mean.

4. DISCUSSION AND CONCLUSIONS

In this work, we aimed at studying how the subjective quality of experience of stereo images displayed on portable auto-stereoscopic display is influenced by varying both the compression artifacts and the level of depth presence in the case of no motion. The results are in accordance with the results obtained in a similar test but for 3D moving scenes (video) of the same content with comparable sample of participants [11]. This is to show that in typical 3D scenery, the most dominating factor for the acceptance and satisfaction of stereo imagery is the compression quality. Both in the cases of video and still images, the presence of stereo was ranked superior against mono for the case of low level of compression artifacts. The presence or lack of moderate motion (videos versus still images) made no difference in the quality evaluation. This result is instructive for the design of objective metrics for video quality evaluation, where different attributes in the scenery (e.g. motion, structure, color, stereo) and different artifacts (e.g. compression) have to be properly weighted.

REFERENCES

[1] D. Strohmeier, S. Jumisko-Pyykkö, and K. Kunze, "Open profiling of quality – A mixed method approach to understand multimodal quality perception," *Advances in Multimedia*, 28 pages, 2010.

[2] D. Strohmeier, and G, Tech, "Sharp, bright, three-dimensional: open profiling of quality for mobile 3DTV coding methods," *Proceedings of SPIE Multimedia on Mobile Devices*, 2010.

[3] A. Boev, D. Hollosi, A. Gotchev, and Egiazarian, K., "Classification and simulation of stereoscopic artifacts in mobile 3DTV content," *Electronic Imaging Symposium Stereoscopic Displays and Applications*, 2009.

[4] A. Gotchev, A. Smolic, S. Jumisko-Pyykkö, D. Strohmeier, G.B. Akar, P. Merkle, and N. Daskalov, "Mobile 3D television: Development of core technological elements and user-centered evaluation methods toward an optimized system," *Proceedings of SPIE 7256 3D Video Delivery for Mobile Devices*, 2009.

[5] N. Hur, J. Kim, "A backward-compatible, mobile, personalized 3DTV broadcasting system based on T-DMB", *in Three-Dimensional Television, Capture, Transmission and Display* (Onural and Ozatkas, eds.), Springer, 2008.

[6] A. Boev, and A. Gotchev, "Comparative Study of Autostereoscopic Displays for Mobile Devices," *Proceedings of SPIE 7881A*, 2011.

[7] S. Jumisko-Pyykkö, D. Strohmeier, T. Utriainen, and K. Kunze, "Descriptive quality of experience for mobile 3D television," *Proceedings of NordiCHI*, pp. 1–10, 2010.

[8] M. Lambooij, W. IJsselsteijn, M. Fortuin, and I. Heynderickx, "Visual discomfort and visual fatigue of stereoscopic displays: A review," *J. Imaging and Technology* 53(3), 2009.

[9] S. Jumisko-Pyykkö, T. Utriainen, D. Strohmeier, A. Boev, and K. Kunze, "Simulator sickness – Five experiments using autostereoscopic mid-sized or small mobile screens," *Proceedings of 3DTV Conference,* pp- 1–4, 2010.

[10] J. Häkkinen, M. Liinasuo, J. Takatalo, and G, Nyman, "Visual comfort with stereoscopic mobile gaming," *Proceedings of the IS&T/SPIE's International Symposium on Electronic Imaging 2005: Stereoscopic Displays and Applications XVII*, pp. 15-19, 2005.

[11] S. Jumisko-Pyykkö, T. Haustola, A. Boev, and A. Gotchev, "Subjective evaluation of mobile 3D video content: depth range versus compression artifacts," *Proceedings of SPIE 7881A*, 2011.

[12] ITU-T P.910 Recommendation, "Subjective video quality assessment methods for multimedia applications," International Telecommunication Union (ITU) – Telecommunication sector, 1999.

[13] R.S. Kennedy, N.E. Lane, K.S. Berbaum, and M.G. Lilienthal, "Simulator Sickness Questionnaire: an Enhanced Method for Quantifying Simulator Sickness," *International Journal of Aviation Psychology*, 1993.

[14] S. Jumisko-Pyykkö, V.K. Malamal Vadakital, & M. Hannuksela, "Acceptance threshold: Bidimensional research method for user-oriented quality evaluation studies," *International Journal of Digital Multimedia Broadcasting,* 20 pages, 2008.

[15] A. Smolic, G. Tech and H. Brust, "Report on generation of stereo video data base," Mobile3DTV technical report D2.1, 44 pages, 2010.

[16] D. Hoffman, A. Girshick, K. Akeley, and M. Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of Vision* 8(3), pp. 1–30, 2008.

[17] A. Gotchev, "Spline and Wavelet Methods for Signal and Image Processing," Doctor of Technology Thesis, Tampere University of Technology, 2003.

[18] ITU-R BT.500-11, "Methodology for the subjective assessment of the quality of television pictures," Telecommunications Union – Radiocommunication sector, 2002.