Subjective evaluation of mobile 3D video content: depth range versus compression artifacts

Satu Jumisko-Pyykkö[a], Tomi Haustola[a], Atanas Boev[b], Atanas Gotchev[b]
[a]Unit of Human-Centered Technology, Tampere University of Technology,
P.O. BOX 589, FI-33101 Tampere, Finland;
[b]Dept. of Signal Processing, Tampere University of Technology,
P.O. Box 553, FI-33101 Tampere, Finland

ABSTRACT

Mobile 3D television is a new form of media experience, which combines the freedom of mobility with the greater realism of presenting visual scenes in 3D. Achieving this combination is a challenging task as greater viewing experience has to be achieved with the limited resources of the mobile delivery channel such as limited bandwidth and power constrained handheld player. This challenge sets need for tight optimization of the overall mobile 3DTV system. Presence of depth and compression artifacts in the played 3D video are two major factors that influence viewer's subjective quality of experience and satisfaction. The primary goal of this study has been to examine the influence of varying depth and compression artifacts on the subjective quality of experience for mobile 3D video content. In addition, the influence of the studied variables on simulator sickness symptoms has been studied and vocabulary-based descriptive quality of experience has been conducted for a sub-set of variables in order to understand the perceptual characteristics in detail. In the experiment, 30 participants have evaluated the overall quality of different 3D video contents with varying depth ranges and compressed with varying quantization parameters. The test video content has been presented on a portable autostereoscopic LCD display with horizontal double density pixel arrangement. The results of the psychometric study indicate that compression artifacts are a dominant factor determining the quality of experience compared to varying depth range. More specifically, contents with strong compression has been rejected by the viewers and deemed unacceptable. The results of descriptive study confirm the dominance of visible spatial artifacts along the added value of depth for artifact-free content. The level of visual discomfort has been determined as not offending.

Keywords: mobile, 3D, television, subjective evaluation, quality of experience, quality perception

1. INTRODUCTION

The challenge for mobile 3D video and television is to deliver highly satisfying viewing experience under tight technical requirements. The optimization of the system resources is required in order to harmonize the high amount of 3D media data with the limited bandwidth of the error-prone transmission channel and power-constrained handheld. Previous studies have shown that three-dimensional video presented on a portable auto- stereoscopic display can be superior over the existing two-dimensional video [1]. Experiments exploring different components of the end-to-end system chain have concluded that '*Experienced added value of depth is visible only if the level of artifacts is low*' [1]. However, the juxtaposition between the added value of the depth and the level of visible artifacts has not been systematically addressed. Furthermore, related works have underlined that 3D video quality of experience on small display is more multifaceted phenomenon than the excellence of visual quality. It includes components such as low-level visual characteristics of video (e.g. depth, spatial and motion quality) and high-level aspects of viewing experience (e.g. ease of viewing, visual discomfort, enhanced immersion) [1], [2], [3], [4]. To address the quality of experience for novel multimedia services in depth, a multi-methodological research approach is needed.

The primary goal of this study is to examine the influence of depth and compression artifacts on quality of experience of 3D video for mobile devices. The supplementary aims of this study are to explore visual discomfort in the settings of this study, and operationalize and apply the descriptive model of quality of experience for 3D video on mobile devices to the evaluation of a subset of stimuli in order to draw deeper understanding on quality characteristics. Related work on 3D video quality of experience for mobile devices is overviewed in Section 2. The characteristics of the experimental research method are given in Section 3. Section 4 summarizes the main results, while Section 5 emphasizes some discussion issues and concludes the study.

2. 3D VIDEO QUALITY OF EXPERIENCE FOR MOBILE DEVICES

Multiple factors, such as the high amount of (3D) audiovisual data, limited bandwidth, error-prone transmission channel, and constraints of receiving devices (e.g. screen size, computational power, battery life-time) set specific requirements for multimedia produced quality on mobile devices. In the mobile television broadcasting scenario, the TV content is captured, encoded, and transmitted over wireless broadcasting channel to be received, decoded and played back on the small handheld's screen. The above stages remain practically the same for the delivery of mobile 3DTV, however substantial modifications are required, mainly related with the representation, encoding and displaying of 3D video content [5]. Driven by the restrictions of the terminal device, stereo video has been considered as the most suitable representation format, while autostereoscopic displays have been considered as the most suitable displays for 3D scene visualization on mobile devices [5], [6], [7]. Stereoscopic 3D is created without wearable glasses but by an additional optical layer placed on the surface of the screen to divide the view into (two or more) fields shown to the right and left eyes separately [6]. The stages of the mobile delivery chain inevitably introduce processing artifacts in the video to be displayed. Here, artifacts refer to impairments, introduced through a process that are not naturally present [8]. They can occur independently or jointly, influencing experienced quality in the end. Examples include blockiness, packet losses, temporal mismatch of stereo frames and display-specific artifacts, such as cross-talk and pseudo-stereo [9].

However, the quality of experience is a more complex phenomenon than the simple perception of change in erroneous versus error-free multimedia presentation. It has been proposed that 3D quality of experience is composed of naturalness, depth perception, image quality, and visual comfort [2]. Recent studies conducted with 3D video on mobile device have also underlined that the ease of viewing is a central requirement for stereoscopic video in natural contexts of use [3]. Subsequently, the need of a multimethodological research approach has been identified in order to understand the quality of experience in depth [1], [3].

Perceivable added value of the depth is among the most essential requirements for the success of stereoscopic media format in customer services [10]. Therefore, comparisons between 2D and 3D presentation modes have been highly relevant in the recent research. Shibata *et al.* [11] conducted a comparison between monoscopic and steroscopic presentation modes for images on a mobile device. The results, concluded from very small sample size (9 participants), initially indicated that the 3D presentation mode can improve visual image quality of experience over 2D, but this experience is not only associated with positive impressions (real-life likeness, presence, perceivable depth) but also with a negative consequence of stereoscopic quality (troublesome feeling while watching and impression of weirdness). In more extensive and relevant studies for stereoscopic video on mobile video, the excellence of 3D has been strongly influenced by the display technique used. Utriainen and Jumisko-Pyykkö [12] compared presentation modes (2D, 3D) and coding parameters for audio and video for mobile video using parallax barrier display technique. They results showed that independently on other studied parameters stereoscopic presentation provided highly unacceptable and lower quality than 2D video. Their qualitative results confirmed that display artifacts as well as associated visual discomfort reduced quality of experience for 3D [13], [14]. In a more recent study based on the use of lenticular display technology, Strohmeier *et al.* [1] concluded strong preference and highly acceptable quality towards stereoscopic presentation mode when the level of visible artifacts was low. This conclusion was further confirmed in a study comparing the 3D video coding formats in limited bandwidth scenarios [15]. Furthermore, the preference toward 3D video is not influenced significantly by varying monoscopic or stereoscopic audio quality [1]. In sum, the above-referred studies indicate that the 3D quality of experience for mobile 3D video can be enhanced by the use of stereoscopic presentation conditionally to the existence of visible impairments. However, the previous studies have not systematically examined the relation between the positive quality of experience introduced by the depth presence and the negative quality of experience caused by the varying intensity of visible artifacts.

Visual discomfort as an essential part of 3D viewing experience can be influenced by multiple factors – not only stereoscopic presentation. Firstly, it is influenced by the accommodation-convergence conflict, visible artifacts and blur (see an extensive overview in [16]). Secondly, in the temporal domain, the increase in immersive time can increase symptoms and introduce adaptation while typically these symptoms start to reduce during the short period after immersion (e.g. after 4-10 min)[14], [17]. Thirdly, fast motion combined with stereoscopic presentation in content, use of small display size and intensive task type can increase the experienced visual discomfort where large individual differences are common [18], [19]. In a recent study by Jumisko-Pyykkö *et al.* the tendency of visual discomfort for stereoscopic video on small screen was concluded based on five extensive experiments with different display technologies, variable depth levels and system parameters [14]. The results in that study showed that significantly lower level of symptoms was registered in the tests using lenticular display technology compared to the parallax barrier display technology. For the former, the intensity of the symptoms was equal to or lower than those recorded after 40 min lasting fast-speed gaming on the CRT or head-mounted displays [18]. This indicates that a short term video viewing (less than 30 min), on a good autostereoscopic display is not offending.

To go beyond the comparisons of excellence of quality and visual discomfort, the descriptive model of 3D quality of experience for video on mobile device was developed to create deeper understanding on broad characteristics of experienced quality [4]. The model was constructed based on five studies using either interview or written attribute description tasks with more than 90 participants. The experiments contained an extensive and heterogeneous set of produced quality by varying content, level of depth, compression and transmission parameters, and audio and display factors for 3D. The model contains two main components – visual quality and viewing experience, and two minor components described as content and quality of other modalities and their interaction. Visual quality represents the directly detectable, data-driven or low-level characteristics of video and is divided into depth, spatial and motion quality including a total of 11 subcomponents. Viewing experience, as composed of ease of viewing, pleasantness of viewing, enhanced immersion, visual discomfort, comparison to existing technology and overall quality, is characterized by higher-level interpretation of quality where human goal-oriented actions, emotions, attitude and knowledge are essentially involved. Although the extensive model has been presented, it has not been further made operational to act as a vocabulary in the sensorial profiling of stimuli.

The main goal of this study is to examine the influence of depth and compression artifacts on quality of experience. The supplementary aims of this study are to explore visual discomfort in the settings of this study and operationalize the descriptive model of 3D quality of experience for video on mobile device and utilize it to the evaluation of the subset of stimuli to draw deeper understanding on quality characteristics.

3. RESEARCH METHOD

3.1 Participants

The test included a total of 30 participants equally stratified by gender and age group (18-45 years). In the analysis, one outlier participant was excluded. 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 [25]). In addition, according to the attitude towards technology, the maximum 20% of the participants were categorized as innovators and early adopters, and the laggards with extreme negative technology attitude were screened out [26]. All participants had normal or corrected-to-normal vision and were native Finnish speakers.

3.2 Test Procedure

The test procedure was divided into three phases: pre-test, test and post test phase. In the pre-test session, sensorial tests (visual acuity - Landot chart 20/40; color vision -Ishihara test, and acuity of stereo vision (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) [20]. Finally, aim of the anchoring and training was to familiarize the participants to the contents used in the study, the extremes of quality range of stimuli, and evaluation task.

The test method applied Absolute Category Rating method and Bidimensional research method of acceptance [25], [27]. The stimuli were presented one by one, rated independently and retrospectively [25]. After each stimulus video clip, the overall quality was rated using 11-point unlabelled scale and the domain specific acceptance of quality for viewing mobile three-dimensional television on a nominal (yes/no) scale [25], [27]. The duration of the stimuli was ten seconds followed by five second lasting answer time when mid-grey still image was presented on a screen on a zero disparity level.

During the post-test session, SSQ survey, demo and psychographic survey, and descriptive quality evaluation task were conducted, respectively. The post-immersive SSQ survey was collected after viewing of approximately 22 minutes excluding the answer time [20]. The goal of the descriptive quality evaluation task within this study was to pilot the measurement tool for the subset of stimuli. The data-collection was based on the model of Descriptive Quality of Experience for 3D video on mobile device [4]. In the data-collection, 17 components of the model characterizing visual quality and viewing experience were used. The components were rated independently on an 11-point scale where the anchor terms per component were marked at the extremes Table 1. Prior to the actual descriptive evaluation, the ease of understanding the terms were tested and improved with three naïve participants. During the descriptive task a subset of stimuli were evaluated in random order.

Table 1 Components of descriptive quality of experience for visual quality and viewing experience with the related anchor terms [4].

	Component	Anchor term (min)	Anchor term (max)
Visual quality	Perceivable depth	not perceivable	perceivable
	Impression of depth	artificial	natural
	Foreground-background layers	separate	smoothly combined
	Balance of foreground-background quality	unbalanced	balanced
	Block-free image	visible errors	error free
	Clarity of image	blur	clear
	Color, brightness and contrast	unpleasant	pleasant
	Objects and edges	inaccurate	accurate
	Fluency of motion	influent	fluent
	Nature of motion	static	dynamic
	Clarity of motion	blurry	clear
Viewing experience	Ease of viewing	difficult	easy
	(Visual) discomfort	not experienced	experienced
	Enhanced immersion	not enhanced	enhanced
	Overall quality	low	high
	Pleasantness of viewing	Unpleasant	pleasant
	Comparison to existing technology	not improved	improved

3.3 Stimuli –content, parameters, and preparations

The variables of the experiment were: Depth levels (3) x Quantization parameters (5) x Contents (4).

Content - Four test contents with variable visual characteristics were used in the experiment (Figure 1). The characteristics varied in the terms of spatial details, temporal motion, amount of depth and depth dynamism. There were neither scene cuts nor audio presented in the selected contents.

Parameters – 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, 45, respectively. The goal of the selection of these parameters was to systematically tackle the juxtaposition between the positive influence of depth and negative influence of artifacts on experienced quality based on previous work. It is known that under the perceptually error-free scenarios, stereoscopic video on small screen has been experienced to provide higher visual quality over conventional monosopic video presentation [1]. However, this positive effect of stereoscopic depth is overridden when detectable artifacts (e.g. by compression, rendering, display) are part of visual quality, and then the mono presentation is more pleasurable [1]. When examining consumer services, such mobile 3D video, visible independently or jointly occurring artifacts from capture to vulnerable wireless transmission and visualization on display can potentially reduce visual quality of experience (e.g. review of artifacts [9]).

Content preparation – The source content was selected from the multi-view videos available to MPEG community and also made available to the Mobile3DTV stereo-video database [21]. The reason of using multi-view videos was determined by the aim to vary the depth range in the test videos. The preparation of variable depth levels contained the following steps: for each video clip, multiple stereoscopic versions were prepared, by selecting different camera pairs from the available multi-view video tracks. The left camera of all sequences was kept the same. The 3D-effect of the video was controlled by the position of the right camera, as follows: 1) Monoscopic (2D), video, where left and right camera correspond to the same camera number in the multiview sequence; 2) Short baseline – a camera baseline, which produces 3D scene with a limited disparity range and less-pronounced, but visible 3D effect; 3) Wide baseline – camera baseline, where the camera baseline is selected to provide the optimal disparity range for the chosen stereoscopic display. By the notation of short baseline, we consider the use scenario where a high-resolution video content is repurposed for mobile use by a direct linear down-scaling. In the general case, this would result in a 'shallow' depth. By notation of wide baseline we consider the scenario where the content is specifically adapted to the viewing conditions of the portable display. In the ideal case, this would result in full utilization of the mobile display's comfort zone [22][23].

Each stereoscopic sequence was converted from its original resolution to the resolution of the target display by using a four-step procedure. 1) The disparity range of each stereo-pair was analyzed. 2) The left and right channels of the video were cropped from the sides with the aim to shift the disparity range and equalize the absolute positive and negative disparity values, as well as to avoid frame violence. The position of the first cropping window in the left channel was kept intact, while varying it in the right channel relatively for the different camera pairs. 3) Both channels were down scaled with respect to the smaller target dimension while maintaining the source aspect ratio. 4) The extra pixels of the larger target dimension were cropped to achieve the display (target) aspect ratio. The position of the second cropping window was the same for all channels and all frames, and was selected manually based on the movie content. For the cropping operations, cubic spline interpolation was applied, while for the resizing (down-scaling), least-squares cubic projection was applied [24]. Following these steps, three depth levels were created for the selected contents. Figure 1 B-C, presents an illustration of the output disparity levels of the test stereoscopic videos. Zero disparity refers to a scene on the display plane, while negative values indicate scenes in front and positive values indicate scenes behind the display plane.

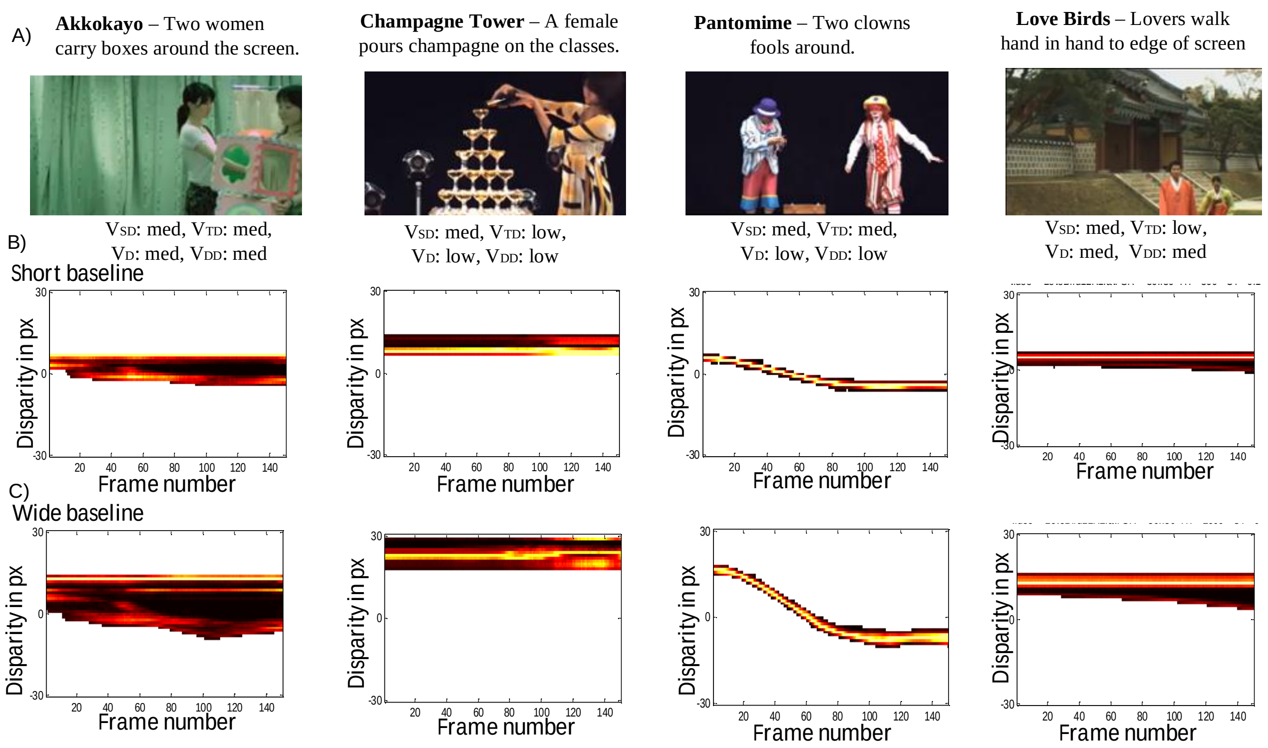


Figure 1 A) Stimuli content descriptions and visual characteristics VSD=visual spatial details, VTD=temporal motion, VD=amount of depth, VDD=depth dynamism, B) Short baseline – disparity in pixels per frames of content, C) Wide baseline - disparity in pixels per frames of content.

After downscaling, each test video was compressed by H.264 reference encoder in intra-frame mode, applied independently to the left and right channels (no inter-view prediction). Five values of the quantization parameter (QP), namely QP=[25, 30, 35, 40, 45] were used to introduce different levels of blocky compression artifacts (Figure 2).

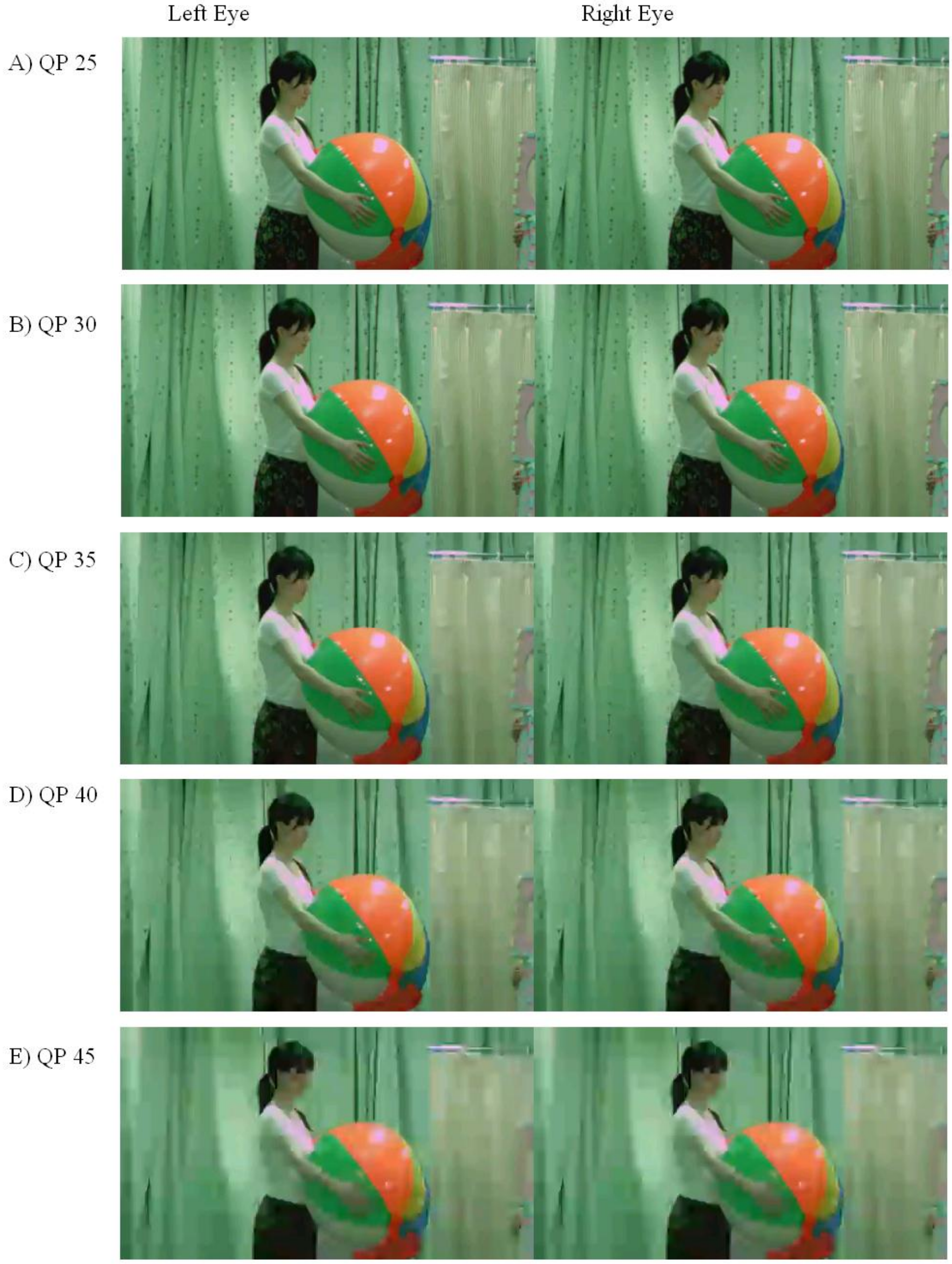


Figure 2 Example of the content Akkokayo with five different quantization parameter values (A-E).

3.4 Viewing conditions and presentation of the stimuli

The experiments were conducted in the controlled laboratory conditions, set according to the ITU-T P.910 Recommendation [25]. The stimuli were presented on a portable autostereoscopic LCD display with horizontal double density pixel (HDDP) arrangement produced by NEC LCD [30]. The physical size of the display is 3,5'' and the resolution is 427x240px at 155DPI. 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 [23]. 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 between the display and the viewer were set to 40 cm. The stimuli were presented twice in pseudo random-order during the quality evaluation task [29].

4. RESULTS

4.1 Acceptance of quality

Analysis – Cochran Q-test conducted to measure the differences between more than two related nominal data sets [28].

The results show that variable combinations of QP and depth level had significant influence on acceptance of quality ($Q(14)=1538$, $p<.001$; Figure 3). Highly acceptable quality (at least 70%) at the two lowest QP levels (25,30) in all depth levels in three contents indicating that low level of artifacts can be acceptable for viewing 3D. However, the general tendency shows that quality becomes unacceptable (below 50% threshold) when QP values are higher than 35 in all depth levels.

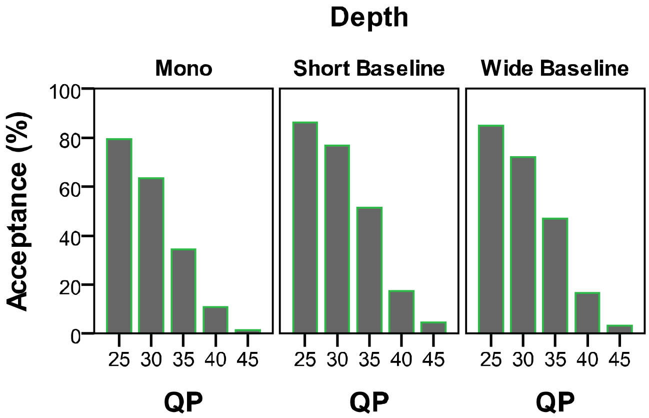
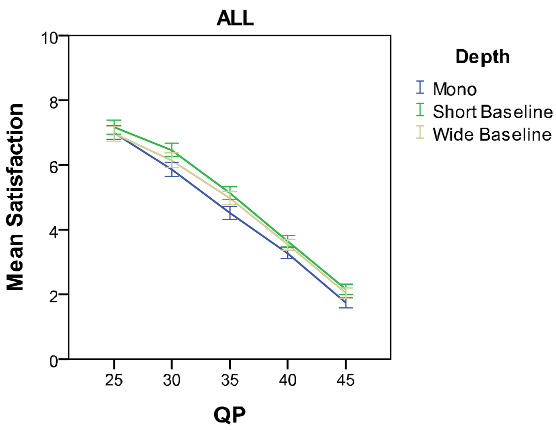


Figure 3 Acceptance of quality for depth and QP parameter combinations averaged across the contents.

4.2 Overall quality satisfaction

Analysis – Analysis was conducted using Friedman's test and Wilcoxon matched-pair signed-ranks test as the normality requirements of parametric methods was not met (Kolmogorov-Smirnow $p < .05$). Friedman's test is applicable to measure differences between more than two related and ordinal datasets and Wilcoxon's test between two of them [28].

The results show that variable combinations of QP and depth level had significant influence on quality satisfaction when averaged across the contents ($Fr=2448$, $df=14$, $p<.001$; Figure 4). The results show that experienced quality decreases significantly when the quantity of QP increases (all pair-wise comparisons, Wilcoxon: $p<.05$). Influence of depth show superiority of stereoscopic presentation over mono with small differences over the quantization quality range. At the highest quality level (QP 25), depth does not have significant influence on perceived quality ($p>.05$, *ns*). In all other studied quantization intensities, both stereoscopic presentation modes provides higher quality over mono presentation (mono vs. other all pair-wise comparisons Wilcoxon: $p<.001$). Finally, the preferences between the short and wide baseline depend on the quantization quantity. At the low QP level (30), short baseline is preferred over the wide baseline ($p<.05$) while on the higher quantization levels (QP35-45) there is no significant difference between them ($p>.05$, *ns*). Content by content analysis follows this main tendency of influence of quantization and depth on experienced quality.



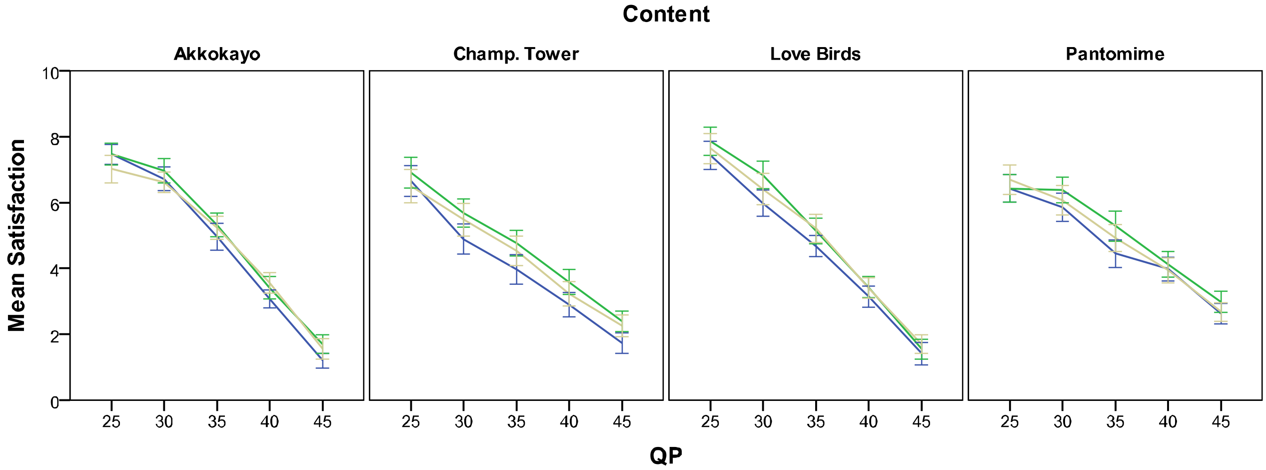


Figure 4 Mean quality satisfaction score for all (averaged across the contents) and content by content. The error bars show 95% CI of mean.

4.3 Supplementary results

Pre and post-immersive simulator sickness

Analysis - The total SSQ scores of the four groups are calculated by summing the ratings of related symptoms in each group [20]. Each sum is then multiplied by a weighting score which has been defined by varimax factor weights during SSQ development [20].The weight scores are: Nausea = 9.54, Oculomotor= 7.58, Disorientation = 13.92 and total score = 3.74. This paper presents the absolute values. Wilcoxon test was used to compare the pre-and post immersive 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=-2.858, p<.05), oculomotor (Z=-2.464, p<.05), and disorientation (Z=-3.974, p<.01) while the difference was not announced for nausea (Z=-0.504, p=*ns*). The overall level of these symptoms is similar to our several previous studies with same display with variable visual stimuli [14].

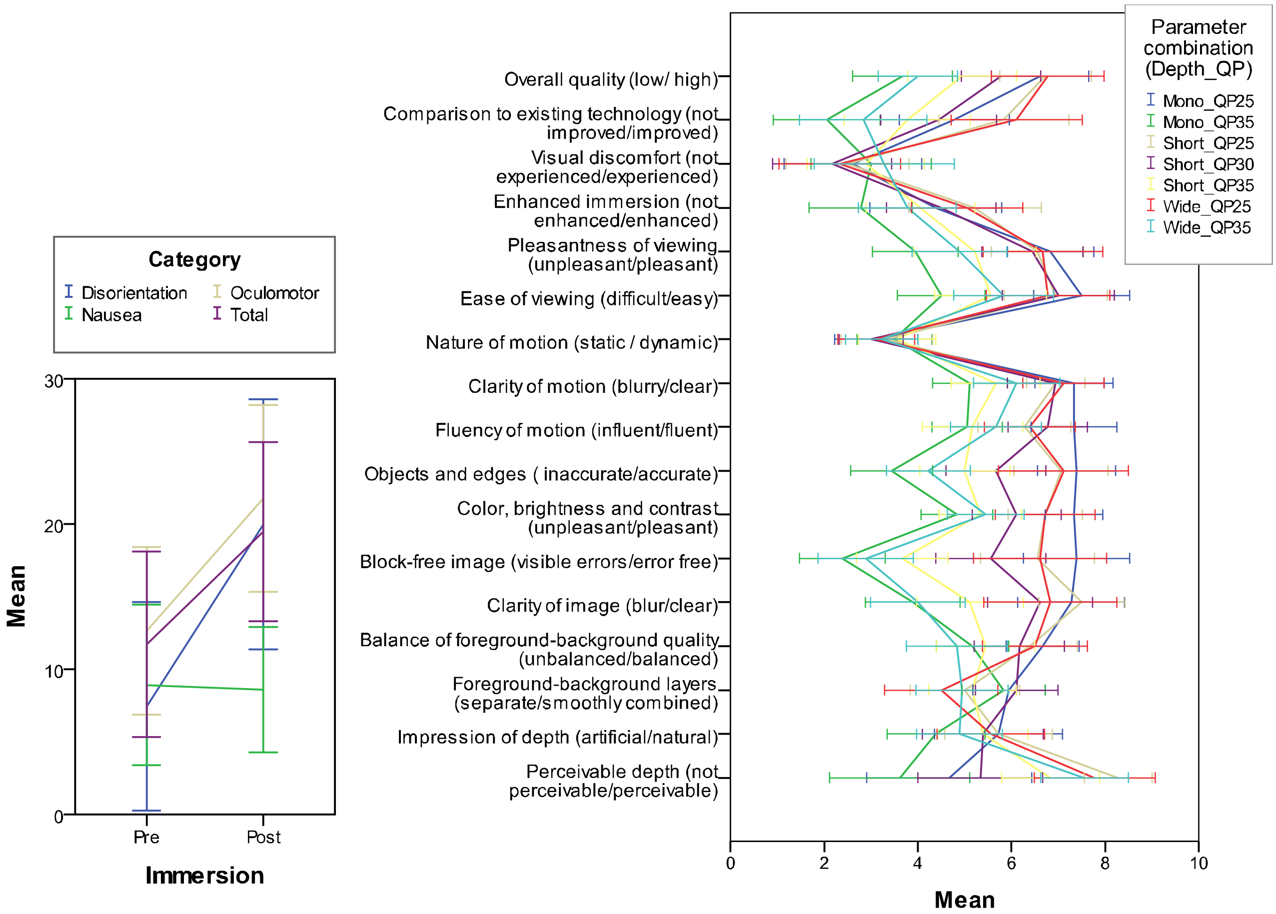


Figure 5 A) Simulator sickness – pre-and post-immersive symptoms of the quality evaluation task B) Descriptive quality - Influence of parameter combinations of components of experience for content pantomime. The bars show 95% CI of mean.

Descriptive quality of experience

Analysis – The analysis included 17 participants with complete data-sets. One-dimensional analysis is conducted to examine the influence of the subset of parameter combinations on multiple components of experienced quality with the content 'Pantomine'. The non-parametric analysis using Friedman and Wilcoxon's test was conducted.

The parameter combinations had significant influence on multiple (13/17) descriptive components (Friedman: $p<.05$) excluding the components called nature of motion, visual discomfort, impression of depth and foreground-background layers (Friedman: ($p>.05$, *ns*). In overall, the results show the tendency that the lowest level of quantization (QP 25, 30) has the most positive evaluation over the multiple components. The components called clarity of image, block-free image, objects and edges are strongly separating between the high and low quantization parameters (Wilcoxon: QP 35 vs. other QPs $p<.05$; all comparisons between depth levels $p >.05$, *ns*). To demonstrate this effect the stimuli with high QP (35) is considered to contain visible errors while the lower OPs (25, 30) are experienced as error-free. In addition to these interpreted data-driven features of stimuli, the influence of quantization is detectable in ease of viewing, affective pleasantness of viewing and finally in overall quality where also the low level of quantization results in more positive experience (Wilcoxon: QP 35 vs. other QPs $p<.05$; all comparisons between depth levels $p >.05$, *ns*).

Finally, the results show that the depth is detectable between mono and stereoscopic modes giving the further support to the results of satisfaction. However, the components expected to strongly convey aspects of excellence of 3D (impression of depth, foreground-background layers) does not show or show only small influence on the higher-level components (e.g. visual discomfort, enhanced immersion, comparison to exiting technology).

5. DISCUSSION AND CONCLUSIONS

The primary goal of this study was to examine the influence of depth and compression artifacts on the quality of experience of 3D video on mobile devices. As previous studies identified the compression artifacts as the most annoying in 3D video experience, the current study aimed at investigating these artifacts on a denser set of quantization parameters so to find informative levels of user acceptance. In addition, depth presence was varied through varying the baseline between the two cameras in the stereo pair and resizing the test content accordingly. The goal of this depth variability approach was two-fold. First, it was interesting to put compression artifacts against different depth levels and to draw some conclusions about the visibility of blockiness in difference presence of depth. Second, it was interesting to test two scenarios about preparation of 3D video for mobiles. In one scenario, the content is considered repurposed directly from higher resolution content - this process manifested by the short baseline stereo video. In another scenario, the content is considered optimal for the resolution and size of the mobile display – this process manifested by the wide baseline stereo video. The secondary aims of this study were to explore visual discomfort in the settings, and make operational and apply the descriptive model of 3D quality of experience for video on mobile device to the evaluation of the subset of stimuli in order to draw deeper understanding on quality characteristics.

The results of the study lead to three main conclusions. The first conclusion underlines the strong dominance of the studied compression over the varied depth levels. The acceptable quality can be reached when the level of compression is low (quantization 30 at the maximum). This conclusion confirms the results of our previous studies indicating that compression is the primary source of artifacts and quality impairments and should be thoroughly addressed in the case of 3D video for mobiles. The result is also instructive for practitioners developing perceptually-driven objective metrics for 3D video quality assessment.

The second conclusion is that the stereoscopic 3D provides higher quality of experience over monosopic presentation in most of the cases however the differences between stereoscopic short and wide baselines are non-significant. This result indicates that optimization of the depth presence through camera selection for the portable stereoscopic display does not necessarily provide pleasurable 3D quality. There are two possible reasons for this observation. First, the wide baseline was selected following mainly a geometrical rationale about the display's comfort zone [22]. This has lead to somehow 'theoretically' optimal baseline, which in practice turned out to generate slightly exaggerated (though still well perceived) depth. This phenomenon was registered also during the descriptive sessions and was pointed out by some of the subjects. While the stereo images were still fuseable in terms of the accommodation-vergence conflict, the depth in some stereo frames was going beyond the Percival zone of comfort. It seems that the actual Percival zone of comfort is narrower than 0.3/3 diopters for the given portable display possibly due to optical effects such as display cross-talk [23]. The second reason for the sub-optimal performance of wide baseline stereo is that ghosting effects are enhanced with the highly pronounced depth. The presence of noticeable ghosting was also commented by some of the test subjects. As s conclusion: we attempted to create stereo video with an optimal baseline separation fitting the full disparity range for the given display. This attempt was based on geometrical and optical considerations [23][22]. The results from the tests showed that the wide baseline is indeed creating fuseable images, but in some scenes go beyond the comfort disparity of the display. In contrary, the short baseline, though shallower, was always comfortable and acceptable for the observers.

The third conclusion, based on the vocabulary descriptive results, concluded from a set of all stimuli, complements the main results of this study. The increased compression had influence not only on the multiple low-level components of spatial visual quality (e.g. clarity of image, block-free image, objects and edges), but also to the higher level components of viewing experience (ease of viewing, overall quality and pleasantness of viewing). Furthermore, depth has been detectable in the most of the cases between the presentation modes without influence on components characterizing nether impression of depth nor visual discomfort. These initial results are encouraging about the use of the vocabulary-based descriptive evaluation for 3D video to explain the main results. However, further work has to address the broader set of heterogeneous stimuli to be able to draw strong conclusions and probe systematically the reliability of the method with naïve participants. Finally, the results of simulators sickness questionnaire show only small difference between pre- and post immersive evaluations. Our results are in line with previous comparable studies concluding non-offending influence on quality of experience [14].

The results of this study are limited to one type of artifacts, used display technique and controlled circumstances. Further work has to address the tolerance towards different type of artifacts and their joint influence with relevance to 3D video. For example, the temporal artifacts might have more interruptive nature towards to user's viewing task while these spatially emphasized artifacts might cause more impression of inaccuracy, as shown in studies with impaired 2D video on small screens [31]. Furthermore, in the noisy natural viewing conditions with actively shared attention (e.g. busses) the people can be more tolerant towards visual artifacts and less sensitive between different system qualities [3], [32]. In the long term, the goal is to identify the acceptance threshold for the multiple types of artifacts and viewing conditions, and develop scalable solutions for guaranteeing a pleasurable quality of experience for 3D video viewing under variable system resources and contexts of use.

ACKNOWLEDGEMENTS

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

REFERENCES

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

[2] Seuntïens, P. J. H., “Visual experience of 3D TV,” PhD Thesis, Technische Universiteit Eindhoven, (2006).

[3] Jumisko-Pyykkö, S. and Utriainen, T., “A hybrid method for the context of use: Evaluation of user-centered quality of experience for mobile (3D) television,” International Journal of Multimedia Tools and Applications: Special issue on Mobile Media Delivery, 1-41 (2010).

[4] Jumisko-Pyykkö, S., Strohmeier, D., Utriainen, T. and Kunze, K., “Descriptive quality of experience for mobile 3D television,” Proceedings of NordiCHI, 1–10 (2010).

[5] Gotchev, A., Smolic, A., Jumisko-Pyykkö, S., Strohmeier, D., Akar, G. B., Merkle, P. and Daskalov, N., “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).

[6] Gotchev, A., Tikanmäki, A., Boev, A., Egiazarian, K., Pushkarov, I., Daskalov, N., “Mobile 3DTV technology demonstrator based on OMAP 3430”, Proc. of 16th International Conference on Digital Signal Processing, (2009).

[7] Willner, K., Ugur, K., Salmimaa, M., Hallapuro, A. and Lainema, J., “Mobile 3D video using MVC and N800 internet tablet,” 3DTV Conference: The True Vision – Capture, Transmission and Display of 3D Video 2008, 69–72 (2008).

[8] “MOT Oxford Dictionary of English 1.0”, Oxford University Press., (2005)

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

[10] Jumisko-Pyykkö, S., Weitzel, M. and Strohmeier, D., “Designing for user experience: What to expect from mobile 3D TV and video?,” The First International Conference on Designing Interactive User Experiences for TV and Video: UXTV, 183–192 (2008).

[11] Shibata, T., Kurihara, S., Kawai, T., Takahashi, T., Shimizu, T., Kawada, R., Ito, A., Häkkinen, J., Takatalo, J. and Nyman, G.,”Evaluation of stereoscopic image quality for mobile devices using interpretation based quality methodology,” Proceedings SPIE Stereoscopic Displays and Applications XX 7237(72371E), (2009).

[12] Utriainen, T. and Jumisko-Pyykkö, S., “Experienced audiovisual quality for mobile 3D television,” Proceedings of 3DTV Conference, 1–4 (2010).

[13] Jumisko-Pyykkö, S., & Utriainen, T., “User-centered quality of experience: Is mobile 3D video good enough in the actual context of use?,” Proceedings of Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics: VPQM, 1–5 (2010).

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

[15] Strohmeier, D., & Tech, G., “Sharp, bright, three-dimensional: open profiling of quality for mobile 3DTV coding methods,” Proceedings of SPIE Multimedia on Mobile Devices, (2010).

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

[17] Kennedy, R., Lane,N., Berbaum, K., and Lilienthal, M., “Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness,” Int. J. Aviation Psychology 3(3), 203-220 (1993).

[18] Pölönen, M. and Häkkinen, J., “Near-to-Eye Display - An accessory for handheld multimedia devices: Subjective studies,” Journal of Display Technology, 5(9), 358-367 (2009).
[19] Häkkinen, J., Liinasuo, M., Takatalo, J. & Nyman, G., “Visual comfort with stereoscopic mobile gaming,” Proceedings of the IS&T/SPIE's International Symposium on Electronic Imaging 2005: Stereoscopic Displays and Applications XVII, 6055, 15-19 (2005).
[20] Kennedy, R., S., Lane, N., E., Berbaum, K., S. and Lilienthal, M., G., “Simulator Sickness Questionnaire: an Enhanced Method for Quantifying Simulator Sickness”. International Journal of Aviation Psychology, (1993).
[21] Smolic, A., Tech, G. and Brust, H., "Report on generation of stereo video data base," Mobile3DTV technical report D2.1, 44 (2010).
[22] Hoffman, D., Girshick, A., Akeley, K. and Banks, M., "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," Journal of Vision 8(3), 1–30 (2008).
[23] Boev, A., and Gotchev, A., “Comparative Study of Autostereoscopic Displays for Mobile Devices,” Proceedings of SPIE 7881A, (2011).
[24] Gotchev, A., “Spline and Wavelet Methods for Signal and Image Processing,” Doctor of Technology Thesis, Tampere University of Technology, (2003).
[25] ITU-T P.910 Recommendation, ”Subjective video quality assessment methods for multimedia applications,” International Telecommunications Union (ITU) – Telecommunication sector, (1999).
[26] Rogers, E., M., [Diffusion of Innovations], Free Press, New York, (2003).
[27] Jumisko-Pyykkö, S., Malamal Vadakital, V., K. & Hannuksela, M., M., ”Acceptance threshold: Bidimensional research method for user-oriented quality evaluation studies,” International Journal of Digital Multimedia Broadcasting, 20 pages (2008).
[28] Coolican, H., [Research Methods and Statistics in Psychology], Hodder & Stoughton, London, (2004).
[29] ITU-R BT.500-11, “Methodology for the subjective assessment of the quality of television pictures,” Telecommunications Union – Radiocommunication sector, (2002).
[30] Uehara, S., Hiroya, T., Kusanagi, H., Shigemura, K and Asada, H., “1-inch diagonal transflective 2D and 3D LCD with HDDP arrangement,” Proc. SPIE-IS&T Electronic Imaging 2008 Stereoscopic Displays and Applications XIX, (2008).
[31] Jumisko-Pyykkö, S., “User-Centered Quality of Experience and its Evaluation Methods for Mobile Television,” PHD thesis Tampere University of Technology, (under review).
[32] Knoche, H. and Sasse, M. A., “The big picture on small screens delivering acceptable video quality in mobile TV," ACM Trans. Multimedia Computing, Communications and Applications TOMCCAP 5(3), 1–27 (2009).