VALIDATION OF A NEW FULL REFERENCE METRIC FOR QUALITY ASSESSMENT OF MOBILE 3DTV CONTENT

Lina Jin, Atanas Gotchev, Atanas Boev, Karen Egiazarian

Department of Signal Processing, Tampere University of Technology
Korkeakoulunkatu 10, FI-33720, Tampere, Finland
email: firstname.lastname@.tut.fi

ABSTRACT

3D video is expected to provide an enhanced user experience bringing more realism through depth impression. Therefore, quality assessment plays an important role when designing and optimizing 3D video processing methods and applications. In this paper, we propose and validate a novel quality metric based on binocular human visual system. Both visual distortions in the cyclopean view and perceptibility of depth are considered. Visual distortions are analysed in 3D-DCT domain taking into account the masking effects of the contrast sensitive function (CSF) and depth variability. The metric is especially tailored for mobile 3DTV applications as it takes into account the target display size and resolution and calculates disparity variations in local areas with relevant foveal size. The metric is designed by matching the mean opinion scores resulted from large-scale subjective tests with 3D videos with varying depth presence and amount of compression artefacts. It is further validated over another 3D video database with wider types of compression artefacts. The results show that the metric outperforms current metrics over different 3D video formats and compression methods.

1. INTRODUCTION

Modern 3D video systems are expected to deliver high quality video plus a natural sensation of depth. Such a system mainly includes the stages of 3D scene capture, format conversion, encoding, transmission, possible post-processing at the receiver side, and rendering on 3D display. Each stage may cause degradation of the 3D visual quality and errors occurred at certain step may propagate through the chain. Consequently, quality assessment (QA) is a key factor when designing or optimizing of 3D video processing systems.

Objective 3D video QA is more complex than its conventional 2D counterpart. Along with trivial 2D visual distortions it has to take into account 3D visual effects related with the 3D display size and displaying technology which in turn are related with the effects of visual discomfort, being characteristic for stereo video [1], [2]. Furthermore, the quality is closely related with the representation format and corresponding coding methods. Currently, within the scope of Mobile3DTV project several format have been considered including simulcast, multi-view video coding (MVC), mixed resolution stereo coding(MRSC), and video plus depth (V+P) [3], [4].

Early attempts to quantify 3D videos had been based on the use of 2D metrics. That is, each channel of a stereo video is evaluated by some 2D metric and then the overall 3D video quality is calculated as the mean of two video channels. This approach, however, hardly corresponds to the binocular mechanisms of the human visual system (HVS) and thus weakly correlate with subjective quality scores. Therefore, inclusion of some 3D factors to the quality evaluation has been attempted. In [5], a monoscopic quality component and stereoscopic quality component for measuring stereoscopic image quality have been combined. The former component assesses the trivial monoscopic perceived distortions caused by blur, noise, contrast change etc; while the latter assesses the perceived degradation of binocular depth cues only. In [6], the popular 2D image quality metric called structural similarity index (SSIM) [7] has been applied for 3D images in the form of view plus depth, where information about depth has been added to the metric using a local or global approach. In [8], an overall quality metric has been suggested by combining image quality with disparity quality using a nonlinear function. In [9], a quality metric for color stereo images has been proposed based o the use of binocular energy contained in the left and right retinal images calculated by complex wavelet transform (CWT) and Bandelet transform.

In this paper, we propose a novel 3D quality assessment metric taking into account HVS properties. The metric aims at modeling the effects forming a cyclopean view. The combined effects of binocular vision and saccades in forming the cyclopean view are modeled through the 3D-DCT block structure. The CSF masking and disparity masking are applied to model the corresponding processing in HVS. In addition, the effect of depth presence is quantified by adding a term assessing the depth variability.

2. PROPOSED QUALTIY METRIC

The proposed stereo quality assessment scheme is given in Figure 1. We consider 3D video represented in the form of two channels – left and right – forming a stereo-pair of views. The depth perception is created by the slightly different perspective of the two views manifested as disparity. The original (reference) video is distorted by some processing stage, e.g. compression, and the perceived level of distortions (or quality) between the reference and distorted videos has to be determined by means of a full-reference metric.

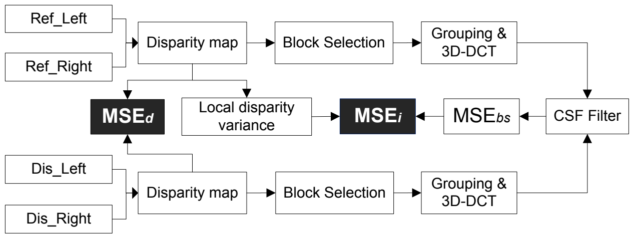

Figure 1. Flow chart of proposed model

Both the reference and distorted videos are processed to find the disparity map between the left and right views. The difference of reference and distorted disparity map is calculated by Mean Square error (MSE) denoted as MSE_d in Figure 1. Furthermore, the local disparity variance of the reference pair is found. The assessment runs on blocks. For each reference block in the left reference view, three similar blocks are found in both left and right images by block-matching (BM). These four blocks (i.e. one reference and three similar blocks in left and right views) are stacked in a 3D structure which undergoes 3D-DCT. The same is done for the structure associated with each block in the left distorted image. Both 3D-DCT domain structures are then corrected to account for the influence of the CSF. Transform-domain MSE denoted as MSE_{bs} is computed between the two set of coefficients to get a measure about the difference in cyclopean views modelled in transform domain. This error is corrected by the local disparity variance for the reference block to emphasize the differences between flat (2D) and pronounced-depth areas, measured by MSE_i, as in Figure 1. The eventual quality metric for the whole image is obtained in terms of dB.

2.1 Disparity map and local disparity variance

The disparity (or parallax) observed between the right and the left frame is inversely proportional to the distance to the object [15]. Stereo matching is to search for a point in an image that corresponds to the point specified in the other image in terms of associated features. Stereo matching plays a key role in the structure-from-stereo algorithms, which aim at getting an image (a map) being indicative for the distance to the object. In our approach, we calculate a dense disparity map between the left and right frames using a colour-weighted local search [10]. Considering rectified images, a window of size 9x9 from the left frame is run in horizontal direction to find similarity in terms of block matching, weighed by the colour difference in a bilateral manner [10]. Holes and mismatches are corrected by means of time-consistent post-filtering [11] and remaining unconfident disparity estimates are marked as holes.

The disparity map is then normalized by the disparity range of the target display – the so-called comfort zone [12]. The comfort zone determines the range of disparities which are allowed for avoiding accommodation-convergence rivalry and divergent parallax, with the former factor strongly dominating in portable 3D displays [12].

The comfort zone of the target display is determined with respect to the Persival's zone of comfort [13] and can be defined in terms of cm or in number of pixels for the target display [12]. Denote it as $C = D_{max} - D_{min}$, where D_{min} and D_{max} are the minimum and maximum disparity of the comfort zone in number of pixels, denoted the estimated reference disparity and the distorted disparity by $\{r_{i,j}\}, i = 1,..M; j = 1..N$ and $\{d_{i,j}\}, i = 1,..M; j = 1..N$, respectively. The normalized disparities are given by $R_{ij} = r_{ij}/C$ and $D_{ij} = d_{ij}/C$. The global disparity difference between reference and distorted scenes in terms of MSE is calculated as follows:

$$MSE_d = \frac{1}{\#\Omega}\sum_{i,j\in\Omega}\left(R_{ij} - D_{ij}\right)^2, \qquad (1)$$

where Ω is the domain of confidently estimated disparities and $\#\Omega$ is its cardinality.

The local disparity variance is calculated for a block of size $N_f \times M_f$ closer to the size fully projected into the eye fovea while looking at a typical viewing distance. For the mobile resolution and typical viewing distances of 30-40 cm, a block of size 28x28 pixels is a good choice. Pixel positions marked as holes are excluded from the estimate. The local variance is calculated as follows:

$$\sigma_d^2(k) = \frac{1}{\#\Omega_k - 1}\sum_{i,j\in\Omega_k}\left(\mu(k) - R_{ij}(k)\right)^2 , \qquad (2)$$

where $\mu(k)$ is the mean disparity calculated over the domain of confidently estimated pixels Ω_k within the foveal area with index k. The local variance characterizes the depth changes around the central block of size 4x4 which is needed for correcting the visual impairment measured for that block. Figure 2 illustrates the block position in the centre of the foveal area of disparity estimate.

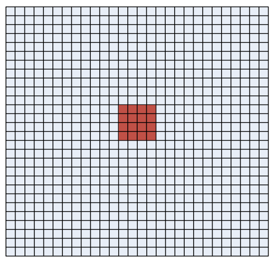

Figure 2. Central block 4x4 with respect to block for local disparity estimation

2.2 Assessment of visual artefacts in a transform-domain cyclopean view model

DCT plays a key role in our approach. We rely on the capabilities of DCT to decorrelate data and achieve highly sparse representation. In our approach, by the use of DCT we aim at modelling two processes taking place in the HVS. First, we model the binocular vision and forming the cyclopean view by combining together corresponding blocks from the left and right views [14]. Furthermore, we simultaneously model the saccades – the pseudo-random movements the eyes are performing while processing spatial information [15]. Saccades are modelled by searching for similar blocks in the spatial neighbourhood of the corresponding blocks in both views. Similar blocks are stacked together in a 3D structure to be jointly projected on the DCT basis. The resulting set of coefficients is expected to be informative about similarities across views and in spatial vicinity.

2.2.1 Block Selection and 3D-DCT Transform

Block-Matching (BM) is applied to find similar blocks around the reference block A_0 and its stereoscopic correspondence within a search range region as shown in Figure 3. MSE is used as dissimilarity measure between reference block and searched blocks. Here, the block size is defined as 4x4 and the search region around the reference block and its stereo correspondence has been fixed to 28x28 pixels.

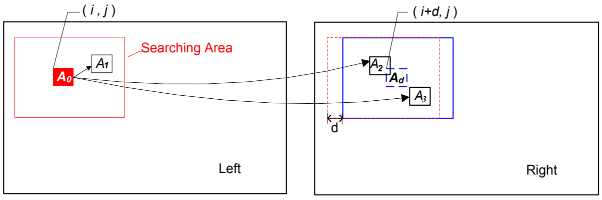

Figure 3. An example of block selection in reference stereoscopic image

Using BM, one best matched block in the left view and two similar blocks in the right view are found as shown in Figure 3. In the figure, A_0 is the reference block with index k in the left view frame (the index is omitted in the notations for simplicity), A_1 is the most similar block to A_0 in the same channel, A_d is the corresponding block in right view frame found through the stereo-correspondence search. Assume the coordinate of left upper corner of A_0 is (i, j), then the coordinate of left upper corner of A_d is $(i + d, j)$ in right view; d is the horizontal disparity of reference block A_0 between left and right view, which is taken as the median within the 4x4 window with index k: $d(k) = \left\lceil median\left(R_{ij}(k)\right)_{4x4} \right\rceil$. In Figure 3, A_2 and A_3 are the two best matching blocks to A_0 which are searched within search region around A_d in the right channel. Note that the similarity between A_0 and A_d has been found through another similarity mechanism an eventually A_d could be or could not be one of the selected blocks A_n, $n = 2, 3$. The four blocks in the reference stereoscopic image $A = \{A_0, A_1, A_2, A_3\}$ and the respective four blocks in the distorted stereoscopic image $B = \{B_0, B_1, B_2, B_3\}$ are found and grouped into two 3D arrays respectively. A 3D-DCT transform is applied to these two 3D arrays, In our setting, for nice symmetry and fast processing the size of the blocks is fixed to 4x4, thus fixing the size of the 3D structure to a cube of ridge length 4. However, any other block size is possible and also more blocks can be collected on the base of similarity, thus forming a 3D structure of bigger size. Our previous experiments with 2D images and videos have shown that finding one similar block is sufficient to account for this type of similarity around the reference block.

2.2.2 Contrast Sensitivity and Disparity Masking

In [16], a perceptually-driven metric for 2D image quality assessment has been suggested. It calculates an MSE between the 8x8 DCT coefficients of the reference and distorted block modified by coefficients reflecting the masking effect of the CSF [16]. While the original approach was developed for DCT of size 8x8 and using masking coefficients determined by the JPEG quantization table, we have modified it to work with DCT of size 4x4 by down-sampling the masking coefficients. The new set of masking coefficients is given in Table I.

Table I. Set of CSF masking coefficients for 4x4 DCT

1.6084	2.5735	1.0723	0.5046
1.8382	1.6084	0.6434	0.3730
1.4297	0.6955	0.3785	0.2499
0.5252	0.3299	0.2499	0.2145

Denote by $U = \{U_{i,j,n}\}, i, j = 1..4, n = 0..3$ and $V = \{V_{i,j,n}\}, i, j = 1..4, n = 0..3$ the 3D-DCT coefficients of the cubes A and B. The coefficients of the top layer $n = 0$ concentrate most of the energy of the grouped similar blocks, and the coefficients of lower layers $n = 1,2,3$ indicate the differences between grouped similar blocks and reference block. Thus, the structure models the formation of a cyclopean view in some transform domain. The layers are masked by the CSF masking coefficients by $T_{i,j}, i, j = 1..4$ scaled down by factors ω_n determined by the energy distribution for each transform layer. The MSE for the block structure is then calculated as follows:

$$MSE_{bs} = \frac{1}{64}\sum_{i,j=1}^{4}\sum_{n=0}^{3}\omega_n T_{i,j}^2\left(U_{i,j,n} - V_{i,j,n}\right)^2. \quad (3)$$

The MSE is then corrected with the information about the local disparity variation. Extensive subjective tests have demonstrated that 3D information plays a role in the case of low level of 2D distortions (such as compression artefacts) [18]. In cases of low level of impairments, 3D scenes have been favoured with respect to 2D scenes. Correspondingly, we use the local disparity variance to correct the block MSE for the block with index k as follows

$$MSE_{3dbs}(k) = MSE_{bs}(k)\frac{MSE_{bs}(k)}{MSE_{bs}(k)+\alpha\sigma_d^2(k)}, \quad (4)$$

where $\sigma_d^2(k)$ which is calculated according to Eq.(2), and α is weighting parameter.

2.3 Composite quality measure

The total MSE for the image stereo pair MSE_i is calculated as average over all disparity-corrected block MSE_{3dbs}. For efficient calculation, we have used non-overlapping reference blocks. The overall assessment of the quality of stereoscopic image takes also into account global changes of the disparity and is transformed from error to PSNR-type of measure as

$$MSE_H = (1 - \varepsilon)MSE_i + \varepsilon MSE_d\ , \quad (5)$$

$$PHSD = 10log(255^2/MSE_H)\ , \quad (6)$$

where ε is weighting factor of MSE_d, MSE_H is final MSE which is combined by MSE_i and MSE_d.

3. TEST SEQUENCES AND SUBJECTIVE TESTS

The design and the validation of the metric have been performed against two data bases of test 3D videos annotated with the results of subjective tests. [17], [18].

3.1 Metric design

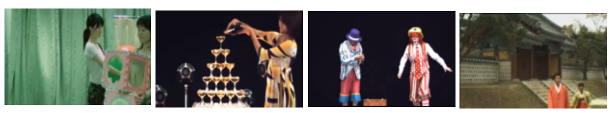

Figure 4. Contents of 3D Video Database I

The design of the metric has been performed using four multi-view video sequences: Akko&Kayo, Champagne Tower, Pantomime, and LoveBirds1 (Database I). Their thumbnails are shown in Figure 4. Different camera baselines have been selected to get stereo pairs with different types of depth: short baseline (3D) and wide baseline (3D). The sequences have been cut into 10 seconds and coded using the simulcast MPEG-4 standard encoder. Five different quantization parameters QP (=25, 30, 35, 40, 45) have been applied to each processed sequence. Thus, a total of 8 reference sequences and 40 distorted sequences have been obtained [17].

The video sequences have been annotated with subjective quality attributes obtained by psychometric tests [17]. The test group included 32 persons equally stratified by gender and age between 18 and 45. The visualization was done on an auto-stereoscopic display provided by NEC [17]. The tests collected the opinion in term of quality (11 point scale) and acceptance (binary scale). The overall ratings of stereoscopic videos have been ranked in terms of mean opinion score (MOS). The results of the subjective tests were used to tune the scaling parameters in Eq. (4) and Eq. (5).

3.2 Metric validation

Figure 5. Contents of 3D Video Database II

The metric was validated against another 3D video test database (Database II). It contains six different contents spanning different genders of mobile 3DTV and video: Bullinger, Butterfly, Car, Horse, Mountain, and Soccer2 as illustrated in Figure 5. Each video sequence has an approximate length of 10 seconds. The sequences have been encoded with four different coding methods including H.264/AVC Simulcast, H.264/AVC MVC, MRSC, and Video plus Depth (V+D) with varying codec profiles and quality parameters shown in Table II [18]. There are six reference 3D video sequence and 96 distorted 3D video sequences in total.

Table II Codec setting of two profiles

Profile	Baseline	High
GOP Size	1(IPPP)	8(Hierarchical B Frames)
Symbol Mode	CAVLC	CABAC
Search Range	48	48
Intra Period	16	16
Quality Level	QP (30, 37)	QP (30, 37)

Subjective tests were carried out with 87 participants equally stratified by gender and age between 16 and 37 years. The visualization was done on the same auto-stereoscopic display with resolution of 428x240 provided by NEC as in the previous subjective tests. The overall ratings of stereoscopic videos have also been ranked in terms of mean opinion score (MOS).

4. RESULTS

The results of the proposed approach are compared with several state-of-art quality metrics: PSNR, MSSIM [19], SSIM [7], UQI [20], NRMSE [21], PSNR-HVS [16], PSNR-HVS-M [22], which are all 2D metrics and the 3D metric from [8]. For the latter, we have set SSIM to measure the image quality and UQI to measure the disparity quality [8]. All algorithms compare the luminance component only. The 2D metrics have been run on the left and right channels separately and the results have been averaged.

Table III Spearman correlations on 3D Video Database I

Metric	Simlucast
PSNR	0.8638
MSSIM[19]	0.6895
SSIM[7]	0.6447
UQI[20]	0.5248
NRMSE[21]	0.8512
PSNR-HVS[16]	0.8901
PSNR-HVS-M[22]	0.8632
Global comb.[8]	0.6570
PHSD	**0.9368**

Table IV Spearman correlations on 3D Video Database II

	MRSC	MVC	Simlu-cast	V+D	ALL
PSNR	0.0757	0.3278	0.3678	0.0504	0.2542
MSSIM	**0.7313**	0.7270	0.7130	0.6557	0.6327
SSIM	0.3983	0.5870	0.5426	0.2052	0.4427
UQI	0.4226	0.4513	0.3635	0.1739	0.2998
NRMSE	0.4522	0.6078	0.5609	0.3157	0.4128
PSNR-HVS	0.3991	0.4226	0.4183	0.0939	0.2266
PSNR-HVS-M	0.5209	0.5052	0.4765	0.1061	0.2449
Global Comb.	0.5617	0.8026	0.6991	0.5217	0.6455
PHSD	0.6487	**0.8104**	**0.7783**	**0.7374**	**0.7999**

Table III and IV show the Spearman correlations for each metric calculated on the 3D video database I and the 3D video database II respectively. As seen in the first group of results, 2D quality metrics such as MSSIM, SSIM and UQI, show lower correlation to MOS. The same can be observed for 3D quality metric denoted as Global combination method [8]. The standard PSNR, NRMSE, along with PSNR-HVS and PSNR-HVS-M suit the visual perception remarkably better. The proposed metric performs substantially better gaining advantage of using 3D structure and disparity information the quality quantification. The second group of results includes tests with much diverse processing methods but encoded at only two different compression levels. For that case, the Global combination method [8] performs better than most of 2D quality metrics for each encoding method. To MRSC method, MSSIM gives a higher correlation to MOS which should be attributed to the multi-scale behaviour of this metric. PSNR gives quite poor performance which Spearman correlation is only 0.0757. The Spearman correlation for the best performing 2D metric (MSSIM) for the whole database is close to the one for the 3D Global combination metric. Most quality metrics fail in the case of V+D encoded sequences. This should be attributed to the presence of a number of artefacts caused by view rendering based on estimated depth. Those artefacts are barely visible but well captured by the feature-based metrics. Finally, it can be seen that the proposed metrics PHSD outperforms the other considered metrics. While it has been designed specifically to quantify quality of stereo video represented by two equal-resolution channels, it performs fairly well for other type of 3D video compression, e.g. V+D, compared to other metrics.

5. CONCLUSION

In this paper, a novel full-reference 3D video quality metric has been validated. The metric aims at modelling the effects forming a cyclopean view, taking also into account the 3D perceptibility of visual artefacts. The combined effects of binocular vision and saccades in forming the cyclopean view are modelled through the 3D-DCT block structure. The CSF masking and disparity masking are applied to model the corresponding processing in the HVS. For the latter, the target display resolution and its corresponding comfort zone are taken into account, which substantially differentiate the proposed metric from 2D metrics and other 3D video metric attempts. Furthermore, masked visual distortions and geometrical errors are combined to form a compound error. The approach is simple yet quite effective as demonstrated by testing two groups of 3D videos annotated by quality attributes from subjective tests. The experimental results have shown that our metric outperforms current state-of-the-art quality metrics. We have to note that our implementation does not take into account masking effects created by motion. This is left for future investigations. However, our experiments so-far have shown that this masking plays minor role in estimating the quality. This observation has been confirmed by subjective tests on still images from the same database, which resulted in the same MOS as in the case of the respective videos [23].

REFERENCES

[1] M. T. M. Lambooij, W. A. Ijsselsteijn, and I. Heynderickx, "Visual discomfort in stereoscopic displays: a review," *in Stereoscopic Displays and Virtual Reality Systems XIV of Proc. of SPIE*, San Jose, USA, pp. 1-13, Jan. 2007.

[2] A. Boev, D. Hollosi, A. Gotchev, K. Egiazarian, "Classification and simulation of stereoscopoic artifacts in mobile 3DTV content, " *Electronic Imaging Symp. 2009*, San Jose, USA, Jan. 2009

[3] P. Merkle, Y. Wang, K. Muller, A. Smolic, T. Wiegand , "Stereo Video Compression For Mobile 3D Services", *3DTV-Conf.*, Potsdam, Germany, May 2009.

[4] P. Merkle, Y. Wang, K. Muller, A. Smolic, T. Wiegand , "Video Plus Depth Compression For Mobile 3D Services", *3DTV-Conf.*, Potsdam, Germany, May 2009.

[5] A. Boev, A. Gotchev, K. Egiazarian, A. Aksay, and G.B. Akar, "Towards compound stereo-video quality metric: a specific encoder-based framework, " *IEEE Southwest Symp. on Image Analysis and Interpretation*, Denver, USA, pp. 218-222, June 2006

[6] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images, " *EURASIP Journal on Image and Video Processing, special issue on 3D Image and Video Process.,* vol. 2008

[7] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity, " *IEEE Trans. on Image Process.*, vol. 13, issue 4, pp. 600-612, April 2004.

[8] J. You, L. Xing, A. Perkis, and X. Wang, "Perceptual Quality Assessment for Stereoscopic Images Based on 2D Image Quality Metrics and Disparity Analysis, " *Int. Workshop on Video Process. and Quality Metrics for Consumer Electronics -VPQM*, Scottsdale, AZ, USA, 2010.

[9] R. Bensalma, and M.C. Larabi, "Towards A Perceptual Quality Metric For Color Stereo Images," *IEEE 17th Int. Conf. on Image Processing*, Hong Kong, Sep. 26-29, 2010.

[10] S. Smirnov, A. Gotchev, and M. Hannuksela, "Comparative analysis of local binocular and trinocular depth estimation approaches", *Proc. of SPIE*, vol. 7724, 2010

[11] S. Smirnov, A. Gotchev, K. Egiazarian "A Memory-efficient and Time-consistent Filtering of Depth Map Sequences", *Proc. of SPIE,* vol. 7532, 2010.

[12] A. Boev, and A. Gotchev, "Comparative Study of Autostereoscopic Displays for Mobile Devices," *Proc. of SPIE 7881A*, 2011.

[13] 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).

[14] Brian A. Wandell, "Foundations of Vision", *Sinauer Associates*, 1995.

[15] Martin J. Tovée, "An Introduction to the Visual System", 2nd ed. Cambridge, 2008.

[16] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli, "New full-reference quality metrics based on HVS, " *Int. Workshop on Video Proc. and Quality Metrics*, Scottsdale, USA, 2006

[17] S. Jumisko-Pyykkö, T. Haustola, A. Boev, and A. Gotchev, "Subjective Evaluation of mobile 3D content: depth range versus compression artifacts", *Proc. of SPIE*, vol. 7881A, 2011.

[18] D. Strohmeier, S. Jumisko-Pyykkö, K. Kunze, G. Tech, D. Bugdayci, M. Oguz Bici, "Results of quality attributes of coding, transmission and their combinations", *Technical report D4.3*, February 2010,

[19] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," *IEEE Asilomar Conf. on Signals, Systems and Computers,* Nov. 2003.

[20] Z. Wang, and A. Bovik, "A universal image quality index, " *IEEE Signal Processing Letters*, vol. 9, pp. 81–84, March, 2002

[21] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M.J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow, " *IEEE Int'l Conf. on Computer Vision*, Crete, Greece, pp. 243-246, Dec. 2007

[22] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin, "On between-coefficient contrast masking of DCT basis functions, " *Int. Workshop on Video Proc. and Quality Metrics*, USA, 2007

[23] S. Jumisko-Pyykkö, T. Haustola, A. Boev, A. Gotchev, "Juxtaposition between compression and depth for stereoscopic image quality on portable auto-stereoscopic display", *Proc.of 3DTV Conference*, Antalya, Turkey, 16-18 May, 2011