MERL – A MITSUBISHI ELECTRIC RESEARCH LABORATORY
http://www.merl.com

Error Resilient Methods for Real-Time MPEG-4 Video Streaming

Wei-Ying Kung, Hao-Song Kong, Anthony Vetro and Huifang Sun

TR-2004-049 June 2004

Abstract

Two novel error resilient approaches are proposed respectively, to replenish intra and inter frame in this work. First we propose a spatial-domain error concealment method to conceal corrupted still images and intra-coded (I) frames. It can recursively restore pixels from the previously concealed pixels, which are selected by a proposed evaluation method. In such method, object edges can be better replenished with low complexity. Second, we propose a novel error resilient approach for P-frame. With our proposed approach, only a small amount of side information is extracted and packed at the end of every frame. Decoder can easily restore motion vector from the side information for concealment. Experimental results show that the proposed algorithms can provide better visual quality in comparison with the existing approaches.

Publication History:

1. First printing, TR-2004-049, June 2004

ERROR RESILIENT METHODS FOR REAL-TIME MPEG-4 VIDEO STREAMING

Wei-Ying Kung, Hao-Song Kong, Anthony Vetro and Huifang Sun

Mitsubishi Electric Research Laboratories
201 Broadway Cambridge, MA 02139
E-mail: weiyingk@sipi.usc.edu

ABSTRACT

Two novel error resilient approaches are proposed respectively, to replenish intra and inter frame in this work. First we propose a spatial-domain error concealment method to conceal corrupted still images and intra-coded (I) frames. It can recursively restore pixels from the previously concealed pixels, which are selected by a proposed evaluation method. In such method, object edges can be better replenished with low complexity. Second, we propose a novel error resilient approach for P-frame. With our proposed approach, only a small amount of side information is extracted and packed at the end of every frame. Decoder can easily restore motion vector from the side information for concealment. Experimental results show that the proposed algorithms can provide better visual quality in comparison with the existing approaches.

1. INTRODUCTION

It is inevitable to have packet loss or bit errors in today's Internet or wireless channels. Also, compressed video signals are very sensitive to transmission errors. Even a single bit error may crash the decoding process due to variable length coding. In addition, the motion compensated prediction scheme cause error propagation in subsequent frames after a corrupted one. Therefore, error resilience tools are essential to achieve robust video transmission [1].

1.1. Related work on I-frame concealment

In low bit rate video transmission system, intra-coded frames (I-frame) are only inserted at scene cuts to save bits. However, once errors occur in those I-frames, it will be very difficult to conceal the loss information due to lacking of correlated temporal previous information. Therefore, just as we deal with loss in still images, only neighboring spatial information are manipulated to replenish the loss parts.

A typical I-frame concealment method is to interpolate each pixel p in a lost macroblock (MB) from intact pixels in adjacent MBs [2]. Let p_i $(i = 1, 2, 3, 4)$ denote the closest pixel to p in the upper, lower, left, and right MBs, respectively. Then, the reconstruction value $\hat{p}$ of p is given by $\hat{p} = \frac{\sum_{i=1}^{4} p_i(16-d_i)}{\sum_{i=1}^{4}(16-d_i)}$, where d_i is the distance between p_i and p. This linear interpolation scheme is a simple yet effective method for smooth images. However, it may result in a blurred image if the lost MB contains high frequency components such as object edges.

More sophisticated methods [3, 4, 5] were proposed to recover high as well as low frequency components. However, those approaches are computationally expensive for real-time applications because of the use of iterative procedures.

In this work, we propose a novel spatial error concealment algorithm for I-frame. It can recursively restore pixels from the selected concealed pixels. The selection is based on proposed evaluation rules, which take distance and pixel value into account. In such method, object edges can be better restored while smooth area can be replenished as well as traditional spatial interpolation method [2]. In addition, it costs approximately same computational complexity as [2].

1.2. Related work on P-frame concealment

It is relative easier to conceal lost information in P-frames since both spatial and temporal information can be utilized. Due to the higher correlation in temporal domain than spatial domain, most concealment methods focus on motion vector recovery so that a lost MB can be motion-compensated from the reference frame. The lost motion vectors can be replaced by several existing motion vectors, such as zero vector, one of the spatially or temporally adjacent MBs, the average or median of motion vectors of adjacent MBs. All above approaches suffer poor concealed quality once false recovery happens for a missing motion vector. Even when a corrupted MB is surrounded by all corrected decoded MBs, the decoder has no clue that which neighboring motion vector provides a better concealment.

Some researchers proposed to choose a best motion vector among candidates by side matching [6] or boundary matching [7] algorithms. However, there are two shortcomings in these matching algorithms. One is increased complexity cost at the decoder for searching and mean square/absolute error calculation, which will lower the decoding speed for real-time system. The other is that it does not guarantee the selected motion vector is the best in terms of mean square/absolute error of the concealed MB.

In this work, we propose a novel error resilient approach for P-frame. It achieves better concealed quality based on additional motion information. With our proposed approach, only a small amount of side information (about 5% overhead for 64kbps video with QCIF size)is extracted and packed at the end of every frame. In many cases, the last RTP packet of a frame is smaller than other packets, which cause inefficient usage of bandwidth. Hence, it won't increase number of transmitted packets nor network loading to piggyback side information.

The rest of this paper is organized as follows. An error concealment algorithm for I-frames is presented in section 2 while another error resilient algorithm for P-frames is discussed in section 3. Simulation results are included in each section to show the performance of our proposed methods, respectively.

2. I-FRAME OUTER-TO-INNER ERROR CONCEALMENT WITH CONDITIONAL REJECTION

2.1. Proposed I-frame concealment method

In this work, a low complexity algorithm for I-frame concealment is proposed and tested in a real-time MPEG-4 video streaming system. Two major contributions are:

- Conceal pixels from error free boundaries to center.
- Evaluation rules are designed to eliminate less correlated candidate pixels.

As we know, an usual I-frame concealment for a lost MB is performed in raster order of pixels. In other words, lost pixels are estimated from upper row to lower row of a MB, and from left corner to right corner of each row. By this way, pixels in center of a MB are concealed earlier than those in bottom of a MB. We argue that in a smooth area, closer intact or concealed pixels always gives a better estimation of the lost pixel to achieve smoothness. It implies that the lost pixels close to intact boundaries can be concealed better than those ones in center of the MB. So it might be advantageous to conceal the outer pixels first, and then estimate inner pixels using the concealed pixels. Note, concealment error will grow along this procedure. However, each step on missing pixel estimation is better compensated by closer candidate pixels as long as they are faithfully concealed.

Figure 1 shows an illustration of I-frame concealment. Suppose a MB is lost during transmission, and its upper, lower, left and right MBs are error-free reconstructed or concealed. Missing pixels in that MB are estimated one by one from outer to inner of the MB. When concealing a specific pixel, four candidate pixels that concealed in last step were retrieved. Simply interpolating those candidate pixels will give a similar results as traditional spatial interpolation method. In fact, not all candidate pixels have consistent values to the pixel to be estimated. In case that candidate pixels vary in a wide range, that must result in a blur effect after spatial interpolating all candidate pixels to obtain the estimated pixel value. Therefore, a conditional rejection algorithm is developed to evaluate each candidate pixels and exclude those with poor scores. Then apply spatial interpolation to the left consistent candidate pixels to get the estimated value for the missing pixel. On one hand, spatial interpolation provides smooth compensation. On the other hand, some edges are preserved due to the rejection of the outlier.

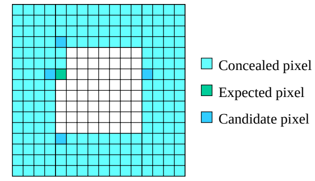

Fig. 1. Illustration of I-frame concealment for a lost MB

The next question is how do we select the right candidates? We have following observations. When the distance between the expected pixel and the candidate pixel is getting smaller, the estimation value becomes more reliable. In addition, if the candidate pixel value is getting farther away from the median value among all candidate pixels, this value will be less representative for the expected pixel. Based on the above two observations, evaluation and rejection rules are designed as follows:

1. Order candidate pixels C_i in ascending order of their values.
2. Find median value among the ordered candidate pixels. If the number of candidate pixels is odd, middle pixel is set to be median. Otherwise, average the two middle pixels and set the mean value to be median.
3. Find the difference $Diff_i$ between i^{th} candidate pixel value and the median.
4. Find distance $Dist_i$ between i^{th} candidate pixel and the expected pixel.
5. Calculate evaluation score S_i for i^{th} candidate pixel as sum of $Diff_i$ and $Dist_i$
6. If S_i is greater than some threshold T, the i^{th} candidate pixel is rejected in next interpolation operation.

(a)25.21dB (b)35.19dB (c)40.67dB

Fig. 2. I-frame error concealment

7. If all candidates are rejected, all rejections become invalid.

8. Linearly interpolate left candidate pixels and assign the obtained value to that expected pixel p.

$$p = (\sum_i \frac{C_i}{Dist_i})/(\sum_i \frac{1}{Dist_i})$$

The threshold T is the key factor to control the smoothness and sharpness of the concealed MB. In this work, the value T is set to be 20.

Although evaluation of each candidate pixel slightly increases complexity comparing to tradition spatial interpolation [2], usually not all candidate pixels are going through the final step of interpolation, thus reduces the complexity. Overall, our proposed method does not introduce much overhead complexity than [2].

Note, there are some special cases need to be considered in real world. When the number (n) of the useful neighboring MBs reduce to two, each missing pixel is always interpolated by the two nearest error-free pixels without conditional rejection. When n became one, each missing pixel is replaced with the nearest error-free pixel. In case of no useful neighboring MB exists, we simply filled the erroneous MB with grey pixels. For above special cases, our proposed concealment method just works as conventional spatial interpolation method.

2.2. Simulation results

Figure 2 demonstrates I-frame error concealment for interleaved lost MBs. The leftmost figure shows the corrupted MBs. The middle figure is obtained when spatial interpolation concealment is applied. Note, there is blur occurs in edge area. But with our proposed method, the right figure preserves edges faithfully while achieves smooth compensation for the rest lost areas. PSNR values to error-free reconstruction are shown below each sub figure.

Next we show the loss of consecutive MBs. As demonstrated in left sub-figure in Figure 3, 2 slices (or GOB) were lost. The first MB in each slice is replenished when only two useful neighboring MBs (upper and lower MBs) are available. For the rest MBs in a lost slices, three useful neighboring MBs are available for concealment, which include two error-free MBs (upper and lower MBs) and the concealed MB in left-hand side. Apparently, our proposed method outperforms the spatial interpolation method.

(a)19.20dB (b)34.47dB (c)37.43dB

Fig. 3. I frame error concealment

1	2	3
8	?	4
7	6	5

Fig. 4. Numbering motion vectors of neighboring MBs

3. P-FRAME ERROR CONCEALMENT WITH ADDITIONAL MOTION INFORMATION

3.1. Proposed P-frame concealment method

In this work, we suggest to add a small amount of side information so that the concealment quality can be significantly enhanced in error-prone environment. The procedure at the server side is described as below:

1. For each MB, we give indices for its neighboring MBs, as shown in Figure 4. $\vec{0}$ is assigned to default index 0.

2. Compensate the present MB by every indexed motion vector, and calculate the mean square/absolute error between the compensated MB and its error-free reconstruction.

3. Select the best motion vector that provides minimum mean square/absolute error, and record its index.

4. Repeat the procedure 1-3 for every MB in a frame, and encode those indices as side information.

5. Piggyback the side information to the last RTP packet of each frame. A predefined start code separates the side information from the video bit stream.

Note, although the searching part in above procedure also increases computational complexity at the server, it could be done off-line before real-time packetization. It thus extremely simplifies concealment procedure at the decoder comparing matching algorithms. In addition, it only costs at most extra 4 bits per MB to encode the side information. Moreover, the last RTP packet of a frame is usually small when every packet is encapsulated into the required length. Hence, it is advantageous to piggyback the side information at the end of last RTP packet of a frame without increasing the total number of packets.

A straightforward procedure performs at the decoder as follows:

1. If no error occurs in a frame, ignore the side information.

2. If one MB is lost, retrieve its corresponding index and find the motion vector replacement. Then the motion compensated MB is copy to the location of the missing MB, and mark the concealed MB.

3. If the indexed motion vector is lost as well, leave the MB alone and process to the next corrupted MB.

4. Repeat the steps 2-3 for every erroneous MBs. Record the total number of remaining erroneous MBs N.

5. If $N > 0$, conceal the remaining erroneous MBs again. Update N.

6. If N is decreasing, repeat step 5 until $N = 0$. Otherwise, for each erroneous MB, simply set index to be 0.

7. If the side information is lost, a boundary matching algorithm is initiated to recover the motion vector.

In some cases when the indexed motion vector is lost, but it will be recovered in its turn. Therefore, it is worthwhile to re-scan the unrecovered MBs as step 5 so that more erroneous MBs can be better concealed.

3.2. Simulation results

Figure 5 demonstrates the performance of our proposed P-frame error concealment method. The erroneous MB is localized as shown in Figure 5(a). Three different error concealment approaches have been implemented. First method is shown in Figure 5(b), which set the motion vector of the lost MB to be $\vec{0}$. This approach has very low complexity, but not effective in heavy motion areas. Second method is demonstrated in Figure 5(c). It interpolates an erroneous pixel with predicted pixels using neighboring motion vectors [8]. This is more complicated approach, but with a better performance. As you can see, there is big distortion on the moving head. That's because the decoder does not favor any neighboring motion vectors when utilizing them for concealment. Some motion vector gives poor compensation, which lower the overall performance. With our proposed method, as shown in Figure 5, it has side information to indicate the best motion vector replacement, which leads to a better concealed quality.

(a)16.16dB (b)25.41dB

(c)31.32dB (d)37.48dB

Fig. 5. P frame error concealment

4. REFERENCES

[1] Y. Wang, S. Wengers, J. Wen, and A. K. Katsaggelos, “Error resilient video coding techniques,” *IEEE Signal Processing Mag.*, vol. 17, no. 4, pp. 61–82, July 2000.

[2] P. Salama, N. B. Shroff, and E. J. Delp, “Deterministic spatial approach,” in *Signal Recovery Techniques for Image and Video compression and Transmission*, A. Katsaggelos and N. Galatsanos, Eds., pp. 212–216. Kluwer Academic Publishers, June 1998.

[3] X. Lee, Y.-Q. Zhang, and A. Leon-Garcia, “Information loss recovery for block-based image coding techniques–a fuzzy logic approach,” *IEEE Trans. Image Processing*, vol. 4, no. 3, pp. 259–273, Mar. 1995.

[4] H. Sun and W. Kwok, “Concealment of damaged block transform coded image using projections onto convex sets,” *IEEE Trans. Image Processing*, vol. 4, no. 4, pp. 470–477, Apr. 1995.

[5] S. Belfiore, L. Crisa, M. Grangetto, E. Magli, and G. Olmo, “Robust and edge-preserving video error concealment by coarse-to-fine block replenishment,” in *Proc. ICASSP*, May 2002, vol. 4, pp. 3281–3284.

[6] W. M. Lam, A. R. Reibman, and B. Liu, “Recovery of lost or erroneously received motion vectors,” in *Proc. ICASSP*, Apr. 1993, vol. 5, pp. 417–420.

[7] C. Li, J. Lu, J. Gu, and M. L. Liou, “Error resilience schemes for digital terrestrial TV broadcasting system,” in *Proc. IEEE Workshop on Signal Processing Systems*, Sept. 2001, pp. 247–258.

[8] W.-Y. Kung, C.-S. Kim, and C.-C. J. Kuo, “Runequal error protection for packet video transmission over wireless cdma channel,” in *Proc. of ITCOM'01*, Aug. 2001.