Fast zerotree wavelet depth map encoder for very low bitrates

M.O. Martínez-Rach[1], P. Piñol[1], O. López Granado[1], M.P. Malumbres[1]

Abstract—**When using Depth Image Based Rendering (DIBR) to encode 3D video, proper encoding of depth map information is critical for reducing the overall bitrate. In this work we propose the use of a fast depth map encoder with a non-uniform quantization and a post processing filter that exhibits good PSNR performance at very low bitrates.**

Keywords— **3D video, Depth Map encoding, non-uniform quantization, performance evaluation.**

I. INTRODUCTION

DEPTH Image Based Rendering (DIBR) was introduced in [1]. The main advantage of DIBR is that it provides high quality 3D video with smaller bandwidth than the required for transmission of two monoscopic streams. DIBR uses a depth map frame to generate two virtual views from the same reference view, one for the left eye and the other one for the right eye [2]. The depth map can be encoded more efficiently than the independent encoding of two monoscopic streams.

Depth maps are not viewed by end users, but are used for virtual view generation. Therefore, it is important to compress depth maps in a way that it minimizes distortion in views rendered with them. DIBR algorithms mainly use information extracted from the vertical edges in the depth map image and are very sensible to impairments present at vertical edges and, in the same way, they are also less sensitive to impairments on horizontal and diagonal ones. In that sense, wavelet-based compression approaches are really suitable for the depth map compression process, as they separate the vertical, diagonal and horizontal edge information into different subbands.

In [3] authors performed experiments to measure the final quality of the rendered process when depth maps are reduced to only keep information around the borders or only at the borders. They found that the same ratings in quality were given to rendered images with full depth maps than to the one obtained with the reduced depth maps. When applying the Discrete Wavelet Transform (DWT) to the depth maps, the high frequency information corresponding to borders is allocated at the high frequency subbands, being almost zero all the coefficients corresponding to large uniform depth areas. So, after the DWT the only relevant information in the areas around the borders is present in high frequency subbands. That subbands have different relative importance for the incoming DIBR process, this fact, allows us to use a non-uniform quantization strategy based on the impact of quantization errors into the ultimate DIBR video reconstruction process. So, we propose a non-uniform quantization strategy that penalizes the horizontal and diagonal high frequency orientations measuring the final impact on the rendered view quality.

Rest of the paper is organized as follows. Section II introduces briefly the zerotree-wavelet encoder used in our tests. In Section III the proposed quantization strategy is presented with a discussion and motivation of the present work. In Section IV the experiments carried out are explained. Section V presents the results obtained with the proposed method and finally conclusions are exposed in section VI

II. A FAST NON-EMBEDDED ENCODER

Lower-Tree Wavelet (LTW) is a fast zerotree-based wavelet image encoder, with state of-the-art coding efficiency, but less resource demanding than other encoders in the literature.

The basic idea of this encoder is very simple: after computing a dyadic wavelet transform of an image [4], the wavelet coefficients are first quantized using a variable dead zone uniform quantizer and then the zero-tree symbol map is encoded by means of an arithmetic coder. For the coding stage, if the rounded absolute value of a coefficient and all its descendants (considering the classic quad-tree structure from [5]) is zero, the entire tree is encoded with a single symbol. But if a coefficient is not zero but all its descendants are, that coefficient is encoded with another symbol.

On the other hand, for each non zero wavelet coefficient, we encode a symbol indicating the number of bits needed to represent that coefficient (its magnitude), along with a binary coded representation of its bits and sign. More details about the coding and decoding algorithms, as well as a formal description can be found in [6].

III. NON-UNIFORM QUANTIZATION APPROACH

Fig. 1 shows a block diagram of the steps we follow in the present study. Relevant steps are marked in gray. A color image and its corresponding depth map are the input to the DIBR process which produces the original views (left and right). These original views will be used later as reference for getting the final PSNR.

The original depth map is encoded by the LTW encoder. A six level decomposition wavelet transform is applied. The quantizer strategy has been changed according to the one proposed in this study. The encoder produces the bit-stream. At this point we get the bit-rate used to encode the depth map. The bit-stream enters the decoder which produces the reconstructed depth map.

1-Universidad Miguel Hernández – Spain
[mmrach,pablop,otoniel,mels]@umh.es

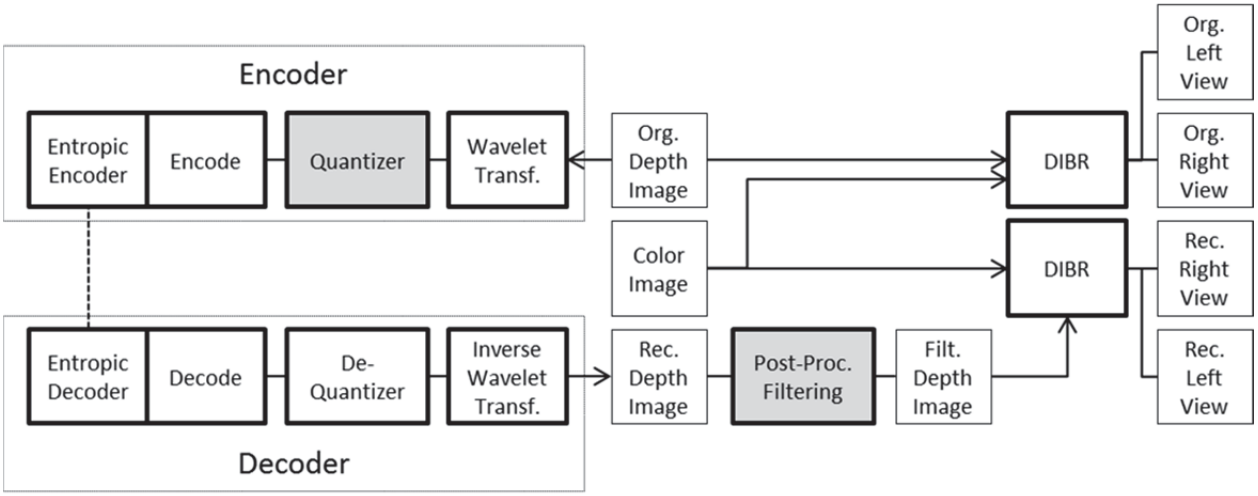

Fig. 1. . Block diagram of proposed approach.

After reconstruction we apply a post-processing filtering stage as we will explain below. The filtered depth map and the original color image are again the input to a DIBR process which produces the reconstructed views (left and right). We get the final left and right PSNR comparing reconstructed views with original ones.

The proposed method is focused only into the compression of the depth map, so we leave the color image uncompressed in order to be able to measure the gains in the final quality caused by the method.

The main drawback of using wavelet compression schemas to encode depth maps at thigh compression rates is the appearance of the ringing distortion effect just close to the borders of the depth image. This distortion becomes higher as the compression rate increases and affects highly to the DIBR rendering process.

Fig. 2. . Ringing effect and result after the application of two dimensional joint bilateral filter to the compressed depth map in a frame of ballet sequence. Left image corresponds to part of the ballet sequence depth map compressed at 0.1 bpp. In the right image the de-ringing filter is applied on the compressed depth map.

In order to eliminate this ringing effect from the compressed depth map we perform a de-ringing post-processing step applying a two dimensional joint bilateral filter [7] to the compressed image, in the same way that DCT based encoders perform a post-processing step for removing blocking artifacts.

Fig. 2 shows the differences between a compressed depth map at 0.1 bpp with the LTW encoder, and the result after the application of the de-ringing filter. The application of this post-processing to the wavelet compressed depth map produces a big gain in the overall quality of the rendered left and right views as explained in [7]. Fig. 3 shows the differences of the rendered views with and without the application of the post-processing filtering to the compressed depth map.

Fig. 3. . Left image corresponds to the rendered left view without post processing in the compressed depth map. For the right image the post processing filter was applied on the compressed depth map at 0.1 bpp.

Some subbands of the wavelet decomposition of the depth map image are more important than others for the incoming DIBR process. Fig. 4 shows the typical wavelet decomposition schema with the different subbands. In our experiments a six level wavelet decomposition is performed. The critical vertical information for the DIBR process resides in the LH subbands and horizontal edges reside in the HL subbands. The horizontal edge information is not as important as the vertical one for the DIBR process [8].

The use of a quantization step, QStep+delta for the HL1 subband while maintaining QStep for the rest of subbands, produces a R/D curve (as QStep increases) that crosses at low bit-rates the R/D curve produced when uniform quantization is applied to all subbands with the same QStep (see Fig.5).

Based in that fact, and in order to improve R/D behavior for very low bit-rates, we can we apply higher quantization to the less important subband types for the DIBR process. The drawback is that for rates above the crossing point quality is reduced because more coefficients are set to zero than in the non-uniform quantized subband. For increasing values of delta the quality for rates above the crossing points decreases while, up to some extent, for rates below the crossing point, the quality increases. We can further increase the

value of delta arriving at a full quantization for that subband (all coefficients are set to zero).

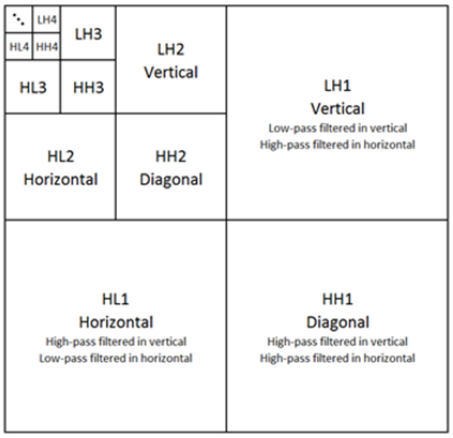

Fig. 4. . Wavelet subband decomposition schema.

In Fig. 5 we can see the R/D curves for the compressed rendered view of the ballet sequence. The UQ (Uniform Quantization) curve corresponds to a uniform quantization schema. The (Non Uniform Quantization) NUQ HL1 curve corresponds to a Uniform Quantization for all the wavelet subbands except for the HL1 subband whose coefficients have been set to zero, full quantized. In the same way the NUQ HH1 HL1 and the NUQ LH1 HH1 HL1 curves apply a full quantization for the corresponding subbands and a uniform quantization for all other subbands.

At very low bitrates the NUQ R/D curves cross the UQ curve. This happens because at the crossing rate, the UQ has reduced the coefficients up to some extent, but it has not set to zero all coefficients located in that “not so important” subbands for the rendering process. If we set them to zero with the NUQ approach, the rate gain is higher than the quality loss (after the rendering process), because this coefficient are not so important for the DIBR process and are still using an important rate budget.

The reason why we do not use some intermediate delta values, for example in the HH1 subband of the NUQ HH1 HL1 curve, is because our tests revealed that using progressively lower values of delta in that subband, causes that in progressively the NUQ HH1 HL1 curve fits the NUQ HL1 curve.

So the problem is to determine what subband or subbands combination, when full quantized, will produce better R/D behavior for very low bit rates, while the uniform quantization is applied to the rest of subbands.

At these low rates, for ballet sequence and working only at the first wavelet decomposition level, the highest gain is for the NUQ HH1 HL1 curve with a gain of 0.79 dB with an average of 0.6 dB bellow crossing point.

As shown in Fig. 5 depending of the combination of the subbands being full quantized, the gain is different.

This happens because the uniform quantization schema reduces practically to zero all the important coefficients for the DIBR process, reducing dramatically the its accuracy. That is, the remaining coefficients are no more useful for the DIBR process, so if we set them to zero the R/D rate will increase, because the rate budget they use is still important.

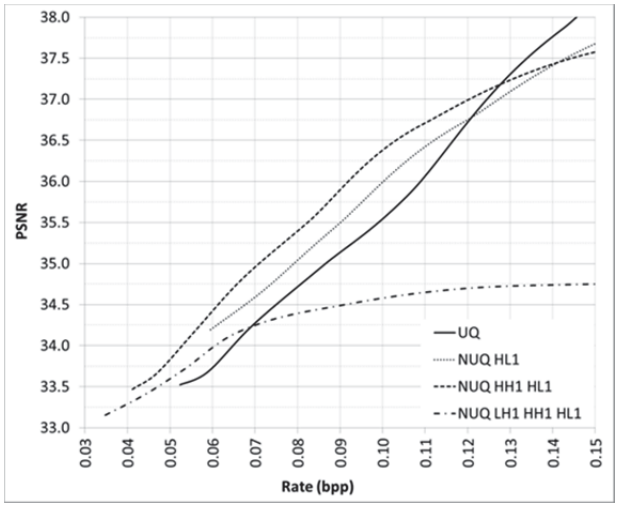

Fig. 5. R/D comparison for Ballet sequence. On the first wavelet decomposition level (higher frequencies) different subbands have been full quantized and a uniform quantization is applied on the rest..

We also see that if the full quantized subband is the LH1, where important high frequency vertical information resides, the initial quality drop is too high.

Results for the first decomposition level for a specific sequence suggest us to perform the following experiment whose objective is to determine which is the best or at least the most suitable candidate subband combination to be full quantized for all sequences.

IV. EXPERIMENTS

A. *Objective*

For these experiments we used five sequences with different frame sizes. For each sequence we pick randomly one of the frames of the sequence. We used Matlab R2009b for the fitting process.

For a six level wavelet decomposition schema there are 18+1 subbands. Three subbands for each level and the LL subband which is not quantized at all.

The objective is to find which combination within all the possible combinations of subbands is the one that when full quantizing its subbands gives the best results in terms of R/D for very low bit-rates, below 0.08 bpp.

As the number of possible combinations chosen from a set of 18 subbands is huge we first decide to narrow the problem. So we perform two experiments, the first one serves to determine the relevant wavelet decomposition levels and the second one exploits all combinations in that relevant levels.

B. *Relevant levels*

The first step is common for both experiments. We perform a UQ for all subbands with increasing values for the QStep. By choosing at least tree QStep values we can perform a curve fitting of the R/D UQ curve with a power2 model. The average goodness of fit by means of rsquare for all the fittings performed in both experiments was 0.9875.

We performed a full quantization for each of the 18 subbands independently and the results of this 18 NUQ curves were analyzed. That is, in each of these 18 combinations, only one subband is full quantized while

on the other a uniform quantization is applied with increasing QStep in order to get the 18 NUQ curves. That curves were fitted in the same way and with the same model as we did for the UQ case.

For each of these NUQ fitted curves we measured the area over the UQ curve, and arrange that measures in a table where each row represents a wavelet decomposition level and each column represents any of the three orientations (subband types) in that level. Values in that tables where normalized dividing by the sum of all areas. So, these values represent the percentage of the overall gain that each full quantized subband produces. We get these values for every sequence in the set, and finally we average the values for each subband across all images. Table I represents the averaged gain that each individual subband produces.

TABLE I

AVERAGED % GAIN FOR INDIVIDUAL FULL QUANTIZED SUBBANDS

	LH	HH	HL
L1	19.00	13.78	38.92
L2	6.97	8.15	9.32
L3	0.03	1.31	1.60
L4	0	0.31	0.37
L5	0	0.06	0.16
L6	0	0	0

As shown in Table I most of the gain is obtained in the first two decomposition levels (96.14%) and practically no gain is obtained for the rest of the wavelet decomposition levels. That means that all curves in those levels where almost below the UQ one. So we decide to reduce the range for the search of the best combination to the first two decomposition levels.

C. *Choosing the best combination*

In this test we search for the combination that gives better results at low rates, within all the possible combinations chosen from the 6 subbands of the first two wavelet decomposition levels. This gives a total of 64 possible combinations. Combinations are labeled in the way shown in Fig. 6. where *f* take the values 0 or 1. A value of 1 indicates that the subband is full quantized and 0 indicates that a uniform quantization will be applied in that subband.

In that way, combination 000_000 corresponds to the UQ combination where no subband is full quantized and combination 111_111 correspond to the combination that full quantizes all subbands. For higher decomposition levels a uniform quantization is performed.

L1				L2		
LH	HH	HL	_	LH	HH	HL
f	*f*	*f*		*f*	*f*	*f*

Fig. 6. . Labeling of the possible combinations

We obtain the fitted curves for all the 63 NUQ combinations. Our results show that there is no combination that best perform for all images and for all bit-rate ranges, so we have to find a trade-off solution. For that, we divide the rates below 0.08 bpp in three intervals. Low interval goes from 0.01-0.03 bpp. Medium interval goes from 0.03-0.05 bpp. And high interval goes from 0.05 to 0.08 bpp.

Table II shows the best and the second best combination for each of the sequences and the selected intervals. We choose a trade-off combination that will work reasonably good bellow crossing point for all sequences. It is the NUQ combination labeled by 011_011 and highlighted in boldface in the Table II. This combination is the best one in the 40% of all possible cases and the second best in the 100% of all other cases bellow crossing point, that is, when the UQ is not the best one.

TABLE II

BEST COMBINATIONS FOR THE BIT-RATE INTERVALS FOR ALL SEQUENCES

Interval	0.01 - 0.03 bpp	
Comb	Best	2nd Best
Ballet	111_011	**011_011**
Breakdance	**011_011**	011_001
Interview	111_011	**011_011**
Badminton	**011_011**	011_001
Flag	111_011	**011_011**
Interval	0.03 - 0.05 bpp	
Comb	Best	2nd Best
Ballet	**011_011**	011_000
Breakdance	**011_011**	011_001
Interview	111_011	**011_011**
Badminton	000_000	001_000
Flag	111_011	**011_011**
Interval	0.05 - 0.08 bpp	
Comb	Best	2nd Best
Ballet	**011_011**	011_000
Breakdance	011_001	**011_011**
Interview	**011_011**	011_001
Badminton	000_000	010_000
Flag	000_000	001_000

D. *Proposed Quantization Schema*

As a result of previous experiments we propose the next quantization schema when using zerotree wavelet compression to encode image depth information that will be used in a DIBR process in order to achieve higher qualities of the rendered views at very low bit rates.

Set to zero all coefficients in the HH1, HL1, HH2 and HL2 subbands (see Fig. 4) and apply a uniform quantization with the proper QStep for achieving the desired low rate.

V. RESULTS

In this section we proceed to show the results obtained in terms of PSNR gain at low rates for all sequences when the proposed quantization strategy is applied. Results are given for left and right rendered views as well as the average of the two views. Some of the final R/D curves will be also shown, as well as some cropped images of the final rendered views.

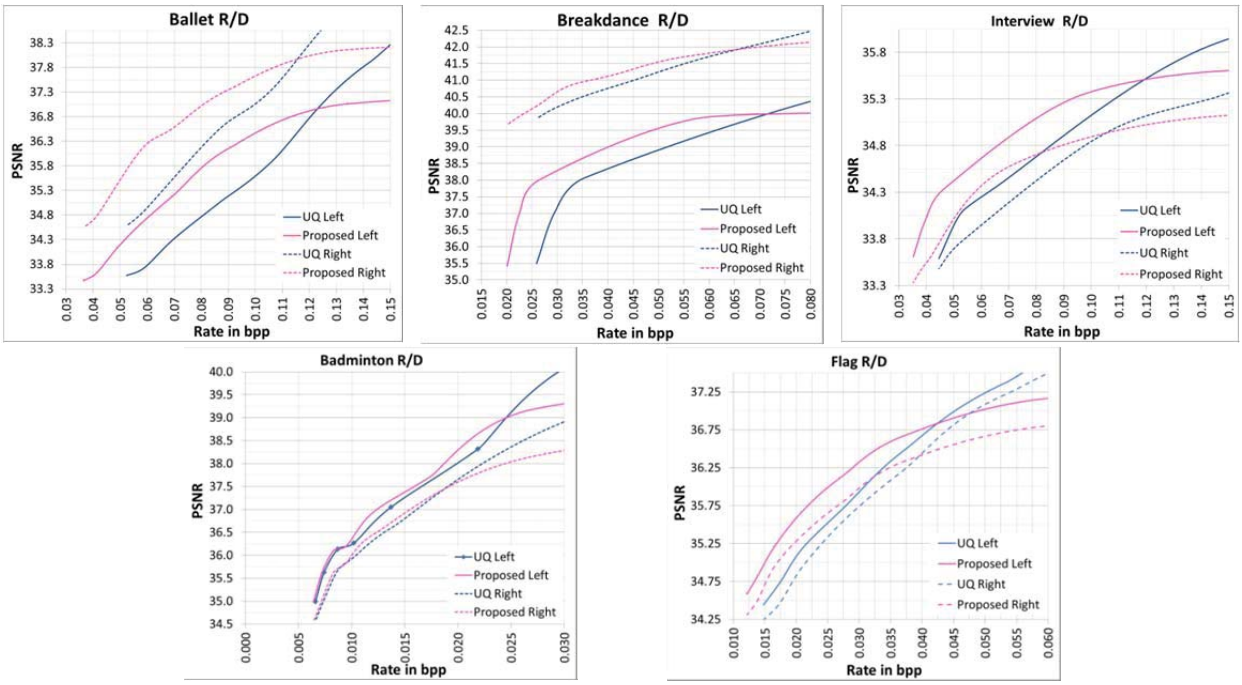

Fig. 7. Rate/Distortion plots for all the sequences and for Left and Right rendered views.

Fig. 7 show the R/D curves for all sequences obtained for the proposed quantization schema compared with the uniform quantization schema. Table III shows the maximum and average PSNR gains for each sequence. Results are given for left rendered view, right rendered view, and averaged for both views. Averages in that table were obtained as proposed in [9]. These values were obtained when the proposed schema is used for rates below the crossing point in each sequence being the lowest considered rate 0.01 bpp.

TABLE III

AVERAGE AND MAXIMUN PSNR GAIN

FOR ALL SEQUENCES AND VIEWS

	Left View		Right View		Both
	Max.	Avg.	Max.	Avg.	Max.
Ballet	0.97	0.62	1.32	0.79	1.14
Breakdance	1.28	0.81	0.55	0.35	0.92
Interview	0.54	0.42	0.29	0.20	0.42
Badminton	0.15	0.11	0.14	0.11	0.14
Flag	0.57	0.43	0.58	0.39	0.57

Fig. 8 to 12 show part of the sequences where some differences in the final quality of the rendered views can be appreciated between the original uniform quantization schema and proposed one. Left images correspond to the UQ schema and Right images corresponds to the proposed quantization schema.

Finally, we have measured the coding delay of the overall process for both, the UQ an NUQ schemas. In general, the NUQ approach is a little bit faster than UQ, mainly due to the fact of saving quantization and coding time of coefficients belonging to full quantized subbands. In Table IV coding delay of both Uniform Quantization (UQ) and Non Uniform Quantization (NUQ) approaches is shown for a random frame of each video sequence (times include I/O processing).

TABLE IV

CODING DELAY OF UQ AND NUQ PROPOSALS

Time / sec.	UQ	NUQ
Ballet	0.59	0.54
Breakdance	0.58	0.52
Interview	0.23	0.18
Badminton	1.04	0.91
Flag	1.17	1.16

VI. CONCLUSIONS

The compression of depth map images with wavelet encoders can be improved in the low bit-rate range with the use of the technics discussed in this study.

First we include a post processing step in the decoder side after decoding the depth map. This step uses a two dimensional joint bilateral filter for alleviating the ringing artifacts produced by the wavelet encoding process that are more evident in the depth map images around the borders of the scene objects as the compression rate increases. Although the depth maps are not viewed by the end user, those ringing artifacts heavily affects to the DIBR process and as shown, the results are improved when applying the filter before the rendering process.

This post-processing of the depth map can be applied independently of the quantization schema used by the encoder. Typically, a uniform quantization over all the frequency subbands in a wavelet decomposed images is applied.

We then proposed a non-uniform quantization schema that improves the R/D performance at very low bit-rates for tree wavelet encoders. This schema discards information in the less relevant subbands for the incoming DIBR process while maintains the coefficients in the high frequency subbands that have the important vertical information used for that rendering process.

Fig. 8. Ballet sequence: Left: UQ, Right: Proposed. The compression rate is 0.075 bpp.

Fig. 9. Interview sequence: Left: UQ, Right: Proposed. The compression rate is 0.071 bpp.

Fig. 10. Breakdance sequence: Left: UQ, Right: Proposed. The compression rate is 0.054 bpp..

Fig. 11. Flag sequence: Left: UQ, Right: Proposed. The compression rate is 0.023 bpp

Fig. 12. . Badminton sequence: Left: UQ, Right: Proposed. The compression rate is 0.020 bpp.

As average for the two rendered views a gain over 0.5 dBs is obtained for the 60% of the test sequences, being this gain lower than 0.4 dB only in one of the sequences. In the worst case this quantization approach for low bit-rates does not produce worst results that the traditional uniform quantization schema.

ACKNOWLEDGMENT

This work was funded by Spanish Ministry of Education and Science under grant DPI2007-66796-C03-03.

REFERENCES

[1] C. Fehn, "Depth-image-based rendering (DIBR), compression and Transmission for a new approach on 3D-TV," Proc. SPIE Stereoscopic Displays and Virtual Reality Systems XI, January 2004, pp. 93-104.

[2] C. Fehn, R. Rarre, and S. Pastoor, "Interactive 3-DTV-Concepts and Key Technologies," Proc. IEEE, vol. 94, no.3, pp. 524-538, 2006.

[3] W. J. Tam, F. Speranza, L. Zhang, R. Renaud, J. Chan, and C. Vazquez "Depth image based rendering for multiview stereoscopic displays: role of information at object boundaries" in Proc. SPIE Three-Dimensional TV, Video, and Display IV. Boston MA. USA. 24 October 2005

[4] J. Oliver, M. P. Malumbres, "Fast and efficient spatial scalable image compression using wavelet lower trees," in Proc. IEEE Data Compression Conference, Snowbird, UT, March 2003

[5] Said, A., & Pearlman, A. "A new, fast and efficient image codec based on set partitioning in hierarchical trees". IEEE Transactions on Circuits, Systems and Video Technology, 1996, 6(3), 243-250.

[6] Oliver, J., & Malumbres, M. P. "Low-complexity multiresolution image compression using wavelet lower trees". IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(11), 1437-1444.

[7] De Silva, D.V.S.X.; Fernando, W.A.C.; Kodikaraarachchi, H.; Worrall, S.T.; Kondoz, A.M.; , "Improved depth map filtering for 3D-TV systems," Consumer Electronics (ICCE), 2011 IEEE International Conference on , vol., no., pp.645-646, 9-12 Jan. 2011

[8] Liang Zhang; Tam, W.J.; , "Stereoscopic image generation based on depth images for 3D TV," Broadcasting, IEEE Transactions on , vol.51, no.2, pp. 191- 199, June 2005

[9] Gisle Bjontegaard "Calculation of average PSNR differences between RD-curves" ITU-Telecomunications Standarization Sector. VCEG - Doc. VCEG-M33, Texas USA 2001.