

Temporal Error Concealment Using Self-Organizing Map*

Yu-Len Huang and Hsiu-Yi Lien

*Department of Computer Science and Information Engineering Tunghai University,
Taichung, Taiwan
ylhuang@mail.thu.edu.tw*

Abstract—The performance of MPEG video transmission over error-prone channels is limited by the channel noise. An efficient error concealment (EC) scheme is essential for diminishing the impact of transmission errors in a compressed video. A number of error concealment (EC) techniques have been developed to combat the transmission errors. However, the techniques are always inefficient when the motions of object in a video are fast or complex. In this paper, a novel adaptive EC algorithm is proposed to conceal the error for block-based image coding systems by using neural network (NN) techniques in the spatial domain. In the proposed algorithm, the self-organizing map (SOM) is used for estimating and reconstructing the motion vectors of damaged blocks. The SOM has a great capacity for visualizing and interpreting high-dimensional data sets. Computer simulation results show that the visual quality and the PSNR evaluation of reconstructed frames are significantly improved by using the proposed EC algorithm. From the experimental results, we find that the proposed algorithm is expected to be a useful EC algorithm for motion vector compressed video in error-prone networks.

Keywords: Error concealment, neural network, SOM, motion vector estimation, MPEG video.

1. Introduction

Most emerging image coding methods, such as H.261, H.263, MPEG-1, MPEG-2 and MPEG-4 [1], use block-based discrete cosine transform (DCT) and motion-compensation to remove spatial and temporal redundancies. In a coded video bitstream, each frame is divided into slices, and each slice is divided into macroblocks. A macroblock can further split into blocks. The content of each block is quantized and encoded using variable length entropy coding. Although these hybrid motion compensation based compression methods successfully reduce the bit rate, they are very sensitive to channel disturbances due to the variable length entropy encoding, especially in

an error-prone environment such as a wireless channel. A single bit error may corrupt all the macroblocks within a slice. In addition, this error may also propagate to succeeding or preceding frames.

Temporal prediction has been shown to be an effective approach to conceal the damaged macroblocks by simply exploiting the temporal correlation between the current decoded frame and previous one [2-4]. The temporal error concealment (EC) methods utilized the temporal redundancy in an image sequence and were intended for inter-coded frames that contain some motion information. However, temporal EC methods always require the knowledge of the motion information in a video. Due to variable length coding, the motion vector of each macroblock may be lost. There are several ways to estimate the lost motion vectors. In [5], the impaired macroblocks are simply replaced by the spatially corresponding macroblock in the previous frame. It works well for stationary areas but always collapsed for moving areas. This approach may cause undesirable artifacts when motion between frames is extensive. Thus a robust method has been proposed by using the motion vectors of the undamaged neighboring macroblocks to estimate the damaged macroblock. This method also work well for the areas with smooth motion, but failed for areas with fast motion and on the object boundaries. Lam [6] proposed a boundary matching algorithm (BMA) to search for the motion vector based on a spatial smoothness constraint on the boundaries of the damaged macroblock. However, the boundary information may not be sufficient to accurately estimate the motion vector if the left and right macroblocks are not available. In addition, slanting edges and rapid gray-level changes may cause an extensive variation and then the BMA method may be flunked. An improved BMA has been proposed to estimate the motion vector by further imposing smoothness properties on the diagonal and anti-diagonal directions for slanting edges and sub-pixel samples for rapid gray-level changes [7]. Feng et al. [8] also proposed a similar method to handle the cases of existing diagonal and anti-diagonal edges on the boundaries. Although the improved BMA approaches can cope with slating edges, they may

* This work was supported by the National Science Council, Taiwan, under Grants NSC-92-2213-E-029-022.

fail in the case where there are edges across the damaged macroblock from left to right because of the lack of information on the left and right boundaries.

We notice that the conventional temporal EC methods utilize different linear statistical relationships to conceal damaged macroblocks. However, if the motions of macroblocks in frames are irregular, the implementation will become inefficient, and then the schemes will produce poor performance for error concealment. Thus, we desire to develop a simple and flexible temporal EC scheme to solve the problem.

The artificial neural network (NN) techniques have been applied to solve complex problems in the field of image processing. The self-organizing map (SOM) has been introduced by T. Kohonen [9] as an effective neural network model for the visualization of high-dimensional data. SOM has been found to be particularly effective for problems that can make use of unsupervised training. The SOM learning algorithm is a type of unsupervised, iterative vector quantization that converts complex, nonlinear statistical data items from a high-dimensional space onto simple reference vectors in a low-dimensional space [10-14]. This study employs the SOM as a predictor to estimate the motion vector of the damaged macroblock. The estimation model proposed herein can exploit the nonlinearity property of the SOM to estimate lost motion vectors more accurately.

The rest of this paper is organized as follows. Section 2 reviews the main features of the SOM neural networks. Section 3 describes the construction of the proposed temporal EC model based on SOM. Experimental results are given in Section 4 for image sequences outside the training set. Finally, conclusions are drawn in Section 5.

2. Motion Vector Estimation Using SOM

An SOM neural network [9] contains an input layer, a single hidden layer, and a mapping array of output, as illustrated in Fig. 2. It refers to the ability of unsupervised learning. The number of input neurons ensures from the dimension of the input vectors. In general, the SOM model defines a mapping from the higher dimension of input data space onto a regular two-dimensional mapping array. With every neuron in the mapping array, a parametric weight vector produced by learning algorithm is associated. An input vector will compare with all weight vectors, and the best match is defined as the SOM response. For a more useful analysis, each neuron in the mapping array may also be marked a class label using training samples. The learning algorithm for SOM is shortly described as the following steps:

1. Initialization of the SOM weights to small random values. Establish the initial radials of the region for weight modification.
2. Presentation of an N -dimension input vector \mathbf{x} .
3. Calculation of the distance d_j between the current input vector and the weight vectors of all mapping array neurons according to the Euclidean distance function

$$d_j = \sum_{i=1}^N (x_i - w_{ji})^2$$

where x_i is the i -th component of the input vector and w_{ji} is the i -th component of the weight vector \mathbf{w}_j .

4. Determination of the output neuron o which is the one with minimal distance
5. Update of all weight vectors of neurons lying within the region of weight modification. Weight modification is given by

$$\Delta w_{ji} = o \eta (x_i - w_{ji})$$

where η is the learning rate parameter.

6. Repeat the steps above by presenting a new input vector until the given stop criterion is achieved.

During the weight modification phase of the learning algorithm, the learning rate parameter and radial of the region of weight modification could be monotonically decreasing over time. This work can be used to minimize the distortion of the SOM. The SOM neural networks perform high unsupervised learning capability and computational efficiency. The SOM models perform high unsupervised learning capability and computational efficiency. The proposed temporal EC system estimates the motion vectors of damaged macroblocks by adding the final SOM mapping array.

3. Adaptive SOM Error Concealment Algorithm

Before any EC techniques can be applied to the compressed images, the locations of damaged macroblocks are necessary first to be found out. In this paper, we only focus on the problem of concealing the error macroblocks for block-based image coding systems. Thus, we assume that the locations of damaged macroblocks are known and discuss techniques for concealing the detected errors. The proposed temporal EC algorithm estimates the motion vectors of the damaged macroblocks by exploiting the information of four adjacent macroblocks (top, bottom, left, and right macroblocks). Let MV_X represents motion vector of the damaged macroblock X . MV_T , MV_L , MV_R and MV_B represent the motion vector of top, left, right, and bottom macroblock surrounding X , respectively. By the spatial redundancy between the motion vectors of adjacent macroblocks, the SOM model utilizes the motion vectors of macroblocks

surrounding the damaged macroblock to estimate the motion vector of the damaged macroblock.

The proposed temporal EC algorithm described previously assumes that the errors are localized to separated macroblocks in the decompressed frames. In fact, errors typically propagate through several consecutive macroblocks for compressed frames and then yield some of adjacent lost macroblocks. In this condition, the conventional temporal EC techniques would fail to accurately estimate the motion vector of the damaged macroblock because there is not enough information. On the contrary, the SOM model has the possibility to estimate the motion vector of damaged macroblock by using incomplete input motion vectors. The SOM will compute the distance calculations and reference vector modification steps using the available data components. For example, partial data can still be used to determine the distribution statistics of the available vector components [9, 15].

Moreover, the proposed method performs a search within the previous frame for the best boundary match surrounding a damaged macroblock. Similar to the BMA, the proposed method selects a motion vector among a set of candidate vectors that minimizes the total variation between the boundaries of the damaged macroblock. Thus candidate vectors of the BMA method are added to the proposed to conceal the damaged macroblocks. Clearly, the proposed temporal EC method is always practical even if the motions in a frame are complex or irregular. The whole proposed method utilizes only one manner of neural network model - the SOM network. Thus, the architecture of the proposed temporal EC algorithm is simple, redressing easily, and suitable for hardware design. Moreover, the SOM neural networks are highly parallel computer architecture and, thus, offer the potential for real-time applications.

4. Experimental Results

In this paper, three image sequences, “bus” (300 frames, CCIR601), “football” (150 frames, CCIR601) and “flower garden” (150 frames, CCIR601) were used in the simulations. The image sequences were encoded by MPEG-2 standard. Each group of picture (GOP) in encoded sequenced consists of 18 frames (1 I-frame, 5 P-frames, 12 B-frames) and the size of macroblock is 16×16 . We perform a two-state Markov model, namely the Gilbert channel model, to simulate the transmission errors. The comparisons for EC methods were made in average macroblock missing rates 1%, 5%, 10% and 20%. The peak signal-to-noise ratio (PSNR) is used to give a quantitative evaluation on the quality of the reconstructed frame.

The proposed method and other popular existing temporal EC approaches are implemented in

this study. In order to prove the precision of the proposed SOM estimation, we compare the results using the proposed EC algorithm and simulation results of temporal replacement (TR) method that replaced all damaged motion vectors by zero and the BMA method proposed in [6]. Table 1 shows the average PSNR values for reconstructed image sequences by using the three EC methods in average packet loss rates 1%, 5%, 10% and 20%. The proposed SOM method achieves the better performance than other methods. Figure 1, Fig. 2 and Fig. 3 show the PSNR values for individual frame of the reconstructed test sequences “bus”, “football” and “flower garden”, respectively.

Table 1. Average PSNR values for the test sequences with different EC methods.

Image sequence	EC method	Average macroblock missing rate			
		1%	5%	10%	20%
Bus	TR	26.5	26.5	23.4	21.0
	BMA	29.9	29.9	26.1	23.6
	The proposed method	30.7	31.3	27.4	25.0
Football	TR	34.4	25.6	23.1	20.2
	BMA	35.8	27.0	24.4	21.5
	The proposed method	37.7	29.0	26.3	23.3
Flower garden	TR	28.2	21.2	18.1	15.6
	BMA	33.4	24.7	21.4	18.8
	The proposed method	35.4	26.1	22.7	20.2

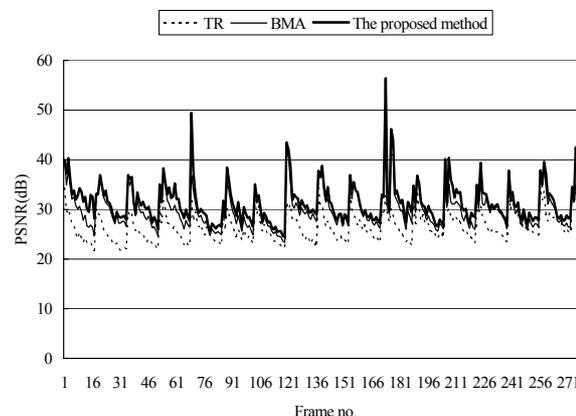


Figure 1. PSNR values for frames of the test sequence “bus.” The sequence is affected by data losses with a 5% of macroblock missing.

Figure 4(a) shows an error-free frame in test sequence “flower garden” and Fig. 4(b) is the damaged frame with no-overlapped stripes lost. The damaged macroblocks in the frame are replaced with gray-level 0. The reconstructed frames using the TR, BMA and the proposed method are shown in Fig. 4(c), (d) and (e), respectively. It was apparent that the reconstructed frame produced by TR was inferior to that of the BMA and the proposed method. The

recovered image using the proposed method shows the improved subjective quality than other recovered images. Similarly, Figs. 5(a) and (b) show the undamaged and damaged frames from test sequence “bus”, respectively. The reconstructed frames using the TR, BMA and the proposed method are shown in Figs. 5(c), (d) and (e), respectively. The BMA shows that can produce acceptable recovered frames. However, it cannot successfully reconstruct the macroblocks that contain edges across the left and right boundaries. That is because the information of left and right macroblocks surrounding the damaged macroblock is not available. We observe that both the smooth and detailed regions in the reconstructed frames can obtain the good visual quality using the proposed EC algorithm. From the simulation results, we find that the proposed algorithm has very good performance.

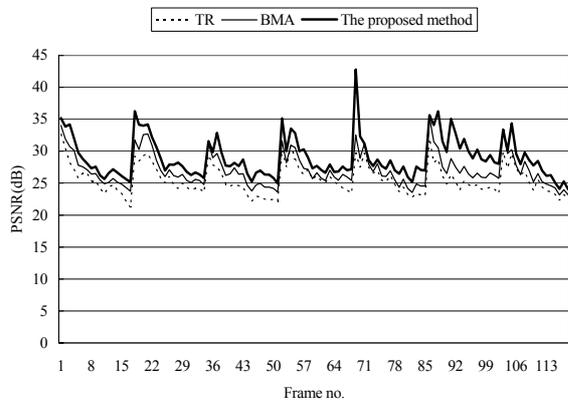


Figure 2. PSNR values for frames of the test sequence “football.” The sequence is affected by data losses with a 5% of macroblock missing.

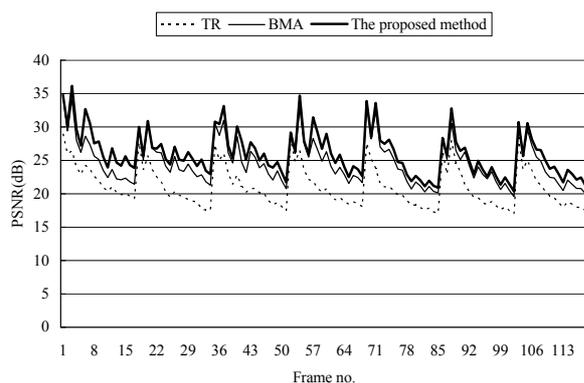


Figure 3. PSNR values for frames of the test sequence “flower garden.” The sequence is affected by data losses with a 5% of macroblock missing.

To quantify the computational complexity of the EC methods, we measured the execution time required for each technique. The average execution time of reconstructed frames using the TR, the BMA

and the proposed method are 0.062, 3.4138 and 3.8578 second, respectively. In the proposed method, the average execution time is slightly larger than other two methods. However, the proposed method achieves the better video quality. We conclude that the proposed robust error concealment technique can be a useful alternative to existing techniques due to its low computational complexity and acceptable PSNR performance.



(a)



(b)



(c)

Figure 4. Subjective video quality comparison for frame 39 in test sequence “Flower Garden”: (a) original frame; (b) corrupted frame; (c) reconstructed frame using the TR method; (d) reconstructed frame using the BMA; and (e) reconstructed frame using the proposed method. (continued)



(d)



(e)

Figure 4. Subjective video quality comparison for frame 39 in test sequence “Flower Garden”: (a) original frame; (b) corrupted frame; (c) reconstructed frame using the TR method; (d) reconstructed frame using the BMA; and (e) reconstructed frame using the proposed method.

5. Conclusions

A new EC technique was presented in this paper, aimed at masking the loss of visual information due to erroneous transmission of coded digital video over unreliable networks. The method using the SOM neural network model overcomes the disadvantage of boundary matching algorithm at fail in areas with unsmooth motion and also for areas with low spatial correlation. The proposed method is able to reconstruct motion vector of the damaged macroblock by using incomplete information and increasing both the subjective and objective quality of the reconstructed frames even under the high packet losses. Even though the motion is extensive, the proposed EC method increase decoded video quality by using the nonlinear statistical relationships feature of the motion. The proposed method can estimate the motion vectors of the corrupted macroblocks even in the sequences including fast moving objects. Thus the high-motion sequences have better performance than the small-motion

sequences. The proposed temporal EC will be applicable to the video decoder requiring low computation complexity and the robust EC to high scene activity.



(a)



(b)



(c)

Figure 5. Subjective video quality comparison for frame 73 in test sequence “Bus”: (a) original frame; (b) corrupted frame; (c) reconstructed frame using the TR method; (d) reconstructed frame using the BMA; and (e) reconstructed frame using the proposed method. (continued)



(d)



(e)

Figure 5. Subjective video quality comparison for frame 73 in test sequence "Bus": (a) original frame; (b) corrupted frame; (c) reconstructed frame using the TR method; (d) reconstructed frame using the BMA; and (e) reconstructed frame using the proposed method.

References

- [1] J.L. Mitchell, W.B. Pennebaker, C.E. Fogg, and D.J. LeGall, MPEG Video Compression Standard. London, U.K.: Chapman & Hall, 1997.
- [2] Y.C. Lee, Y. Altunbasak, R. Mersereau, "A temporal error concealment method for MPEG coded video using a multi-frame boundary matching algorithm", International Conference on Image Processing, Vol. 1, Oct. 2001, pp. 990-993.
- [3] M.H. Jo and W.J. Song, "Error concealment for MPEG-2 video decoders with enhanced coding mode estimation", IEEE Trans. on Consumer Electronics, Vol. 46, Nov. 2000, pp. 962-969.
- [4] Y.W. and Q.F. Zhu, "Error control and concealment for video communication: a review", Proc. IEEE, Vol. 86, May 1998, pp. 974-997.
- [5] P. Haskell and D. Messerschmitt, "Resynchronization of motion compensated video affected by atm cell loss", ICASSP in Proc. IEEE, Vol. 3, March 1992, pp. 545-548.
- [6] W.M. Lam, A.R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors", ICASSP in Proc. IEEE, Vol. 5, Mar. 1993, pp. 417-420.
- [7] C.S. Park, J. Ye, and S.U. Lee, "Lost motion vector recovery algorithm", ISCAS in Proc. IEEE, Vol.3, 1994, pp. 229-232.
- [8] J. Feng, K.T. Lo, H. Mehrpour, and A. E. Karbowiak, "Cell loss concealment method for mpeg video in atm networks", GLOBECOM in Proc. IEEE, Vol. 3, 1995, pp. 1925-1929.
- [9] T. Kohonen, Self-Organizing Maps, Springer-Verlag, Heidelberg, 1995.
- [10] N.V. Swindale and H.U. Bauer, "Application of Kohonen's self-organizing feature map algorithm to cortical maps of orientation and direction preference", The Royal Society, January 1998, pp. 827-838.
- [11] D. Cook, A. Buja, J. Cabrera, and C. Hurley, "Grand tour and projection pursuit," Journal of Computational and Graphical Statistics, Vol. 4, no. 3, 1995, pp. 155-172.
- [12] M.A. Kraaijveld, J. Mao, and A.K. Jain, "A nonlinear projection method based on Kohonen's topology preserving maps," IEEE Transactions on Neural Networks, Vol. 6, May 1995, pp. 548-559.
- [13] J. Vesanto, "SOM-based data visualization methods," Intelligent Data Analysis, Vol. 3, April 1999, pp. 111-126.
- [14] I.S. Dhillon and D.S. Modha, "Concept decompositions for large sparse text data using clustering," Machine Learning, January 2001, Vol. 42, pp. 143-175.
- [15] T. Samad, S.A. Harp. Self-organization with partial data. Network: Computation in Neural Systems, 1992, pp. 205-212.