Perceptual Quality of Reconstructed Medical Images on Projection-Based Light Field Displays

Peter A. Kara[1,4(✉)], Peter T. Kovacs[2,3], Suren Vagharshakyan[3], Maria G. Martini[1], Sandor Imre[4], Attila Barsi[2], Kristof Lackner[2], and Tibor Balogh[2]

[1] Wireless Multimedia and Networking Research Group, Kingston University, London, UK
{p.kara,m.martini}@kingston.ac.uk
[2] Holografika, Budapest, Hungary
{p.kovacs,a.barsi,k.lackner,t.balogh}@holografika.com
[3] Department of Signal Processing, Tampere University of Technology, Tampere, Finland
{peter.t.kovacs,suren.vagharshakyan}@tut.fi
[4] Department of Networked Systems and Services,
Budapest University of Technology and Economics, Budapest, Hungary
{kara,imre}@hit.bme.hu

Abstract. With the appearance of light field displays, users may enjoy a much more natural sensation of 3D experience compared to prior technologies. This type of autostereoscopic, glasses-free visualization allows medical applications to improve both in usability and efficiency. The high angular resolution of medical images is resource-consuming, but can only be reduced while maintaining a sufficient level of overall quality through continuous parallax. A dense image set can also be achieved by applying the synthesis of intermediate views. In this paper we provide the analysis of the effect of reduced angular resolution and image synthesis on Quality of Experience in medical applications. Two separate series of subjective quality assessment measurements were conducted with 20 participants each, one focusing on angular resolution reduction and another one comparing the effect of such reductions with the quality of reconstructed images.

Keywords: Quality of experience · Perceived quality · Medical QoE · Light field display · Angular resolution · Light field reconstruction · View synthesis · Image-based rendering

1 Introduction

Autostereoscopic displays enable 3D experience without any special head gear (e.g., 3D glasses). Unlike stereoscopic display technologies, where the number of views is exactly 2, glasses-free systems do not define such a value. Although it needs to be at least 2, no theoretical upper limit exists; the number of views is only bounded by device capabilities.

Using more views to display a specific content requires more resources but, on the other hand, an insufficient number of views may result in serious degradation of the perceived quality [1] and can completely ruin usability. Immersive visual applications –

K. Giokas et al. (Eds.): eHealth 2016, LNICST 181, pp. 476–483, 2017.
DOI: 10.1007/978-3-319-49655-9_58

such as those with medical content – necessitate high quality at the user side [2] in order to prevent certain incidents, i.e. flawed diagnosis.

In certain cases, it might not be possible to acquire enough visual inputs to support a display with sufficient image density. In case of light field displays, the reconstruction of intermediate views – also known as light field reconstruction – enables continuous motion parallax [3, 4], so that no discrete borders may appear between views.

In this paper we wish to address two research questions. The first one focuses on the relationship between angular resolution and the perceived quality of medical images, and the second one compares the effect of angular resolution reduction with light field reconstruction. While image synthesis directly affects the quality of the image, a lower number of views alters the way the image is displayed, diminishes continuous parallax and thus disturbs perception. Both may degrade the experience and impair usability, but the question is which one of these two makes a higher impact on the Quality of Experience.

The paper is structured as follows: Sect. 2 introduces the research setup of the series of measurements carried out, Sect. 3 presents the obtained results and Sect. 4 concludes our findings.

2 Research Configuration

Although the research was performed in two parts, both series of measurements used the same light field display and the test stimuli were based on the same reference stimulus (see Fig. 1). The display was a HoloVizio C80 [5, 6], a LED-based 3D projection unit with a 40° field of view. The core stimulus was the 3D still image series of a rendered human heart in 1024 × 576 resolution. During the experiment, the test participant had to make at least a slight movement to the left and the right in order to properly observe the stimuli and to witness issues with continuous motion parallax during test cases with lower angular resolutions.

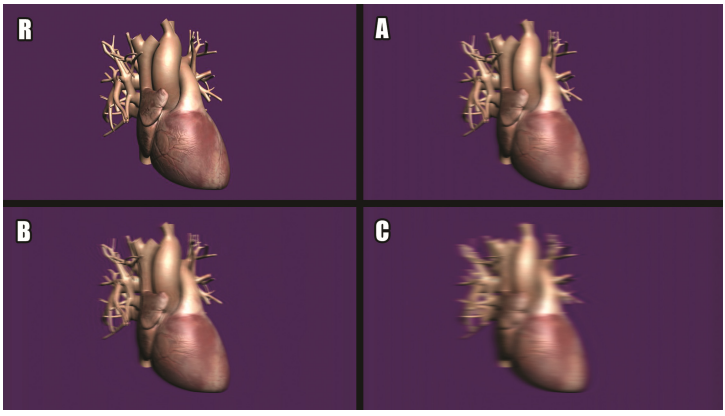

Fig. 1. Test stimuli of the research measurements. Stimulus R is there reference quality, stimulus A, B and C are the reconstructed images.

2.1 Research 1 – Angular Resolution

The first part of the research was the evaluation of angular resolution. For this, we selected 10 views, running from view number 15 to 150 with intervals of 15 (15, 30, 45 etc.). Assessment was performed on a quantitative 10-point Absolute Category Rating (ACR) [7], running from 1 to 10, where 1 was the lowest possible score and 10 was the reference quality. The subjective assessment of quality always began by showing the reference test case – which is just to be observed and not to be rated – followed by a randomized sequence of test cases. The reference test case had 150 views, thus it was identical to one of the test cases, making a hidden reference in the test measurement.

2.2 Research 2 – Light Field Reconstruction

In the second part of the research, there were six test cases in total. Three of them were identical to the ones in the previous experiment, namely those with 30, 60 and 90 number of views. The other three (see Fig. 1) were created with light field reconstruction using shearlet transform [8, 9], based on the reference image. Stimulus A was decimated by a factor of 2 (meaning that every 2^{nd} row was retained while the others were zeroed) and this value was 3 for stimulus B. In case of stimulus C, in order to generate an input with significantly inferior quality, representing a poorly designed reconstruction solution, the setup was tuned so that the maximum disparity between two images was too high for the algorithm, resulting in a high level of blur and distortion.

The algorithm used for reconstruction generated 1024 views, from which every 4^{th} (256 views) was used for the research measurement. Similarly to Research 1, the assessment was performed on a 10-point ACR scale, but the test participant had to simultaneously take into consideration the changing image quality and the number of views. Also, the measurement started with the reference quality, followed by the 6 test cases in random order.

3 Results

In both Research 1 and 2, a total of 20 test participants provided scores for the test cases. 8 of them were medical experts and 12 were non-experts. The average age of the participants was 26.

3.1 Research 1

The results of Research 1 (see Fig. 2) show the breakpoint of excellence to be at 75 views. The three lowest number of views (15, 30 and 45) cannot be considered to provide an acceptable level of quality for medical purposes. Sufficient angular resolution can be obtained at 60 views and above.

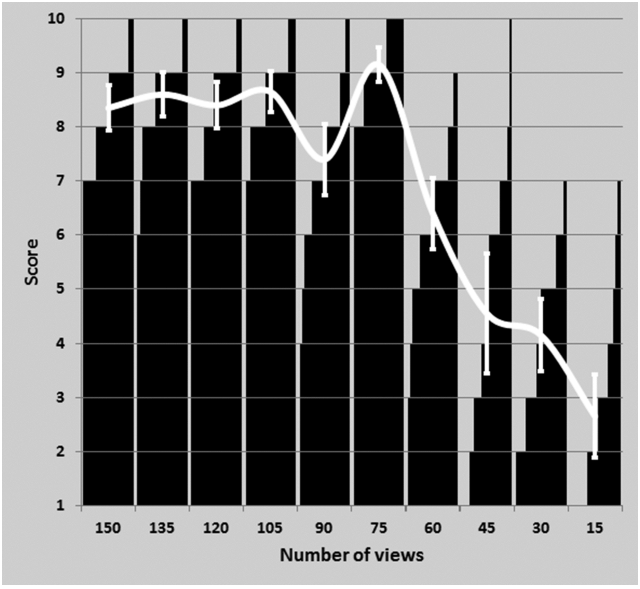

Fig. 2. Results of measurement series Research 1. The black columns provide the actual scores of the test cases, while the points with vertical bars in the white curve represent the Mean Opinion Score with confidence interval.

In case of 45 views, the upmost score extremes can be witnessed, since one particular test subject could not distinguish it from the reference quality, while another deemed it to be absolutely unacceptable. Such outliers can of course distort the mean results, not to mention other types of cognitive bias, like the avoidance of extreme values.

The reason of the sudden high scores of 75 could be the contrast effect if test cases had not been random; while for most people test case 60 provided certain perceptual artefacts, such as ghost image parts or perceivable discrete borders between images, 75

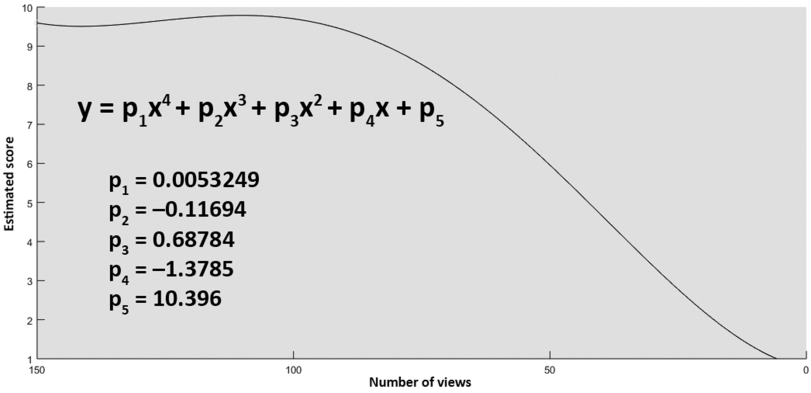

Fig. 3. Estimated Mean Opinion Score model for number of views.

had nearly none. Regarding 90, the lack of Just Noticeable Difference (JND) between 75 and 90 can create cognitive dissonance reduction [10], during which the preconception of notable differences overwrites perception [11, 12].

In case we remove outliers and inconsistent scores, we can acquire a preliminary model that maps between number of views and estimated MOS (see Fig. 3).

3.2 Research 2

Based on the ACR scores obtained for Research 2 (see Fig. 4), we can state that image quality was proven to be a more important aspect than angular resolution. The reason why most test cases of different view numbers (30 and 60) received notably higher scores than identical ones in Research 1 is that image quality was taken into consideration as well, which was the same as the reference quality. Test case 90 was assessed with a very similar quality rating compared to the top 4 number of views in Research 1. This is due to the lack of cognitive bias, because these test cases were evidently distinguishable for most.

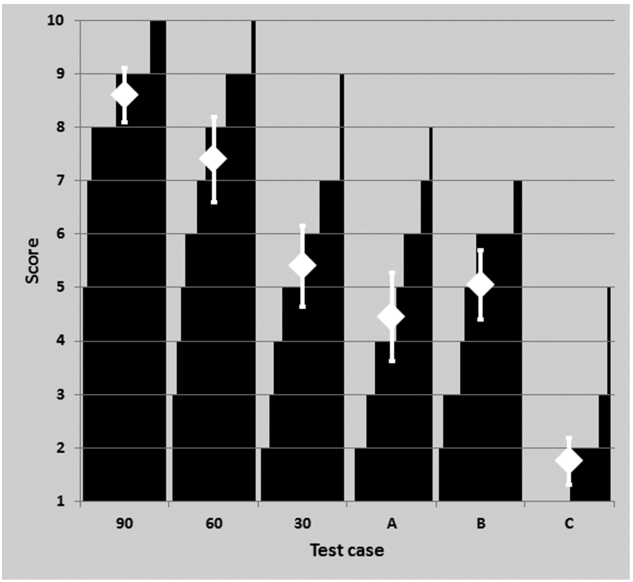

Fig. 4. Results of measurement series Research 2. The black columns provide the actual scores of the test cases, while the white diamonds are the values of the Mean Opinion Score.

As it can also be seen, the properly reconstructed images received slightly lower yet quite alike scores than test case 30, suggesting similar acceptability. However, what is rather curious is that stimulus B received higher scores than stimulus A. As a reminder, stimulus A was decimated by a factor of 2 and stimulus B by a factor of 3, meaning that stimulus A was meant to be the better one.

If we observe the score distribution of test case 30, A and B (see Fig. 5), we can see great diversity among the results. 8 test participants declared test case 30 to have a better quality than the reconstructed light fields, 5 evaluated in the opposite direction, 3 could not distinguish the overall quality of these test cases and 4 provided mixed results. With respect to test case A and B, B received higher scores 9 times, A and B were given the same scores also 9 times, and A was evaluated to be better 2 times. These differences are usually just 1 or 2, but can be even 7, on a scale from 1 to 10.

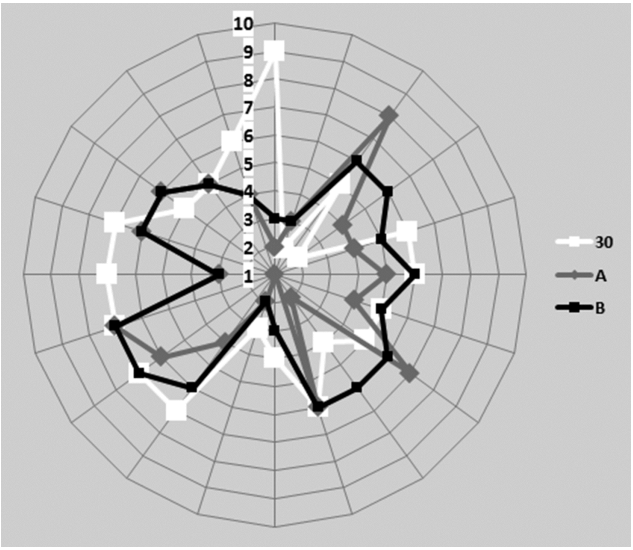

Fig. 5. Score distribution of test case 30, A and B.

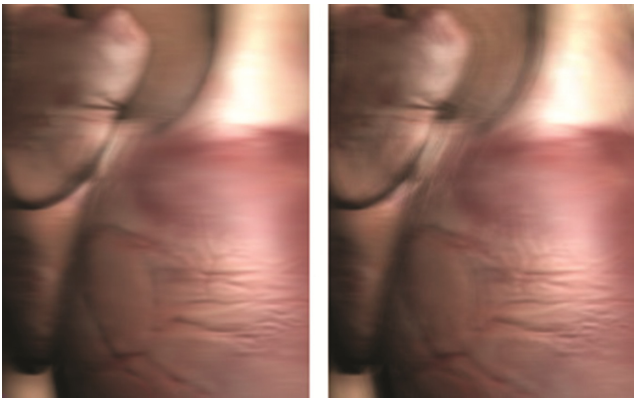

Fig. 6. Comparison of detail of stimulus A (left) and B (right).

The assessments of Research 2 favored test case B, because even though decimation by a factor of 3 provides lesser quality, some points of the stimulus actually appeared to be sharper, less blurry, due to some overlapping visual artefacts, creating higher contrast (see Fig. 6).

4 Conclusions

The paper has presented 2 series of measurements, addressing the topic of sufficient angular resolution and light field reconstruction for medical images. We found that for a light field display for 40 degrees of field of view, 75 or more views are enough to provide excellent quality. According to the collected subjective data, observers are more sensitive to degradations in texture due to view synthesis than to a lower number of views. The investigation also shows that depending on medical content, decimation by a higher factor may actually provide better overall visual experience due to the increased contrast created by overlapping artefacts. Possible continuations of these 2 researches are closely investigating the breakpoint in experienced quality for the number of views, compare them with different view synthesis methods and to utilize actual medical footage in the measurements instead of rendered ones.

Acknowledgments. The work in this paper was funded from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 643072, Network QoE-Net. The research leading to these results has received funding from the PROLIGHT-IAPP Marie Curie Action of the People programme of the European Union's Seventh Framework Programme, REA grant agreement 32449.

References

1. Ijsselsteijn, W., Seuntiens, P., Meesters, L.: Human factors of 3D displays. In: Scheer, O., Kauff, P., Sikora, T. (eds.) 3D Video Communication, pp. 219–233. Wiley (2005)
2. Butt, A., Mahoney, M., Savage, N.W.: The impact of computer display performance on the quality of digital radiographs: a review. Aust. Dent. J. **57**, 16–23 (2012)
3. Ouazan, A., Kovacs, P.T., Balogh, T., Barsi, A.: Rendering multi-view plus depth data on light-field displays. In: 3DTV Conference: The True Vision – Capture, Transmission and Display of 3D Video (3DTV-CON), Antalya, pp. 1–4 (2011)
4. Farre, M., Wang, O., Lang, M., Stefanoski, N., Hornung, A., Smolic, A.: Automatic content creation for multiview autostereoscopic displays using image domain warping. In: IEEE International Conference on Multimedia and Expo (ICME), Barcelona, pp. 1–6 (2011)
5. HoloVizio C80 3D cinema system. http://www.holografika.com/Products/NEW-HoloVizio-C80.html. Accessed February 2016
6. Balogh, T.: The HoloVizio system. In: Proceedings of SPIE, Stereoscopic Displays and Applications XIII, vol. 6055 (2006)
7. ITU-T Rec. P.910: Subjective video quality assessment methods for multimedia applications, September 1999
8. Vagharshakyan, S., Bregovic, R., Gotchev, A.: Image based rendering technique via sparse representation in shearlet domain. In: IEEE International Conference on Image Processing (ICIP), Quebec City, pp. 1379–1383 (2015)

9. Vagharshakyan, S., Bregovic, R., Gotchev, A.: Light field reconstruction using shearlet Transform. CoRR, (2015)
10. Festinger, L.: A Theory of Cognitive Dissonance. Stanford University Press, Stanford (1957)
11. Kara, P.A., Sackl, A., Mourão, M., Bokor, L.: What your phone makes you see: investigation of the effect of end-user devices on the assessment of perceived multimedia quality. In: 7th International Workshop on Quality of Multimedia Experience (QoMEX), Costa Navarino, Messinia, Greece, pp. 1–6 (2015)
12. Sackl, A., Zwickl, P., Egger, S., Reichl, P.: The role of cognitive dissonance for QoE evaluation of multimedia services. In: IEEE Globecom Workshops (GC Wkshps), pp. 1352–1356 (2012)