3D Res. **04**, 02(2013)5
10.1007/3DRes.02(2013)5

3DR EXPRESS

Depth Map Coding using Residual Segmentation for 3D Video System

Cheon Lee • Yo-Sung Ho

Received: 06 October 2012 / Revised: 23 January 2013 / Accepted: 15 April 2013

Abstract Advanced 3D video systems employ multi-view video-plus-depth data to support the free-viewpoint navigation and comfortable 3D viewing; thus efficient depth map coding becomes an important issue. Unlike the color image, the depth map has a property that depth values of the inner part of an object are monotonic, but those of object boundaries change abruptly. Therefore, residual data generated by prediction errors around object boundaries consume many bits in depth map coding. Representing them with segment data can be better than the use of the conventional transformation around the boundary regions. In this paper, we propose an efficient depth map coding method using a residual segmentation instead of using transformation. The proposed residual segmentation divides residual data into two regions with a segment map and two mean values. If the encoder selects the proposed method in terms of rates, two quantized mean values and an index of the segment map are transmitted. Simulation results show significant gains of up to 10 dB compared to the state-of-the-art coders, such as JPEG2000 and H.264/AVC.

Keywords 3D video coding, depth map coding, residual segmentation

1. Introduction

With growing demands for comfortable three-dimensional videos, multi-view generation techniques using depth data become a hot issue in multi-dimensional image processing.

Cheon Lee • Yo-Sung Ho (✉)
Gwangju Institute of Science and Technology (GIST)
123 Cheomdan-gwagiro, Buk-gu, Gwangju, 500-712, Korea
Telephone: +82-62-970-2263
Fax: +82-62-970-3164
E-mail: hoyo@gist.ac.kr

However, the increased number of views requires huge channel capacity in result. Consequently, an efficient coding technique is strongly required. In order to fulfill such demand, many research works have been developed with various data formats and coding methods[1,2]. In particular, MPEG (moving picture experts group) and JVT (joint video team) have developed the multi-view video coding (MVC) which compresses multi-view videos using inter-view correlations[3]. It employs inter-view and temporal prediction scheme that expanded the motion estimation into inter-view direction. During the activities on MVC standardization, many coding techniques had been proposed such as prediction structure[4], view synthesis prediction[5,6], illumination compensation (IC)[7], and motion skip mode[8].

Since the MVC can compress only multi-view color data with limited number of views, it is hard to generate intermediate views at the decoder side. After finalizing the standardization on MVC in 2008, experts raised request of a new coding standard that can provide multi-view video generation at the decoder with low complexity. As a result, the multi-view video-plus-depth data (MVD) format has been employed as a new 3D video format. The problem is how to compress the huge amount of data increased by the number of views including depth data.

According to the *Plenoptic sampling* theory, only few viewpoint videos with depth data can render multi-view images at the decoder[9]. Since the depth data provides geometrical information of a scene, we can reconstruct a virtual viewpoint image using the depth-based image rendering (DIBR) techniques[10,11]. Employing the depth maps as supplementary data, it is possible to reduce the number of views to be transmitted. The 3D video coding is under developing by MPEG with a framework that involves the MVD format[12]. Therefore, one of the main issues in 3D video coding is the depth map coding scheme.

There are many approaches to compress the depth map. Morvan *et al.*[13] used the platelet-based depth map coding with quad-tree decomposition and modeling depth edges. Jäger *et al.*[14] proposed a depth coding method by signaling the location of depth edges for JPEG2000. Milani *et al.*[15] presented a depth map coding using a segmentation method to predict the shape of surfaces using a reconstructed color image. Kang *et al.*[16] used geometrical modeling in intra prediction for depth map coding. Extending to the multi-view video, Lee *et al.*[6] proposed view synthesis prediction method, which employs the view depth synthesis method using the depth map to generate additional reference frames.

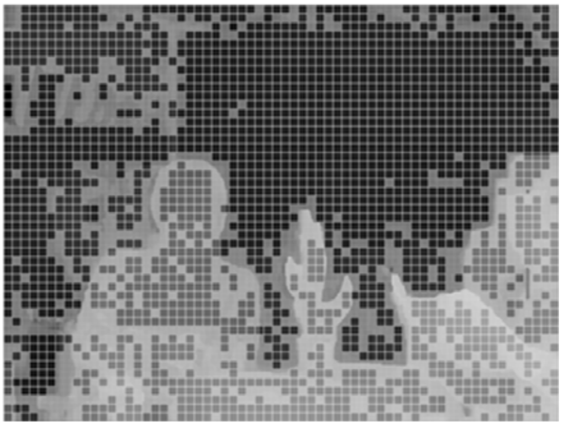
Moreover, several depth coding technologies are under investigation by JCT-3V (joint collaborative team on 3D video coding) group. The coding schemes are classified into two categories as 3D-AVC (AVC compatible 3D video coding) and 3D-HEVC (HEVC compatible 3D video coding)[17]. For 3D-AVC technology, a plane segmentation intra prediction (PSIP) method is proposed for the depth map coding, which uses a segmentation method to divide a block into *k* regions and applies different prediction scheme [18]. For 3D-HEVC technology, a depth modeling mode (DMM) is proposed which uses a wedgelet or a contour to approximate depth values of objects boundaries[19]. In addition, a region boundary chain coding for 3D-HEVC is proposed, which uses a chain coder to describe the shape of object boundaries[20].

In this paper, we aimed at development of an efficient depth map coding utilizing different characteristics of depth data from the color video. The variation of values of the depth map is relatively monotonic compared to that of the color image. In detail, if an object in the color image has many textures such as patterns, abrupt changes of color values are found in both inner and boundary. However, the corresponding depth map contains relatively simple values. The inner part of the object may have similar depth values, whereas the boundaries may have abrupt depth changes due to the different locations of objects. Upon this, residual data generated by intra/inter prediction modes may contain a few high-magnitude error values around object boundaries. Consequently, many transform coefficients are generated. This property inspired us to skip the transformation and to send two representative values of two groups using a segmentation method. In the following Chapter, we will describe the details of the proposed depth map coding using a segmentation method on residual data.

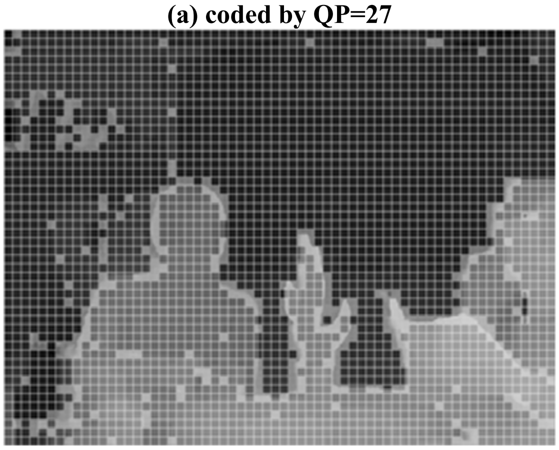
The rest of this paper is organized as follows. In Chapter II, we explain the analysis on depth map coding. In Chapter III, we propose a depth map coding method using residual segmentation. In Chapter IV, we show experimental results for the depth map coding and discussions. Concluding remarks are given in Chapter V.

2. Analysis on Depth Map Coding

Depth maps and color images have different characteristics. Generally, as we mentioned in the introduction, depth values around object boundaries change abruptly while non-boundary regions have monotonic depth values. This abrupt depth change consumes high portion of bits in depth map coding. We visualized this property in figure 1 with depth maps coded by two quantization parameters (QPs) such as QP=27 and QP=37. The brighter regions are coded with Intra4x4 mode and consume many bits due to the transform coefficients. From this result, we observed that object boundaries are coded by Intra4x4 mode, which is the smallest intra block and consumes many bits in general in the H.264/AVC. As a QP goes to higher, the selection rate of Intra4x4 goes to lower. In addition, since depth values around object boundaries change abruptly, the residual data may contain high-magnitude values. These high-magnitude values can generate many transform coefficients that consume many bits in the entropy coding.



(a) coded by QP=27



(b) coded by QP=37

Figure 1 Intra4x4 mode distribution in depth map coding

Here is an analysis on depth map coding that object boundaries consume many bits. By coding 100 frames of 'Newspaper' sequence with intra prediction, we analyzed the coded blocks containing transform coefficients and their consuming bits. Among many prediction modes, we focused on the blocks coded with Intra4x4 modes. Figure 2(a) shows the percentage of the blocks containing non-zero coefficients among whole blocks; the rates of interest blocks are rather high. Figure 2(b) shows the percentage of bits generated by transform coefficients. As can be seen, the percentages are high. Since there are many syntaxes of a block, these values imply that the transform coefficients occupy high portion of whole coded bits. . Figure 2(c) shows the average bits of transform coefficients for a block. The coded bits of the blocks containing transform coefficients are rather high.

The recent coding standard such as the H.264/AVC uses a block-based integer DCT (discrete cosine transform) to use spatial redundancy. The block-based transformation is very effective for color video coding since the residual data induced by prediction still have high spatial correlation. However, it is sometimes inefficient for depth map coding at region of object boundaries. Figure 3 is an example that one high-magnitude residual can generate many transform coefficients in depth map coding; those coefficients consume many bits in entropy coding. Due to the abrupt depth changes, such cases occur frequently. Although most pixels are predicted precisely, one high-magnitude prediction error generates many high-magnitude coefficients. The motivation of this work is that an alternative residual representation method of using transformation can be beneficial for depth map coding.

3. Depth Map Coding using Residual Segmentation

The proposed depth map coding uses a residual segmentation method instead of the block-based transformation. The residual segmentation divides the residual data into two groups. By indicating the meaningful residual values to the decoder, we can reconstruct the block data. Figure 4 describes the overall procedures of the proposed method. It is developed onto the H.264/AVC intra coding since most of residual data are induced by intra prediction. It consists of five steps as follows; 1) the residual data are obtained using the conventional intra prediction, 2) residues are divided into two regions via the proposed residual segmentation method, 3) two mean values of each segment are quantized, 4) the encoder selects the best coding method between the conventional transform-based coding method and the proposed method in terms of coded bits for residual data, 5) the quantized mean values and the index of segment map are coded using the entropy coder.

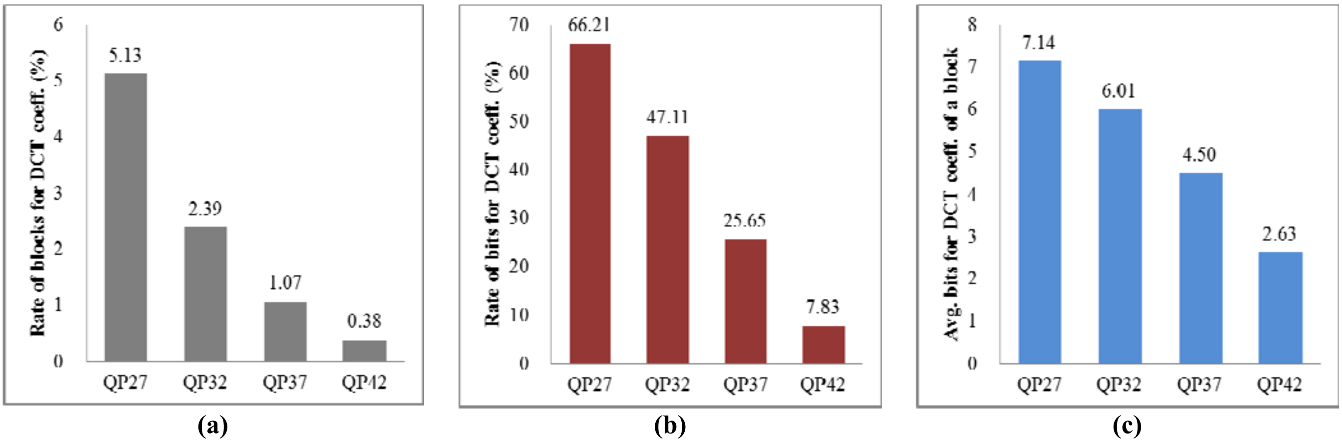


Figure 2 Analysis on depth map coding: (a) percentages of the blocks that coded with transform coefficients, (b) percentages of the bits for transform coefficients compared to the total bits, (c) average bits for blocks containing transform coefficients

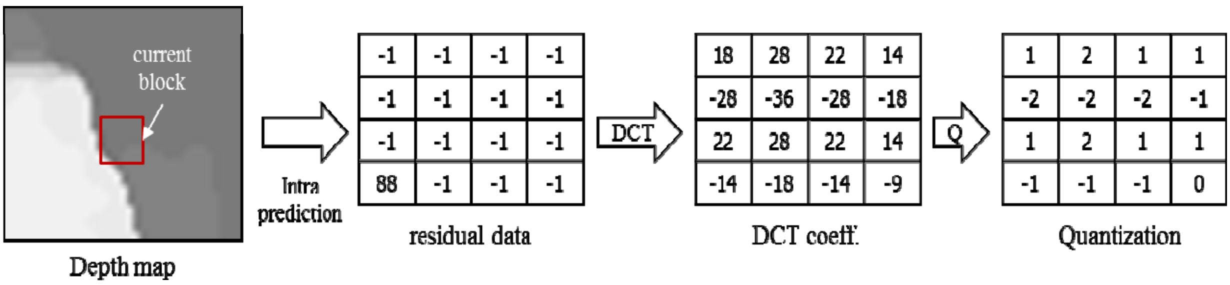


Figure 3 Effect of transformation on object boundary

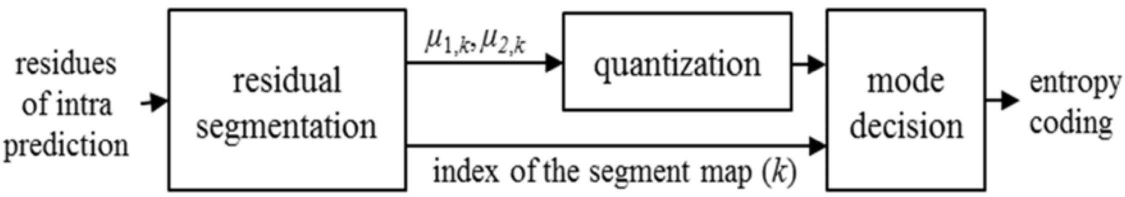


Figure 4 Procedures of residual coding method

A. Residual Segmentation

As we mentioned above, the residual data to be processed are obtained by using the predicted data using the best mode in the conventional modes. Since the objective of this work is to design an alternative coding method of using DCT, we only take blocks containing coefficients into account by checking the number of coefficients of a block. If there is any coefficient in a block, i.e. *coded_block_pattern*≠0, we obtain residual data by subtracting the predicted data of the best mode from the original picture.

The next step is the residual segmentation, which is based on the K-means clustering with K=2. Let $\mathbf{R}=\{R_0, R_1,\ldots, R_{15}\}$ be the residual data of a block arranged by raster scanning, then the algorithm selects the best segment map $\mathbf{G}_k=\{\delta_{0,k}, \delta_{1,k},\ldots, \delta_{15,k}\}$ among the pre-defined 256 segment maps $\mathbf{G}=\{\mathbf{G}_0, \mathbf{G}_1,\ldots, \mathbf{G}_{255}\}$. Since a block contains 16 pixels, the total number of possible segment map is 2^{16}. In order to discard meaningless divisions and to reduce the number of possible segment maps, we selected only 256 segment maps by checking the possibilities of occurrence. Every segment map is consisted of 16 binary values with $\delta\in\{0,1\}$, where 0 indicates the first segment and 1 indicates the other segment. The optimal segment map $\mathbf{G}_k$ is determined by iterative searching as,

$$k = \underset{i\in\{0,1,\ldots,255\}}{\text{argmin}} \sum_{j=0}^{15}\left|R_j - \mu_{1,i}(1-\delta_{j,i}) - \mu_{2,i}\delta_{j,i}\right| \quad (1)$$

where k is the index of the optimal segment map, and $\mu_{1,i}$, $\mu_{2,i}$ are the mean values of each segment with respect to the i-th segment map. Each means are calculated by,

$$\mu_{1,i} = \frac{\sum_{j=0}^{15} R_j(1-\delta_{j,i})}{\sum_{j=0}^{15}(1-\delta_{j,i})} \qquad \mu_{2,i} = \frac{\sum_{j=0}^{15} R_j\delta_{j,i}}{\sum_{j=0}^{15}\delta_{j,i}} \quad (2)$$

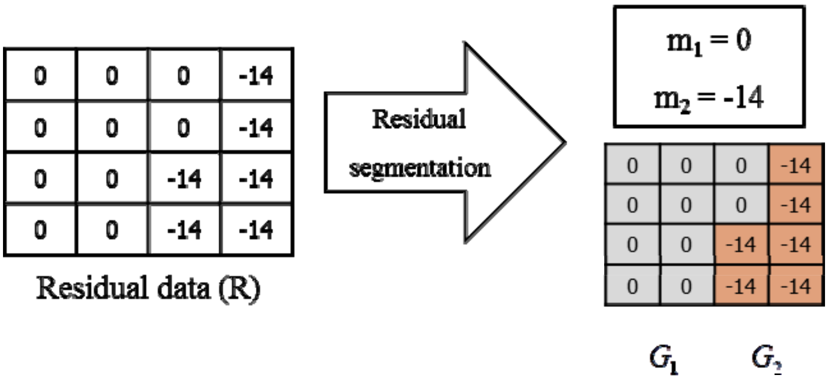


Figure 5 Example of residual segmentation

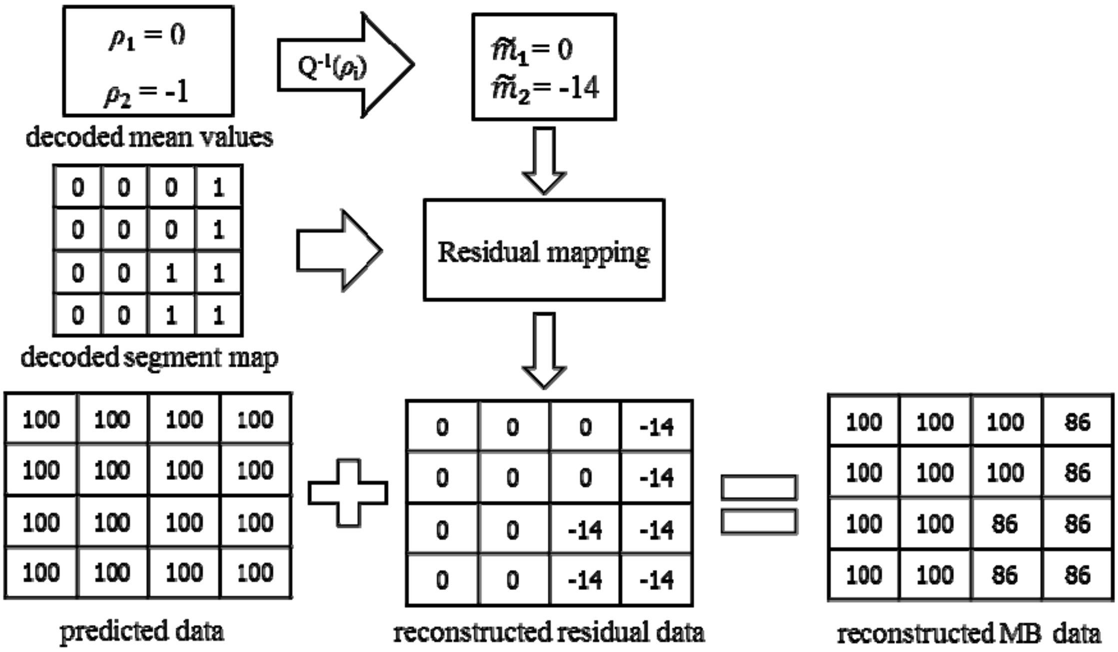


Figure 6 Data reconstruction for a block

Here is an example with figure 5. If residual data consists of two representative values, 0 and -14, we obtain two mean values, m_1=0 and m_2=-14. The segment map consists of binary values. In particular, zeros indicate G_1 group having m_1 while ones indicate G_2 group having m_2. Note that most of the mean values of the first segment, i.e. m_1, are close to zero.

B. Quantization of mean values

Since the proposed method works as an additional mode, we need to select which method between the conventional best mode and the proposed method is efficient for coding. For this, we compare the consuming rates. First, we quantize the resulting mean values of the residual segmentation using a midtread quantizer as described in Eq. (3).

Table 1 Quantization step sizes

QP	0	1	2	3	4	5	6
Q_{step}	0.625	0.6875	0.8125	0.875	1	1.125	1.25
QP	7	8	9	10	11	12	13
Q_{step}	1.375	1.625	1.75	2	2.25	2.5	2.75
…..							
QP	45	46	47	48	49	50	51
Q_{step}	112	128	144	160	176	208	224

$$\rho_1 = round\left(\mu_{1,k} / s \cdot Q_{step} \right)$$
$$\rho_2 = round\left(\mu_{2,k} / s \cdot Q_{step} \right) \quad (3)$$

where Q_{step} is the step sizes referred to the H.264/AVC coder, and s is a scale factor of the quantizer, as presented in Table 1. In addition, we use a controller s for the step size. In practical, we use a fixed value of s at both encoder and decoder. For de-quantization, we use Eq. (4).

$$\widetilde{m}_1 = \widetilde{\rho}_1 \times Q_{step} \times s \quad \widetilde{m}_2 = \widetilde{\rho}_2 \times Q_{step} \times s \quad (3)$$

C. Data reconstruction

The decoding process of the residual data is straightforward. After decoding the transmitted quantized values, we reconstruct the mean values. Since we have decoded the segment pattern by decoding the index of segment pattern, we can assign the binary value for each pixel. Then, the reconstructed residual data can be obtained by,

$$\widetilde{R}_{j\in\{0,1,\ldots,15\}} = (1-\delta_{j,k})\times \widetilde{m}_1 + \delta_{j,k_j} \times \widetilde{m}_2 \quad (5)$$

Figure 6 shows an example of data reconstruction. If the decoded mean values are 0 and -1, the dequantized mean values are 0 and -14. Using Eq. (5), we can obtain the reconstructed residual data. Using the predicted data at the decoder side, we can obtain the final reconstructed block data.

D. Entropy coding for segment data

There are three header data in the proposed coding method: a flag bit for indicating the proposed method, two quantized mean values, and the index for the segment map. The flag bit is coded in MB level; hence each MB has one additional flag bit to indicate the use of the proposed method. The other two data such as quantized mean values and the index of the segment map are coded in the block level. If a block contains transform coefficients, i.e. *coded_block_flag* =1, we encode these two header data instead of transform coefficients. The quantized mean values are coded by means of the signed integer Exp-Golomb-code; it is similar to encoding motion vector difference. After encoding the mean values, we encode the index of the segment map using the fixed-length coding. Since we use only 256 frequent segment maps, we send 8-bits for the index of the segment map to the decoder.

E. Mode decision

The proposed depth coding performs as an additional mode at the encoder; hence we need to select which method is efficient for coding. Here is an assumption for mode decision. Considering that the depth map is used for generation of a virtual view image, minor distortion of depth value is negligible if it does not affect the quality of virtual view image. In this sense, we allow the distortion of the depth values due to the quantization. Since the best mode in the conventional prediction method is, in general, predicted precisely, the distortion of depth values barely deteriorates the quality of the synthesized image.

In order to decide the best mode for the current macroblock, we calculate two costs using Eq. (6) and (7)

$$J_{\text{conv\_mode}} = R(M,C) \quad (6)$$

$$J_{\text{pro\_mode}} = R(M,Z,G) \quad (7)$$

where R represents the actual coding bits for a macroblock. The cost for the conventional mode J_{conv_mode} is the coding bits for the mode type M and the transform coefficients C, and the cost of the proposed mode J_{pro_mode} is the coding bits for the mode type M, two quantized mean values Z, and the index of the segment map G. The encoder selects a lower method for the actual coding method for a macroblock.

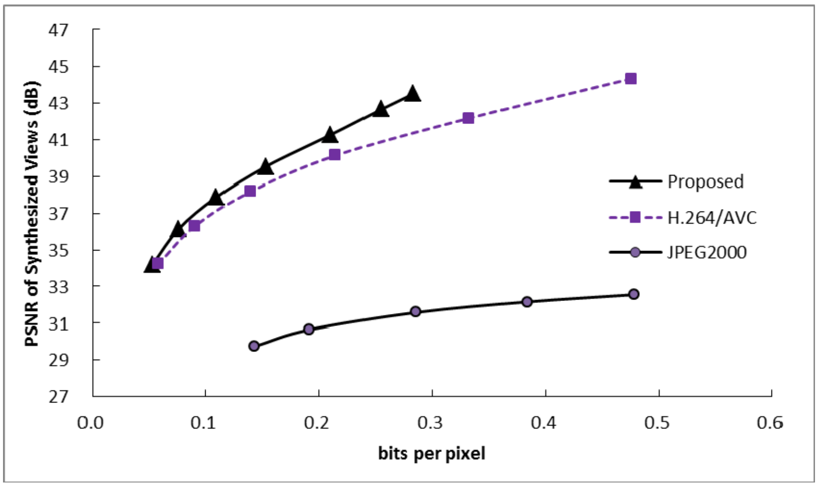
4. Experimental Results

In order to evaluate the coding performance of the proposed depth map coding, we used three test set among the provided 8 sequences from MPEG 3D video coding group [17]. Since there are two classes of test sequences regarding to the resolution, we selected 'Undo_Dancer' sequence for the 1920x1088 resolution sequences and 'Newspaper' and 'Balloons' sequences for the 1024x768 resolution sequences. The depth data of 'Undo_Dancer' sequence has ground truth depth data since those are generated by the computer graphics. However, the depth data of the other two sequences are generated by the depth estimation software; hence the depth may contain erroneous depth values.

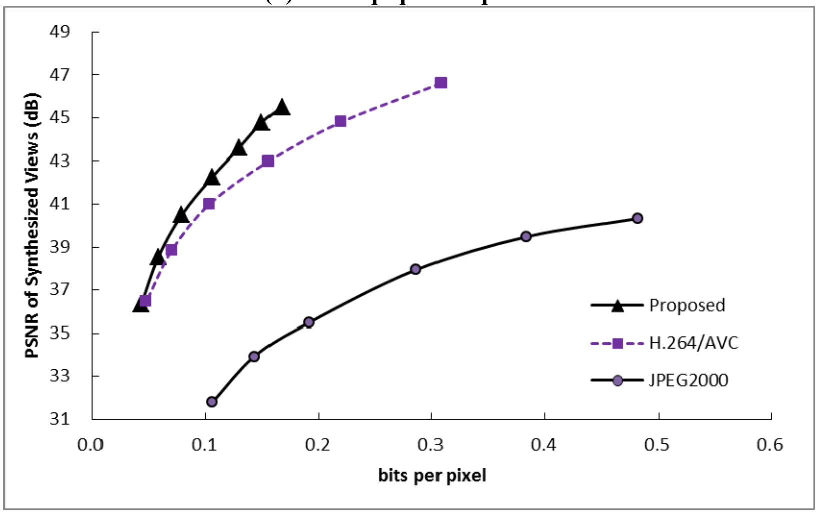
The proposed residual coding is implemented in the reference software JMVC 8.3 (joint multi-view video coding). We applied the proposed coding scheme only to the Intra4x4 blocks. Since even in P- or B- slices the intra prediction is employed as well as the inter prediction, our proposed method can be applied in any slice type. We tested 100 frames for testing with varying QPs (quantization parameters) from 18 to 38. After reconstructing the depth video, we generated three intermediate views using the VSRS-1D-Fast software provided by JCT-3V[21], where the color videos for the virtual view generation were the reconstructed data coded with the same QP set. The quality of the coded depth map was evaluated with the quantity of PSNR (peak signal-to-noise ratio) values of the generated three intermediate views[22]; it is PSNR of synthesized views. The coding performances were compared with the state-of-the-art coders, i.e. JPEG2000, H.264/AVC, and the proposed algorithm. For the JPEG2000, the Jasper software was used with a rate parameter varying from 0.1 to 0.002. The color data for virtual view generation were coded by H.264/AVC INTRA with the same QP set.

Figure 7 shows the experimental results of the INTRA only coding. The RD curves of the proposed method are higher than the H.264/AVC. We observed that the coding gains of high bitrates are greater than that of low bitrates. This can be explained by the fact that the coder at low bitrates loses most of residual data due to the quantization while our method is applied to blocks containing coefficients. In this sense, we observed that the proposed method is hard to contribute the coding gain in predictive coding such as P- or B-picture coding.

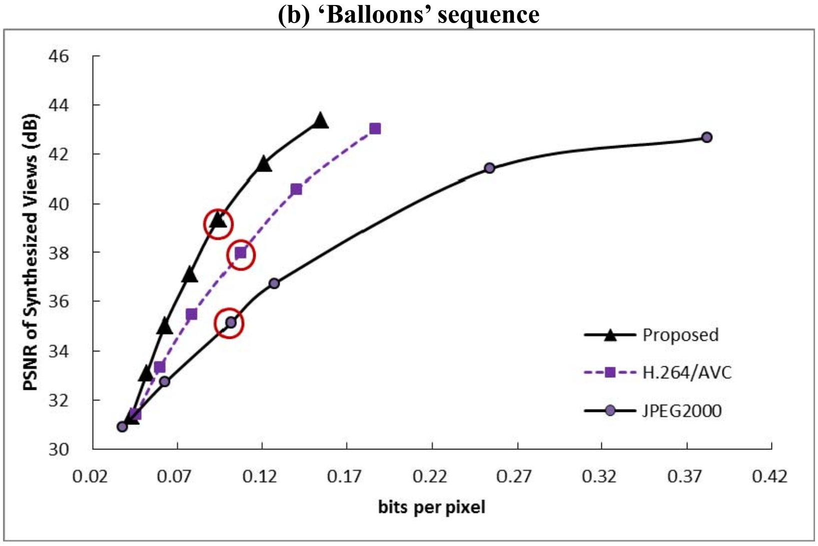
In the proposed method, the scaling factor s is used for quantization of defined mean values in (3) and (4). Since we refer to the fixed quantization step sizes according to the QP values of the H.264/AVC coder, we can control the fidelity of residual data. If it is close to 1, the quantized mean values have low quantization errors.



(a) 'Newspaper' sequence



(b) 'Balloons' sequence



(c) 'Undo_Dancer' sequence

Figure 7 Rate-distortion curves of the depth map coding

Therefore, the reconstructed residual data are close to the original, while the quantized mean values consume more bits for coding. On the other hand, if it is larger than 1, the quantization errors of the mean values become larger. Although the consuming bits for quantized mean values less than that of the previous case, the accuracy of the residual data become lower. Therefore, selecting optimal value of the scaling factor is important task in the proposed method. As shown in figure 8, selection of the scaling factor affects the coding performance. For the case of s=1, the coding gain was close to the results of the H.264/AVC. As the scaling factor goes to higher value, the coding performances were getting better. However, the coding performances of *s*=3 and *s*=4 were very close to each other.

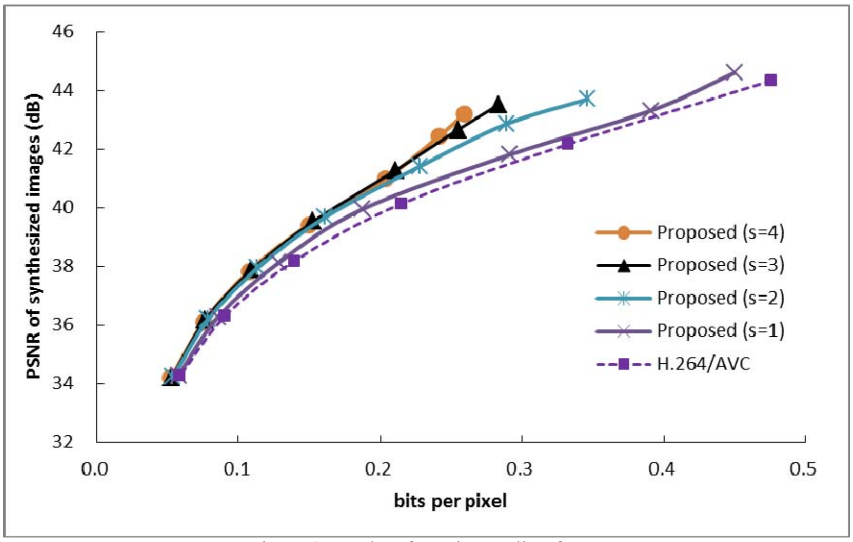


Figure 8 Results of varying scaling factor

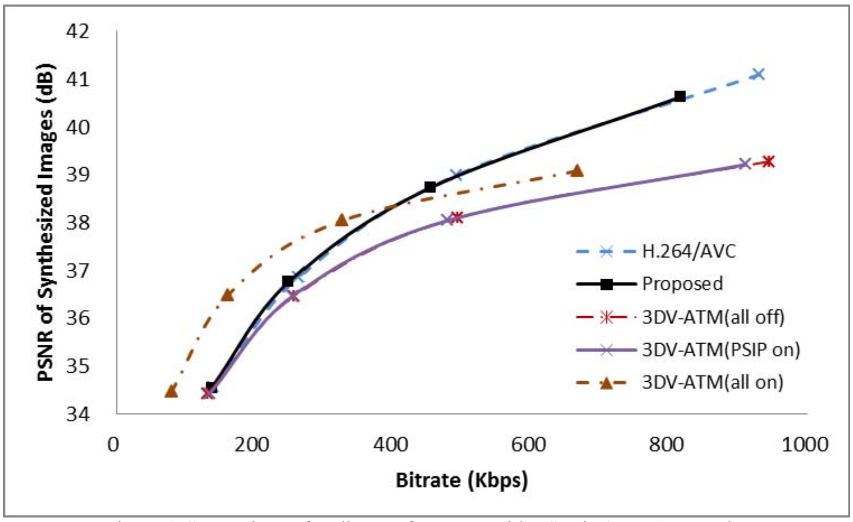


Figure 1 Comparison of coding performance with PSIP in 3DV-ATM coder

To evaluate the coding performance with the recent state-of-the-art 3D-AVC coder developed by JCT-3V, we used 3DV-ATM (ver. 6.1)[22]. Since the proposed method is aimed at the intra coding method in the H.264/AVC coder, we discarded the 3D-HEVC coder. In addition, we synchronized the coding structure of the proposed method with that of 3DV-ATM to make a fair comparison; it is the MVC-like inter-view referable hierarchical B-picture structure, where there is only one I-picture in a GOP (group of pictures). Based on these conditions, we tested the PSIP technique[18] as disabling all depth coding methods (3DV-ATM all off), enabling only the PSIP (3DV-ATM PSIP on), and enabling all depth coding methods (3DV-ATM all). Figure 9 shows the comparison of coding performance with the PSIP implemented in the 3DV-ATM coder. Both the conventional H.264/AVC coder and the proposed method showed almost the same performance since there is only one I-picture is conducted in one GOP. The results of 3D-AVC (all off) and 3D-AVC (PSIP) showed worse coding performance than that of the proposed method in all bitrates. Only the result of the 3D-AVC (all on) showed better coding performance in low-bit rates.

To evaluate the overall coding performance, we employed the Bjontegaard Delta bitrate (BDBR) and PSNR (BDPSNR) measures[23]. By selecting four QP sets, i.e. 18, 22, 26 and 30, we evaluated the coding performance. Additionally, we used the hierarchical B-picture coding by setting GOP (group of pictures) to eight. Table 2 describes the coding performance of the proposed method by comparing with H.264/AVC. The maximum coding gain was the INTRA only case of 'Balloons' sequence, -21.91 % bit-saving or 2.10 dB quality improvements. On average, the INTRA only coding structure showed -20.32 % bit-saving.

Table 2 Coding performance of each sequence

Test data (depth)	INTRA only		Hierarchical B	
	BDBR (%)	BDPSNR (dB)	BDBR (%)	BDPSNR (dB)
Undo_Dancer	-20.46	1.41	-3.84	0.27
Balloons	-21.91	2.10	-1.66	0.07
Newspaper	-18.59	1.86	-5.37	0.22
(average)	-20.32	1.79	-3.62	0.19

Figure 10 demonstrates the selected regions with the proposed method; brighter regions in the figure indicate the selected regions. As can be seen, the proposed method is used frequently at the case of a lower QP. In addition, the selected regions are distributed around the object boundaries since the proposed method is applied to the blocks containing transform coefficients. In higher QP, e.g. QP38, the selected regions are relatively minimal. To evaluate results of reconstructed depth maps, we selected one decoded frame of 'Undo_Dancer' and compared the depth maps, as shown in figure 11. Each depth maps are selected from the results indicated by the red circles in figure 7(c). The result of JPEG2000 shows that depth values around object boundaries have lots of coding errors. However, other two methods were similar each other; but the proposed method consumes less bits. To evaluate the subjective evaluation of the coded depth map, we compared the synthesized images, as shown in figure 12. Similarly with the depth map comparison, the synthesized images of the H.264/AVC and the proposed method show similar subjective quality.

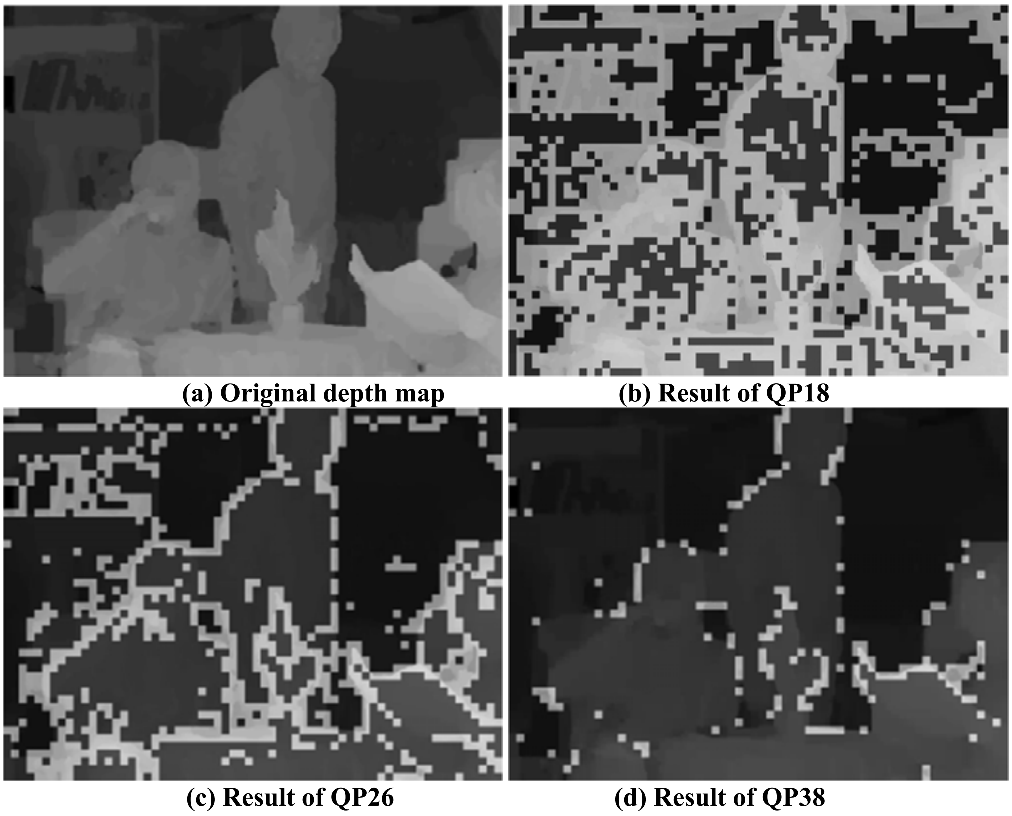


Figure 10 Coded blocks with the proposed coding method

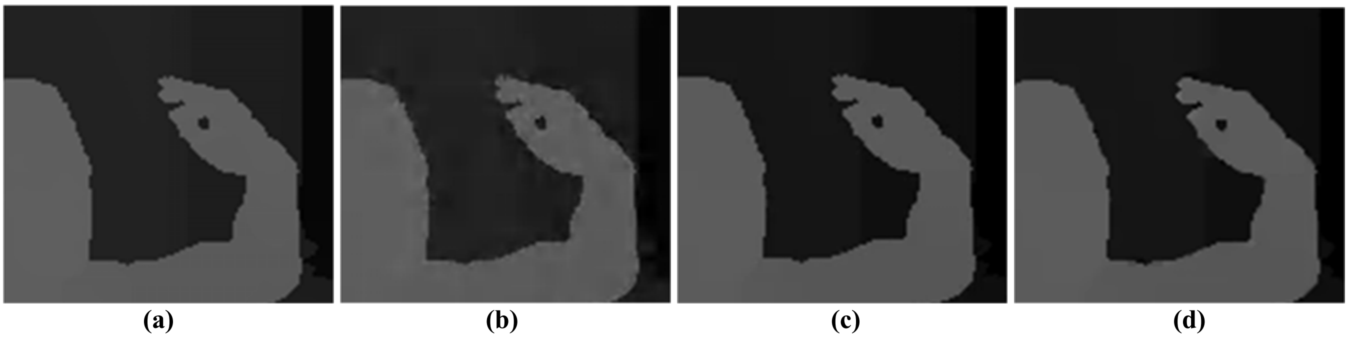


Figure 11 Coded depth maps: (a) original depth map, (b) JPEG2000, (c) H.264/AVC, (d) proposed

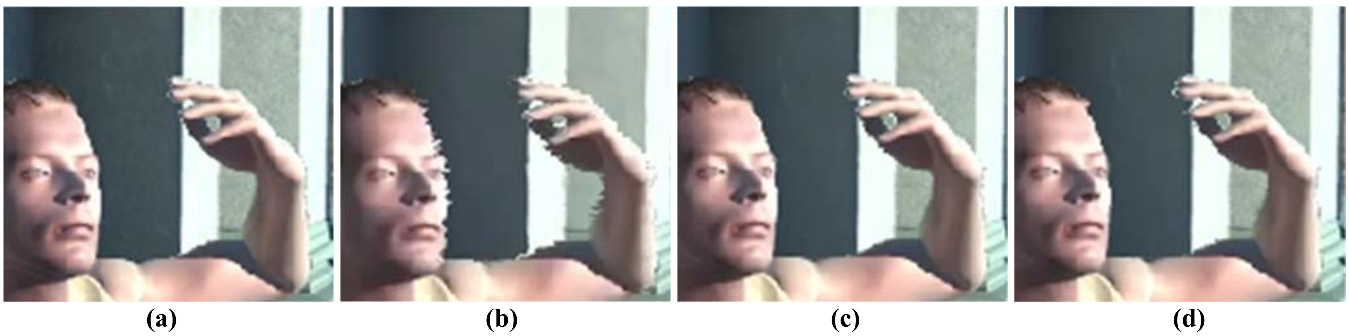


Figure 12 Synthesized images using: (a) original depth maps, (b) coded with JPEG2000, (c) coded H.264/AVC, (d) coded with proposed method

5. Conclusion

In this paper, we proposed an efficient depth map coding using the proposed residual segmentation method instead of using the conventional transformation coding. Since depth values around object boundaries changes abruptly without mixture pixels, the magnitude of residual data is rather high; it results in a lot of transform coefficients consuming many bits. To reduce the bits of transform coefficients, we proposed the residual segmentation method to divide the residual data into two groups. Most of the residual data can be divided into zero-mean and non-zero mean group. By signaling these mean values and the segment information to the decoder, we obtained high coding gains around the object boundaries. By experiments, the intra only coding structure showed -20.32 % bit saving on average. The qualities of the reconstructed depth map were maintained compared to the results of the state-of-the-art H.264/AVC coder, whereas the rates were reduced significantly.

Acknowledgment

This work is supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012-0009228).

References

1. L. Onural (2010) Signal Processing and 3DTV, *IEEE Signal Processing Magazine*, **27**:141-142.
2. A. Smolić, K. Mueller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, and T. Wiegand (2006) 3D Video and Free Viewpoint Video - Technologies, Applications and MPEG Standards, *IEEE International Conference on Multimedia & Expo*, 2161-2164.
3. ISO/IEC JTC1/SC29/WG11 (2009) Study Text of ISO/IEC 14496-10:200x/FPDAM 1 Constrained Baseline Profile and supplementary enhancement information, *MPEG output document N10540*.
4. P. Merkle and K. Müller (2007) Efficient Prediction Structures for Multiview Video Coding, *IEEE Transactions on Circuits and Systems for Video Technology*, **17**:1461-1473.
5. S. Yea and A. Vetro (2009) View Synthesis Prediction for Multiview Video Coding, *Signal Processing: Image Communication*, **24**:89-100.
6. C. Lee, B. H. Choi, and Y. S. Ho (2011) Efficient Multiview Depth Video Coding using Depth Synthesis Prediction, *Optical Engineering*, **20**(077004):1-14.
7. Y. L. Lee, J. H. Hur, Y. K. Lee, K. H. Han, S. H. Cho, N. H. Hur, J. W. Kim, J. H. Kim, P. L. Lai, A. Ortega, Y. Su, P.Yin, and C. Gomila (2006) CE11 : Illumination Compensation, *JVT of ISO/IEC MPEG & ITU-T VCEG JVT-U052*.
8. H. S. Koo, Y. J. Jeon, and B. M. Jeon (2007) MVC Motion Skip Mode, *JVT of ISO/IEC MPEG & ITU-T VCEG JVT-W081*.
9. J. X. Chai, X. Tong, S. C. Chan, and H. Y. Shum (2000) Plenoptic Sampling, *ACM SIGGRAPH*, 307-318.
10. C. Fehn (2004) Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV, *Stereoscopic Displays and Virtual Reality Systems XI*, 93-104.
11. C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski (2004) High-quality Video View Interpolation using a Layered Representation, *ACM SIGGRAPH*, 600-608.
12. ISO/IEC JTC1/SC29/WG11 (2009) Vision on 3D Video, *MPEG output document N10357*.
13. Y. Morvan, P. H. N. De With, and D. Farin (2006) Platelet-based Coding of Depth Maps for the Transmission of Multiview Images, *Stereoscopic Displays and Virtual Reality Systems XIII*, 6055:1-12.
14. F. Jäger (2011) Contour-based Segmentation and Coding for Depth Map Compression, *IEEE Visual Communications and Image Processing,* (6115989):1-4.
15. S. Milani, P. Zanuttigh, M. Zamarin, and S. Forchhammer (2011) Efficient Depth Map Compression Exploiting Segmented Color Data, *IEEE International Conference on Multimedia and Expo*, (6011969):1-6.
16. ISO/IEC JTC1/SC29/WG11 (2011) Description of 3D Video Coding Technology Proposal by Nagoya University, *MPEG input document M22567*.
17. ISO/IEC JTC1/SC29/WG11 (2011) Call for Proposals on 3D Video Coding Technology, *MPEG output document N12036*.
18. B. T. Oh, H. C. Wey, and D. S. Park (2012) Plane Segmentation based Intra Prediction for Depth Map Coding, *Picture Coding Symposium* (6213281) 41-44.
19. P. Merkle, C. Bartnik, K. Muller, D. Marpe, and T. Wiegand (2012) 3D Video: Depth Coding based on Inter-component Prediction of Block Partitions, *Picture Coding Symposium*, (6213308)149-152.
20. JCT-3V (2012) CE6.h Region Boundary Chain Coding for Depth-map, *JCT2-A0070*.
21. ISO/IEC JTC1/SC29/WG11 (2011) Description of 3D Video Technology Proposal by Fraunhofer HHI, *MPEG input document M22570*.
22. ISO/IEC JTC1/SC29/WG11 (2012) Common Test Conditions for AVC and HEVC-based 3DV, *MPEG output document N12560*.
23. G. Bjontegaard (2001) Calculation of Average PSNR Differences between RD Curves, *ITU-T SG16/Q6*, 13th VCEG Meeting VCEG-M33.