Fast Mode Decision For Depth Video Coding Based On Depth Segmentation

Yequn Wang[1], Zongju Peng[1],Gangyi Jiang[1], Mei Yu[1] and Feng Shao[1]
[1] Faculty of Information Science and Engineering, Ningbo University
Ningbo, China.
[e-mail: 364055685@qq.com, {pengzongju, jianggangyi, yumei,shaofeng}@nbu.edu.cn]
*Corresponding author: Zongju Peng

Received December 13, 2011; revised March 13, 2012; accepted April 7, 2012; published April 25, 2012

Abstract

With the development of three-dimensional display and related technologies, depth video coding becomes a new topic and attracts great attention from industries and research institutes. Because (1) the depth video is not a sequence of images for final viewing by end users but an aid for rendering, and (2) depth video is simpler than the corresponding color video, fast algorithm for depth video is necessary and possible to reduce the computational burden of the encoder. This paper proposes a fast mode decision algorithm for depth video coding based on depth segmentation. Firstly, based on depth perception, the depth video is segmented into three regions: edge, foreground and background. Then, different mode candidates are searched to decide the encoding macroblock mode. Finally, encoding time, bit rate and video quality of virtual view of the proposed algorithm are tested. Experimental results show that the proposed algorithm save encoding time ranging from 82.49% to 93.21% with negligible quality degradation of rendered virtual view image and bit rate increment.

Keywords: Depth video coding, MVD, fast mode decision.

This work was supported by Natural Science Foundation of China (61171163, 60872094, 61071120, 60832003), Natural Science Foundation of Zhejiang Province (Y1090752, Y1101240), Zhejiang Scientific and Technical Key Innovation Team of New Generation Mobile Internet Client Software (2010R50009). It is also sponsored by K.C.Wong Magna Fund in Ningbo University.

http://dx.doi.org/10.3837/tiis.2012.04.010

1. Introduction

With the development of three-dimensional (3D) display and related technologies, 3D video applications, such as 3DTV and free viewpoint TV (FTV), have been booming up recently. Multiview plus depth (MVD) data format fulfills the 3D video system's requirements and supports wide angle for 3D displays and auto-stereoscopic displays [1]. Hence, MVD is the main 3D representation format. MVD signals consist of multiview color video and multiple associated depth video. Depth video indicates the distance between the captured scene and cameras and can be used to render continuous virtual views by depth-image-based-rendering (DIBR) technique in combination with stereo or multiview video [2].

The data amount of MVD is tremendous because it is proportional to the number of cameras. It is necessary to develop an efficient encoding method for storage and transmission of the MVD signals. Park proposed an improved distributed multi-view video coding method robust to illumination changes among different views [3]. Joint Video Team (JVT), formed by ISO/IEC Moving Pictures Experts Group and the ITU-T Video Coding Experts Group, developed a joint multiview video coding (JMVC) [4] which is based on Hierarchical B Pictures (HBP) structure [5]. JMVC uses rate distortion (RD) optimization technique [6] to select the optimal coding mode. For each macroblock (MB), the encoder calculates the RD cost of all modes and selects the best mode with the minimum RD cost. Consequently, the computational complexity of mode decision is enormous and the process is time-consuming. In 3D video (3DV) system, depth videos and color videos can share the same encoder. Although depth video is monochromatic and has less texture, the coding time of depth video is as much as color videos. Hence, it is very significant to design a fast mode decision algorithm for compressing depth video.

In recent years, many works on accelerating the mode decision process in color video coding have been proposed [7][8]. However, the main goal of these works lies in speeding up the encoding process with negligible RD performance loss because color videos are directly used for end user viewing. However, these methods are not suitable for depth video coding because (1) depth videos are different from color videos and have some special characteristics, and (2) depth videos are geometric information and used in synthesizing virtual view instead of displaying. In terms of virtual view rendering, the edge region of depth video is more important than other regions [9]. Thereafter, slight distortion in flat region introduced by depth video coding may not affect the visual quality of the rendered views. In addition, because of inaccurate depth estimation, the mode distribution in encoding depth videos shows different rules. Thus, during the process of designing a fast depth video encoding algorithm, the above-mentioned factors should be taken into consideration.

Recently, several fast algorithms for depth video coding have been proposed [9][10][11]. Wang utilized the feature of smoothness and stability of depth videos and proposed an early mode termination strategy based on difference detection [10]. Peng et al. proposed a virtual view rendering oriented fast multi-view depth video algorithm, considering the effect of distortion in depth- continuity and depth-discontinuity regions on virtual view rendering and the MB mode distribution in the two regions [9]. Wang et al. proposed a region of interest (ROI) oriented fast mode decision strategy for depth video coding by making the best use of smoothness and stability [11]. ROI detection was introduced to meet the requirements of higher efficiency and keep the edge quality. By using these methods, the speed of encoding depth video can be improved to a certain extent. However, the encoding time saving ratio is still limited and all of these methods did not consider the MB mode correlation.

In this paper, a novel fast mode decision algorithm is proposed for compressing depth videos. The proposed algorithm is designed based on depth perception and MB mode correlation. The rest of this paper is organized as follows. Section II describes the proposed algorithm. Experimental results are shown in Section III and the work is concluded in Section IV.

2. Proposed Algorithm

This section descibes the proposed algorithm in detail. Firstly, the depth video is segmented into three types of regions based on depth perception. The mask of segmentation result functions as a pre-processor and controller during the encoding process. Then, the characteristics of mode selection are analyzed. Finally, a fast mode decision algorithm is proposed based on the results of depth segmentation and analyses of mode selection.

2.1 Depth Segmentation

Depth videos can provide people with depth perception. Visual physiological and visual psychological studies show that people are interested in regions with small depth value and depth discontinuous regions. The regions with small depth value are often closer to viewer and sometimes pop-out from video screen. As the distance between object and observer decreases, the interest increases. Depth discontinuous regions which mainly locate at the edge of objects make viewer feel strong 3D depth sensation. As far as DIBR is concerned, the edge of objects is more important than other regions. According to the analyses, we divide the depth videos into three types of regions: edge, foreground and background. Fig. 1 shows the framework of region segmentation method which **D** is the original depth video, **M** is the output mask. **E** represents edge, **FF** and **FB** represent the foreground and the background in the non-edge region respectively. **F1** and **F2** are the foreground detected by OTSU method and mean based segmentation method respectively. The region segmentation method is described in detail as follows.

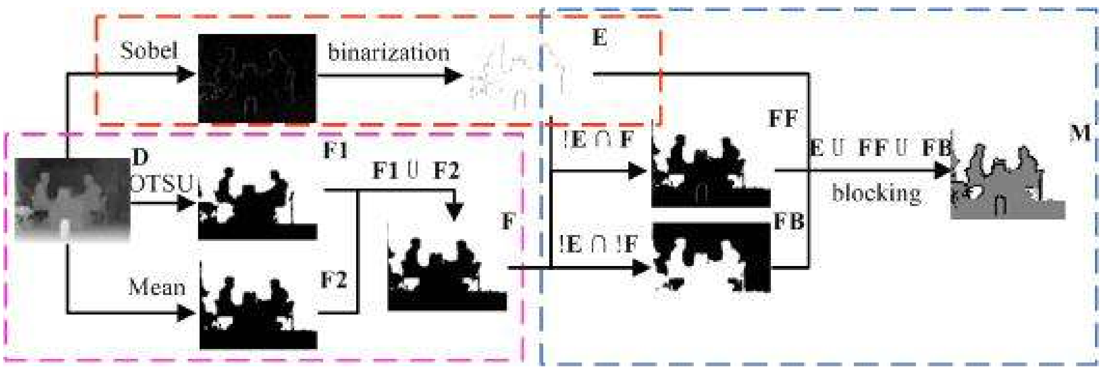


Fig. 1. Framework of region segmentation method

2.1.1 Edge Extraction

The edge is detected by Sobel operator which is simplicity and low complexity. For an image with the resolution of $W \times H$, let $g(x)$ and $g(y)$ be the horizontal and vertical component of gradient. $p(x, y)$ be the edge information detected by Sobel operator, they are calculated by

$$g(x) = | f(x+1, y-1) + 2\times f(x+1, y) + f(x+1, y+1) - f(x-1, y-1) - 2\times f(x-1, y) - f(x-1, y+1) | , \quad (1)$$

$$g(y) = | f(x-1, y+1) + 2\times f(x, y+1) + f(x+1, y+1) - f(x-1, y-1) - 2\times f(x, y-1) - f(x+1, y-1) | \quad (2)$$

and

$$p(x, y) = g(x) + g(y) \quad (3)$$

where $f(x, y)$ is the pixel value of **D** at (x, y).

Adaptive OTSU threshold is adopted to extract the pixels at edges. The OTSU method assumes that an image contains two classes of pixels, edge and non-edge. We exhaustively search for the optimum threshold T_{OTSU} that maximizes the intra-class variance g, defined as a traversing equation

$$g = \omega_0 \omega_1 (\mu_0 - \mu_1)^2 \quad (4)$$

where ω_0 and ω_1 are the probabilities of edge and non-edge, μ_0 and μ_1 are means of foreground and background respectively.

Combining Sobel operator with OTSU threshold, the edge of depth video can be classified by

$$\mathbf{E} = \{(x, y) \mid p(x, y) > T_{OTSU}, x \in [0, W), y \in [0, H)\} \quad (5)$$

As is shown in **Fig. 2-(a)**, the above edge detection method can extract the edge of depth map.

2.1.2 Foreground Extraction

Non-edge region of the video is further classfied into foreground and background. We adopt OTSU method, which is from the perspective of grayscale characteristics.

The foreground detected by OTSU method can be represented by

$$\mathbf{F1} = \{(x, y) \mid f(x, y) > T_{OTSU}, x \in [0, W), y \in [0, H)\} \quad (6)$$

where T_{OTSU} is the adaptive OTSU threshold.

However, due to the inaccuracy of depth estimation and distribution of grayscale, we cannot get a good result by only using the OTSU method. On the other hand, the depth value is the distance between the objects and the camera, it can be used to extract the semantic object in the scene. Hence depth value itself is beneficial to the foreground segmentation. If the pixel value is larger than a threshld which is corresponding to the average pixel value of current depth frame, it belongs to the foreground. The mean based foreground segmentation method can be represented by

$$\mathbf{F2} = \{(x, y) \mid f(x, y) > \alpha \times f_{av}, x \in [0, W), y \in [0, H)\} \quad (7)$$

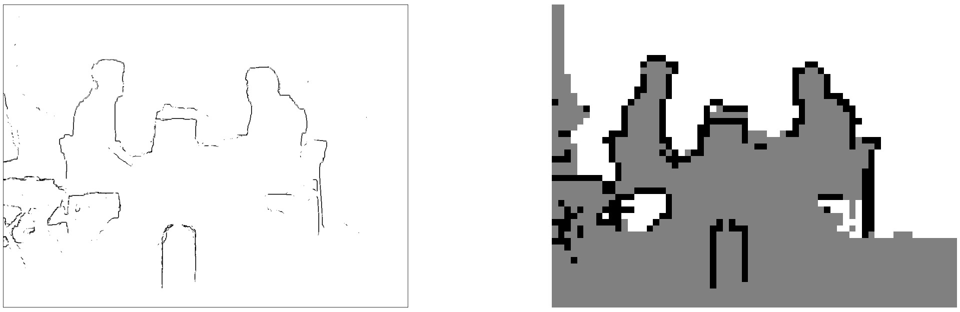
where f_{av} represents the average pixels value of a frame $\frac{1}{W \times H}\sum_{i=0}^{W-1}\sum_{j==0}^{H-1} f(i, j)$, and α is a weighted factor.

Finally, **FF** and **FB** are obtained by

$$f(x, y) \in \begin{cases} \mathbf{FF} & if \quad f(x, y) \in \mathbf{F1} \text{ or } f(x, y) \in \mathbf{F2} \\ \mathbf{FB} & else \end{cases} \quad (8)$$

2.1.3 Segmentation

In order to facilitate the encoder, we finally adjust **E**, **FF** and **FB** into 16×16 block-wise because the MB size is 16×16. If the total number of pixel which belongs to **E** is more than 10, then the block is marked to **E**. If the block is not marked to **E**, and the total number of pixels which belong to **F1** or **F2** is more than 32, the block is marked to **FF**. The other blocks are marked to **FB**. The blocked results were outputted as a mask. Fig. 2 -b) shows the mask of Leave Laptop sequence where the black, gray and white parts represent **E**, **FF** and **FB**.

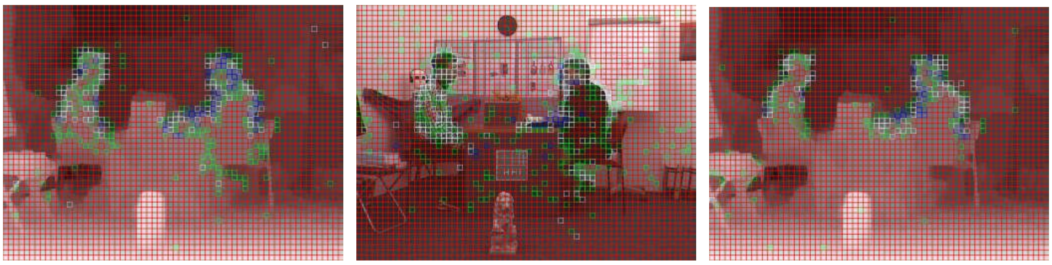


(a) Edge detection result (b) Final mask

Fig. 2. Results of segmentation of Leave Laptop sequence

2.2 Analyses Of Mode Selection

Analyses on MB mode selection is helpful to design a fast MB mode selection algorithm. For convenience of statistics, MB modes are classified into SKIP mode, Intra mode and Inter mode according to the prediction algorithm. Fig. 3 shows the optimal MB mode distribution of Leave Laptop sequence where S_mT_n denotes the n^{th} frame of view m. The blocks with red, green, white and blue borders denote the MB encoded with SKIP, Inter16×16, the other Inter and Intra modes. There are more MBs encoding with SKIP mode in depth video than that of color video due to the simplicity of the depth video. The Inter or Intra modes are mainly distributed in the foreground, moving object regions and object borders.

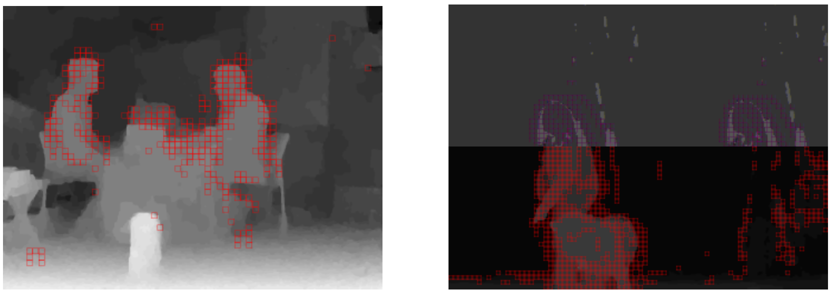


(a) S_7T_8 of depth video (b) S_7T_8 of color video (c) S_7T_4 of depth video

Fig. 3. Optimal mode distribution of Leave Laptop sequence

By comparing Fig. 3-(a) and 3-(c), it can be concluded that strong correlation of SKIP mode exists between current frames and reference frames of depth video. Hence, the optimal mode information of a co-located MB in reference frame can be used to determine the optimal mode of current frame.

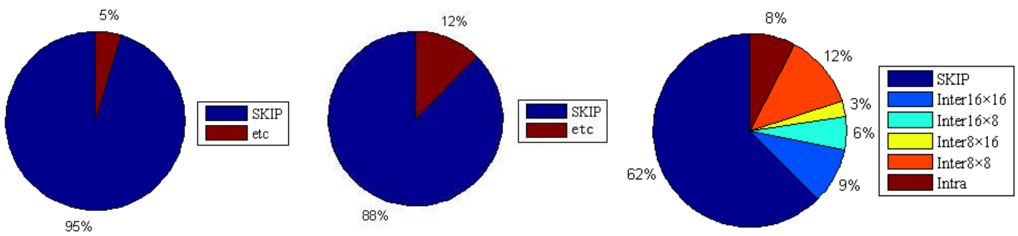
RD cost analyses are also an important cue for designing a fast algorithm. Generally, the RD cost rises when an object's motion becomes violent. Hence, the RD cost can represent the motion characteristics of an MB. For example, if the RD cost of the current MB is similar to the RD cost of the co-located MB in the reference frame, it is highly possible that the optimal modes of current MB is the same as that of the co-located MB. It is especially true for the SKIP mode. Fig. 4 shows the moving areas (red regions) where the RD cost difference between current frame and reference frame is larger than a threshold. Therefore, if the optimal mode of the co-located MB in reference frame is SKIP and the RD cost difference is less than the threshold, we can assume that the current MB is also encoded in SKIP mode.



(a) Leave Laptop (b) Dog

Fig. 4. Moving regions

Fig. 5 shows the mode distribution proportion in three different regions. Where “etc” in **Figs. 5-(a)** and **5-(b)** stands for the other modes except SKIP mode. In the **FB**, almost all the MBs, up to 95%, select SKIP mode. As for the **FF**, there is also a severe imbalance in the mode distribution and nearly 88% MBs are encoded by SKIP mode. However, in the **E**，the modes are distributed balanced relatively, and MBs with SKIP mode is about 62%. In the proposed algorithm, the feature is utilized.



(a) **FB** (b) **FF** (c) **E**

Fig. 5. MB mode distribution proportion in different regions

2.3 Fast Depth Coding Algorithm

The fast depth coding algorithm is proposed based on (1) the depth segmentation and (2) the analyses of mode selection in different regions. **Fig. 6** shows the flowchart of proposed algorithm which is described as follows.

Step 1) If the current MB belongs to first B frame in a GOP, go to Step 2. Otherwise, go to Step 3.

Step 2) Search candidate modes for final encoding MB mode. The candidate modes are determined by the region current MB belongs to. If the MB is in **FB** or **FF**, search the SKIP and Intra mode and choose the optimal mode. If the MB is in **E**, search all modes to obtain the best mode. Then go to Step 1 for next MB.

Step 3) Search SKIP mode and calcuate the RD cost value. The purpose of estimating SKIP mode is to get the motion characteristics of current MB. If the absolute value of RD cost difference between current MB and co-located MB is less than a threshold T, we can assume the MB is static , go to Step 4. Otherwise, go to Step 2.

Step 4) For a static MB, we can further make use of the MB mode correlation. If the MB is in **FB**, the MB is directly set to SKIP mode without any other mode decision work. If the MB is in **FF**, choose the SKIP mode as the optimal mode only on the condition that the co-located MB, the top MB and the left MB are all encoded with SKIP mode.Otherwise, search the SKIP and Intra mode. If the MB is in **E**, search all modes if the co-located MB is encoded with SKIP or Inter16×16, otherwise, search SKIP, Inter16×16 and Intra mode. Then go to Step 1 for next MB.

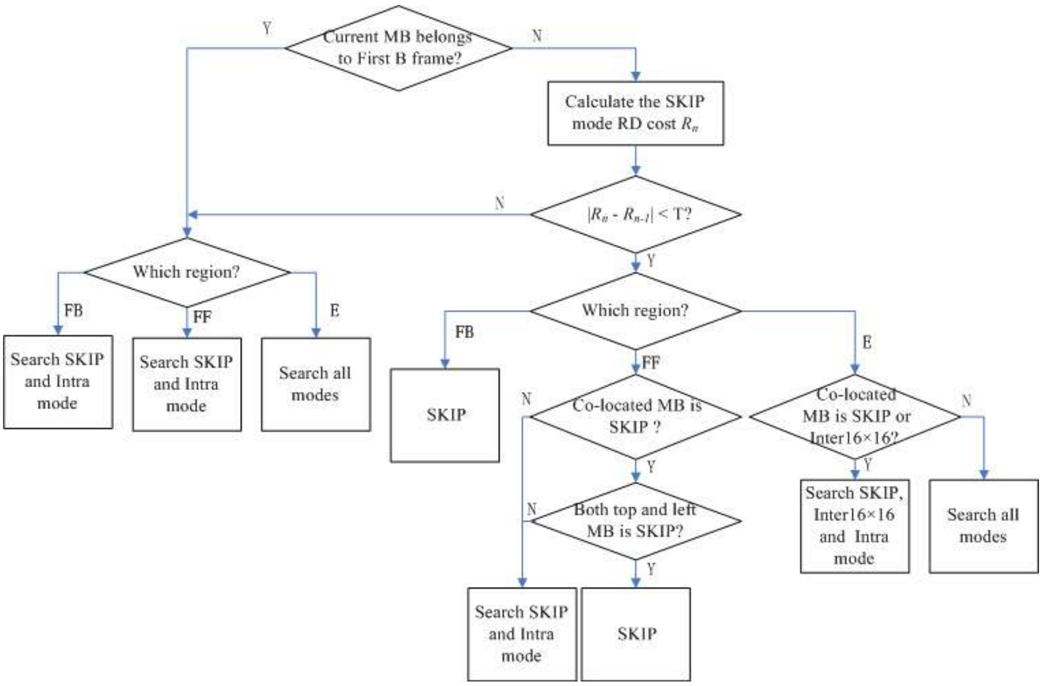


Fig. 6. Flowchart of the proposed fast depth coding algorithm

3. Experimental Results and Analyses

The proposed depth video segmentation method is implemented with C code. The threshold T is set as 200 empirically in the process of extracting moving areas by RD cost. α is set as 0.8 empirically. In order to evaluate the performance of the proposed fast depth coding algorithm, we performed our proposed algorithm on the reference software of JMVC. Five sequences, Book Arrival, Door Flowers, Leave Laptop, Dog and Newspaper were tested under common test conditions. These sequences except Newspaper are evaluated from the multiview color videos by DERS5.0. The detailed parameters are shown in **Table 1.** The performance of the proposed algorithm is overall evaluated by comparing with Peng's algorithm [9] and Wang's reconstructed frame based difference detection (RDD) algorithm [10].

Table 1. Parameters of simulation

software	Basis QP	DeltaLayerXQuant						Structure
JMVC 8.3	22,27, 32,37	0	1	2	3	4	5	HBP
		0	3	4	5	6	7	

Depth sequences		Resolution	views	Properties of array	GOP length	Frames of encoding
HHI	Door Flowers	1024×768	7, 10	1D/parallel	15	61

HHI	Book Arrival	1024×768	7, 10	1D/parallel	15	61
HHI	Leave Laptop	1024×768	7, 10	1D/parallel	15	61
Nagoya	Dog	1280×960	38, 41	1D/parallel	15	61
GIST	Newspaper	1024×768	4, 6	1D/parallel	15	61

Table 2 compares the time reduction ratios of Wang's RDD algorithm, Peng's algorithm and the proposed algorithm where

$$\Delta T_{Fast} = \frac{T_{JMVC} - T_{Fast}}{T_{JMVC}} \times 100\% \ , \tag{9}$$

$$\Delta T_{Peng} = \frac{T_{JMVC} - T_{Peng}}{T_{JMVC}} \times 100\% \tag{10}$$

and

$$\Delta T_{RDD} = \frac{T_{JMVC} - T_{RDD}}{T_{JMVC}} \times 100\% \ . \tag{11}$$

T_{JMVC}, T_{Fast}, T_{Peng} and T_{RDD} are encoding time of the JMVC platform, the proposed algorithm (including time of depth segmentation), Peng's algorithm and Wang's RDD algorithm. The proposed algorithm saves the encoding time ranging from 82.49% to 93.21%, while the time reduction of Peng's algorithm is from 70.75% to 90.50% and RDD algorithm is from 7.59 % to 45.67%. Since the proposed algorithm utilizes not only the statistical properties of mode distribution, but also the MB mode correlation of depth videos, its speedup performance outperforms both RDD algorithm and Peng's algorithm. The speedup performance of Dog sequence is better than that of other sequences because 1) edge region which needs to be protected only amounts to a small proportion, and 2) large flat regions exist in Dog sequence and the correlation of time-spatial MB modes is strong.

Table 2. Comparison of time reduction ratio regarding Wang's RDD algorithm, Peng's algorithm and the proposed algorithm (%)

QP & Seq. / Res. #		Book Arrival	Door Flowers	Leave Laptop	Dog	Newspaper
22	$\triangle T_{RDD}$	11.97	17.08	16.68	45.67	19.72
	$\triangle T_{Peng}$	70.75	72.77	73.85	87.14	77.95
	$\triangle T_{Fast}$	82.60	84.55	84.48	93.21	86.48
27	$\triangle T_{RDD}$	11.20	15.72	15.11	45.53	19.30
	$\triangle T_{Peng}$	71.02	73.04	74.09	87.32	77.88
	$\triangle T_{Fast}$	82.66	84.62	84.57	93.06	86.32
32	$\triangle T_{RDD}$	8.66	12.39	11.83	45.54	17.82
	$\triangle T_{Peng}$	70.93	73.16	74.26	87.44	77.80
	$\triangle T_{Fast}$	82.62	84.57	84.64	92.85	86.19
37	$\triangle T_{RDD}$	7.59	11.61	11.34	20.15	17.14
	$\triangle T_{Peng}$	71.04	73.19	74.36	87.68	77.78
	$\triangle T_{Fast}$	82.49	84.27	84.40	92.60	85.76

For the compression performance, the RDD algorithm does not cause any degradation. **Table 3** lists the bit rate comparison between Peng's algorithm and the proposed algorithm where

$$\Delta BR_{Fast} = \frac{BR_{Fast} - BR_{JMVC}}{BR_{JMVC}} \times 100\% \tag{12}$$

and

$$\Delta BR_{Peng} = \frac{BR_{Peng} - BR_{JMVC}}{BR_{JMVC}} \times 100\% . \tag{13}$$

BR_{JMVC}, BR_{Fast} and BR_{Peng} are bit rate of JMVC platform, the proposed algorithm and Peng's algorithm. Compared with full search algorithm of JMVC platform, the bit rate of Peng's algorithm and the proposed algorithm slightly varies. The general trend is that the bit rate will increase for complex depth video such as Book Arrival or under the case of high bit rate. The underlying reason of the phenomenon is that more MBs will not select real optimal mode in complex depth video or under the case of high bit rate. For Dog sequence, bit rate is saved up to 5.87% and 4.55% for Peng's algorithm and the proposed algorithm, because more MBs are encoding with SKIP modes. **Table 3** also shows that only the negligible bit rate change is introduced by the proposed algorithm while the encoding speedup is greatly accelerated.

Table 3. Bate rate comparison Peng's algorithm and proposed algorithm (%)

Sequences	ΔBR_{Peng}				ΔBR_{Fast}			
	22	**27**	**32**	**37**	**22**	**27**	**32**	**37**
Book Arrival	1.33	-0.72	-0.15	1.91	3.98	2.14	2.2	2.96
Door Flowers	0.86	-1.19	-1.4	-0.57	2.92	-0.01	-1.02	-0.78
Leave Laptop	0.91	-1.85	-1.18	0.54	2.36	0.34	0.2	1.17
Dog	-1.89	-5.87	-5.87	0.28	1.69	-3.87	-4.55	-2.37
Newspaper	-0.14	-1.3	-1.31	-0.09	2.79	0.87	0.15	0.35

Since depth videos are not for final viewing by end users but for virtual view rending. What the end users see is the rendered virtual views. In order to verify the effectiveness of the proposed fast depth coding algorithm, it's suitable to measure the quality of rendered virtual views. We use the MPEG View Synthesis Reference Software (VSRS3.5) to synthesize the intermediate virtual view. The selected left, right and virtual views are tabulated in **Table 4**.

Table 4. Selected views for rendering

Depth sequences	**Left view**	**Right view**	**Virtual view**
Book Arrival, Door Flowers and Leave Laptop	7	10	8
Dog	38	41	39
Newspaper	4	6	5

Table 5 shows the PSNR difference of virtual view video between the proposed algorithm, Wang's RDD algorithm and Peng's algorithm where

$$\Delta PSNR_{Fast} = PSNR_{Fast} - PSNR_{JMVC} , \tag{14}$$

$$\Delta PSNR_{Peng} = PSNR_{Peng} - PSNR_{JMVC} \tag{15}$$

and

$$\Delta PSNR_{RDD} = PSNR_{RDD} - PSNR_{JMVC} . \tag{16}$$

$PSNR_{JMVC}$, $PSNR_{Fast}$, $PSNR_{Peng}$ and $PSNR_{RDD}$ are virtual view PSNR rendered by reconstructed depth video of JMVC platform, the proposed algorithm, Peng's algorithm and RDD algorithm.

The RDD algorithms has not differences while both Peng's and the proposed algorithms have negligible differences in range from -0.04 dB to 0.01 dB, indicating that all the three algorithms maintain rendering quality. **Fig. 7** shows the rendered results of Leave Laptop sequence. Clearly, there is no significant rendering quality degradation between the full search and the proposed algorithm.

Table 5. Virtual view quality comparison between the proposed algorithm and Peng's algorithm (dB)

Sequences	$\triangle PSNR_{RDD}$				$\triangle PSNR_{Peng}$				$\triangle PSNR_{Fast}$			
	22	**27**	**32**	**37**	**22**	**27**	**32**	**37**	**22**	**27**	**32**	**37**
Book Arrival	0.00	0.00	0.00	0.00	0.00	0.01	0.01	-0.01	0.01	0.01	0.00	-0.02
Door Flowers	0.00	0.00	0.00	0.00	0.01	0.01	0.00	-0.02	0.01	0.01	0.00	-0.03
Leave Laptop	0.00	0.00	0.00	0.00	0.00	0.00	0.01	-0.02	0.00	0.01	0.00	-0.03
Dog	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	-0.03	-0.01	-0.03	-0.04	-0.02
Newspaper	0.00	0.00	0.00	0.00	0.01	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01

The structural similarity metrics (SSIM) index is a method for measuring the similarity between two images [12]. **Table 6** shows the SSIM index. $SSIM_{Proposed}$, $SSIM_{Peng}$ and $SSIM_{RDD}$ are the SSIM index of virtual view video rendered by reconstructed depth video of the proposed algorithm, Peng's algorithm and RDD algorithm. The SSIM index of all algorithms are almost the same, which indicates the structure of virtual views for all algorithms are very similar to the original view.

From **Table 5**, **Table 6** and **Fig. 7**, the rendered virtual views have good quality for all three algorithms.

Table 6. SSIM index of Wang's RDD algorithm, Peng's algorithm and the proposed algorithm

	QP	**22**	**27**	**32**	**37**
$SSIM_{RDD}$	Book Arrival	0.9619	0.9619	0.9619	0.9615
	Door Flowers	0.9613	0.9614	0.9615	0.9610
	Leave Laptop	0.9621	0.9623	0.9622	0.9620
	Dog	0.9551	0.9547	0.9542	0.9546
	Newspaper	0.9651	0.9651	0.9653	0.9649
$SSIM_{Peng}$	Book Arrival	0.9619	0.9619	0.9620	0.9615
	Door Flowers	0.9614	0.9615	0.9615	0.9609
	Leave Laptop	0.9621	0.9623	0.9622	0.9619
	Dog	0.9550	0.9547	0.9541	0.9544
	Newspaper	0.9651	0.9652	0.9653	0.9649
$SSIM_{Proposed}$	Book Arrival	0.9619	0.9619	0.9619	0.9614
	Door Flowers	0.9614	0.9615	0.9615	0.9609
	Leave Laptop	0.9621	0.9623	0.9622	0.9619
	Dog	0.9550	0.9546	0.9540	0.9542
	Newspaper	0.9651	0.9652	0.9653	0.9649

4. Conclusion

In this paper we present a novel algorithm for compression of depth video in 3D video, aiming at accelerating the encoding process. Firstly, each frame of depth video is segmented into three types of regions based on depth perception. Then, mode distribution and MB mode correlation are analyzed. Based on region types and the analyses of mode decision, different mode candidates are searched in different region. The experimental results show that the proposed algorithm reduces the encoding time by 82.49% to 93.21% with a negligible quality degradation for virtual view and bit rate increment.


(a) Full search algorithm


(b) Proposed algorithm

Fig. 7. Synthesized results of Leave Laptop sequence

References

[1] K. Müller, P. Merkle, G. Tech and T. Wiegand, "3D video formats and coding methods," In *Proc. of IEEE International Conference on Image Processing*, pp.2389-2392, Sep.2010. Article (CrossRef Link)

[2] K. Müller, P. Merkle and T. Wiegand, "3D video representation using depth maps," in *Proc. of IEEE* , vol.99, no.4, pp.643-656, Apr.2011. Article (CrossRef Link)

[3] S.Park, D.Sim and B.Jeon, "Distributed video coding for illumination compensation of multi-view video," *KSII Transactions on Internet and Information Systems,* vol.4, no.6, pp.1222-1236, Dec.2010. Article (CrossRef Link)

[4] http://iphome.hhi.de/suehring/tml/download/old_jm, "Joint multiview video coding (JMVC)", Version 8.3, May.2010.

[5] H. Schwarz, D. Marpe and T. Wiegand, "Hierarchical B pictures," ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16, Doc. JVT-P014, Jul.2005.

[6] G. Sullivan and T. Wiegand. "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol.15, no.6, pp.74-90, Nov.1998. Article (CrossRef Link)

[7] W. Zhui, W. Jiang and Y. Chen, "A fast inter mode decision for multiview video coding," *IEEE International Conf. on Information Engineering and Computer Science*, pp.1-4, Dec.2009. Article (CrossRef Link)

[8] L. Shen, Z. Liu, T.Yan, et al. "Early SKIP mode decision for MVC using inter-view correlation," *Signal Processing: Image Communication*, vol.25, no.2, pp.88-93, Feb.2010.Article (CrossRef Link)

[9] Z. Peng, M. Yu, J. Jiang, et al, "Virtual view synthesis oriented fast depth video encoding algorithm," *2010 2nd International Conf. on Industrial and Information Systems*, vol.1, pp.204-207, Jul.2010. Article (CrossRef Link)

[10] M. Wang, X. Jin and S. Goto, "Difference detection based early mode termination for depth map coding in MVC," *Picture Coding Symposium*, pp.502-505, Dec.2010. Article (CrossRef Link)

[11] M. Wang, C. Liu, T. Zhang, et al, "Region of interest oriented fast mode decision for depth map

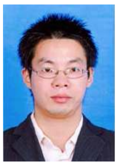
coding in DIBR," *2011 IEEE 7th International Colloquium on Signal Processing and its Applications*, pp.177-180, Mar.2011. Article (CrossRef Link)
[12] Z. Wang, A. Bovik, H. Sheikh and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol.13, no.4, pp.600-612, 2004. Article (CrossRef Link)

Yequn Wang is currently a MS candidate with the College of Information Science and Engineering. Her research interests mainly include digital video compression and communications, multi-view video coding.

Zongju Peng received his B.S. degree from Sichuan Normal College, China, in 1995, and M.S. degree from Sichuan University, China, in 1998, and received his Ph.D. degree from Institute of Computing Technology, Chinese Academy of Science in 2010. He is now an associate professor in Faculty of Information Science and Engineering, Ningbo University, China. His research interests mainly include image/video compression, multi-view video coding and video perception.

Gangyi Jiang received his M.S. degree from Hangzhou University in 1992, and received his Ph.D. degree from Ajou University, Korea, in 2000. He is now a professor in Faculty of Information Science and Engineering, Ningbo University, China. His research interests mainly include digital video compression and communications, multi-view video coding and image processing.

Mei Yu received her M.S. degree from Hangzhou Institute of Electronics Engineering, China, in 1993, and Ph.D. degree form Ajou University, Korea, in 2000. She is now a professor in Faculty of Information Science and Engineering, Ningbo University, China. Her research interests include image/video coding and video perception.



Feng Shao received the B.S. and Ph.D degrees from Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively, all in Electronic Science and Technology. He is currently an associate professor in Faculty of Information Science and Engineering, Ningbo University, China. His research interests include image/video coding, image processing, and image perception.