ERROR CONCEALMENT FOR STEREOSCOPIC VIEW PLUS DEPTH BASED ON K NEAREST NEIGHBOR METHOD

[1]WAHIBA ISMAIEL, [2]NIEMA OSMAN, [3]IS-HAKA MKWAWA

[1,2,3]Sudan University of Science & Technology (SUST), SUST, University of Plymouth, Plymouth, UK
E-mail: Wahiba2008123@yahoo.com, niema.osman@gmail.com, Is-Haka.Mkwawa@Plymouth.ac.uk

Abstract— This paper proposes a novel hybrid concealment scheme to recover packet loss on a slice during stereoscopic color plus depth or view plus depth transmission over error prone channel. The proposed scheme aims at minimizing quality degradation on both color and depth maps sequences. The scheme uses two resilience tools in video coding H264/AVC, which are slice groups also known as Flexible Macro-block Ordering (FMO) and Redundant slices (RS). FMO adds flexibility to control errors that are spread spatially. This paper also exploits the K nearest neighbors (KNN) regression method to conceal missing packets. Initial results have shown that the proposed scheme significantly minimizes quality degradation for 3D video transmission over error prone channels.

Index Terms— View plus depth, concealment scheme, KNN, slice groups (FMO), redundant slices (RS) , H264/AVC, subjective quality.

I. INTRODUCTION

Stereoscopic view plus depth is popular because of some features such as its accuracy, easy data acquisition and it requires less memory and bandwidth. 3D video transmitting suffers from different aspects such as channel condition and network parameters (delay, jitter and packet loss). Packet loss has an impact on the overall quality. To solve this problem, concealment and resilience methods could be used. This paper proposes a solution at both encoder and decoder sides which use FMO to reduce error propagation on the frame and redundant slices (RS) to recover error when the whole slice is missing. These H264/AVC techniques are used for both color and depth maps sequences. The proposed hybrid concealment scheme investigates slice loss by using proposed k Nearest Neighbors method (KNN) based on spatial concealment using neighboring MBs. A number of previous studies in this area used methods such as Depth-Assisted in [1] and a hybrid error concealment for intra frame in [2] to conceal the loss. Presented methods for slice loss for intra frame 3D video which is called depth-assisted error concealment method of 2D+depth video; such a method is not suitable for color sequences. In addition, [1] uses motion information from depth directly without considering the similarities in the motion information, while [2] used the motion correlation of view and depth sequences The Depth-Offset Error Concealment is extract elect (Motion Vector)MV from depth sequence by repeated the first frame in depth one time to obtain depth sequence one frame offset from the view sequence, this solution is suffer from the 1- frame delay of the depth sequence and extra bits for the added frame also it uses additional MV candidates and it is not suitable with color frame lost also if P-frame or B-frame in depth when are lost too ,in addition to this algorithm require to obtain with subjective quality. Another intra frame concealment mode is considered by [3]. However, it is not better with heterogeneous regions and color sequences. Other papers depend on decision making on a pixel value such as in [4], however, it is not efficient in the low motion sequence color sequences. Data partition for resilience method in [5], does not take color into account. The paper is organized as follows: Section 1 provides an introduction. Section 2 discusses the tools and K Nearest Neighbors Regression. Section 3 proposes the scheme while Section 4 presents the experimental setup, results and discussion. Section 5 concludes the paper and suggests future work. .

II. H264/AVC RESILIENCE TOOLS AND K NEAREST NEIGHBORS REGRESSION METHOD

This section presents the resilience techniques which are used in this research; the KNN method was also used to extract the spatial information of nearest neighbors' macroblock, this information is used to predict the packet:

A. H264/AVC Resilience Tools

H264/AVC video coding in [6] is widely used in multimedia communication for coding and transmitting video over IP network. It is initiated by Video Coding Experts Group (VCEG) in early 1998 approved by the collaboration between the ISO/IEC Moving Picture Experts Group (MPEG) and ITU-T recommended as H264 VCEG. It is also referred to International Standard 14496-10 or MPEG-4-part 10 Advance Video Coding (AVC) of ISO/IEC. H264/AVC supports video quality adaptation on the robustness of video quality such as Flexible Macro-block Ordering (FMO).

FMO is one of the resilience techniques allowed within the base line and extended profile. Slice Groups (SGs) which contain more than one slice are used in this research. SGs are the result of splitting the picture into several regions or part called slices and each slice includes a sequence of the macroblock. The Macroblock Allocation maps (MBAmap) is the data structure used in FMO to allow each macroblock mapped to the specific slice by specific address for each macroblock. FMO allows macroblock to be grouped and set in any direction order which stops error propagation by limiting the error inside slice due to burst packet losses to the area of the frame. FMO has seven patterns to form macroblock in the slice as shown in the Figure 1. Type 0 denotes that the Interleaved consists of row of macroblock and each belongs to different slice Intra leaved slice group.
Type1 shows that the scattered or dispersed slice group of similar checker. Type 2 illustrates foreground and background, this includes top left and bottom right rectangle named as Regions of Interest (ROI), this mode is used in this research for mapping macroblock to slice groups. Type 3 to type 5 (box- out, raster scan and wipe) are dynamic types that allow the SGs increase and shrink over the different picture in acyclic way while Type 6 is the most flexible compared to other patterns mentioned earlier.

Redundant Slices (RS) is the second resilience tool which is used in this research for recovering the loss of the whole slice. H264/AVC is enabling the encoder to send redundant of the primary slice to be replaced with the primary slice at the decoder side [7 -8].

B. K Nearest Neighbors (KNN) Regression Method

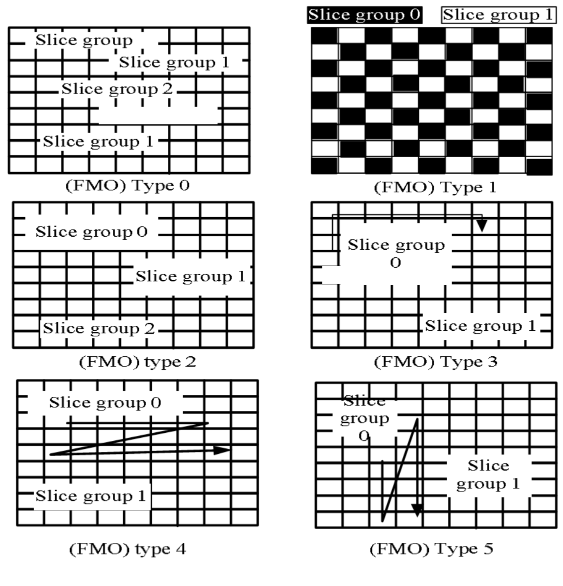


Figure 1. Shows FMO patterns types

According to nearest neighbors' techniques, we use KNN to select nearest neighbors for each macroblock in SGs It is used in several applications such as data mining, statistical pattern recognition, image processing. KNN stores all available cases and then classify new cases depending on a similarity measure of local neighbors to prediction the loss. To measure distance, it uses different functions such as Euclidean distance, Manhattan distance and Minkowski distance. The prediction can be measured by different ways such as Simple Interpolation, Averaging, Local Linear Regression and Locally Weighted Regression [10].

III. THE PROPOSED SCHEME

The proposed scheme contains two techniques based on H264/AVC codec standard. It focuses on spatial prediction algorithm using nearest neighboring MBs. Figure 2 presents the proposed scheme, the process of the proposal conducted at the encoder side by exploit the information of nearest neighbors of MB according to pattern type 2 of FMO and first calculate the distance between target MB and neighbors using KNN method and find distance using Euclidean distance function in equation 1 then select three neighbors for each MBs and calculate luma and Chroma sample pixel values from 16x16 macroblock then encode this information and transmitted so for prediction value of loss macroblock at the decode side we calculate values according to k prediction equation 2.
Also (RS) was sent too at the decoder side before decode RS will be used to recover the loss of the whole slice while spatial information was used to conceal the loss of MBs according to the proposed concealment algorithm at the decoder side by calculating the average of spatial information to predict the pixel value .
Euclidean distance is given by the following equation,

$$\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{1}$$

Where, n denotes the number of coordinates, i denotes each coordinate. X and y indicates width and height, respectively.
The k predictor formula is presented as,

$$y^{\wedge} = \frac{1}{k}(y_{nn1} + y_{nn2} + \cdots y_{nnk})$$
(2)

Where, $y^{\wedge}$ denote to prediction value y_{nn1} and y_{nn2} denotes the nearest neighbor while y_{nnk} is the last nearest neighbor according to specific k, k is specific number of the nearest neighbor.

IV. EXPERIMENTAL RESULT AND DISCUSSION

A. Materials

The experiment was carried out using three different sequences which are Interview in figure.2 , Orbi in figure.3 and Cg sequences. Each sequence has resolution of 720x576 and frame rate of 25 fps. The video duration is 10 second for the Interview while Orbi and Cg sequences are of 5 seconds long. The Interview sequence has 250 frames while Orbi and Cg sequences have 150 frames, Cg sequence was used for a training session while Interview and Orbi sequences for testing session. Each color and depth map sequences are encoded separately using H264/AVC with three different packet loss rate (PLR) 5%, 10%, 20%. The quantization parameters (QP) are 28 and 40 for both color and depth sequences. The Group of Picture (GoP) structure followed the format of IPPPIPPPIPPP and the sequences format is YUV4:2:0, the number of slice group for each sequence is 3 slices and the pattern used for each macroblock allocated slice is type2 and macroblock has the size of 16x16.

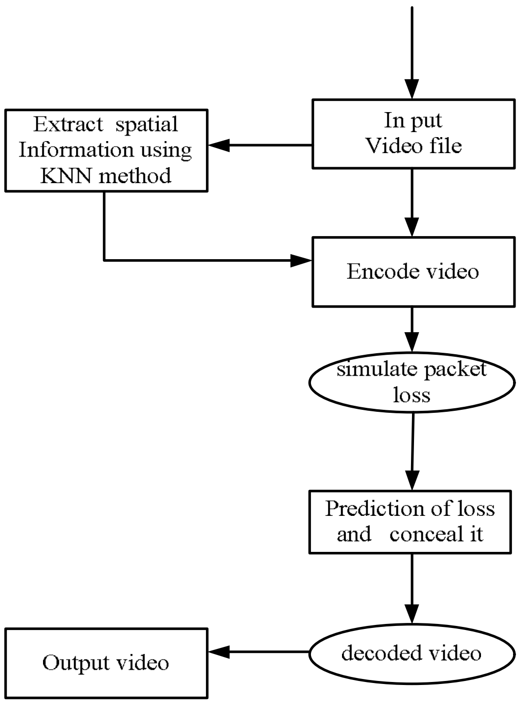
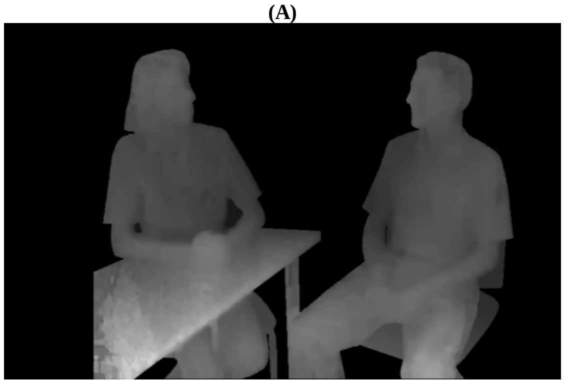


Figure 2. Presents proposed scheme

B. Experimental Setup

We conducted two human observer tests, first for JM frame copy and second for the proposed scheme. Experimental setup was according to ITU-T P.913 [11] and ITU-R- BT.1438 [12]. Two sessions were done in each experiment, first for training and the second for testing methods. After each training session assessors had 10 minutes break before the next session. The experiment was using screen resolution of 1920 x540 HDTV and 32 inches for displaying video sequences. The viewing distance was 1.75m with the picture height of 6.3H. 41 assessors participated in all experiments 26 males and 15 females (4 expert and 37 non expert of 3D video), all have good visual vision and their age was between 20 to 55 years, the average age was 36 year.

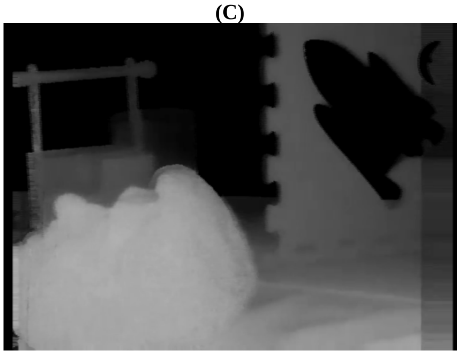
(A)



(B)

Figure 3. Interview sequences color (A) and Depth (B)

(C)



(D)

Figure 4.shown Orbi sequence color(C) and depth(D)

C. Subject methodology and results

Assessors were asked to assess color and depth maps in the impaired video sequences by using a questionnaire. Subjective evaluation methodology was using ACR according to ITU-R BT-500 [13]. The ACR has a score ranging from 1 to 5 denoted as bad to excellent. The Mean Opinion Score (MOS) result are presented in Figure 4-7 for JM frame copy and Figure 8-11 for the proposed scheme. The results for 28 quantization parameters 28 give higher MOS score than QP of 40 as expected for all methods, however, the proposed scheme achieves better quality than the JM frame copy.

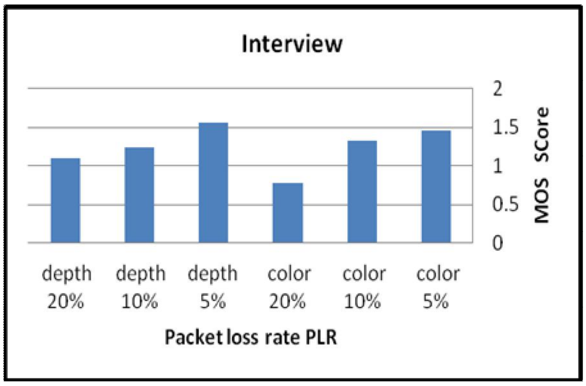


Figure 4. MOS Result of JM frame copy with QP 28

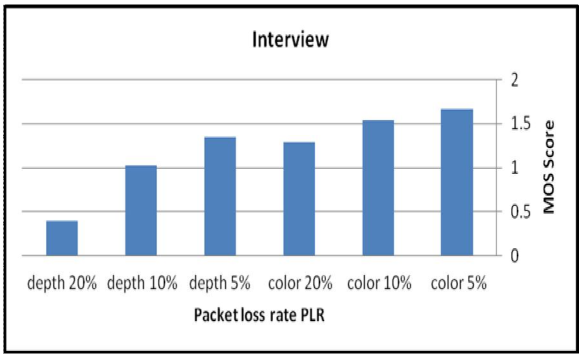


Figure 5. MOS result of JM frame copy with QP 40

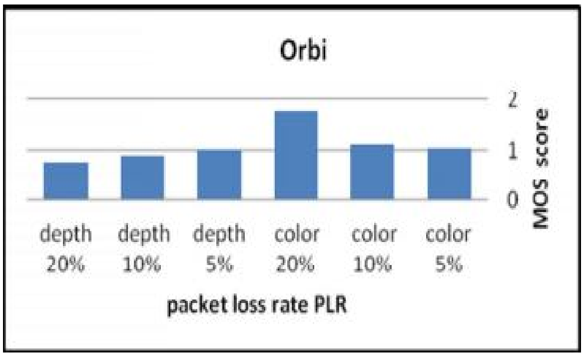


Figure 6. MOS result of JM Frame copy with QP28

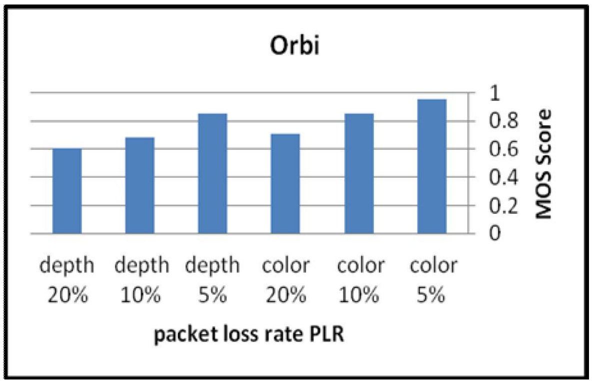


Figure 7. MOS result of JM Frame copy with QP40

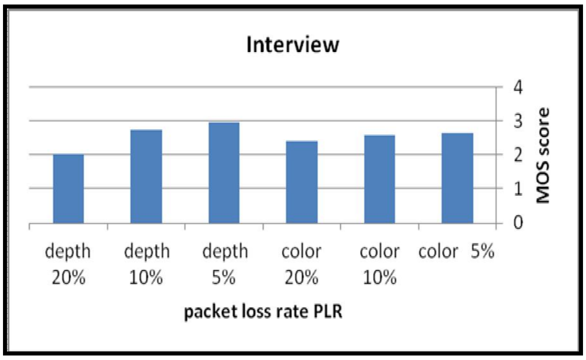


Figure 8. MOS Result of proposal scheme with QP28

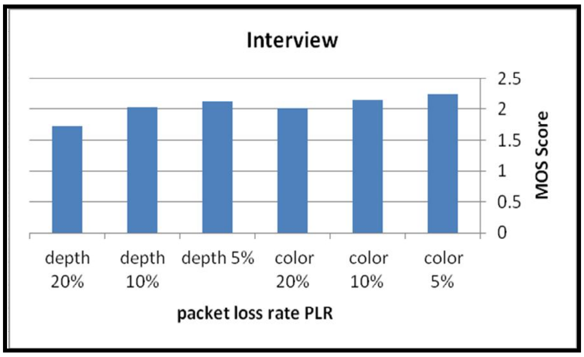


Figure 9. MOS Result of proposal scheme with QP 40

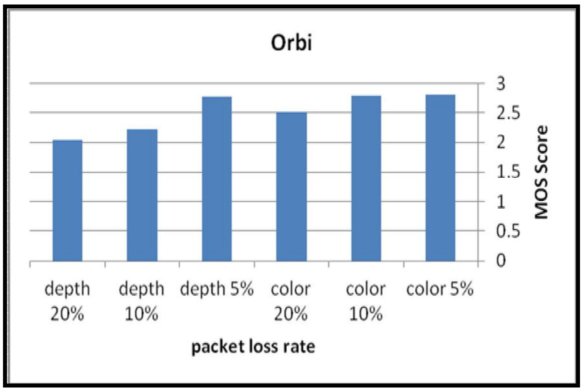


Figure10 MOS Result of proposal scheme with QP 28

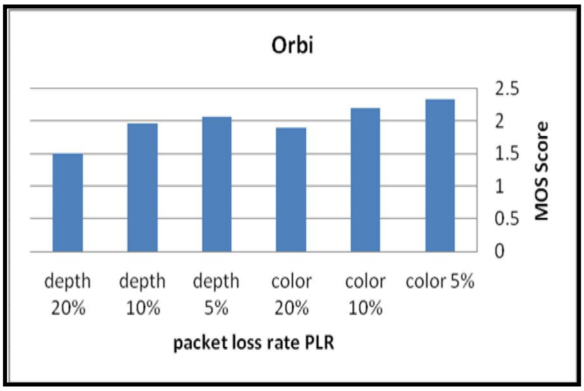


Figure 11 MOS Result of proposal scheme with QP 40

D. Experiment results and discussion

We conducted an experiment using H264/AVC reference software (JM19.0) adapted on baseline profile level 3. Table 1, 2, 3 and 4 show the results of proposed method and algorithm compared to two previous methods. The first method is the proposed concealment by H264/AVC, first we conduct an experiment by using frame copy and the second method is Depth Assisted [1] by using one slice and recovers only depth sequence. We present Average SNRY dB results of different methods.

CONCLUSION

This paper aims to develop an error concealment scheme to recover packet loss in the stereoscopic 3D video in order to support quality of experience. The proposed a novel scheme used spatial information which is generated by KNN method in order to predict the loss of macroblock based on concealment algorithm. The resulting outcome of the proposed scheme under quantization of 28 and 40 were higher in MOS and SNRY values for color and depth sequences than with frame copy and Depth assisted methods. Future work will look at proposing the objective model based on the subjective results.

Table 1:SNRY DB result of interview sequence with 28 QP

METHOD	INTERVIEW COLOR 28QP			INTERVIEW DEPTH28QP		
	5%	10%	20%	5%	10%	20%
COPY FRAME	32.41	32.39	32.41	33.38	33.81	33.80
PROPESD METHOD	39.55	39.12	38.70	38.99	38.89	38.02

Table 2:SNRY DB result of interview sequence with 40 QP

METHOD	INTERVIEW COLOR40QP			INTERVIEW DEPTH40QP		
	5%	10%	20%	5%	10%	20%
COPY FRAME	27.31	27.30	27.29	34.18	34.16	34.10
PROPESD METHOD	35.37	35.19	35.04	35.98	35.76	35.14

Table 3:SNRY DB result of Orbi sequence with 28QP

METHOD	ORBI COLOR 28QP			ORBI DEPTH 28QP		
	5%	10%	20%	5%	10%	20%
COPY FRAME	29.57	29.56	29.53	33.18	33.14	33.04
Depth ssisted 2014				38.82	38.63	37.48
PROPESD METHOD	39.38	39.24	38.67	39.31	39.11	38.87

Table 4:SNRY DB result of Orbi sequence with 40QP

METHOD	ORBI COLOR 40QP			ORBI DEPTH 40QP		
	5%	10%	20%	5%		20%
COPY FRAME	27.96	27.28	27.26	33.77	33.72	33.73
Depth Assisted 2014				32.69	32.18	31.3
PROPESD METHOD	35.36	35.18	34.79	35.04	34.58	34.13

REFERENCES

[1] Meng Yang, Xuguang Lan, Nanning Zheng, Cosman.P, Depth-Assisted Temporal Error Concealment for Intra Frame Slices in 3-D Video, IEEE Trans. on Broadcasting , vol. 60, PP. 385- 393,2014.

[2] Yuhong Yang, Meng Yang, Pamela Cosman, “a hybrid error concealment for intra frame in stereoscopic video”, Signal Processing Conference (EUSIPCO), Proceedings of the 20th European 2012,PP.1104- 1114.

[3] Xue Zhang, Yao Zhao, Chunyu Lin, Huihui Bai,Chao Yao, and Anhong Wang ,”Warping-Driven Mode Selection for Depth Error Concealment”,in IEEE Signal and Information Processing (GlobalSIP) on 3-5.Dec. 2014,PP. 302–306.

[4] M. Ranjbari , A. Sali , H. A. Karim ,F. Hashim ,” Depth Error Concealment Based on Decision Making” , IEEE on Signal and Image Processing Applications (ICSIPA),on 8–10.Oct. 2013,PP. 193–196

[5] H. Abdul Karim, Nor Azhar Mohd Arif , A. Sali 3, S. Worrall, A. H. Sadka,“ An Error Resilience Method for Depth in Stereoscopic 3D Video”, Signal and Image Processing Applications (ICSIPA),IEEE International Conference on ,2009, PP.140-144.

[6] IOS/IEC 14496-10:2003, Coding of audiovisual Object –part 10:Advanced Video Coding ," 2003 also ITU-T Recommendation H.264 "Advance Video Coding for generic audiovisual services".

[7] P. Lambert, W. D. Neve, Y. Dhondt, and R.V. de Walle,“Flexible Macroblock Ordering in H.264/AVC,”Journal of Visual Communication and Image Representation, Vol. 17, pp.358-375, April 2006.

[8] Y. Dhondt, P. Lambert, R. Van der Walle, “A Flexible Macroblock scheme for Unequal Error Protection,” IEEE Interconference on Image Processing(ICIP), October 2006.

[9] Friedman (1994): Flexible metric nearest neighbor classification; Technical Report 113, Stanford University, Statistics Department

[10] l. c. zhao,"exponential bounds of mean error for the nearest neighbor estimates of regression functions" journal of multivariate analysis vol.21, pp. 168-178 , 1987.

[11] Recommendation ITU-T P.913 (01/2014) ,“ Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment”, 2014.

[12] Recommendation REC. ITU-R BT.1438, “subjective assessment of stereoscopic television pictures”, 2000.

[13] Recommendation ITU-T P.910, “Subjective video quality assessment methods for multimedia applications”, 2008.

★★★