Methodology for 3D Video Subjective Quality Evaluation

Filip Lewandowski, Mateusz Paluszkiewicz, Tomasz Grajek, and Krzysztof Wegner

Abstract—The paper presents the methodology for 3D video subjective quality evaluation. Described methodology was designed to compare different 3D video compression technologies without an influence of any particular displaying or rendering technology. In addition detailed step by step description of test session design and preparation is provided. Experimental results for state-of-the-art 3D encoders are also included. All test were conducted on two 3D monitors (polarization and autostereoscopic) thus influence of different displaying technologies on 3D video quality assessments has also been evaluated.

Keywords—3D video, video compression, subjective quality assessment.

I. Introduction

CURRENTLY, a rapid development of various kinds of 3D television services can be observed. Stereoscopic television in which a viewer can watch 3D images is already being deployed on the market. Stereoscopic monitors are widely available to the customers. Moreover, autostereoscopic (i.e. without a need to use dedicated glasses) displays are under extensive development and some of them are also available on the market. Even at this moment users are using mobile devices with 3D glassesless displays. First freeview television services are currently under study. These are services in which a user can choose the point from which the scene will be seen.

For all the abovementioned 3D television services rendering or synthesis of an intermediate views based on 3D scene representation is required. Currently, the most commonly used technique for such purposes is multiview and depth representation along with Depth Image Based Rendering (DIBR) [1]. However, in the literature many different formats were proposed, for example Layer Depth Images, Warps [2].

Essential issue in developing new compression technology is reliable quality assessment. Quality assessments are very important because it is highly desirable to balance compression performance versus provided quality.

Image quality can be described in many ways e.g as an integrated set of factors determining the overall degree of image perfection. The application used strongly affects the factors to be evaluated. For instance, medical images will be judged for fidelity to the original one. However, in case of television services, quality is understood in the context of the possible occurrence of distortion in an image.

The most simple and straightforward way to assess quality of an image is to ask viewers about their opinion (to score

an image). Because this kind of quality assessment depends on users' opinion, it is called subjective. In order to eliminate influence of individual user deviations on the assessment, e.g. likes and dislikes, opinions (scores) should be averaged on a wide group of people. Such a measure is called Mean Opinion Score (MOS). Many different procedures of subjective quality assessments have been developed over the years. They differ in:

- an object of the assessment (e.g. quality, distortion, fidelity of the image),
- test conditions (e.g. with or without reference),
- data processing (statistical analysis).

In order to obtain reliable results, the procedure of assessments should be precisely designed and described. In practice, commonly used methodology for the subjective assessment of 2D image quality is the one from recommendation BT.500 [3]. Subjective quality assessments base on real user experience therefore they are the most reliable approach to judge the real quality of an image. Unfortunately, they require involvement of many people and are very time and cost consuming. In order to eliminate the need of people participation in the assessments, many automatic quality metrics have been developed. Because automatic assessment are independent from individual user opinion, there are called objective. The most simple and most commonly used objective quality metrics are Peak Signal to Noise Ratio of the luminance (PSNR-Y) and Mean Square Error (MSE), also calculated on luminance. Their main advantage is simplicity (i.e. they can be easily computed, processed and compared). However, they can be considered only as a rough approximation of the real quality of the image. In literature, some more sophisticated objective metrics can be found, e.g. SSIM - Structure Similarity [4] or JND - Just Noticeable Difference [5]. But they still have many weaknesses (e.g. a limited number of analyzed distortion, high computational complexity).

Because we observe the rapid development of various kinds of 3D television services, there is a strong need for reliable quality measurement procedure for 3D sequences.

Our main goal was to compare two or more compression technologies for 3D sequences. Due to the lack of good and reliable procedure of quality assessments for 3D video content, we had to develop appropriate quality assessment methodology for such a case.

II. Proposal

3D video quality assessment is still an open issue. Currently many different representation of 3D video data are

Filip Lewandowski, Mateusz Paluszkiewicz, Tomasz Grajek, and Krzysztof Wegner are with the Chair of Multimedia Telecommunications and Microelectronics, Poznan University of Technology, Poznan, Poland (e-mail: tgrajek; kwegner@multimedia.edu.pl)

used. Some of them use additional supplementary data to describe 3D scene, for example depth maps, invisible to a user but strongly affecting 3D video quality. Additionally, various technologies for displaying 3D content are currently in use. Therefore, a question arises what should be assessed in order to get reliable quality comparison of various 3D video compression technology. We propose, with the spirit of subjective quality assessment, to judge quality of views/images presented directly to the end user. Many new 3D video systems utilize synthesized views, therefore, we propose to assess quality of synthesized views displayed to the viewer.

We have developed original methodology for 3D video quality assessment based on BT.500 recommendation. BT.500 is a series of recommendations for 2D video quality assessment. A 3D material presented to the viewer is a composition of two or more 2D images.

BT.500 describes a wide variety of methods that may be used in video quality assessments. They can be divided into two basic categories:

- without a reference image - where viewer never sees the original undistorted images,
- with the reference images - where viewer is asked to compare distorted images with undistorted original.

Methods that use the original undistorted image (if available) always give better results. The definition of undistorted original image applicable in case of synthesized view has two meanings:

- Reference image - obtained from real camera at spatial position of a synthesized view. Obtaining such a view is not always possible, for example when cameras are placed too close to each other there is no possibility to put another one between them. Also in case of already recorded material there is no possibility of capturing additional views.
- Rendered reference image - it means rendered from undistorted original data. It is always possible to use undistorted original data to synthesize a given view and use it as a reference. Additionally, in this way we abstract from possible distortions caused by rendering technology and we are able to focus only on assessing the influence of the used compression technology on 3D video quality.

In practice we want to assess 3D compression technology without the influence of particular displaying or rendering technology used, therefore the case with rendered reference is more appropriate and so it has been used.

BT.500 recommends to use 5-point grading scale to express video quality. We have found that in case of 3D video it is insufficient to reflect full spectrum of viewer experience, therefore we propose to use 11-point scale which better differentiates 3D image quality.

In our experiments 3D video fidelity to the reference video was measured.

III. SUBJECT SELECTION

WHEN overall subjective quality of the image is to be measured, it is preferable to perform the tests on a whole population, but of course it is impossible. To overcome this problem, usually statistical analysis is incorporated, where only a limited number of subjects (observers) is involved. The main assumption is that this limited number of subjects is a representation of the entire population. Therefore, the proper sampling (i.e. selection of subjects) should be ensured. According to BT.500 subjects should have no expertise in assessing the images quality and in digital image/video compression. Moreover, subjects should be in age of 18-30 years, because vision system of people in this age range is in optimal condition.

Standard BT.500 recommends also that, prior to a test session subjects should be screened for normal visual acuity (Snellen or Landolt charts) and proper color perception by Ishihara plates for instance. However, BT.500, which is dedicated for evaluation of 2D images, does not include depth perception test. Such a test is crucial when stereoscopic or 3D image/video is evaluated. The following depth perception test is proposed. Two squares of exactly the same size (subjectively) and color at different depths are shown on the screen. An observer has to point out the closer one. Because squares are the same and the only difference is the depth, a subject's depth perception is examined. Depth perception test should be repeated several times with squares randomly placed in depth direction. Failure in any of the abovementioned tests excludes the subject from participation in quality assessment.

Concluding, a subject should be rather young person, who will assess the presented video sequences according to personal feelings about their quality.

IV. SESSION CONSTRUCTION

IN order to obtain statistically reliable results, each test session/examination has to be precisely designed and carried out. One test session consists of some number of test points which are presented one by one to the subjects. As a test point a pair of video sequences is considered. In our case we have the reference view rendered from uncompressed data and the processed sequence, which is the object of study.

Based on recommendation BT.500 each test point should be presented to the subject in the following manner (see Table I):

- First, for about 3 sec, the number of evaluated test point (i.e. first, second, etc.) at mid-grey background is presented - T1.
- Second, a reference sequence (in our case views rendered from uncompressed data) is shown - T2.
- Then, for about 3 sec, mid-grey screen is displayed - T3.
- Next, the assessed sequence is shown (views of one of the sequences rendered from data compressed using evaluated technology) - T4.
- Finally, for about 5 sec, again mid-grey screen is presented. This is the time for voting (an observer gives score for the viewed test point) - T5.

Concluding, to assess one test point at least $T_P = 31$ seconds is needed.

In practice, the obtained results are considered as statistically significant if level of significance α is below 5%. To fulfill this requirement the appropriate number of scores for

TABLE I
SINGLE TEST POINT DESIGN

Name	Length [s]	Type of sequence
T1	3	Grey screen
T2	at least 10	Reference
T3	3	Grey screen
T4	at least 10	Tested
T5	5	Grey screen

one test point has to be collected. This number of votes may be estimated based on the following equation:

$$n = \frac{t_\alpha^2 s^2}{d^2} + 1 \qquad (1)$$

where n is the needed number of scores for a single test point, t_α is a quantile of Student's $t-$distribution, s is a standard deviation of scores and d is the confidence interval. Even though the standard deviation s is unknown, it may be easily estimated based on small preliminary viewing session. Finally, the necessary number of scores for assumed confidence level may be estimated.

At the beginning and at the end of each test session additional k test point should be introduced. Scores for those test points have to be discarded due to a fact that at the beginning subjects learn how to assess the quality of the presented material whereas at the end subjects start being bored, distracted and their scores again may fluctuate.

Moreover, additional l test points have to be added to test session to check how repetitive the scores given by a single subject on the same test point are (the so called consistency test). It means that one selected test point is repeated l - times at random positions in the test session. This is done because individual scores of a subject may deviate significantly from one test point to the other. If confidence intervals of the average scores of the same test point overlap, results for this subject are considered consistent. Otherwise, results of the subject must be rejected.

According to recommendation BT. 500 each person should assess video material individually, but it is very time and cost consuming. To cppe that, quality assessments may be also conducted in groups. However, it should be noted that each person in the group has to be offered identical conditions of observations. This is the main reason why, especially in small rooms, the number of people in such a group is limited. Therefore, if there are more subjects than test room limitation allows, they have to be divided into smaller groups and test session must be repeated for each group separately. In such situation there is a risk that the contextual effect will be observed (i.e. one of the test points in a given order may affect assessment of the next test point). To eliminate this undesirable contextual effect, an order in which test points are presented should be different in each session.

Another very important issue which can significantly affect the evaluation is the human eye fatigue and loss of subject's attention due to watching sequences of similar content. Thus, if the total duration time of the single test session exceeds 30 minutes (so called people focus time T_f), it has to be divided into shorter subsessions. In this case we can not be sure that results gathered in one subsession may be compared with all others separately. In order to check that, m additional test points have to be added to each subsession (each repeating test point from a different test subsession). Thus, some number of test points from one subsession will be presented and scored in another one and it will give the opportunity to check whether viewers give the same score to this repeated test point. This kind of test is called session overlapping test. Analyzed test sessions shall be considered consistent if the confidence intervals of the average scores of the same test point from different test sessions overlap.

Knowing the maximum people focus time T_f and the number of different test points N we developed a formula to calculate necessary number of test session x which satisfies all of the abovementioned conditions.

$$x > \frac{n \cdot T_P}{T_f - (m + l + 2 \cdot k) \cdot T_P} \qquad (2)$$

where T_P is a single test point duration time and m, k and l are numbers of additional test points added as described above.

Prior to each session, subjects should be carefully introduced to the method of assessment, the grading scale, the sequence and timing (reference picture, grey, test picture, voting period). Also the type and range of the impairments to be assessed should be illustrated on images (rather different from those used in the tests, but of comparable sensitivity). It must not be implied that the lowest subjective grade corresponds to the worst quality seen during the introduction. Viewers should be asked to assess the overall impression given by the image/video and express their judgments with words used to define the subjective scale [3].

V. EXPERIMENTS

A. Test sequences

All the tests have been conducted on four 3D FullHD test sequences [6]–[8]. These sequences are recommended by an international expert group MPEG (Moving Picture Experts Group) which develops standards for coding audio and video as the official multiview test sequences. They are also used worldwide in researches on processing, compression and quality evaluating. Table II provides a brief summary of the sequences used in our subjective quality assessment tests. Sequences have an average duration of 10 sec. which means on average T_P = 31 seconds for each test point to present. The data was taken from the Call for Proposals (CfP) on 3D Video Coding Technology [9] announced by International Organisation for Standardisation (ISO) in March 2011. This CfP was an invitation to propose 3D Video Coding (3DVC) technology providing efficient compression and high quality view reconstruction of an arbitrary number of dense views.

B. 3D video compression technology used

We have used six different 3D video compression algorithms and associated coders in order to evaluate the proposed methodology.

TABLE II
3D TEST SEQUENCES

Name	Length	Type of sequence	Supplier
PoznanHall2	8s	natural	Poznan University of Technology
PoznanStreet	10s	natural	Poznan University of Technology
Dancer	10s	synthetic	Nokia Corporation
GTFly	10s	synthetic	Nokia Corporation

We have chosen a state-of-the-art techniques based on HEVC (High Efficiency Video Coding) technology. HEVC is a draft standard for 2D video compression, a successor of MPEG-4 AVC/H.264 (Advanced Video Coding) [10], developed currently jointly by ISO/IEC Moving Picture Experts Group (MPEG) and ITU-T Video Coding Experts Group (VCEG).

Among those techniques we used five coders developed by tream from Poznan University of Technology [12], which are various modifications of their 3D codec prepared in response to Call for Proposals document on 3D Video Coding Technology [9]. Poznan University of Technology 3D codec was one of two top ranked proposals. Currently this coding technology is under standardization within the ISO/IEC and ITU-T [14]. First of the coders used is the original codec submitted in response to CfP and it is further referred to as Poznan 3D Coder. We have used also four modifications of the original Poznan proposal:

- Poznan 3D Coder with Residual Layer Coding off,
- Poznan 3D Coder without Residual Layer added,
- MV-HEVC + Disoccluded Region Coding [13],
- HEVC + Nonlinear Depth Representation.

The last codec used was 2D HEVC codec used for 3D Video in simulcast mode and referred to as HEVC Simulcast.

C. Used monitors

We have chosen two most popular 3D display technologies used nowadays: polarization stereoscopic display and glassless autosteroscopic display. We have chosen best available 3D monitors on the market:

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

The polarization monitor (fig. 2) displays two interlaced views of the scene, each with different polarization of the light. This way each of the shown views have only half of the vertical resolution. The autostereoscopic monitor (fig. 1) shows 28 views at once on Full HD matrix, resulting in the possibility of viewing 3D images without any glasses but every view is subsampled by factor of square root of 28 at each direction. Such two different display technologies give us an ability to evaluate influence of display technology on 3D video quality assessment.

Fig. 1.   28-view DIMENCO autostereoscopic monitor

Fig. 2.   Hyundai polarization monitor

D. Coded material preparation

If the coding efficiency of the investigated technology/-ies is to be evaluated, results for wide range of bitrates have to be gathered. Therefore, in our tests, three views along with three depth maps of each test sequences were encoded with all six of the encoders at some predefined bitrates. The ranges of the bitrates were chosen in a way that visual quality is equally distributed from low to high. A given sequence coded with a given encoder at a given bitrate defines single test point.

For each test point, based on the decoded material, the stereo pair at a spatial position located in between of spatial positions of the compressed views were rendered (fig. 3). In order to avoid optimization of the encoding technology on a given stereo pair, exact spatial position was selected randomly. For the autostereoscopic display 28 dense spaced views were rendered at exactly the same spatial position (center of 28

TABLE III
BITRATES USED FOR TESTS

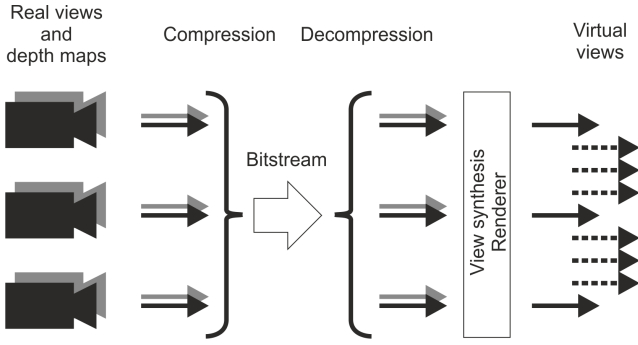| Name	Bitrate[kbps]			
PoznanHall2	140	210	320	520
PoznanStreet	280	480	800	1310
Dancer	290	430	710	1000
GTFly	230	400	730	1100

Fig. 3.   3D video compression scenario used in our experiments

views) as randomly selected stereo pair. In order to properly display rendered sequences on the autostereoscopic monitor, those 28 views were interleaved with software provided by display manufacturer.

Finally, a video file ready to display on appropriate display device was obtained for each test point.

E. Conducted tests

Prior to the tests, the necessary number of subjects have to be estimated using formula 1. Therefore, preliminary test session on 16 subjects were conducted in order to estimate population variance. Based on this session the variance was estimated as $s^2 = 6.693$. The confidence intervals of $d = 0.55$ were assumed in our tests as a trade off between reliability and necessary number of subjects. It was calculated as a 5% of 11-point scale. Finally, for assumed significance level $\alpha = 0.05$ (see Section IV), the necessary number of subject $n = 60$ was estimated.

In our tests we have $I = 4$ sequences, $C = 6$ various encoders and $B = 4$ bitrates which gave $N = 96$ different test points with an average duration of 31 seconds each. Thus, the total presentation time for all test points was approximately 49min 36sec ($N \cdot T_P$). This is much more than people focus time T_f ($T_f = 30min$ as mentioned earlier).

Therefore, in order to ensure equal test conditions for all subjects and gather $n = 60$ scores for each test point (see Section IV), the following steps were taken:

1) At the beginning, the number of test sessions has been calculated using equation 2. For the given data $x > 1.92$ was calculated, so it was decided that the number of test sessions should be equal 2.

2) All test points have been randomly divided into two test sessions. This resulted in 48 test points per test session. Each session lasted on average 25 minutes.

3) For each test session $2 \cdot k = 4$ test points were randomly selected from all test points available (N) and half of them was put at the beginning and the other haft at the end of each test session.

4) For consistency test, $l = 2$ test points have been randomly selected from each test session and repeated at random positions in the same test session.

5) For overlapping test, $m = 2$ test points have been randomly selected from all test points and added to all test sessions.

6) Because our test room can accommodate only 10 people assuring identical viewing conditions for all viewers, subjects have been randomly divided into 6 groups. Each group viewed its own version of tests sessions (test points have been randomly ordered in each session). In other words steps from 2 to 4 have been redone 6 times. This resulted in 6 groups of tests sessions with two test subsessions in each group.

7) Finally, all tests sessions were repeated separately on 2 different 3D monitors (polarization and autostereoscopic). Concluding, a single subject has taken part in 4 test sessions (two sessions on two monitors).

Prior to each test session, training of subjects has place. It contained an explanation of the session's structure, together with exemplary sequences with the high and low quality. The instruction how to fill specially prepared sheets for assessing sequences was presented. The subjects were also informed about the moment for giving the score and how much time they have to evaluate the image quality.

VI. RESULTS

AFTER collecting all of the scores from the subjects participating in test sessions, we have performed session overlapping and session consistency tests. Both tests showed
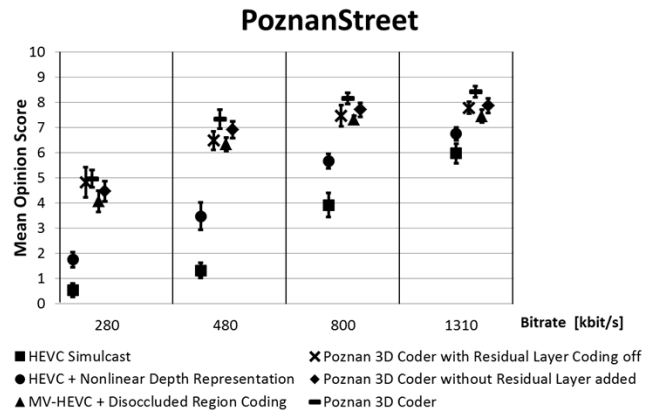

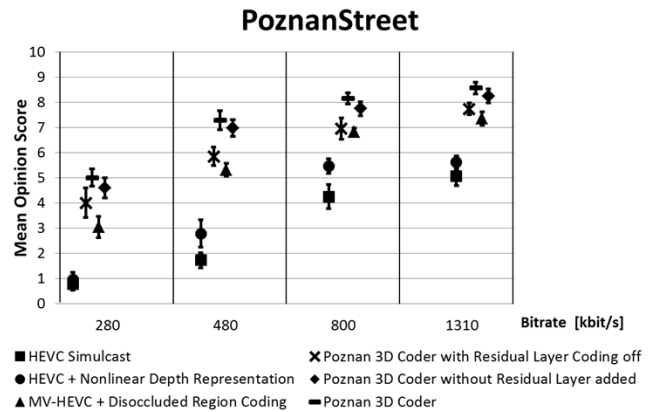Fig. 4.   Results for Poznan Street sequence obtained on polarization monitor

Fig. 5.   Results for Poznan Street sequence obtained on autostereoscopic monitor
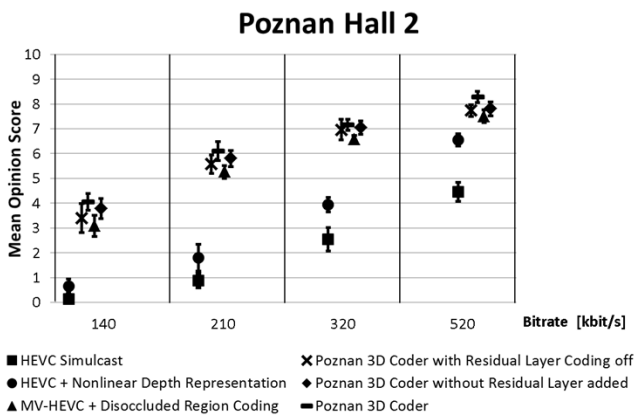
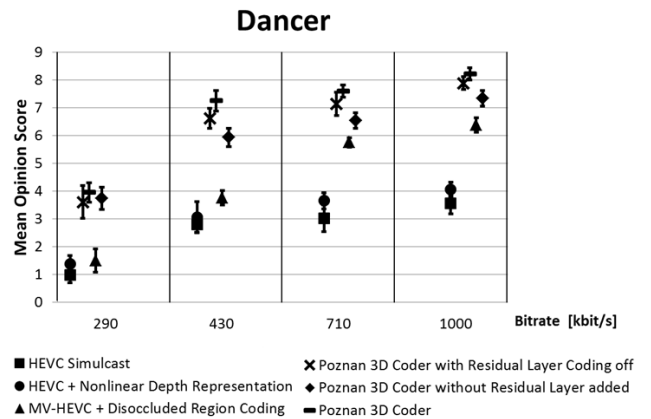Fig. 6. Results for Poznan Hall 2 sequence obtained on polarization monitor

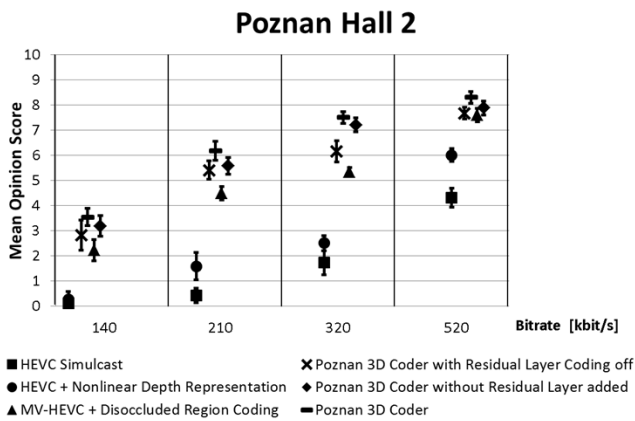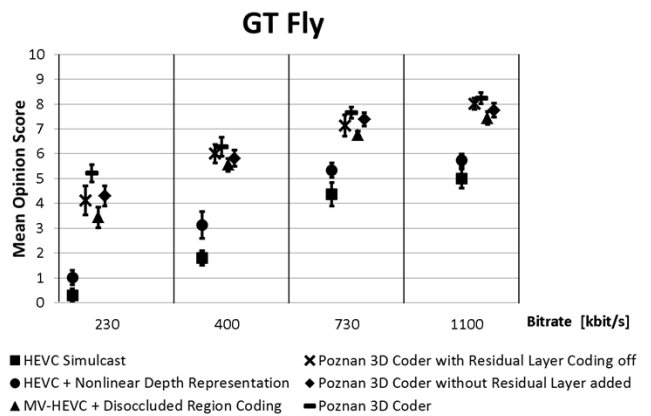Fig. 7. Results for Poznan Hall 2 sequence obtained on autostereoscopic monitor

Fig. 8. Results for Dancer sequence obtained on polarization monitor
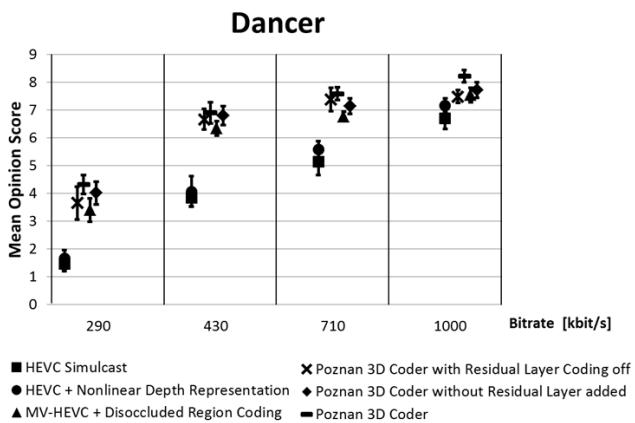

Fig. 9. Results for Dancer sequence obtained on autostereoscopic monitor

Fig. 10. Results for GT Fly sequence obtained on polarization monitor
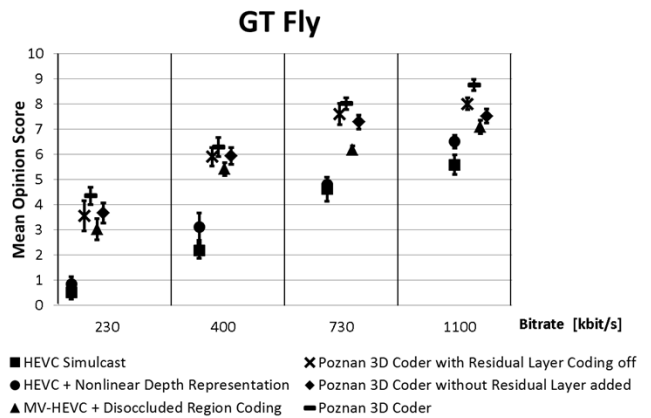

Fig. 11. Results for GT Fly sequence obtained on autostereoscopic monitor

no need to reject any outlier scores. Based on the obtained scores we have calculated average score (mean opinion score) and confidence interval for every test point. We have assumed significance level $\alpha = 0.05$. Designing the test sessions we have assumed confidence interval of 0.550. Once all of the scores were summarized the average confidence interval was 0.337 (0.335 on polarization monitor and 0.339 on autostereoscopic), which is much better than we have expected.

Results for particular sequences are presented on Fig. 4, 6,

8, 10 (obtained on polarization monitor) and on Fig. 5, 7, 9, 11 (obtained on autostereoscopic monitor).

Drawing conclusions based on raw MOS data can lead to misleading results. It is quite common that for some sequences one 3D codec is better while for different ones it is superb. The confidence intervals also have to be taken into account. Only if confidence intervals do not overlap it can be concluded that one result is better that another one, which means that it is statistically significantly better. For all those reasons we
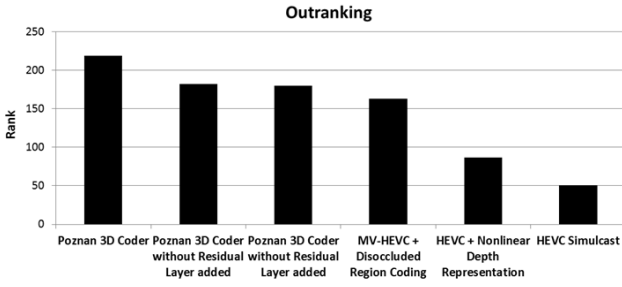
Fig. 12. Outranking chart for results obtained on polarization monitor
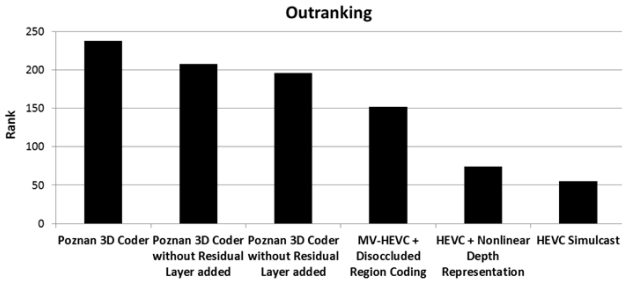

Fig. 13. Outranking chart for results obtained on autostereoscopic monitor

Fig. 14. Correlation bettween results obtained on polarization and autostereoscopic monitor

propose to use the so-called outranking charts [15] for results summary. Outranking chart is constructed in a way that all the codecs at each test point are compared to each other and if one result is statistically significantly better (it means its confidence intervals do not overlap) than the other result it is ranked 1. All ranks for a given codec are summed up and plotted on a chart, so that outranking chart informs how many times a given codec/technology is statically significantly better than the all others.

The summary of our experiments in a form of an outranking chart is shown on Fig. 12 and Fig. 13 for polarization and autostereoscopic monitors respectively. From charts it can be easily seen that Poznan 3D Coder outperformes all others used in our evaluation. Next 3 coders are comparable to each other because their ranks are similar. Finally the worst performing codec in our evaluation is HEVC Simulcast as one could have expected.

It is worth to mention that this ranking of the codecs used is independent from display technology used. In order to prove that results obtained on various monitors are consistent with each other we have computed correlation between results obtained on polarization monitor and results obtained on autostereoscopic monitor. Fig. 14 shows a chart where on one axis we have MOS obtained on the first monitor while on the another MOS obtained on the latter monitor. All test points were marked with 95% confidence intervals. We have fitted linear regression to all results

$$MOS_{autostereo} = a \cdot MOS_{polarization} + b \qquad (3)$$

It can be seen that all of the results are well correlated with each other, and the Pearson correlation coefficient is $r = 0.953$. In order to estimate monotony and consistency,
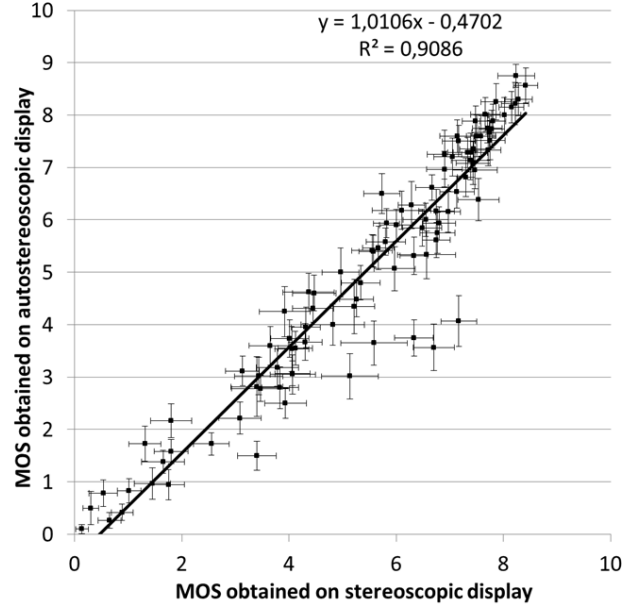
we have ranked our results (fig. 15) and the Spearman rank order correlation coefficient have been calculated $\rho = 0.957$

It proves that the proposed methodology is independent from a display technology used and gives the same results on wide range of 3D monitors regardless a display technology used.

VII. Conclusion

We have proposed the methodology of subjective quality assessment for 3D video sequences derived from BT.500 recommendation. The proposed methodology was designed to compare different 3D video compression technologies without an influence of any particular displaying or rendering technology. In addition detailed step by step description of test session design and preparation was provided. The proposed methodology assures lowest session time possible to obtain results with the assumed accuracy (in terms of confidence intervals).

We proposed to summarize subjective quality results with outranking chart, which gives clear ranking of the cases under comparison.

Experimental results performed with the use of state-of-the-art 3D coders proves high accuracy of the proposed methodology. High correlation of the results obtained on two different monitors representing currently commonly used 3D

TABLE IV
SUMMARY OF CORELATION BETWEEN RESULTS OBTAINED ON
POLARIZATION AND AUTOSTEREOSCOPIC MONITORS

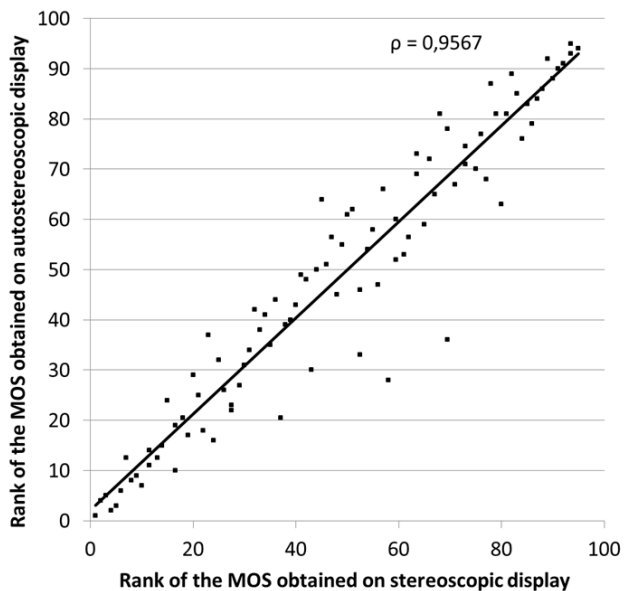| Corelattion metric	
Pearson correlation coefficient	0.953
Spearman rank order correlation	0.957
Regression coefficient	1.011

Fig. 15. Spearman rank order correlation bettween results obtained on polarization and autostereoscopic monitor

display technology proves that the presented methodology is independent from any particular rendering or displaying technology.

ACKNOWLEDGMENT

REFERENCES

[1] C. Fehn, *Depth-image-based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV*, in Proc. SPIE Conf. 5291, CA, U.S.A., pp. 93-104, Jan. 2004.

[2] J. W. Shade, S. J. Gortler, L.-W. He, and R. Szeliski, *Layered depth images*, in Computer Graphics, vol. 32, Annual Conference Series, pp. 231-242, Jul. 1998.

[3] ITU-R Rec. BT.500-11, *Methodology for the Subjective Assessment of the Quality of Television Pictures*, 2002.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, *Image quality assessment: From error visibility to structural similarity*, IEEE Trans. on Image Processing, vol. 13, pp. 600-612, Apr. 2004.

[5] D.-F. Shen and L.-S. Yan, *JND Measurements and Wavelet-based Image Coding*, Proceedings of SPIE on Input/Output and Imaging Technologies, Vol. 3422-19, July, 1998.

[6] M. Domanski, T. Grajek, K. Klimaszewski, M. Kurc, O. Stankiewicz, J. Stankowski, and K. Wegner, *Poznan Multiview Video Test Sequences and Camera Parameters*, MPEG 2009/M17050, Xian, China, October 2009.

[7] J. Zhang, R. Li, H. Li, D. Rusanovskyy, M. M. Hannuksela, *Ghost Town Fly 3DV Sequence for Purposes of 3DV Standardization*, MPEG 2011/M20027, Geneve, Switzerland, March 2011.

[8] P. Aflaki, D. Rusanovskyy, M. M. Hannuksela, *Undo Dancer 3DV Sequence for Purposes of 3DV Standardization*, MPEG 2011/M20028, Geneve, Switzerland, March 2011.

[9] *Call for Proposals on 3D Video Coding Technology*, ISO/IEC JTC1/SC29/WG11 MPEG2011/N12036, Geneva, Switzerland, March 2011.

[10] ITU-T and ISO/IEC JTC 1, *Advanced Video Coding for Generic Audiovisual Services*, ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG-4 AVC), 2011.

[11] M. Domanski, J. Konieczny, M. Kurc, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, *3D Video Compression by Coding of Disoccluded Regions*, IEEE The International Conference on Image Processing (ICIP), Orlando, USA, 2012 - in print.

[12] M. Domanski, T. Grajek, D. Karwowski, K. Klimaszewski, J. Konieczny, M. Kurc, A. Luczak, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, *New Coding Technology For 3D Video With Depth Maps As Proposed For Standardization Within Mpeg*, 19th International Conference on Systems, Signals and Image Processing (IWSSIP), Vienna, Austria, 11-13 April 2012.

[13] M. Domanski, J. Konieczny, M. Kurc, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, *3D Video Compression by Coding of Disoccluded Regions*, IEEE The International Conference on Image Processing (ICIP), Orlando, USA, 2012.

[14] G. Tech, K. Wegner, Y. Chen, S. Yea, *3D-HEVC Test Model Description draft 1*, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 Doc. JTC2-A1005, 1st Meeting: Stockholm, Sweden, 16-20 July 2012.

[15] C.A. Floudas, P.M. Pardalos (eds.), Kluwer *Outranking Methods*, in *Encyclopedia of Optimization*, 2001, vol. 4, 249255.

Krzysztof Wegner received the M.Sc. degree from Poznan University of Technology in 2008. Currently he is working toward his Ph.D. there. He is co-author of several papers on free view television, depth estimation and view synthesis. His professional interests include video compression in multi viewpoint systems, depth estimation, view synthesis for free view television, face detection and recognition. He is involved in ISO standardization activities where he contributes to the development of the future 3D video coding standards.

Tomasz Grajek received his M.Sc. and Ph.D. degrees from Poznan University of Technology in 2004 and 2010 respectively. At present he is an assistant at the Chair of Multimedia Telecommunications and Microelectronics. He is author and co-author of several papers on digital video compression, entropy coding and modeling of advanced video encoders. He has been taking part in several projects for industrial research and development projects. His current researches encompass implementing and optimizing video codecs and modeling of advanced video coders.

Filip Lewandowski received the B.Sc. degree from Poznan University of Technology in 2011. Currently he is working toward his M.Sc. there. He is interested in multiview video compression.

Mateusz Paluszkiwicz received the B.Sc. degree from Poznan University of Technology in 2011. Currently he is working toward his M.Sc. there. His professional interests include video compression.