ICSE 2009

PROCEEDINGS OF THE FIRST INTERNATIONAL CONFERENCE ON SCIENCE AND ENGINEERING

Volume - 1

Electronics
Electrical Power
Information Technology
Engineering Physics

Sedona Hotel, Yangon, Myanmar
December 4-5, 2009

PROCEEDINGS OF THE FIRST INTERNATIONAL CONFERENCE ON SCIENCE AND ENGINEERING

Volume - 1

Electronics
Electrical Power
Information Technology
Engineering Physics

Organized by
Ministry of Science and Technology

DECEMBER 4-5, 2009
SEDONA HOTEL, YANGON, MYANMAR

ELECTRICAL POWER ENGINEERING

Error Concealment of Whole Frame Loss in H.264/AVC Using Improved Motion Vector Estimation

Simon Jude Q. Lam[#1], Supavadee Aramvith[*2]

[#]*Department of Electrical Engineering, Chulalongkorn University, 54 Phayathai Road, Patumwan, Bangkok, Thailand. 10330*

[1]simonjudelam@gmail.com

[2]supavadee.a@chula.ac.th

Abstract— **H.264/AVC provides the capability to encode highly compressed videos by exploiting spatiotemporal redundancy. High compression is needed to transmit videos over bandwidth restricted channels, but also increases its vulnerability to transmission errors. The loss of one packet may lead to the loss of an entire frame, and subsequently degrading the quality of the following frames. We propose an error concealment algorithm for H.264/AVC that will estimate the lost frame based on motion information from the previous frames. We give importance to the frames close to the lost frame, but also consider the motion of other previous frames. The results show some improvement over conventional methods of concealing frame losses such as frame copy and motion copy methods.**

Keywords— **H.264, Error Concealment, Motion Vector Estimation, Frame Loss, Transmission Error**

I. INTRODUCTION

Video transmission over lossy networks, such as a wireless network, may degrade the visual quality of the decoded video. Of special interest is the highly compressed video using the latest video coding standard, H.264/AVC. H.264/AVC video can suffer from entire frame loss in the event of a transmission error. This is because an entire video frame can be encoded into a single network packet [3]. In addition, the succeeding frames also become affected by this error due to the spatiotemporal dependencies among neighboring frames in H.264/AVC [5].

Error concealment at the decoder can be used in order to reduce the effects of such errors. It requires no additional information nor modification at the encoder side [6]. Although the recent versions of JM reference software internally support error concealment algorithms, their performance are often considered inadequate.

In the scenario of a frame loss, one cannot simply apply simple error concealment techniques. Let us assume that frame *n* of a video sequence was lost during transmission. It is not possible to perform spatial error concealment on the macroblocks of frame n, since it requires the data from neighboring macroblocks on the same frame. This is due to the fact that the whole frame is lost. One of the common techniques in temporal error concealment is the Frame Copy (FC) method. In this technique, the pixels of the frame *n-1* are duplicated onto the lost frame. The motion vectors for this frame are set to zero. This technique may work for stationary regions of the frame, but fails on the opposite scenario [4]. Another technique is the Motion Copy (MC) method. In this case, the motion vectors for each macroblock of the previous frame are projected onto the lost frame. The lost frame is estimated using these motion vectors and an assigned reference frame. However, MC method often results to blocking artifacts in the concealed frame, due to the block-based prediction.

In [2], Belfore, et al. proposed an error concealment scheme for MPEG-2 that utilizes the motion information of up to *L* past frames. The underlying assumption of their work characterizes the motion of each pixel following a linear trajectory. For each pixel in frame *n-1*, a motion vector history is generated. Starting from frame *n-1*, the motion vectors of each pixel between frame *n-1* and frame *n-2* are recorded. The process is repeated until the *L*th frame or until the motion vectors point to a valid location. Afterwards, a motion vector estimate is formed in order to predict frame *n* from frame *n-1*. The predicted motion vector for each pixel in frame *n-1* is taken as the average motion vector of its motion vector history. Any pixels on frame *n* with no projected motion vector will be filled by applying a median filter around its neighborhood. The linear motion assumption has proven to be sufficient for their purposes [2]. In [1], the motion vectors are estimated in a similar manner with [2], but the estimation is performed at the block level in order to reduce computational complexity.

Several researches have used [1] and [2] as a benchmarks for their error concealment schemes, and are commonly referred to as CA_B (concealment algorithm on blocks) and CA_P (concealment algorithm on pixels) respectively. For example, in [8], recognizes the effective performance of CA_P but also points out that the method proposed by Belfiore, et al. is computationally expensive, mainly because of repeated median filtering operations. One of the proposed ideas in [8] also estimated the motion vectors at the pixel level, however, it is done by taking a weighted average of the motion vectors of the macroblocks in a neighborhood. With CA_B, the number of computations is reduced since motion vectors are estimated on the block level.

While the above algorithms mainly consider the motion vectors on the same position in the previous frames, other algorithms estimate the motion vectors whose projection falls on the currently concealed macroblock. Such an example can be seen in [9]. Motion vectors are estimated from the temporal neighbors of the frame to be concealed. For instance, the macroblocks of frame *n-1* are projected onto the lost frame *n*; it is possible that a macroblock in frame *n* may have multiple macroblock projections from frame *n-1*. The motion vectors corresponding to these macroblocks will be used to generate an estimated motion vector. The forward motion vector estimate is calculated as a weighted average of the pertinent motion vectors. The weights correspond to the area in pixels covered by the macroblock's projection. [10] shows how the coverage of the projection of the extrapolated motion vectors influence a damaged area of a corrupted frame.

In this paper, we propose an error concealment scheme that uses the motion information in multiple previous frames. However, unlike in [2], we also consider the influence of the neighboring macroblocks in our motion estimation method, since there is spatial correlation among neighboring blocks. Motion vectors from the collocated macroblock, as well as the macroblock on top and to the left are assigned weights relative to their displacements in the horizontal and vertical directions, instead of the overlapped area.

The rest of the paper is organized as follows: Section II describes the proposed motion estimation scheme, while Section III presents the experiments and results. Finally, a conclusion is found in Section IV.

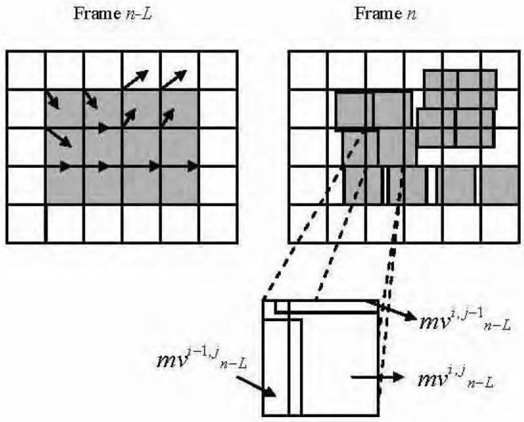
II. PROPOSED BLOCK-BASED WEIGHTED MOTION ESTIMATION

In H.264/AVC, it is possible to use multiple reference frames in video encoding. For each macroblock, a best match can be searched from frames other than the previous frame. We want to exploit this feature in order to generate a more accurate estimate of the motion vectors. For our experiments, we have set the number of reference frames *L* equal to 5. Block-based approach is also used to lessen the computational cost. We tried to compare the performance of simply taking the average motion vector of the past *L* frames and assigning different weights to each frame. Our assumption is that the correlation of motion information is higher for adjacent frames than of frames with further temporal distance. We assign the weights of the motion vectors with linear decay as shown in eq. (1). Our experiments show that assigning different weights is better than assigning equal weights to the motion vectors.

$$mv^{i,j}{}_{n} = \frac{1}{15}\left(\begin{array}{l} 5\times(m\hat{v}^{i,j}{}_{n-1}) + 4\times(m\hat{v}^{i,j}{}_{n-2}) + \\ 3\times(m\hat{v}^{i,j}{}_{n-3}) + 2\times(m\hat{v}^{i,j}{}_{n-4}) + \\ (m\hat{v}^{i,j}{}_{n-5}) \end{array} \right) \quad (1)$$

Where $mv^{i,j}{}_{n}$ is the motion vector used for predicting the macroblock at (i, j) in the lost frame. $m\hat{v}^{i,j}{}_{n-L}$ is calculated as a weighted average of $mv^{i,j}{}_{n-L}$, $mv^{i-1,j}{}_{n-L}$, *and* $mv^{i-1,j}{}_{n-L}$.

Let $m\hat{v}_x{}^{i,j}{}_{n-L}$ and $m\hat{v}_y{}^{i,j}{}_{n-L}$ be the horizontal and vertical components of $m\hat{v}^{i,j}{}_{n-L}$. $m\hat{v}_x{}^{i,j}{}_{n-L}$ and $m\hat{v}_y{}^{i,j}{}_{n-L}$ in the *L*th previous frame is computed as shown in eqs. (2) and (3):



$$\begin{aligned} m\hat{v}_x{}^{i,j}{}_{n-L} &= \alpha_{LEFT} mv_x{}^{i-1,j}{}_{n-L} + \alpha_{TOP} mv_x{}^{i,j-1}{}_{n-L} \\ &+ \alpha_{CENTER} mv_x{}^{i,j}{}_{n-L} \end{aligned} \quad (2)$$

$$\begin{aligned} m\hat{v}_y{}^{i,j}{}_{n-L} &= \alpha_{LEFT} mv_y{}^{i-1,j}{}_{n-L} + \alpha_{TOP} mv_y{}^{i,j-1}{}_{n-L} \\ &+ \alpha_{CENTER} mv_y{}^{i,j}{}_{n-L} \end{aligned} \quad (3)$$

Where α is a weighting factor computed as follows:

$$\alpha_{:LEFT} = \frac{mv_c{}^{i-1,j}{}_{n-L}}{mv_c{}^{i-1,j}{}_{n-L} + mv_c{}^{i,j-1}{}_{n-L} + mv_c{}^{i,j}{}_{n-L}} \quad (4)$$

$$\alpha_{:TOP} = \frac{mv_c{}^{i,j-1}{}_{n-L}}{mv_c{}^{i-1,j}{}_{n-L} + mv_c{}^{i,j-1}{}_{n-L} + mv_c{}^{i,j}{}_{n-L}} \quad (5)$$

$$\alpha_{:CENTER} = \frac{MB\_size - mv_c{}^{i,j}{}_{n-L}}{mv_c{}^{i-1,j}{}_{n-L} + mv_c{}^{i,j-1}{}_{n-L} + mv_c{}^{i,j}{}_{n-L}} \quad (6)$$

In eqs. (4), (5), and (6), the subscript *c* indicates for which component (x or y) the weight is computed for. Hence, the horizontal and vertical components are weighed differently. MB_size indicates the macroblock size which is kept to 16x16 in this paper. We consider the motion of the top, left, and collocated macroblocks wince the spatial correlation among them is high. The estimated motion vectors are used for motion compensated prediction of frame *n* from frame *n-1*.

III. EXPERIMENTS AND RESULTS

In order to evaluate the performance of the proposed algorithm, we ran a test on the following conditions. The encoded videos are in QCIF resolution at 30 fps. The number of reference frames is set to 5, and using only 16x16 blocks for motion compensated prediction. Finally, we encapsulate each video frame into one slice group and set the bitrate at 128000 bps. These videos were then subject to a packet loss rate of 5%. The video sequences used for this test are Akiyo, Carphone, Coastguard, Foreman, and Mobile.

The proposed concealment algorithm is implemented in JM 14.2. We compare the performance of our proposed algorithm with the motion copy method.

Figure 1. PSNR plot of Carphone sequence

Fig. 2 Frame 33 of Carphone sequence. (a) No Error, (b) Motion Copy, (c). Proposed method.

As shown in the results, there is generally an improvement in the PSNR. Particularly in the Carphone sequence, the proposed algorithm was able to realize an improvement of around 2 dB for the lost frame. In addition, the succeeding frames also exhibited around 1dB improvement, showing our method's potential to suppress error propagation.

Fig. 2 shows a comparison among the error-free frame (a), the lost frame concealed by motion copy (b), and the same frame concealed by our proposed method (c). In the Motion Copy method, we can see some blocking artifacts typically arising from conventional block-based concealment methods. IN our method, the blocking artifacts are less visible since we consider the spatial correlation of neighboring macroblocks.

TABLE I
AVERAGE PSNR FOR CONCEALED SEQUENCES

Sequence	Motion Copy	Proposed
AKIYO	37.06 dB	**37.20 dB**
CARPHONE	29.21 dB	**29.74 dB**
MOBILE	21.96 dB	**22.00 dB**

IV. CONCLUSION

In this paper, we have seen the potential of improving the performance of error concealment by reformulating the manner in which motion is estimated from a correctly received frame to a lost frame. We have found that by considering the temporal distance, a better motion estimate can be formulated. In addition, we found that the motion information found in the macroblock neighborhood can be used to further improve the estimated motion vector for the lost frame. Furthermore, the same technique may find potential application in the scalable extension of H.264. H.264/SVC is also capable of performing motion vector refinement at quarter-pel resolution which may improve the accuracy of the estimated motion vectors, potentially for intra layer and inter layer prediction modes.

REFERENCES

[1] P. Baccichet, D. Bagni, A. Chimienti, L. Pezzoni, and S. Rovati, "*Frame concealment for H.264/AVC decoders*," IEEE Trans. Consumer Electron., vol. 51, no. 1, pp. 227-233, Feb. 2005.

[2] S. Belfiore, M. Grangetto, E. Magli, G. Olmo, "*An error concealment algorithm for streaming video*," in Proc. ICIP 2003, 2003

[3] Y. Liu, J. Bu, C. Chen, L. Mo, K. He, "*Multiframe Error Concealment for Whole-Frame Loss in H.264/AVC*." ICIP (4) 2007: 281-284

[4] T. Tang, J. Wang, Y. Liu, Y. Gao, S. Yu, "*Signal Processing Systems*", 2007 IEEE Workshop, Oct. 2007 Page(s):692 – 697

[5] Y. Wang, J. Ostermann, and Y.Q. Zhang, *Video Processing and Communication.*, Signal Processing Series, NJ: Prentice Hall, 2002.

[6] Y. Wang, S. Wenger, J. Wen and A. Katsaggelos, "*Error Resilient Video Coding Techniques*," IEEE Signal Processing Magazine, pp. 61-82, July 2000.

[7] P. Baccichet, A. Chimienti, "*A low complexity concealment algorithm for the whole-frame loss in H.264/AVC*," Multimedia Signal Processing, 2004 IEEE 6th Workshop on , vol., no., pp. 279-282, 29 Sept.-1 Oct. 2004

[8] K. Song, T. Chung, Y. Kim, Y. Oh, Chang-Su Kim, "*Error Concealment of H.264/AVC Video Frames for Mobile Video Broadcasting*," Consumer Electronics, IEEE Transactions on , vol.53, no.2, pp.704-711, May 2007

[9] J.H. Kim, M.C. Hwang, H.Y. Yang; S.J. Ko; Morales, A.W., "*Frame Error Concealment Technique Using Adaptive Inter-Mode Estimation for H.264/AVC*," Consumer Electronics, 2008. ICCE 2008. Digest of Technical Papers. International Conference on , vol., no., pp.1-2, 9-13 Jan. 2008

[10] Q. Peng, T. Yang, C. Zhu, "*Block-based temporal error concealment for video packet using motion vector extrapolation*," Communications, Circuits and Systems and West Sino Expositions, IEEE 2002 International Conference on , vol.1, no., pp. 10-14 vol.1, 29 June-1 July 2002