

A New Error Concealment algorithm for H.264 Video Transmission

*Pei-Jun Lee, †Homer H. Chen, and ‡Liang-Gee Chen

*Department of Electronic Engineering, St. John's & St. Mary's Institute of Technology, Taiwan, R.O.C.

†Graduate Institute of Communication Engineering, Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.

‡Graduate Institute of Electronics Engineering, Department of Electrical Engineering, National Taiwan University, Taiwan, R.O.C.

ABSTRACT

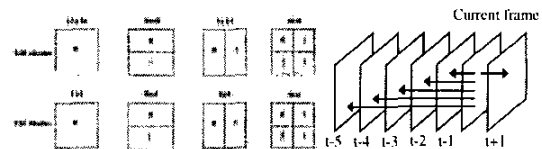
In this paper, a new error concealment algorithm for the new coding standard H.264 is presented. The algorithm consists of a block size determination step to determine the size type of the lost block and a motion vector recovery step to find the lost motion vector from multiple reference frames. The main features of this algorithm are as follows. In the block size determination step, we propose a criterion to determine the size type of the lost block from the current frame. In the motion vector recovery step, the optimal motion vector for the lost block chosen from multiple previous reference frames with the minimum value of the side match distortion. The proposed algorithm not only can determine the most correct mode for the lost block, but also can save much more computation time for motion vector recovery. Experimental results show that the proposed algorithm achieves 0.4~7 dB improvement than conventional VM method does.

1. INTRODUCTION

Joint Video Team (JVT) gathered experts from ISO/IEC MPEG-4 Advanced Video Coding (AVC) and ITU-T H.264 to develop the latest standard [1] in which some new coding schemes are adopted. One of the major differences between H.264 and the previous coding standards is that the motion estimation scheme is changed. In H.264, the motion estimation with variable block sizes and multiple reference frames greatly reduce prediction errors. Fig. 1 shows the motion estimation prediction with multiple reference frames in which the block sizes have the seven different block division modes (16*16, 16*8, 8*16, 8*8, 8*4, 4*8, and 4*4).

The H.264 standard aims at better compression efficiency, more flexible network adaptation, and enhanced error robustness. In order to have a realistic view of H.264 operation in error-prone environment, basic error resilience tools have to be implemented in the JM reference software [2]. The structure of compressed

bitstream is encoded by H.264. Since the macroblock header of the bitstream is not protected enough, the error in bitstream will affect all slice data. Therefore, the error concealment algorithm is necessary on video transmission. H.264 adopts new motion estimation scheme in which most traditional motion vector recovery algorithms designed for 16*16 macroblocks are not suitable for it. Therefore Wang *et al.* [3] have presented the weighted pixel value average to conceal the damaged block in Intra picture and a modified boundary matching algorithm by neighborhoods motion vector to conceal 16*16 macroblock motion vector in Inter pictures, respectively. Their algorithm has been implemented in the test model of the draft ITU-T video coding standard H.26L.



(a) The different block division (b) Multiple reference frame modes in H.264

Fig. 1 H.264 motion estimation with multiple reference frame

H.264 standard provides the variable block-size and multiple reference frames motion estimation/compensation to achieve improved compression performance. However, such features and functionality also entail additional complexity in encoding and decoding. In order to avoid this disadvantage, [4] analyzes the available information after intra prediction and motion estimation from previous one frame to determine whether it is necessary to search more frames.

Based on the analysis result of [4], this paper proposes an algorithm to determine the type of damaged block sizes and the optimal prediction reference frame such that the error concealment is more effective to conceal the lost motion vectors.

The paper is organized as follows: Section 2 presents the proposed algorithm, in which it includes the selection of the size type for the lost block and the expression of error concealment with multiple reference frames. Extensive experimental results are shown in Section 3. Finally a conclusion is drawn in the final section.

2. VARIABLE BLOCK SIZE ERROR CONCEALMENT WITH MULTIPLE-HYPOTHESIS

Based on the analysis of [4], in order to reduce the computational complexity, several characteristics for motion estimation scheme of H.264 coding are summarized. In this paper, five of those characteristics of [4]

- (a) 76.82% of MBs need only previous one reference frame to get the optimal mode;
- (b) if a boundary crosses through an MB, the MB is split into smaller blocks;
- (c) 80% optimal motion vectors are determined by nearest reference frame;
- (d) if 16*16 mode is selected, the optimal reference frame tend to be unchanged;
- (e) if smaller blocks are selected, searching more frames tends to be helpful;

are adopted to determine the type of damaged block sizes and the optimal reference frame for motion vector recovery.

The characteristics (a) and (b) are used to determine the block mode of damaged block. The characteristics (c), (d) and (e) are used to select reference frame to improve error concealment performance.

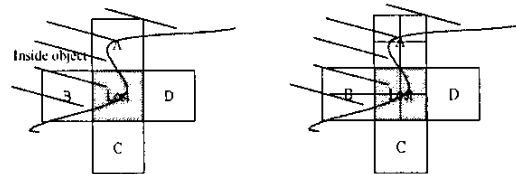
2.1 Size type selection of the lost block

This section discusses the selection of the lost block type. In this paper, four size types of the block, such as 16*16, 16*8, 8*16 and 8*8, are considered. The selection depends on the reliability degree of the neighborhood motion vector. It is defined as

$$RD = \sum_{i=A}^D \left| \frac{1}{4} * (\sum_{i=A}^D MV_i) - MV_i \right| \tag{1}$$

where MV_i is the neighborhood block motion vector of the lost block and i denotes the neighborhood block of the lost block shown in Fig. 2. A threshold (T) for RD is pre-assigned to determine the lost block type from either current or previous frame. If $RD < T$, the motion vector of the lost block is similar to the neighborhood block's motion vector. The size type of the lost block is the same as that in the previous one frame. It coincides with the characteristic (a) for H.264 coding. If $RD \geq T$, it means

that the motion vector between the lost block and the neighborhood blocks is discontinuous. In other words, it may have a boundary crossing through the lost block. The previous one frame is not reliable for us to estimate the lost block type. We select the block size type from the information of the current frame.



(a) The neighborhood block (b) An example for mode determination

Fig.2 Illustration of the mode determination for the lost block

If there is a boundary crossing through the upper or the bottom block (A or C in Fig. 2) of the lost block and where A or C is of the mode 8*8 or 8*16, there may be vertical direction boundary in the lost block. Then according to the characteristic (b), the lost block size is suggested to split into smaller blocks. Similarly, if there is a boundary crossing through the block B or D in Fig. 2, and where B or D is of the mode 8*8 or 16*8, there may be horizontal direction boundary in the lost block. Then, the lost block size is also suggested to split into smaller blocks.

Let us explain it in more detail. In the following description, we use "flag" to present the mode of the neighborhood blocks. If the neighborhood block A (or C) is of the mode 8*8 or 8*16, then "the flag of A" is set to 1 (or "the flag of C" is set to 1). Otherwise, they are set to 0. If the neighborhood block B (or D) is of the mode 8*8 or 16*8, we set "the flag of B" to 1 (or set "the flag of D" to 1). Otherwise, they are set to 0. Therefore, there are 16 cases being needed to determine the mode of the lost block which is list in Table 1. The bottom row of Table 1 shows the mode of the lost block determined by the flags of the neighborhood four MB. The symbols of the mode in the bottom row are defined as α, β, γ and λ , which denote the modes 16*16, 8*8, 8*16 and 16*8, respectively.

Table 1 Flag status for the mode determination of the lost block

MB	flag			
A	0	0	0	0
B	0	0	0	1
C	0	0	1	0
D	0	1	0	1
mode	α	β	γ	λ

mode decision α :16*16 β :16*8 γ :8*16 λ :8*8

For instance, in Fig. 2(b), we have the flags of A and B are 1, and the flag of C and D are 0, then the mode of the lost block is set of the mode λ , i.e., 8*8. It means, the boundary may cross through the lost block with the direction as in Fig. 2(b). Further, in the 11th column, the flag of A and C are 1, and the flag of B and D are 0, the lost block is set of the mode 8*16.

2.2 Error concealment with multiple reference frames

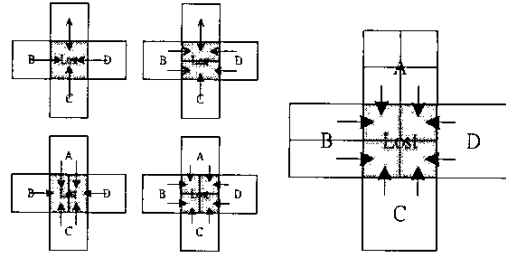
After the lost block mode is determined, some error concealment algorithms should be used to conceal the lost block with different block types. The adopted error concealment algorithm is described in this session.

In general, the motions between spatially neighboring areas are highly correlated, the motion of the lost block is predicted from a spatial neighborhood MB's motion vector. Fig. 3 shows the candidate selection of the predicted motion vector for the lost block. Fig. 3(b) is an illustration for predicting the motion vector to Fig. 2(b), in the neighborhood of the lost block, the motion vectors of the left 16*8 block and the lower 16*8 block are the candidates of the left bottom 8*8 block in the lost block. The selection of the optimal candidate is based on the smoothness of the concealed image in a search region. In other words, the selected motion vector is the one that minimizes the side match distortion d_{sm} as follows.

$$d_{sm} = \sum_{j=1}^N \left| Y_j(MB_{in}^{MV^{pred}+SR}) - Y_j(MB_{out}) \right| \quad (2)$$

$MB_{in}^{MV^{pred}+SR}$ is the locations of the inside boundary pixel of the candidate block which is adopted from the predicted candidate MV (MV^{pred}) in the searching region(SR). MB_{out} is the location of the outside boundary pixels of the candidate block. $Y_j(MB_{out})$ is the pixel value of j-th location of MB_{in} or MB_{out} . N is the number of total pixels of MB_{in} or MB_{out} .

It should be noted that, in some cases such as larger motion occurring etc., the motion vector selected from the previous one frame may be not the optimal one. Error concealment with multiple reference frames may be able to improve the performance of image reconstruction [5]. However, the disadvantage of that is the computational complexity being increased. In order to avoid the computational complexity and keep good image reconstruction performance, the characteristics (c), (d) and (e) for H.264 video coding are utilized to decide the optimal reference frame.



(a) Motion vector candidate (b) Motion vector prediction for Fig.2 (b)

Fig.3 Motion vector candidate selection

Based on the block type decision in Section 2.1, we can determine the mode of the lost block. If the determined mode is 16*16, we only use the previous frame to conceal the lost block. Otherwise, searching more frames to find the optimal frame is helpful for estimating the optimal motion vectors. Because of the characteristic (c), 80% optimal motion vectors are determined by the nearest reference frame. Let a threshold (T_d) be set for d_{sm} . If $d_{sm} > T_d$, the side matching candidate block is selected from the previous five frames in error concealment algorithm; otherwise, it is selected from the previous one frame.

3. SIMULATION RESULT

The simulation video sequences are encoded and decoded by JM61d program that is a standard codec VM program of H.264. The slice interleaving packetization mechanism was used. By changing the quantization factor, the videos are encoded at different bitrates. To improve the performance of the proposed algorithm, the VM program adopted method is selected for comparison. The boundary matching algorithm used in the simulation was presented by Wang *et al.* [3], which is adopted to H.264. In their algorithms, the motion vectors of 8*8 blocks, which are adjacent to the lost macroblock, are calculated as the average of the motion vectors of the spatially corresponding 4*4 or other shaped (e.g. 4*8) blocks. And each of the candidate motion vectors, that cover neighboring 8*8 blocks, will be used to reconstruct the lost macroblock. The motion vector, that can minimize the sum of absolute boundary differences between reconstructed macroblock and its neighboring macroblocks, will be used to predict the motion vector of lost macroblock.

Tables 2 show the performance comparison from the objective view for variable bitrate test sequences with 3%, 5%, 10%, 20% error rate. It is seen that proposed algorithm achieves PSNR 0.3 ~7dB improvement than

the VM methods. Fig. 5 shows the PSNR comparison with 10% lost macroblock and QP is 20. The proposed algorithm has an explicit improvement in larger motion. Fig. 6 shows the 76th frame of the "foreman" with a 10% error rate and QP equals to 24 where the significant errors occur in its wall edge. The simulation results show that the proposed algorithm can efficiently improve the quality of corrupted video.

Table 2. The simulation results of proposed algorithm

Video Sequence	QP	Bitrate (kbits/s)	Original PSNR	Method	PSNR of different Macroblock lost rate			
					3%	5%	10%	20%
Foreman	24	32.72	38.46	NE	29.71	27.28	24.08	20.99
				VM	42.43	39.36	35.25	30.70
				Proposed	43.96	41.30	37.40	34.50
	20	59.50	41.48	NE	29.68	23.29	24.07	20.98
				VM	42.13	39.21	35.18	30.63
				Proposed	43.72	41.02	37.34	33.26
	16	106.96	44.76	NE	29.68	27.26	24.06	20.97
				VM	41.99	39.16	35.05	30.57
				Proposed	43.75	41.04	42.03	33.34
Sefan	24	89.37	37.48	NE	31.31	29.01	25.97	22.91
				VM	35.09	31.96	28.53	25.04
				Proposed	36.49	33.47	30.11	26.71
	20	139.62	40.97	NE	31.28	28.97	25.94	22.87
				VM	34.92	31.87	28.46	24.91
				Proposed	36.31	33.36	30.02	26.62
	16	208.01	44.50	NE	31.27	28.96	25.92	22.85
				VM	34.97	31.86	29.15	24.86
				Proposed	36.31	35.32	29.43	26.51

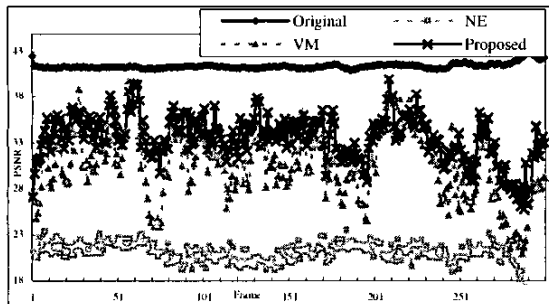


Fig. 5 The PSNR comparison with 10% error rate and QP is 20.



Fig. 6 Reconstructed image: (a) the original picture coding by QP=24; (b) the damaged picture with a 10% error rate. (c), and (d) the recovered pictures by VM and proposed algorithm, respectively.

3. CONCLUSION

In this paper, a new error concealment algorithm for the new coding standard H.264 has been presented. The algorithm includes two processes. The first process of the algorithm is to determine the size type of the lost block. The second is utilizing the multiple reference frames to find the optimal lost motion vector for the variable block size. The first process includes a main contribution of this paper, that is, a criterion is proposed to determine the size type of the lost block from the current frame. This operation not only can determine the most correct mode for the lost block under boundary crossing, but also can save much more computation time to find the optimal motion vector for the lost block from multiple reference frames. The experimental simulation results have shown that the proposed algorithm has indeed improved the quality of corrupted video images, especially, which has larger motion and edge.

4. REFERENCES

- [1] Joint Final Committee Draft (JFCD) of Joint Video Specification (ITU-T Rec. H.264 j ISO/IEC 14 496-10 AVC), T. Wiegand, Ed., July 2002.
- [2] Joint Video Team (JVT) software JM61d.
- [3] Y. Wang, M. M. Hannuksela, and V. Varsa, "The error concealment feature in the H.26L test model", Proceeding of the *Int. Conf. Image Processing*, vol. 2, pp. 729-732, Sept. 2002.
- [4] Y. W. Huang, B. Y. Hsieh, T. C. Wang, S. Y. Chien, S. Y. Ma, C.F. Shen, and L. G. Chen, "Analysis and reduce of reference frames for motion estimation in MEG-4 AVC/JVT/H.264", Proceeding of the *Int. Conf. On Acoustics, Speech, & Signal Processing*, vol. 2, pp. 809-812, Sept. 2003.
- [5] M. Flierl, T. Wiegand, B. Girod, "Rate-constrained multihypothesis prediction for motion-compensated video compression", *IEEE Trans. Circuits System Video Technology*, vol. 12, pp. 957-969, Nov. 2002.