A LOW COMPLEXITY ERROR CONCEALMENT SCHEME FOR MPEG-4 CODED VIDEO SEQUENCES

Arvind Raman and Murali Babu

Ittiam Systems Pvt. Ltd, 1 Richmond Road, Bangalore-560025, India.
{www.ittiam.com}

1. ABSTRACT

MPEG coded bit streams are very sensitive to channel disturbances due to variable length coding. A single bit error can lead to very severe degradation in that part or entire slice of the image. This is of particular concern in video over wireless networks. Error concealments methods, implemented at the decoder side, present one way of dealing with this problem. In this paper a low complexity error concealment scheme for MPEG4 video sequences is proposed. Spatial and temporal information is used to reconstruct the corrupted slice/frame. The proposed schemes are very simple to implement and are well suited for real time video decoding applications that are used in handheld devices. Performance of the schemes for various BER (Bit Error Rates) is compared with the existing schemes.

Key Words: Error Concealment, MPEG 4, Bit error rate, motion vectors, spatial error concealment, temporal error concealment, handheld devices, wireless networks.

2. INTRODUCTION

Transmission of compressed video over the data network poses a challenge of dealing with the channel errors. This is particularly a problem with data loss or erasure. These errors can irreversibly damage the decoded picture leading to unacceptable degradation in video quality. Also, most of the video transmission is over a band-limited channel with bandwidth varying over time. Therefore, there can be packet losses and video retransmission under such conditions can only degrade the real time video throughput of the system. Hence, a robust error concealment scheme is key to producing a good quality video. Also, the errors can be random bit errors, burst errors or packet losses. Random bit errors do not produce undesirable degradation in the case of decoding schemes, which use fixed length codes. However, in MPEG-4 type schemes variable length codes (VLC) are used to get better compression ratios and these one-bit errors in the VLC can lead to several contiguous bits being decoded erroneously. Therefore, in this scenario the burst error and the random bit errors have very similar degrading effects on the video quality. MPEG-4 has a built in packet technique where in several macroblocks (a 16X16 pixel block) are grouped together such that there is no data dependency on the previous packet [1]. This helps in localizing the errors. Numerous schemes have been proposed to combat the effects of data loss in video decoding [2], [3], [4]. Some of the proposed schemes achieved macroblock reconstruction by estimating the low frequency DCT coefficients from the DCT coefficients of the neighboring undamaged blocks [5] or by using projections onto convex sets [6], [7] or by using MAP estimates [8]. The DCT based scheme [5] can lead to blocking artifacts. However, most of the other schemes are very computationally intensive or require large amount of data to be stored specifically for the purpose of error concealment. Since, most video decoding applications have to meet real time constraints, the proposed error concealment schemes will have to be computationally less intensive.

In this paper, we describe schemes for error concealment suited for various type of picture coding modes. The proposed schemes are well suited for low complexity decoders used in handheld devices like PDA etc. Also since these schemes work only on the received data, they are well suited to handle video packet losses over low bandwidth wireless networks. The proposed schemes are categorized as being spatial or temporal. In the spatial case the pixel data is used to reconstruct the lost data, while in the temporal case the motion vectors are used to reconstruct the lost data.

This correspondence is organized as follows, the problem of spatial error concealment of lost macroblocks and the proposed solution is addressed in section 3 and the proposed schemes for temporal error concealment is detailed in section 4 and the conclusions, results and future work are tabulated in section 5.

3. SPATIAL ERROR CONCEALMENT

Spatial error concealment techniques are used for I – pictures for which no motion information exists. Spatial error concealment techniques make use of the spatial similarity in a picture. Several techniques listed in the literature use the MAP estimates [8] such that the blocking artifacts are minimized. These schemes are iterative and are very computationally intensive. One simple scheme of spatially interpolating the pixels using the adjacent undamaged blocks is presented in [9]. However, this scheme utilizes only those macroblocks that are located spatially above and below the damaged macroblock. This scheme works well if the error is limited to one slice and that the adjacent rows are intact. This is a very unrealistic assumption and may not work very well for bigger slices that have several macroblock rows. Also this interpolates the horizontal edges but the interpolation scheme involving the vertical edges is very computationally intensive. The new scheme is aimed at solving the above said drawbacks. Here, both the horizontal and vertical edges are smoothed.

Each pixel of the current macroblock will be concealed by weighted interpolation of the surrounding four pixels. Here the surrounding four pixels are as defined in fig 1. The equation describing the process is

$$p_{x,y}(i,j) = \{(2N-i+1).p_{x-1,y}(2N,j) + i.p_{x+1,y}(1,j) + (2N-j+1).p_{x,y-1}(i,2N) + j.p_{x,y+1}(i,1)\} / (2(2N+1)) \quad \textbf{(1)}$$

where, {x, y} represents the block location and {x±i, y±i} for i $\in$ [0,1,2…] represent the positions of the neighboring blocks. As is shown in figure 1, the pixel to be interpolated (a circle inside the 8X8 block with thick edges) is estimated from the four pixels, which are on the edges of the adjacent blocks. Only the available macroblocks are used. Though, this scheme has the disadvantage of the error being propagated through adjacent blocks being spatially interpolated our simulations suggest that this works well to smooth the image across the damaged slice, thereby producing a visually better image quality. It can be seen from equation (1) that the scheme is very easy to implement. The resulting image PSNR value for the spatially reconstructed image using equation (1) is listed in table 1 and 2. The table also compares our scheme with the scheme given in [9]. It can be seen from column 3 of table 1 that our scheme performed better than the algorithm in [9] even without significantly increasing the computational complexity.

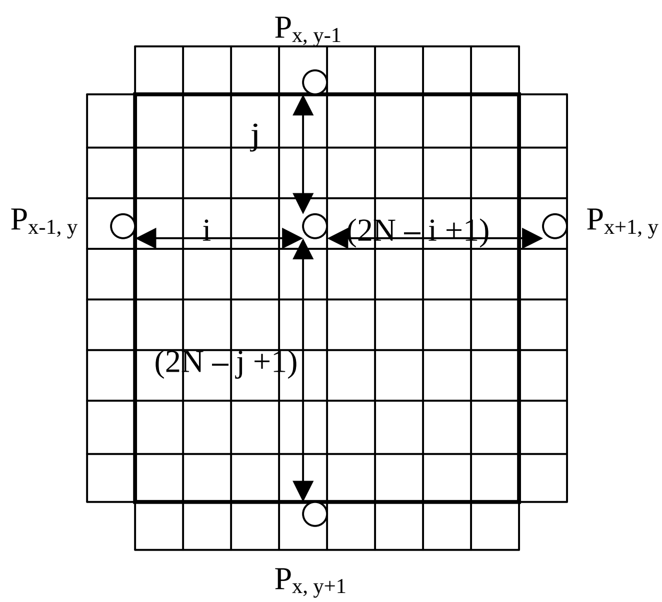

Figure 1

4. TEMPORAL ERROR CONCEALMENT

Temporal error concealment is a technique by which errors in P – pictures (predictive coded using the previous frame) are concealed. Macroblocks in P- pictures are coded by locating the best matching block in the previous frame, using which a displacement vector and transform coded difference signal are transmitted. Though the difference signal enhances the block to be coded, it is the displacement (or the motion vector) that codes most of the information in the block. Hence the recovery of motion vectors in case of errors is very crucial for the error concealment of P-pictures. The simplest temporal error concealment technique is to copy the corresponding macroblock from the previous frame [9]. This assumes that the motion is zero in that macroblock, which is not a very good assumption for video coding. Some of the techniques listed in the literature [10], [8] include finding the average, median and MAP estimate of the motion vectors using the adjacent motion vectors. However, the averaging and the median techniques have the problem of producing a wrong estimate if the adjacent non-error macroblocks are fewer in number. The MAP estimate has the problem of being computationally intensive, though it has been shown to give a good PSNR value.

The proposed scheme is based on the fact that the motion vector of the macroblock in error differs from the adjacent macroblocks in proportion to the distance between them. It was also shown in [9] that using the top and bottom row motion vectors to predict the damaged motion vector gives better results. However, this technique works well if the damaged slice includes only one row of macroblocks. Hence this scheme was modified to take into account several rows being damaged.

$$V_{x,y} = \alpha V_{x-i,y} + (1-\alpha) V_{x+i,y} \qquad \textbf{(2)}$$

Where $V_{x,y}$ is the estimated motion vector for the block at {x, y} and $V_{x-i,y}$ is the motion vector of the block at {x-i, y}, and α is such that $0 < \alpha < 1.0$ is the smoothing constant which gives a smooth estimate of the motion vector. The value of α is inversely proportional to the distance between the estimated block and the block whose motion vector is used for computation. As before, {x, y} represents the block location and {x±i, y±i} for $i \in [0,1,2\ldots]$ represent the positions of the neighboring blocks.

Since, in MPEG-4 encoding the motion vectors are predictive coded, this assumption though may not be valid for all cases when there is localized movement, it holds good for majority of the macroblocks. Also this assumption smoothens the image when subsequent frames are predicted from this error-concealed image. The results obtained by using this scheme are compared against the other schemes in table 1 & 2. Though the results from MAP estimates are better than our scheme for the last two sequences in table 1 and for the first sequence in table 2, our scheme is very easy to implement and the PSNR values are very close to the MAP estimates.

5. RESULTS, CONCLUSIONS, FUTURE WORK

The proposed algorithms have been tested on standard MPEG 4 video test sequences, assuming a received byte was in error, for Bit Error Rates of 10^{-3}, 10^{-4} and 10^{-5}. It can be seen from table 1 and table 2 that our proposed algorithms worked better or as good as the other algorithms in the literature, while taking fewer cycles of operations. It can also be seen from the Figure 2, 3 and 4 where the error concealment of Akiyo, Foreman and Silent sequences are done using the algorithms. The recovered sequences have acceptable image quality. A demo of the proposed scheme is available from the authors upon request.

In this correspondence it was shown that, a new and efficient scheme can be derived from the existing complex schemes and that the proposed scheme performs very well while being computationally less expensive. It was also shown that these schemes provide visually acceptable error concealment for very minimal increase in decoder complexity. Hence this scheme is well suited for low complexity video applications on handheld devices and over low bandwidth wireless networks. Though this scheme has been tested on MPEG-4 bit stream, the generality of the scheme suggests that it can be used for error concealment in video compression standards supporting packetized structure like MPEG-2 and H.263 version 2. The proposed algorithms have one disadvantage, that of not being able to recover from errors in the header. This can be handled by providing unequal error protection for header information and using the above said techniques for error concealment. There are several upcoming additions to the existing multimedia transport standards that provide support for sending additional data in error conditions. This can be effectively used to achieve error protection for crucial header information.

6. REFERENCES

[1] ISO/IEC 14496-2 MPEG-4 Video Coding Standard, ISO/IEC 14496-2, *Information technology – Coding of audio-visual objects – Part2: Visual*. ISO, 2000.

[2] Y. Wang and Q. Zhu, "Error control and concealment for video communication: A review", *Proc. Of the IEEE, Vol. 86, No. 5, May'98*, pp. 974-997.

[3] S. Tsekeridou, I. Pitas and C. L. Buhan," An Error Concealment Scheme for MPEG-2 Coded Video Sequences", *in Proc. of ISCAS'97, 9-12 June 1997, vol. 2, pp. 1289--1292*.

[4] S. Tsekeridou and I. Pitas, "MPEG-2 Error Concealment based on Block Matching Principles", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, no. 4, pp. 646-658, June 2000.

[5] Y. Wang, Q. Zhu and L. Shaw, "Maximally Smooth Image Recovery in Transform Coding", *IEEE Trans. on Comm., Vol. 41, No. 10, Oct'93, pp 1544-1551*.

[6] Y. Yang and N. P. Galatsanos, " Removal of Compression Artifacts Using Projections onto Convex sets and Line Process Modeling", *IEEE Trans. on Image Processing, Vol.6, No.10, Oct. 1997, pp: 1345-1357*.

[7] Y. Yang, M. Choi and N. P. Galatsanos," Image Recovery from Compressed Video Using Multi Channel Regularization", in *Signal Recovery Techniques for Image and Video Compression and Transmission, Kluwer Academic Publishers, 1998, pp 69-98.*

[8] P. Salama, N. B. Shroff and E. J. Delp," Error Concealment in Encoded Video Streams", in *Signal Recovery Techniques for Image and Video Compression and Transmission, Kluwer Academic Publishers, 1998, pp 199-234.*

[9] S. Aign, "Error Concealment for MPEG-2 Video", in *Signal Recovery Techniques for Image and Video Compression and Transmission, Kluwer Academic Publishers, 1998, pp 235-268.*

[10] M. Wada, "Selective Recovery of Video Packet Loss using Error Concealment", *IEEE Journal on Selected Areas in Comm., Vol. 7, Jun'89, pp.807-814*.

MPEG-4 Sequence	Average PSNR for a Bit stream with $8X10^{-4}$ BER	Spatial Concealment (PSNR)		Temporal Concealment (PSNR)		Average PSNR	
		Algorithm in [9]	Proposed algorithm	MAP estimate	Proposed algorithm	Algorithm in [9] + MAP	Proposed spatial + temporal
Foreman	16.313	27.931	28.045	29.547	29.753	29.201	29.444
Silent	23.49	31.154	31.195	31.190	31.000	31.184	31.000
Akiyo	20.193	30.038	30.038	31.583	31.529	31.256	31.214

Table 1: The 3 video sequences are 1.5 sec of video data each at 15 frames per second and at a bit rate of 64kbps with a BER of $8X10^{-4}$

MPEG-4 Sequence	Average PSNR for a Bit stream with $8X10^{-5}$ BER	Spatial Concealment (PSNR)		Temporal Concealment (PSNR)		Average PSNR	
		Algorithm in [9]	Proposed algorithm	MAP estimate	Proposed algorithm	Algorithm in [9] + MAP	Proposed spatial + temporal
Foreman	28.449	28.994	28.994	31.708	31.698	31.165	31.156
Silent	28.267	32.412	32.412	32.959	32.959	32.868	32.868
Akiyo	31.422	32.8	32.8	36.939	36.939	36.396	36.396

Table 2: The 3 video sequences are 1.5 sec of video data each at 15 frames per second and at a bit rate of 64kbps with a BER of $8X10^{-5}$

Figure 2a: Degraded Akiyo Image

Figure 2b: Error Concealed Image

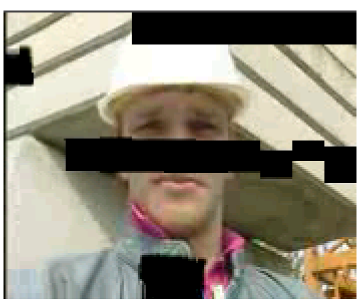

Figure 3a: Degraded Foreman Image

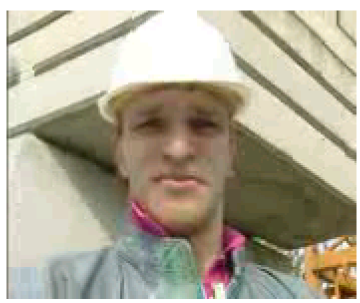

Figure 3b: Error Concealed Image

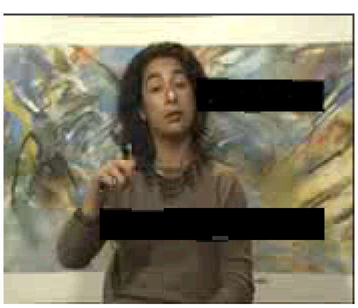

Figure 4a: Degraded silent Image

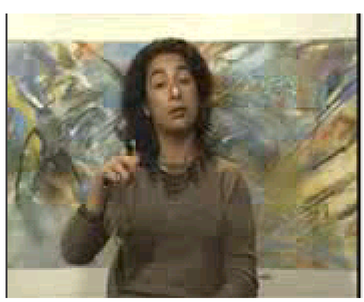

Figure 4b: Error Concealed Image