# Comparison of 3D video subjective quality evaluated using polarization and autostereoscopic displays[1]

K. Wegner, T. Grajek and M. Domański

This paper answers an important question whether subjective evaluation of the 3D video can be performed on one type of 3D monitor only or multiple types of displays are required in order to get reliable results. The experimental material used includes MOS data collected on two types of monitors: polarization and autostereoscopic. The paper reports analysis of this data. Observed high correlation of the gathered quality metrics clearly indicates that results obtained using polarization monitor are as statistically significant as the results obtained using autostereoscopic one. Therefore, in order to save time, only one type of monitor can be used.

*Introduction:* Currently we observe rapid development of various 3D video services. Especially MVD [1] representation of 3D video gains a lot of attention. Even current works on 3D video standardization for example within ISO/IEC Moving Pictures Experts Group (MPEG) are focused on MVD representation. As far as 3D video is concerned, there is a problem with quality evaluation. Currently there is no reliable and widely recognized objective 3D video quality measurement, like PSNR in case of 2D video. Therefore, we are forced to use subjective quality evaluation procedure. In such procedure viewers observe 3D video on 3D monitor and express their opinion. An average opinion is then used as quality metric. There are many different 3D monitors available on the market. Polarization stereo monitors and glassesless autostereoscopic monitors are the examples. Each type of 3D monitor uses different technology to provide 3D experience. The polarization monitor presents only a stereo pair whereas the autostereoscopic monitor presents many subsampled views. There is a question about influence of the given displayed technology on the obtained quality measurement. Therefore, in one of the first big worldwide experiment on 3D video subjective quality conducted on behalf of ISO/IEC Moving Pictures Experts Group (MPEG), it was decided to use as many display technologies as possible. After deep consideration two representative display technologies were chosen: polarization and autostereoscopic. Results obtained using both types of monitors were used to evaluate the quality of the given 3D video.

*Problem:* Subjective evaluation of 3D video quality on both polarization and autostereoscopic monitors is time and effort consuming. Therefore, a question arises whether evaluation of a 3D video on polarization display is comparable (gives the same statistically significant results) to evaluation of the same material on autostereoscopic display.

*Experimental material:* In order to get reliable answer to the question stated above, a wide set of 3D videos containing different kinds of distortion should be considered. Therefore, we decided to analyse data from the above mentioned worldwide experiments on 3D video subjective quality conducted on behalf of ISO/IEC Moving Pictures Experts Group (MPEG). In this experiment, 22 different 3D video coding technologies from all over the world submitted in response to Call for Proposal [2] were evaluated. In other words, 22 codecs significantly different in compression performance and introduced distortion, have been analysed. Fig 1 shows the 3-view configuration used in the evaluation process of the CfP responds, where three views along with corresponding three depth maps were fed into 3D Video Encoder (Fig. 1). Next, 3D Video Decoder produced a required number of virtual views depending on monitor type used:

- polarization stereoscopic: Hyundai, model S465D,
- autostereoscopic: 28-view Dimenco, model BDL5231V3D.

[1] This paper is a postprint of a paper submitted to and accepted for publication in Electronics Letters and is subject to Institution of Engineering and Technology Copyright. The copy of record is available at IET Digital Library

Exact positions of views produced by each decoder and fed to the displays are specified in CfP and presented in Table 1. Views generated for polarization monitor were centred over the middle coded view with a baseline of 1/4th of spacing between two neighbouring inputs/transmitted views. On the other hand, views for the autostereoscopic display were spaced every 1/16th of baseline of two neighbouring inputs/transmitted views, resulting in 28 views covering a much wider range. However, each view was heavily decimated. Quality of 3D video material produced by each proposal was subjectively evaluated [3] by over 600 people on both: polarization and autostereoscopic monitors during the formal evaluations of the proposals. It should be kept in mind that for both monitors the same coded data was used (see Fig. 1). A number of MOS (Mean Opinion Score) data was collected [4] for each proposal for HD (1920x1080) and ED (1024x768) sequences [5-10].

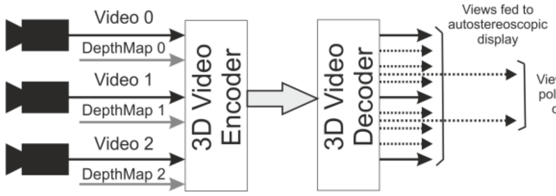


**Fig. 1** *Polarization and autostereoscopic data for 3-view configuration.*

**Table 1:** Specification of views for polarization and autostereoscopic displays.

| Test Sequence | Test Class | 3-view configuration | | |
|---|---|---|---|---|
| | | Input views | Views for stereo pair | Views for autostereoscopic display |
| Poznan_Hall2 | HD | 7-6-5 | 6.125-5.875 | All 1/16 positions between 7 and 5 |
| Poznan_Street | | 5-4-3 | 4.125-3.875 | All 1/16 positions between 5 and 3 |
| Undo_Dancer | | 1-5-9 | 4.5-5.5 | All 1/4 positions between 1 and 9 |
| GT_Fly | | 9-5-1 | 5.5-4.5 | All 1/4 positions between 9 and 1 |
| Kendo | ED | 1-3-5 | 2.75-3.25 | All 1/8 positions between 1 and 5 |
| Balloons | | 1-3-5 | 2.75-3.25 | All 1/8 positions between 1 and 5 |
| Lovebird1 | | 4-6-8 | 5.75-6.25 | All 1/12 positions between 4 and 8 |
| Newspaper | | 2-4-6 | 3.75-4.25 | All 1/12 positions between 2 and 6 |

*Methodology:* As mentioned before, we want to check if there is any relation between subjective quality assessment conducted on both types of monitors. Therefore, we have analysed correlation between results obtained on polarization monitor and results obtained on autostereoscopic monitor. For each sequence class (ED/HD) we have fitted linear regression to the MOS obtained on both monitors

$$MOS_{autostereoscopic} = a \cdot MOS_{polarization} + b \qquad (1)$$

and calculated Pearson correlation coefficient [11]. The same was done for all MOS data considered as one class (HD&ED).

However, even high correlation is not enough to claim that results obtained on both monitors are the same (give the same statistically significant results). To make such a statement, we also need to know if the relation between scores of two randomly selected test points is the same on both monitors (see Fig. 2). In other words: if the test point A outperforms the test point B on polarization monitor, will it outperform the point B on the autostereoscopic monitor as well.

The answer for such a question is given by Spearman rank order correlation coefficient [11]. Therefore, we have ranked all the results and calculated the Spearman rank order correlation coefficient.

*Results:* Fig. 3 shows a chart with MOS obtained on the polarization monitor shown on one axis, and MOS obtained on the autostereoscopic monitor shown on the other axis, both for HD & ED sequences. Additionally, 95% confidence intervals have been marked for all the test points. In Table 2 calculated correlation coefficients of MOS data collected on both types of monitors have been gathered.

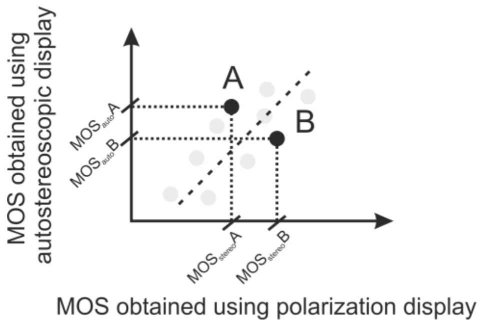


**Fig. 2** *Linear correlation problem.*

**Table 2:** Pearson correlation, Spearman rank order correlation and regression coefficients for the analysed data.

| | HD | ED | HD & ED |
|---|---|---|---|
| Pearson correlation coefficient | 0.955 | 0.975 | 0.961 |
| Spearman rank order correlation coefficient | 0.943 | 0.969 | 0.951 |
| Regression coefficient | 1.052 | 0.975 | 1.028 |

It can be noticed that results obtained on polarization monitor are well correlated with those obtained on the autostereoscopic one (Pearson correlation coefficient higher than 0.95 in all cases). Additionally, it has been proved that the relation between scores gained by two randomly selected test points is the same on both monitors (Spearman rank order coefficient higher than 0.94 – see Fig. 4).

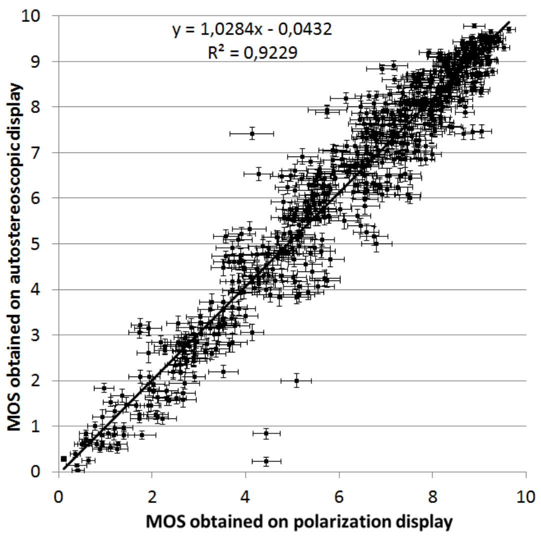


**Fig. 3** *Linear regression fitted for data for polarization and autostereoscopic monitors for all sequences.*

*Conclusion:* Our analysis showed that the results of subjective evaluation of 3D video performed on polarization monitor are highly correlated with those obtained using autostereoscopic one, even though on autostereoscopic monitor the spatial resolution of rendered views is lower. Therefore, it is worth considering to perform 3D subjective tests on one type of 3D monitor only (huge time and effort savings).

*Acknowledgments:* Research project was supported by National Science Centre, Poland according to the decision DEC-2012/05/N/ST6/03378.

K. Wegner, T. Grajek and M. Domański (Chair of Multimedia Telecommunication and Microelectronics, Poznan University of Technology, Poland)

E-mail: kwegner@multimedia.edu.pl

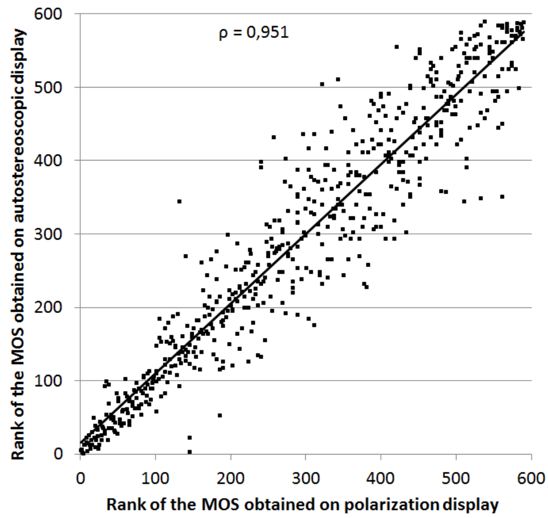


**Fig. 4** *Spearman rank order correlation for all sequences.*

**References**

1. Merkle, P., Smolic, A., Muller, K., Wiegand, T.: 'Multi-View Video Plus Depth Representation and Coding', IEEE International Conference on Image Processing, San Antonio, USA, October 2007, pp. 201-204, doi: 10.1109/ICIP.2007.4378926

2. ISO/IEC JTC1/SC29/WG11: 'Call Proposals on 3D Video Compression Technology', Doc. N12036, Geneva, Switzerland, March 2011

3. ITU-T Recommendation BT.500: 'Methodology for the subjective assessment of the quality of television pictures', 2009

4. ISO/IEC JTC1/SC29/WG11: 'Report of Subjective Test Results from the Call for Proposals on 3D Video Coding Technology', Doc. N12347, Geneva, Switzerland, November 2011

5. Domanski, M., Grajek, T., Klimaszewski, K., Kurc, M., Stankiewicz, O., Stankowski, J., Wegner, K.: 'Poznan Multiview Video Test Sequences and Camera Parameters', ISO/IEC JTC1/SC29/WG11 , Doc. M17050, Xian, China, October 2009

6. Zhang, J., Li, R., Li, H., Rusanovskyy, D., Hannuksela, M.M.: 'Ghost Town Fly 3DV sequence for purposes of 3DV standardization', ISO/IEC JTC1/SC29/WG11, Doc. M20027, Geneva, Switzerland, March 2011

7 Rusanovskyy, D., Aflaki, P., Hannuksela, M.M.: 'Undo Dancer 3DV sequence for purposes of 3DV standardization', ISO/IEC JTC1/SC29/WG11, Doc. M20028, Geneva, Switzerland, March 2011

8. Tanimoto, M., Fujii, T., Fukushima, N.: '1D Parallel Test Sequences for MPEG-FTV', MPEG M15378, Archamps, France, April 2008

9. Ho, Y.-S., Lee, E.-K., Lee, C.: 'Multiview Video Test Sequence and Camera Parameters', ISO/IEC JTC1/SC29/WG11, Doc. M15419, Archamps, France, April 2008

10. Um G.-M. , Bang G , Hur N , Kim J , Ho Y.-S. . '3D video test material of outdoor scene', ISO/IEC JTC1/SC29/WG11, Doc. M15371, Archamps, France, April 2008

11. Press, W.H., et al.: 'Numerical Recipes: The Art of Scientific Computing' (Cambridge University Press, New York, 2007, third edition)