3D Video Multiple Description Coding Considering Region of Interest

Ehsan Rahimi and Chris Joslin
Department of Systems and Computer Engineering, Carleton University, 1125 Colonel By Dr., Ottawa, ON, Canada
{*ehsan.rahimi, chris.joslin*}*@carleton.ca*

Keywords: Multimedia Communication, 3D Video, Colorful 2D Image, Gray Scale Depth Map Image, Multiple Description Coding, Error Prone Environment, Packet Failure, Spatial Decimation, Compression Efficiency.

Abstract: 3D video is becoming a most favorable video and attracting researcher's mind to provide robust methods of streaming since packet failure has always been the inseparable characteristic of wired or wireless networks. This paper aims to provide a new multiple description coding for 3D video considering objects existed in the scene. To this end, a low complex algorithm for realizing objects in 3D scene will provided and then a non-identical decimation method with respect to objects will be utilized to produce descriptions of MDC approach. Also, in point of depth map image, a new non-identical MDC algorithm will be be introduced to stream depth map image saving bandwidth without affecting the quality of decoded video in the receiver side.

1 INTRODUCTION

Todays, multimedia communication including broadcast TV, video conference, TV on demand, etc are one of the most favorable methods of communication among public. Due to accessibility of the multimedia communication ubiquitously, multimedia consumers and consequently, demands for bandwidth in the last 5 years have been increased dramatically so that current communication technologies cannot keep up with such huge demands. For example, according to CBCnews (2015) downloading of video content from Netflix in North America has been doubled in five years (from 35% in 2010 to 70% in 2015)(CBCnews, 2015); however the video traffic is exponentially increasing, the linear rate of the increment in the video traffic load (being twice in next five years) is difficult for operators to support due to physical restrictions of the communication systems. To make the matter of insufficient resource to stream video worse, 3D videos are now becoming more and more popular among public and video marketing. In addition, according to CTVnews (2016), Cineplex , Canada's largest chain of movie theatres, has announced the new multi screen service in Toronto, Edmonton, and Vancouver (CTVnews, 2016). With the new type of display observers are provided 270 degree of space by one screen in the front and two side screens. To represent 3D video on client's device, depth information needs to be transmitted toward receivers in addition to the color 2D video. Hence, need for higher bandwidth to provide multimedia services are more highlighted in near future.

Smolic and Kimata defined 3D video as ”geometrically calibrated and temporally synchronized video data” (Smolic and Kimata, 2003a); Which means that more memory and bandwidth are required to store or stream 3D video, respectively. Even though, the technology for producing memory has been developed in the last decade, it is still a challenge to save the enormous volume of 3D video data effectively. More importantly, There are some restrictions to stream immersive videos. There are still quite a few challenges to stream 2D HD, or Ultra HD video dynamically, efficiently, and reliably. Since, there is no 3D video coding standard specifically and the main core of current 3D video encoding methods relies on 2D video coding standards highly, those 2D video challenges can also be applicable for 3D video transmission.

As described by MPEG-3DAV (Hewage, 2014; Smolic and Kimata, 2003b), 3D videos can be represented in three ways: panoramic video, stereoscopic video, and multiview video. It can be said that stereoscopic videos are a subset of multiview videos as they capture only two adjacent views just as human visionary system. Therefore, stereoscopic video needs less bandwidth, processing power and storage than multiview video and is more common compared to the other two representations.

Stereoscopic video can be generated in three ways: dual camera configuration, 3D/Depth-range cameras, 2D to 3D conversion algorithms (Hewage,

Rahimi E. and Joslin C.
3D Video Multiple Description Coding Considering Region of Interest.
DOI: 10.5220/0006151802080215
In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017)*, pages 208-215
ISBN: 978-989-758-225-7

2014; Meesters et al., 2004). This paper mainly focuses on the color image plus depth map representation which can be derived from either left and right views or depth cameras.

One of the major problems that occur when delivering videos to users is packet failure. This happens during video streaming in wired and wireless networks. Basically, wireless networks suffer from unreliability due to the noise and interference that exist in the environment. In wired networks error can occur because of packet loss, corruption, and large packet delay. Therefore, it can be said that packet failure is common for wired or wireless networks, however the variances of the error for various channels are different. Such errors may generally produce unacceptable impacts on the delivered video and reduce the quality of experience. To avoid such disappointment by users, error resilient methods of streaming need to be applied by transmitters. In communication systems there are usually three methods to tackle packet failure: Automatic Repeat reQuest (ARQ), Forward Error Correction (FEC), Error Resilient Coding (ERC). The ARQ approach requires a network with feedback capability, so it is not beneficial for real-time or broadcast applications. FEC methods are designed to cope with a specific amount of error and this makes them impractical in environments that the variance of noise fluctuates and exceed the acceptable threshold. As can be understood from the term ”error resilient coding”, it produces an error robust video stream which is done in the source layer. By an ERC method, redundancy bits are added to the stream to enhance its resistance against packet corruption. There are quite a few methods that redundancy can be introduced to the stream such as Reversible Variable Length Coding (RVLC), Intra refreshment, Flexible Macrobloc Ordering (FMO), layered Coding (LC), Multiple Description Coding (MDC), etc. Among these methods, RVLC used in H.263 suffers from the low coding efficiency and Intra refreshment and FMO (used in H.264) are beneficial for the channels with low variance of noise. In layered coding, the layers are not separately decodable and the performance depends on receiving the lower layers without error. Therefore, this method is also less advantageous for the error prone environment. For channel with large noise power, multiple description coding is more beneficial. MDC avoids packet failure happens in the network by creating multiple complimentary and separately-decodable descriptions.

Using MDC, a video stream is partitioned into several separately decodable descriptions and transmitted. In contrast to the error resiliency aspect of the MDC technique, each description needs to include header independently (in order to be separately decodable) and this reduces coding efficiency. MDC also affects performance of Differential Pulse Code Modulation (DPCM) in the hybrid video encoding algorithm used by all current video coding standards. DPCM in the hybrid video encoding algorithm tends to eliminate neighbor pixels' dependency, while MDC deteriorate correlation of neighbor pixels. For example, in a spatial MDC approach a frame is partitioned into distinct subimages so that two neighbor pixels are not assigned to a same subimage necessarily. Compared to its drawbacks, by the MDC technique if enough resource (such as bandwidth) is not available, a subset of all descriptions can be received and decoded; Also when an error occurs in a descriptions, it may be fixed using other error free descriptions. For error prone environment, these two advantages outweigh the disadvantages and make MDC a powerful strategy of video streaming for peer to peer communications, cooperative network, or heterogeneous network (Kazemi, 2012; Padmanabhan et al., 2003).

Different MDC approaches can be classified based on the type of data which is divided among descriptions. It can be temporal, spatial, frequency, or compressed type or it can be a hybrid approach. For example, by a temporal MDC approach with two descriptions, one description can include only odd frames while the other description includes just even frames. With a spatial MDC, each video frames is partitioned into several subimages with lower resolution, briefly. Frequency MDC tends to divide frequency components of video frames among descriptions.

However, multiple description coding approach for 2D videos has been investigated throughly, more investigation is required to apply MDC to 3D video specifically. For 3D video, since there is one more dimension (depth) we have more degree of freedom in order to partition video data. This paper aims to introduce a new 3D spatial MDC considering objects of the scene.

This paper is organized as follows: a background of spatial multiple description coding is provided in the Section 2 and challenge and opportunities will be discussed. Afterward, the proposed system model described in Section 3. Finally, some simulation result will be provided and argued in Section 4.

2 STATE OF THE ART

The first aspect of a single video description that can be considered for the purpose of MDC can be the spatial domain. This way, each video frame is partitioned by poly phase subsampling (PSS) into several images with a lower resolution called subimages (Shirani et al., 2001; Gallant et al., 2001; Kazemi, 2012). Each description is encoded separately and sent to the receiver. Two types of decoders, called the central decoder and the side decoder, are utilized by receiver to decode the received data stream. Based on availability of the descriptions in the receiver, the central or side decoder decodes the received video descriptions. If decoder receives all descriptions, the central decoder decodes the received data stream; otherwise, the side decoder will be used, however it may produce some distortion. The central decoder combines all descriptions and reconstructs original image but side decoder tries to interpolate image using available descriptions. Although, the central decoder provides a better quality of resolution, the side decoder can provide a better quality of experience if current rate is not enough to support all descriptions or transmission channel is so noisy that some descriptions may receive unsuccessfully. Such advantage is achieved at the expense of compression efficiency because pixel subsampling deteriorates pixel correlation and also each description must have its header data. This means some redundancy can be added without any advantage to improve side quality.

It is worth of mentioning that with a simple spatial MDC, there is no precise adjustment tools over redundancy to control side quality(Shirani et al., 2001; Gallant et al., 2001; Kazemi, 2012). This means that there is no way to increase redundancy specifically to improve resistivity against noise. For example, it is impossible to make three, six, or seven symmetric descriptions.

To improve MDC performance, Tillo and Olmo introduced a new MDC algorithm called ”least predictable vector directional multiple descriptions coding”(Tillo and Olmo, 2007). This approach basically copies the least predictable part of the frame in all descriptions. The simulation result shows that this method improves side quality compared to previous simple PSS approach although the new method provides more redundancy. They also argued that this algorithm is more complex as it needs to detect least predictable data.

In another work, Shirani presented a non-linear PSS approach and analysed its performance in case of missing one or more descriptions (Shirani, 2006). According to Shirani’s work, some pixels (called region of interest (ROI)) are sampled with greater rate than those are not important based on an exponential equation. On the other hand, descriptions include more information regarding the ROI and this enhances the side quality in the side decoder. Since the human visionary system is more sensitive to objects rather than pixels, this method can provide better performance in point of subjective assessment, significantly. Although, this method provides a better subjective evaluation, his paper hasnot discussed how to obtain the ROI. This problem is more sensible for applications involving with fast video content.

To extend the MDC algorithm for 3D video, it also needs to apply MDC approach to the depth map image. Clearly, the depth map image mainly contains depth information of scene objects. Because of the nature of real objects, depth information of 3D scenes rarely contain high frequency contents. Hence depth information can be compressed effectively and consequently bandwidth and disk space will be saved much more compared to dual camera capturing (Fehn, 2004; Hewage, 2014). In another research done by Karim et al. (Karim et al., 2008), a new MDC algorithm has been introduced for 3D video. They carried out their experiments using color plus depth map image representation. To save bandwidth, they showed that the down sampled version of depth map image is enough for an acceptable reconstruction in the decoder. To this end, they used a down sampled version of depth map image in their investigation on scalable video coding. They compared the quality of the reconstructed 3D videos using the original depth map image and the down sampled version of the depth map image and concluded that decimation of the depth map image does not cause a considerable degradation in the decoded 3D video. They also checked the result for a scalable multiple description coding approach and observe the same result. Therefore, down sampling of the depth map image does not affect the quality of reconstructed image and this is because, the depth map image rarely includes high frequency contents and also the depth values of adjacent pixels are usually similar. This fact that the depth values of pixels for an object are very closed to each other has been used in the research, done by Liu et al. (Liu et al., 2015), and the variance of the depth values are utilized to do ”texture block partitioning”.

This paper combines the facts used by Tillo and Olmo (Tillo and Olmo, 2007), Shirani (Shirani, 2006), Karim et al. (Karim et al., 2008), and Liu et al. (Liu et al., 2015) and introduce a new MDC method for 3D videos. More explanation about the proposed method will be provided in the next section.

3 PROPOSED METHOD

This section describes the proposed method for 3D video multiple description coding. Figure 1 shows an overview of this method. As can be seen in this figure, first 3D raw frames are split into 2D color frames and gray scale depth map frames. Then the region of interest is identified from the depth map image. The process of extracting the ROI will be described in the next part of this section. After detecting the important pixels (ROI), both color and depth frames are partitioned separately into four subimages using poly phase subsampling. Afterward, two opposite algorithms are used for color and depth map subimages. For the color image, the whole resolution of color pixel values for interesting components will be added to the stream in order to increase the resolution of the ROI; this helps to recover the ROI perfectly in cased of missing descriptions. For the depth map subimages, a contrary procedure is applied. As discussed earlier, depth values of pixels for an object are very similar and there is no need to stream depth information of all pixels of an object. Since the ROI identification algorithm detects objects as the ROI, a decimated version of the ROI in each depth map subimage is enough to recover an acceptable reconstruction of the original depth map image for the ROI. Instead, as the non-important parts of the subimages contain very different depth values, they need to be streamed in original resolution to have a proper reconstruction in case of packet loss. In other words, more bits are assigned to those blocks of depth map image that have diverse depth pixel values because pixels of these blocks cannot be estimated from adjacent pixels in the decoder.

As discussed in the previous section, Karim et al. showed that the depth map image can be de decimated without causing significant distortion in the decoder (Karim et al., 2008). It can be said that the proposed method uses this (concluded from Karim et al. work) and enhances its performance by changing the identical decimation to a non-identical decimation of the depth map image. Since the depth values of important pixels are very closed to each other compared to the remainder of the depth map image, the non-identical decimation algorithm would provide less error in the decoder having only one description (with an objective assessment like PSNR or SSIM as can be seen by simulation results). For the subjective assessment, the proposed method can provide much better performance since the human visionary system is more sensitive to objects rather than pixels and also the fact that the new method focus on objects.

One important issue in this process is its requirement to a low complexity operation to realize interesting objects. In other words, it is very crucial to find ROI in real time as some applications dealing with live video (such as video conferencing) and since complex operations cause delay, a complicated algorithm is impractical for the real time applications. Therefore, the need for a low complexity ROI identification algorithm is highlighted in the new method. To this end, we utilize the ratio of the standard deviation σ to the mean μ, also known as coefficient of variation (CoV), in a block wise manner:

$$c_v = \frac{\sigma}{\mu}, \qquad (1)$$

where c_v is CoV. σ and μ are the standard deviation and mean of a block in the depth map image. Indeed, CoV shows the normalized variation of pixel values. The CoV ratio is applied on depth map image in a block wise manner and it can be argued that if the CoV of a block is very big, the block is related to several objects since depth values of pixels for an object are usually similar. Also, if it is very small (close to zero) it is related to an absolutely straight (vertically flat) objects (we are not looking for); If the CoV ratio of a block is around 1, the block is probably related to an interesting object. It can be justified that in terms of the depth values the important objects usually contain low frequency contents, neither zero nor high frequency contents.

The algorithm identifying objects is as follows (see Figure 2):for the first step, the entire depth map frame is considered as one block. The CoV ratio of the block is calculated and checked to determine whether it is around 1 or not. If not, the block is partitioned into four equal-size blocks and then each block is consider as a new block. This hierarchical process continues until there are no more blocks with the CoV ratio greater than 1 (3 is used for the simulation). In this algorithm it is assumed that the minimum acceptable size of a block is 2×2 pixels. As an example, this process has been depicted in Figure 3. Numbers inside blocks in this figure represent typical CoV values. For this figure, it also has been assumed that the resolution of the depth map image is 16×16 pixels and the highlighted blocks include the important pixels that the proposed algorithm is looking for. As can be seen in 3, CoV values of highlighted blocks are around 1 while others are much larger than 1; but since the minimum acceptable size of a block is 2×2, the algorithm does not continues the hierarchical division algorithm for those blocks which have a large CoV values.

As an example, Figure 4 shows the performance of ROI identification algorithm for the videos ”Interview” and ”Orbi”. As can be seen in this image, interesting objects have been identified very well. Further performance evaluations are shown in the next section.

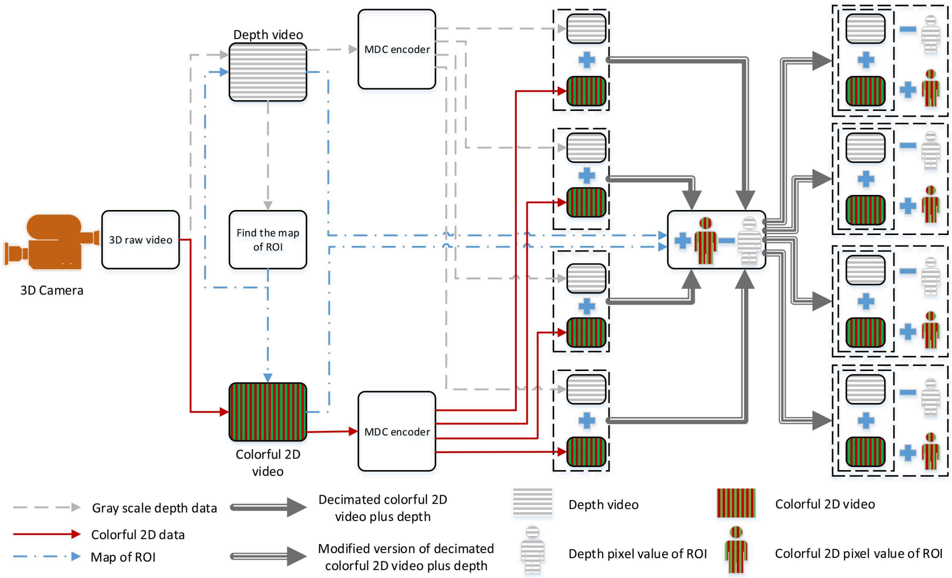


Figure 1: Block diagram of the proposed method.

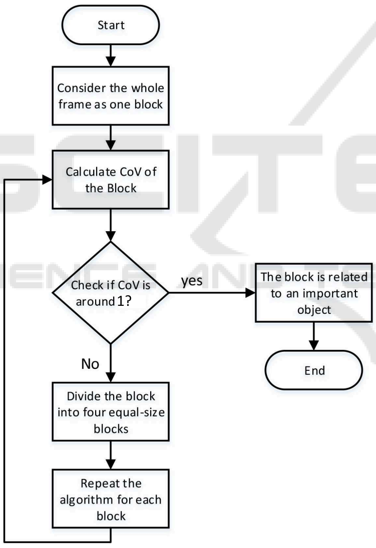


Figure 2: The algorithm of identifying important pixels.

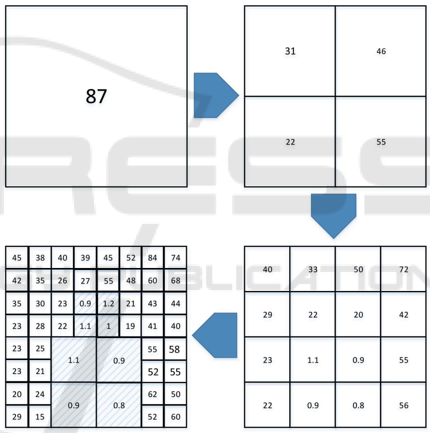
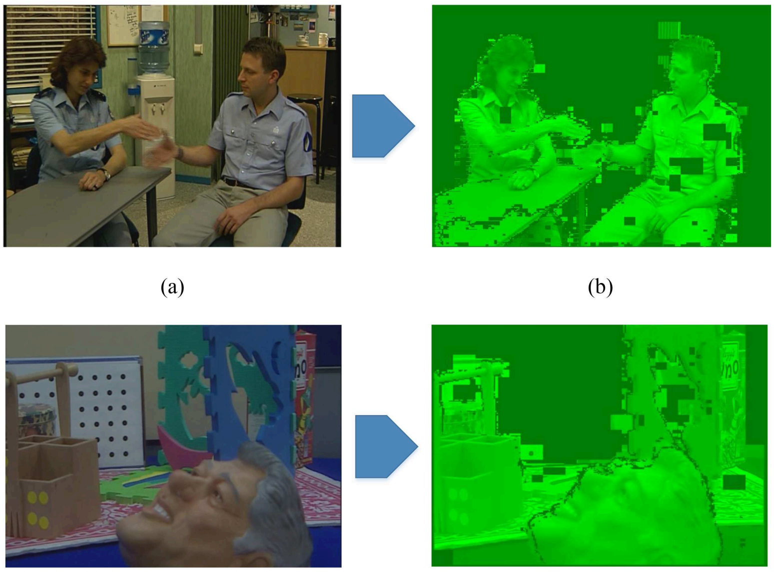


Figure 3: Hierarchical division to identify ROI.

4 SIMULATION RESULT AND DISCUSSION

For evaluation of the proposed method, this paper carried out tests using two stereoscopic test sequences with the format of DVD-Video PAL (720×576), called ”Interview” and ”Orbi” videos. The chroma and depth subsampling format is 4: 2: 2: 4 (the last 4 shows that the resolution of the depth map image is the same as the Y image or in other words the total frame resolution is 1440×576). Each video is 90 frames and the frame rate is 30 frames per second (fps). I frames are repeated every 16 frames and only P frames are used between I frames. The new method is implemented using H.264/AVC reference software, JM 19.0 (Institut, 2015). As described in the previous section, the hierarchical division ROI identification algorithm halves both width and height in each iteration to make smaller blocks. Therefore, as the width of the depth map frame ($720 = 2^4 \times 3^2 \times 5$) is not dividable after the 4th iteration, to have better resolution we assumed that the width of depth map frame for the ROI identification algorithm is 768 ($= 2^8 \times 3$) (we add zeros to the left side of the depth map image). With the same argument, the height of the depth map frame has been assumed to be 512 ($= 2^8 \times 2$). Therefore, the acceptable minimum size of block in the hierarchical division ROI identification algorithm



(a) (b) (c) (d)

Figure 4: A sample performance of ROI identification: (a) Original 2D video (video ”Interview”). (b) Detected ROI (video ”Interview”). (c) Original 2D video (video ”Orbi”). (b) Detected ROI (video ”Orbi”).

after eight iterations is 2×3 (this means that the minimum block size after 8th iteration in the hierarchical division algorithm is 6 pixels (see Table 1)).

As mentioned earlier, Figure 4 shows the original 2D video frames (the 87th frame of video ”Interview” and the 1st frame of video ”Orbi”) and their identified ROI. Detecting pixels related to the hands of objects in the video ”Interview” during handshaking (moving objects) can show acceptable performance of the algorithm for realizing ROI. This figure also shows a good performance for the second test sequence, i.e. video ”Orbi”. As can be seen objects in this video has been identified very well by the proposed algorithm.

Table 1 shows the number of blocks with different sizes after hierarchical division algorithm. As can be seen, there is one block with the size of 24576 ($= 128 \times 192$). This means that about 6% of entire depth map image is excluded of being partitioned more and stopped after the second iteration. Considering the second large block size in Table 1, i.e. 6144 ($= 64 \times 96$), it can be said that the hierarchical division process will be stopped for more than one third of the entire depth map image after the third iteration. This result can show that the algorithm does not have high load of calculation and it is not so complex.

Table 1: Number of block after hierarchical division algorithm.

Blocks' size	Number of Blocks	percent(%)
6	3188	4.86
24	1099	6.70
96	514	12.55
384	213	20.80
1536	49	19.14
6144	19	29.69
24576	1	6.25

Table 2 shows the number of blocks with different CoV values. As can be seen by this table, about 36% of the depth map image have CoV values less than 1. On the other hand, more than one third of the depth map image have very closed depth values. This can be the reason that decimation of the depth map image does not affect its quality when it is reconstructed in the decoder; as discussed earlier, Karim et al. showed by simulation results that the decimation of depth map image does not cause any considerable degradation in the decoder (Karim et al., 2008). Table 2 also shows that about 96% of the depth map image have the CoV values less than 3. The fact that about 96% of the depth map image have similar depth value and there is no need to be sent with the original resolution, can justify why the non-identical decimation is more advantageous than the identical decimation. This means that only about 4% of the depth map image need to be encoded with the original resolution and the remainder can be decimated to save bandwidth or storage.

In Figure 5 and 6, PSNR and SSIM measurements of the color 2D video using the proposed method and the basic Poly phase SubSampling MDC (PSS-MDC) are compared. For both graphs, it has been assumed that the decoder receives only one description among

Table 2: Number of block with different Cov values after hierarchical division algorithm.

		Blocks' size							Percent of blocks with specific CoV(%)
		6	24	96	384	1536	6144	24576	
CoV	≤ 1	799	471	228	90	24	17	1	60.64
	1 ∼ 3	833	627	286	123	25	2	0	36.98
	3 ∼ 10	602	1	0	0	0	0	0	0.92
	10 ∼ 20	276	0	0	0	0	0	0	0.42
	20 ∼ 03	320	0	0	0	0	0	0	0.49
	30 ∼ 40	171	0	0	0	0	0	0	0.26
	40 ∼ 50	13	0	0	0	0	0	0	0.02
	50 ∼ 60	2	0	0	0	0	0	0	0.003
	60 ∼ 70	16	0	0	0	0	0	0	0.02
	70 ∼ 80	46	0	0	0	0	0	0	0.07
	80 ∼ 100	0	0	0	0	0	0	0	0
	≥ 100	110	0	0	0	0	0	0	0.17
Percent of blocks with specific size(%)		4.86	6.70	12.55	20.80	19.14	29.69	6.25	

four descriptions. As can be seen by the first figure, more than 1 dB improvement for the video ”Interview” and more than 2 dB for the video ”Orbi” with a same rate can be achieved by the proposed MDC method. Regarding to SSIM, the proposed method provides about 0.3 improvement for the high rate streaming. It should be noted that since human visionary system is more sensitive to objects rather than pixels, subjective assessment can highlight more the improved performance of the proposed algorithm compared to the previous methods.

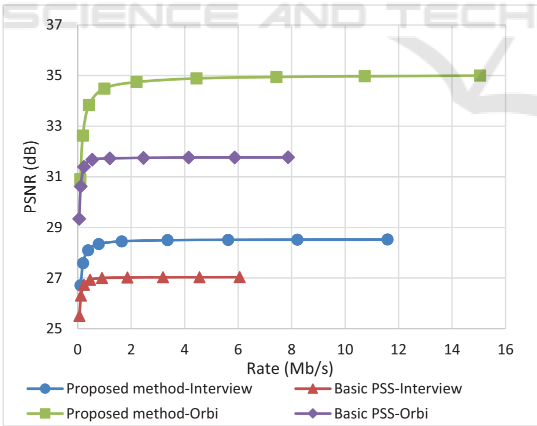


Figure 5: PSNR assessment of color image.

When it comes to the assessment of the proposed method for the depth map image, it shows the better performance more clearly. As can be seen in Figure 7 and 8, the improvement between the proposed method and the PSS subsampling is larger than the improvement achieved in the assessment of the color image. For the PSNR evaluation of the depth map image, the proposed method outperforms about 8 dB for the sequence ”interview” and about 9 dB for the sequence

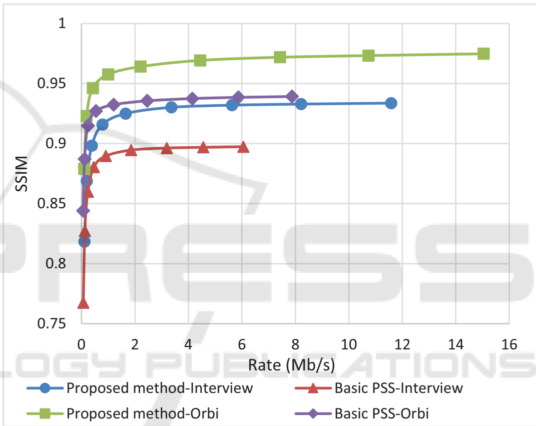


Figure 6: SSIM evaluation of color image.

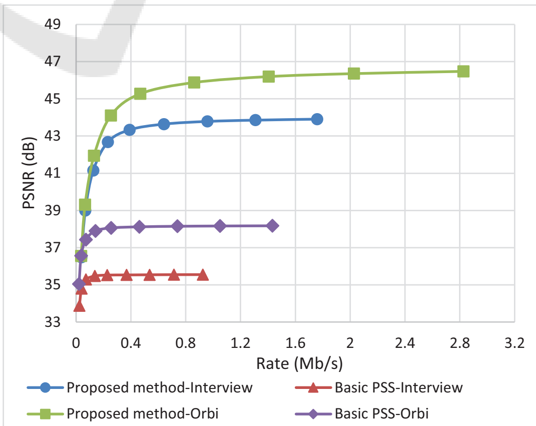


Figure 7: PSNR evaluation of depth map image.

”Orbi”. About the SSIM assessment, the proposed method outweighs about 0.01 compared to PSS sumsampling method.

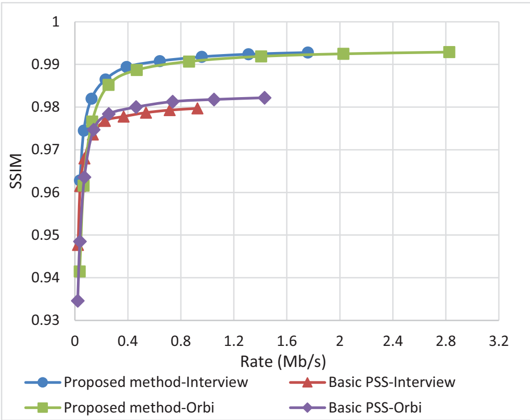


Figure 8: SSIM evaluation of depth map image.

5 CONCLUSION

Multimedia streaming suffers from packet failure in the network due to the packet loss, packet corruption, and large packet delay. An appropriate solution against packet failure in the error prone environment can be multiple description coding (MDC). With MDC, one video description is partitioned into several separately decodable descriptions. Missing a descriptions during transmission, decoder is capable to estimate the lost description from other error free description. To improve the basic spatial partitioning MDC algorithm and also be applicable for 3D videos, a non identical decimation algorithm for the stereoscopic videos has been provided in this paper. Such algorithm works based on objects existing in the scene and assigns more bandwidth to those region of interest. Since human eyes are more sensitive to objects rather than pixels, the proposed algorithm provides better performance than PSS MDC. In point of depth map image, the proposed algorithm enhance the current basic decimation to non identical decimation. As shown earlier, most parts of the depth map have similar depth value and therefore decimation in those part can save bandwidth or storage without degrading in quality considerably. Therefore, by the new algorithm,those parts of the depth map image that have large variances is encoded with the original resolution.

ACKNOWLEDGEMENT

The authors would like to acknowledge that this research was supported by NSERC Strategic Project Grant: ”Hi-Fit: High Fidelity Telepresence over Best-Effort Networks”.

REFERENCES

CBCnews (2015). Netflix, youtube video streaming dominate internet traffic in north america.

CTVnews (2016). Cineplex to open panoramic theatres in toronto, edmonton, vancouver.

Fehn, C. (2004). Depth-image-based rendering (dibr), compression and transmission for a new approach on 3d-tv. *SPIE: Stereoscopic Displays and Virtual Reality Systems*, 5291:93– 104.

Gallant, M., Shirani, S., and Kossentini, F. (2001). Standard-compliant multiple description video coding. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 1, pages 946–949 vol.1.

Hewage, C. (2014). *3D Video Processing and Transmission Fundamentals*. Chaminda Hewage and bookboon.com.

Institut, H.-H. (2015). H.264/avc reference software.

Karim, H., Hewage, C., Worrall, S., and Kondoz, A. (2008). Scalable multiple description video coding for stereoscopic 3d. *Consumer Electronics, IEEE Transactions on*, 54(2):745–752.

Kazemi, M. (2012). *Multiple description video coding based on base and enhancement layers of SVC and channel adaptive optimization*. PhD thesis, Sharif University of Technology, Tehran, Iran.

Liu, Z., Cheung, G., Chakareski, J., and Ji, Y. (2015). Multiple description coding and recovery of free viewpoint video for wireless multi-path streaming. *IEEE Journal of Selected Topics in Signal Processing*, 9(1):151–164.

Meesters, L., IJsselsteijn, W., and Seuntiens, P. (2004). A survey of perceptual evaluations and requirements of three-dimensional tv. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(3):381–391.

Padmanabhan, V., Wang, H., and Chou, P. (2003). Resilient peer-to-peer streaming. In *Network Protocols, 2003. Proceedings. 11th IEEE International Conference on*, pages 16–27.

Shirani, S. (2006). Content-based multiple description image coding. *Multimedia, IEEE Transactions on*, 8(2):411–419.

Shirani, S., Gallant, M., and Kossentini, F. (2001). Multiple description image coding using pre- and post-processing. In *Information Technology: Coding and Computing, 2001. Proceedings. International Conference on*, pages 35–39.

Smolic, A. and Kimata, H. (2003a). Report on status of 3dav exploration. Technical Report N5558, ISO/IEC JTC1/SC29/WG11, Thailand.

Smolic, A. and Kimata, H. (2003b). Report on status of 3dav exploration. Technical Report W5877, ISO/IEC JTC1/SC29/WG11, Norway.

Tillo, T. and Olmo, G. (2007). Data-dependent pre- and postprocessing multiple description coding of images. *Image Processing, IEEE Transactions on*, 16(5):1269–1280.