Quality of Experience of Stereoscopic Content on Displays of Different Sizes: A Comprehensive Subjective Evaluation

Lino E. CORIA, *Senior Member, IEEE*, Di XU, *Student Member, IEEE*, and Panos NASIOPOULOS, *Member, IEEE*

***Abstract*-3D-capable devices have entered the consumer market. The depth that is perceived with stereoscopic content is strongly linked to the size of the screen in which it is displayed. We tested the effect that different display sizes have on the quality of experience of 3D content and we offer some recommendations.**

I. INTRODUCTION

Devices that are capable of displaying stereoscopic content such as 3D TVs and 3D-enabled mobile phones are presently available in the consumer market. Because of this, there is an increasing demand for 3D content. Creation of high-quality 3D content, however, is technically challenging and content producers are still trying to understand the intricacies of this medium. Subjective tests have been performed [1], [2], [3] in order to study the influence that the acquisition parameters have on the perceived quality of stereoscopic images.

While a good understanding of the acquisition process of stereoscopic content is essential for providing a high quality of experience to viewers, the display aspect should not be overlooked. It is important to acknowledge that the depth that is perceived in stereoscopic content is strongly linked to the size of the screen on which it is displayed. This is illustrated in Fig. 1 where the same content is shown on a large display (Fig. 1a) and on a smaller one that is half the size of the first display (Fig. 1b). The blue squares (A_L, B_L) and red squares (A_R, B_R) indicate the pixels from the left view and right view, respectively, while the gray squares (A, B) denote the pixel's perceived position in 3D. Because the second display is smaller, the content occupies less space. The viewer's interpupillary distance, however, remains constant. Besides, the viewer will generally get closer to the smaller screen. Because of this, the perceived depth is different for each display. The larger the screen and the farther away the viewer is from it, the larger the illusion of depth. This means that content that can be pleasing for a large screen might become unnatural for small displays and vice versa.

In this study, we tested the effect that different display sizes have on the quality of experience of 3D content. We created our own image database, comprised of scenes depicting people and landscapes. Several viewers of different ages watched these images on three displays (a 2.8" 3D hand-held device, a 22" 3D computer monitor, and a 55" 3D TV) and rated them. We used the outcome of these subjective tests to produce a set of recommendations for the display of high-quality 3D content.

II. ACQUISITION

We used two identical HD cameras (Sony HDR-XR500V 1080 60i NTSC) aligned in parallel and attached to a bar that was secured to a tripod. Only the extreme end of the zoom range was used, i.e., a focal distance of 5.5 mm. A single remote control was employed to synchronize the cameras. The distance between the cameras was kept constant at 77 mm. However, we varied the distance between the cameras and the closest point captured in the stereoscopic image pair, d_{min}, and the distance between the cameras and the background.

Even though the cameras were carefully lined up, it is virtually impossible to do this task perfectly. Therefore, each stereoscopic image pair was aligned both vertically and horizontally before being displayed. The two larger displays have a 16:9 aspect ratio while the small one has a 4:3 aspect ratio. Since the original images on our database have an aspect ratio of 16:9, we cropped the left and right sides of each image to achieve a 4:3 aspect ratio for the small display.

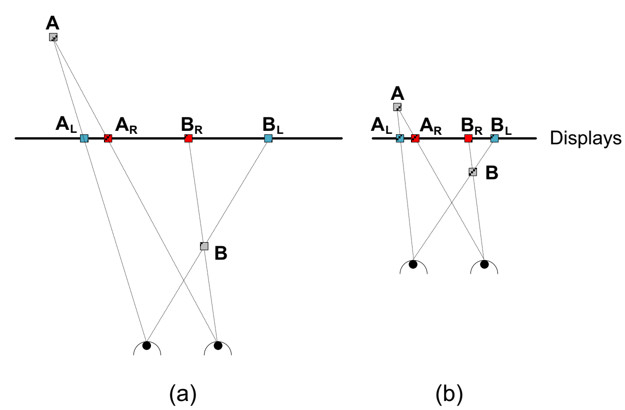

Fig. 1. The same content is shown on (a) a large display and (b) a smaller one. The perceived depth changes due to the size of the screen and the distance between the viewer and the screen.

III. SUBJECTIVE EVALUATIONS

The subjective tests were conducted on three different sizes of stereoscopic displays, namely, a 55" 3D LED TV, a 22" 3D LCD display, and a 2.8" Fujifilm 3D camera display. The detailed specifications of the three stereoscopic displays are listed in Table I.

Nineteen observers participated in our subjective tests, six females and thirteen males. Their ages ranged from 23 to 59, with an average age of 33. All subjects have none to marginal 3D image and video viewing experience. In the test, each subject was asked to view the same set of 3D images on all three displays. Following the ITU-R BT.710 recommendation [4], the viewers were seated in line with the center of the display, and at the distances that were approximately three times the height of the displays. One to three viewers

978-1-4244-8712-7/11/$26.00©2011 IEEE

participated in each viewing session. The viewers were asked to rate each image on a discrete scale from one to ten (ten being the highest rating) based on the level of quality of their experience watching the 3D content.

We set up the viewing conditions for the subjective assessment according to Section 2.1 of the ITU-R BT.500-11 [5]. Before the subjective evaluation of each display, we ran a training session to show to the observers the quality range of our stereoscopic images, without imposing the quality of the images. Images used in the training session were different from the test images. Twenty-four test images were used for the subjective tests and were shown in random order. In addition, the order in which displays were viewed was also random. The test images included indoor and outdoor scenes, different backgrounds, and various values of d_{min} (0.5, 1, 2, and 3 meters). During the test, each stereoscopic image was shown for five seconds followed by a five-second interval of a 2D mid-gray image with the image index as a grading and relaxation period.

TABLE I
PROPERTIES OF THE 3D DISPLAYS USED IN OUR TEST

Size	Type	Resolution	Glasses
55”	Samsung UN55C7000 TV	1920 × 1080	3D shutter glasses
22”	Samsung 2233RZ monitor	1680 × 1050	NVIDA GeForce 3D Vision glasses
2.8”	Fujifilm LCD screen	Approx. 230,000 dots	No glasses needed

IV. ANALYSIS AND RESULTS

Before analyzing the scores provided by the observers, we first screen the outliers according to the subjective scores they gave. The screening process is based on the guidelines provided in Section 2.3.1 of Annex 2 of the ITU-R BT.500-11 recommendation [5]. One out of nineteen observers was detected as the outlier, and all the scores of this observer were eliminated. Therefore, the data analysis that follows is based on the scores provided by the eighteen valid observers.

We compare the image qualities of the same image set on different sizes of 3D displays. We take the average score across all valid observers and across all images with the same d_{min} as the mean opinion score. To assess the credibility of the mean opinion score, we compute confidence intervals using the Student's t-tests with a 95% significance level.

Fig. 2 shows the bar graph of the quality score versus the display size for d_{min} equal to 0.5, 1, 2, and 3 meters. From Fig. 2, we observe that the 3D subjective quality on the 2.8-inch display is the worst among the three displays at all levels of d_{min}. The 22-inch monitor provides the best 3D quality when d_{min} is large (2 and 3 meters), and it has slightly inferior quality to the 55-inch TV when d_{min} is relatively small (0.5 and 1 meter). It is also worth noting that the 22-inch monitor has the smallest confidence interval among all three sizes of displays at various levels of d_{min}, meaning that the high scores for the 22-inch monitor are most consistent.

Based on the results from the questionnaire asked after each subjective test session (a variation from the one used in [6]), we notice that the majority of observers prefer to watch 3D content on medium and large size displays, i.e., the 22-inch and 55-inch displays. In fact, 83% of them explicitly said that they would not watch 3D content on a mobile device.

When asked if their opinion of 3D presentations had changed because of our test, 17% of the viewers said that their opinion was now somewhat more negative, while 33% said that their opinion had become more positive. Half the viewers said their opinion of 3D content had not changed after our test.

Our future research includes developing post processing schemes tailored for each particular display in order to significantly improve the subjective quality of 3D content. We will verify the success of our schemes with the subjective assessment of new observers.

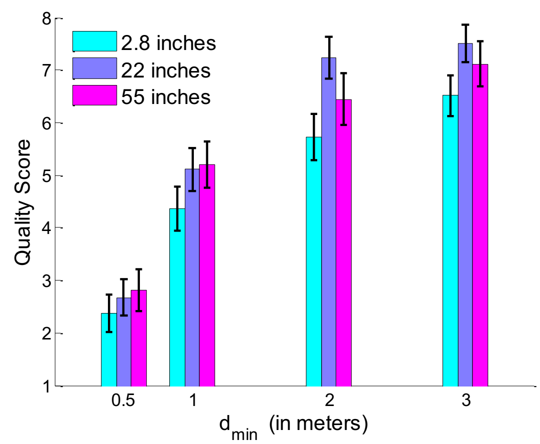

Fig. 2. The mean opinion scores and their confidence intervals versus different sizes of 3D displays at various levels of d_{min}.

REFERENCES

[1] Lachlan D. Pockett and Marja P. Salmimaa, “Methods for improving the quality of user created stereoscopic content,” *Stereoscopic Displays and Applications XIX, Proc. of SPIE-IS&T Electronic Imaging*, SPIE Vol. 6803, 680306, 2008.

[2] Lutz Goldmann, Francesca De Simone, and Touradj Ebrahimi, "Impact of acquisition distortion on the quality of stereoscopic images," *International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, Scottsdale, Arizona, USA, January 13-15, 2010.

[3] Di Xu, Lino Coria, Panos Nasiopoulos, "Guidelines for Capturing High Quality Stereoscopic Content Based on a Systematic Subjective Evaluation," *IEEE International Conference on Electronics, Circuits, and Systems, ICECS 2010*, Athens, Greece, December 12-15, 2010.

[4] ITU-R, “Subjective assessment methods for image quality in high-definition television,” Tech. Rep. BT.710-4, ITU-R, 1998.

[5] ITU-R, “Methodology for the subjective assessment of the quality of television pictures,” Tech. Rep. BT.500-11, ITU-R, 2002.

[6] Monika Pölönen, Marja Salmimaa, Viljakaisa Aaltonen, Jukka Häkkinen, and Jari Takatalo, “Subjective measures of presence and discomfort in viewers of color-separation-based stereoscopic cinema,” *Journal of the Society for Information Display*, Volume 17, Issue 5, pp. 459-466, May 2009.