Spatial and Temporal Error Concealment Techniques for Video Transmission Over Noisy Channels

Wei-Ying Kung, *Member, IEEE*, Chang-Su Kim, *Senior Member, IEEE*, and C.-C. Jay Kuo, *Fellow, IEEE*

Abstract—Two novel error concealment techniques are proposed for video transmission over noisy channels in this work. First, we present a spatial error concealment method to compensate a lost macroblock in intra-coded frames, in which no useful temporal information is available. Based on selective directional interpolation, our method can recover both smooth and edge areas efficiently. Second, we examine a dynamic mode-weighted error concealment method for replenishing missing pixels in a lost macroblock of inter-coded frames. Our method adopts a decoder-based error tracking model and combines several concealment modes adaptively to minimize the mean square error of each pixel. The method is capable of concealing lost packets as well as reducing the error propagation effect. Extensive simulations have been performed to demonstrate the performance of the proposed methods in error-prone environments.

Index Terms—Directional interpolation, minimum mean square error (MMSE) decoding, robust video transmission, spatial error concealment, temporal error concealment.

I. INTRODUCTION

VIDEO compression technologies have been extensively studied in recent years. The basic concept of video compression is to reduce the amount of bits for video representation by exploiting spatial and temporal correlations in image sequences. In general, the discrete cosine transform (DCT) is employed to transform time domain signals to frequency domain coefficients so that signal energies are concentrated in low frequency regions. Then, those frequency components can be effectively encoded with quantization and variable length coding (VLC) due to energy compaction and long consecutive zeros. Moreover, the compression performance can be further enhanced by employing motion-compensated prediction, which predicts each frame blockwise from the previous frame. The prediction error can be more effectively compressed than the original frame data.

Unfortunately, most channels such as wireless channels and the Internet are not reliable enough to guarantee error-free transmission. Wireless channels have the path loss, long-term fading effects, and short-term fading effects which result in fast fluctuation and unreliability. Also, packet loss and delay are inevitable in the Internet. Compressed video signals are very sensitive to transmission errors. In VLC, synchronization between the encoder and the decoder is required for correct decoding. Even a single bit error may cause the loss of synchronization so that the remaining bit stream cannot be decoded properly. The motion compensated prediction scheme is also vulnerable, since transmission errors in a frame tend to propagate to subsequent frames.

Error resilience is needed to achieve robust video transmission [1], [2]. One strategy is to use a feedback channel to request retransmission or adjust encoding modes according to channel conditions [3]. It is efficient in stopping error propagation but introduces extra delay, which is not acceptable in many interactive applications. Another way to achieve robustness is to insert redundant information systematically into compressed video signals so that the decoder can compensate transmission errors. The redundant information can be error correction codes [4], [5] or multiple descriptions [6], [7]. The former one combined with layered coding can provide good performance in prioritized networks while the latter is suitable for delivery over multiple channels to enhance reliability. However, error resilience is achieved at the expense of coding efficiency in both methods.

Error concealment techniques at the decoder attempt to conceal erroneous blocks using the correctly decoded information without modifying source and channel coding schemes [8], [9]. They are hence suitable for a wide range of applications. Depending on the available information, different error concealment methods can be developed to exploit the information effectively. Typical video codecs, such as MPEG-4, H.263 and H.264, classify video frames into three types: the intra (I), the predictive (P) and the bidirectional (B) frames. Erroneous B-frames can be simply dropped, since they are not referenced by subsequent frames. In contrast, erroneous I- or P-frames may result in error propagation to subsequent frames and have to be concealed in some way.

In this work, we propose novel spatial and temporal error concealment algorithms for I- and P-frames. The algorithm for I-frame concealment can restore edge components as well as low frequency information by employing edge detection and directional interpolation. The algorithm for P-frame concealment adaptively fills in erroneous blocks with the information in previous frames based on a dynamic error tracking model.

W.-Y. Kung is with Motorola Advanced Technology, San Diego, CA 92121 USA (e-mail: vivian.kung@motorola.com).

C.-S. Kim is with the Department of Electronics and Computer Engineering, Korea University, Seoul 136-701, Korea (e-mail: cskim@ieee.org).

C.-C. J. Kuo is with the Department of Electrical Engineering and Integrated Media Systems Center, University of Southern California. Los Angeles, CA 90089-2564 USA (e-mail: cckuo@sipi.usc.edu).

It is demonstrated by simulation results that the proposed algorithms can suppress error propagation as well as conceal erroneous blocks effectively.

The rest of this paper is organized as follows. Previous work on error concealment is reviewed in Section II. An error concealment algorithm for the I-frame is presented in Section III while another error concealment algorithm for the P-frame is discussed in Sections IV and V. A few implementation issues are examined in Section VI, and experimental results are presented in Section VII. Finally, concluding remarks are given in Section VIII.

II. PREVIOUS WORK ON ERROR CONCEALMENT

A. I-Frame Concealment

In many low bitrate applications, the I-frame mode is used only for the frames at the beginning of a sequence or a scene cut, for which no temporal information can be exploited to reduce the bit rate. Various algorithms have been proposed for the concealment of errors in I-frames based on the spatial information.

A typical method is to interpolate each pixel p in a lost macroblock (MB) from intact pixels in adjacent MBs [10], [11]. 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

$$\hat{p} = \frac{\sum_{i=1}^{4} p_i (W - d_i)}{\sum_{i=1}^{4} (W - d_i)} \qquad (1)$$

where W is the horizontal or vertical size of an MB, and d_i is the distance between p_i and p. This linear interpolation scheme is a simple yet effective method for smooth images. Note that the weighting coefficient $(W - d_i)$ is selected to be inversely proportional to distance d_i. In [12], [13], a more advanced technique was proposed to perform the interpolation adaptively to achieve the maximum smoothness. Generally speaking, these methods attempt to reconstruct a lost MB as a smooth interpolated surface from its neighbors. However, they may result in a blurred image if the lost MB contains high frequency components such as object edges.

The fuzzy logic reasoning approach [14], [15] uses a vague similarity relationship between a lost MB and its neighbors to recover high as well as low frequency information. It first recovers the low frequency information with surface fitting. Then, it uses fuzzy logic reasoning to coarsely interpret high frequency information such as complicated textures and edges. Finally, a sliding window iteration is performed to integrate results in the previous two steps to get the optimal output in terms of surface continuity and a set of inference rules. In [16], another iterative error concealment algorithm was proposed. It uses a block classifier to determine edge directions based on the gradient data. Then, instead of imposing a smoothness constraint only, an iterative procedure called "projections onto convex sets (POCS)" is adopted to restore lost MBs with an additional directional constraint. This approach provides satisfactory results when the missing MB is characterized by a single dominant edge direction. In [17], the coarse-to-fine block replenishment (CFBR) algorithm was proposed, which first recovers a smooth large-scale pattern, then a large-scale structure, and finally local edges in a lost MB. The fuzzy logic, POCS, and CFBR approaches are, however, computationally expensive for real-time applications because of the use of iterative procedures.

In [18], a computationally efficient algorithm was proposed based on directional interpolation. First, it infers the geometric structure of a lost MB from the surrounding intact pixels. Specifically, the surrounding pixels are converted into a binary pattern and one or more edges are retrieved by connecting transition points within the binary pattern. Then, the lost MB is directionally interpolated along edge directions so that it is smoothly connected to its neighbors with consistent edges. However, the transition points are selected heuristically and connected using only the angle information. Thus, the retrieved edges may not be faithful to the original ones. In Section III, we will propose an algorithm for I-frame concealment, which is computationally as efficient as [18] but employs a more robust edge detection scheme.

B. P-Frame Concealment

For the error concealment of P-frames, temporal, as well as spatial, information is available. In fact, temporal correlation is much higher than spatial correlation in real world image sequences so that P-frames can be more effectively concealed than I-frames. In P-frames, the compressed data for an MB consist of one or more motion vectors and residual DCT coefficients. If only DCT coefficients are lost, a motion-compensated MB still provides acceptable visual quality. However, if both the motion vector and DCT coefficients are lost, the motion vector is recovered using the information in adjacent MBs, and the lost MB is motion-compensated using the recovered motion vector. There are several approaches to recover lost motion vectors.

1) Set the lost motion vector to zero. Thus, this approach replaces a lost MB by the MB at the same spatial location in the previous frame.
2) Use the motion vector of one of the spatially or temporally adjacent MBs.
3) Use the average or median of motion vectors of adjacent MBs.
4) Choose the motion vector based on the side matching criterion [19], [20]. Among the set of candidate motion vectors, this approach selects the vector minimizing the side matching distortion so that the concealed MB is smoothly connected to the surrounding pixels.
5) Estimate the motion vector with block matching [21]–[23]. This approach estimates the motion vector for the set of the surrounding pixels, and applies that vector to the lost MB.

It was shown that the error concealment performance can be improved by employing advanced motion compensation techniques such as the overlapped block motion compensation [20] and the affine motion compensation [24] after motion vector recovery.

Another method for P-frame concealment [25] interpolates damaged regions adaptively to achieve the maximum smoothness in the spatial, temporal and frequency domains. Statistical methods [26]–[28] model image pixels or motion fields as Markov random fields, and then estimate the lost content
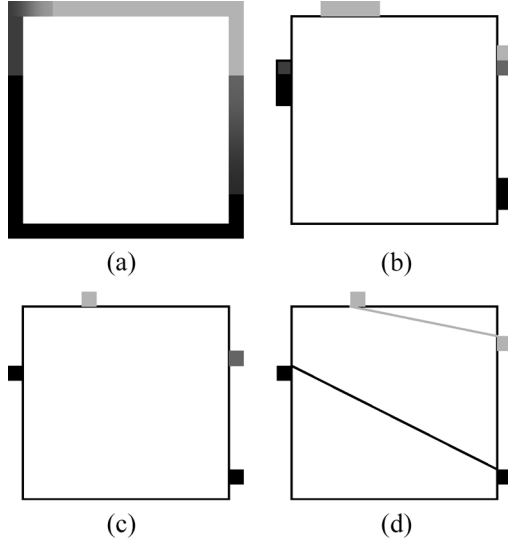
Fig. 1. Edge recovery process. (a) Edge detection on boundary pixels. (b) Detected edge points. (c) Obtaining representative edge points. (d) Edge matching and linking.

using maximum *a posteriori* (MAP) estimators. Alternatively, a model-based method [29] builds a model for the region of interest (e.g. the face) during the decoding of image sequences and recover the corrupted data by projecting it onto the model. Lee *et al.* [30] proposed a hybrid method that models video as a mixture of Markov processes and conceals erroneous blocks by combining both spatial and temporal information seamlessly.

All the above methods focus on the concealment of erroneous blocks only. However, the concealment effect is not complete, and concealment errors tend to propagate to subsequent frames because of motion compensated prediction. In Sections IV and V, we will propose a novel P-frame error concealment method, which attempts not only to conceal erroneous blocks but also to suppress the error propagation phenomenon.

III. DIRECTIONAL INTERPOLATION FOR I-FRAME CONCEALMENT

In this section, we propose an algorithm for I-frame concealment, which can restore edge components as well as low frequency information. The proposed algorithm first detects edges components in neighboring boundary pixels, and connects broken edges in the lost MB via linear approximation. Then, the lost MB is partitioned into segments based on the recovered edge information. Finally, each pixel in a segment is directionally interpolated from the boundary pixels that are adjacent to the segment.

A. Edge Recovery

Edges, which mean sharp changes or discontinuities in luminance values, play an important role in human perception of images. Generally, an image with blurred edges is annoying to human eyes. In this work, edges in missing MBs are recovered by the scheme illustrated in Fig. 1.

Suppose that a missing MB is surrounded by four correctly decoded MBs. First, edges are detected by calculating the gradient field on the boundary pixels in neighboring MBs. The gradient at pixel (x, y), denoted by $(G_r(x, y), G_c(x, y))$, can be computed by the convolution of the image $P(x, y)$ with row and column impulse arrays as

$$G_r(x, y) = P(x, y) * H_r(x, y) \tag{2}$$
$$G_c(x, y) = P(x, y) * H_c(x, y). \tag{3}$$

The following Sobel operator is adopted in this work:

$$H_r = \frac{1}{4} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, \quad H_c = \frac{1}{4} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \tag{4}$$

Note that if the Sobel operators are directly applied to boundary pixels, the gradient calculation involves corrupted pixel values, which leads to inaccurate edge detection. Instead, we apply the Sobel operators to the second boundary lines from the top, bottom, left and right of the corrupted MB. The amplitude and angle of the gradient are then defined as

$$A(x, y) = \sqrt{G_r^2(x, y) + G_c^2(x, y)} \tag{5}$$
$$\theta(x, y) = \arctan \frac{G_r(x, y)}{G_c(x, y)}. \tag{6}$$

If the amplitude $A(x, y)$ is larger than a pre-specified threshold, pixel (x, y) is said to lie on an edge. The threshold is set to the variance of pixel values here. Several consecutive pixels are often detected as edge points as shown in Fig. 1(b). Among them, only one pixel with the largest gradient amplitude is selected as the true edge point as shown in Fig. 1(c).

It is assumed that there are two cases when an edge enters a lost MB through an edge point. The first case is that the edge exits the MB via another edge point. The second case is that the edge meets another edge within the MB and, as a result, does not exit the MB. Based on this assumption, we should compare the edge points to find the matched pairs. The attribute vector of an edge point at (x, y) is defined as

$$\alpha(x, y) = (G_r(x, y), G_c(x, y), \theta(x, y), P(x, y)). \tag{7}$$

Each element in $\alpha(x, y)$ gives similar contribution for an edge point. So, by setting the normalized factor to be 1, a simple attribute distance between two edge points can be calculated via

$$
\begin{aligned}
d\left(\alpha(x_1, y_1), \alpha(x_2, y_2)\right) \\
= |G_r(x_1, y_1) - G_r(x_2, y_2)| + |G_c(x_1, y_1) - G_c(x_2, y_2)| \\
+ |\theta(x_1, y_1) - \theta_{1,2}| + |\theta(x_2, y_2) - \theta_{1,2}| \\
+ |P(x_1, y_1) - P(x_2, y_2)|
\end{aligned} \tag{8}
$$

where $\theta_{1,2}$ is the slant angle of the line connecting (x_1, y_1) and (x_2, y_2). A pair of edge points is deemed to be a match if
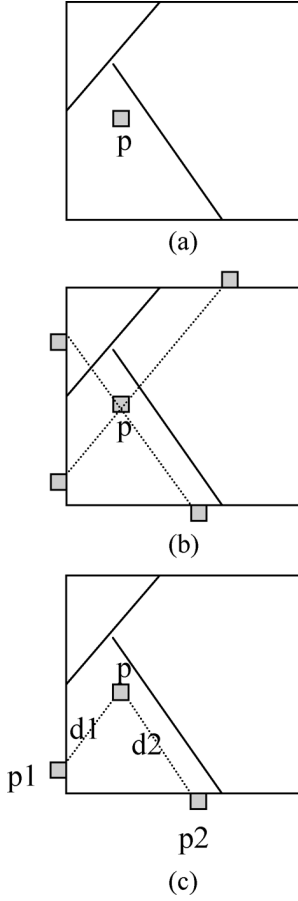
Fig. 2. Selective directional interpolation: $p = (p_1/d_1 + p_2/d_2)/(1/d_1 + 1/d_2)$. (a) A lost MB with two edges linked. (b) Two reference pixels are determined along each edge direction. (c) Select reference pixels within the same region of p.

their attribute distance is the smallest among all. Thus, we will label them as a pair and treat the remaining edge points as a new group. The same matching process is performed iteratively until all points are matched or the attribute distance between two edge points is still above a certain threshold. Finally, each matched pair is linked together to recover a broken edge. After edge linking of all pairs, if there is still some unmatched edge point, it is extended into the lost MB along its gradient until it reaches an edge line.

B. Selective Directional Interpolation

After edges are recovered in a missing MB, the resulting edge lines partition the 2-D plane into several regions. As shown in Fig. 2, pixel p in the missing MB is interpolated using only boundary pixels in the same region to smoothly recover the lost information in that region.

Let us assume that there are n edges in a missing MB. Each edge can be represented by a line equation

$$y - y_i - m_i(x - x_i) = 0, \qquad 1 \le i \le n \qquad (9)$$

where m_i is the edge slope and (x_i, y_i) is the coordinate of an edge point of the ith edge. If this edge is recovered by a matching pair of edge points (x_i, y_i) and (x_j, y_j), $m_i = (y_j - y_i)/(x_j - x_i)$. Otherwise, $m_i = (G_r(x_i, y_i)/G_c(x_i, y_i))$. That is, it is determined by the gradient of the unmatched edge point.

For each lost pixel p, we find its reference pixels to be used in the interpolation process. Along each edge direction, the reference pixels in neighboring MBs are obtained as shown in Fig. 2(b). Note that only those reference pixels within the same region as p are reliable due to discontinuities caused by edges. Thus, sign tests are performed for the line equation of each edge to eliminate unreliable reference pixels. Specifically, let (p_x, p_y) denote the coordinate of the lost pixel p, and (r_x, r_y) the coordinate of a reference pixel. The reference pixel is within the same region as p, if and only if $[p_y - y_i - m_i(p_x - x_i)]$ and $[r_y - y_i - m_i(r_x - x_i)]$ have the same sign for each i.

After eliminating unreliable reference pixels, the missing pixel p can be directionally interpolated via

$$p = \frac{\sum_k \frac{p_k}{d_k}}{\sum_k \frac{1}{d_k}} \qquad (10)$$

where p_k is the kth reliable reference pixel, and d_k is the distance between p_k and p. Fig. 2(c) shows an example when two reference pixels are available. If a lost pixel is enclosed by edges, then no reference pixel is available. In such a case, p is interpolated from the nearest pixels along those edges.

IV. MMSE DECODING FOR P-FRAME CONCEALMENT

In this section, we propose a novel error concealment algorithm based on the minimum mean square error (MMSE) criterion by improving the original scheme presented in [31]. This algorithm attempts to conceal erroneous blocks as well as to suppress the error propagation effect. To be more specific, the decoder adopts an error propagation model to estimate and track the mean square error (MSE) of each reconstructed pixel value. Several modes are developed to conceal erroneous MBs, where each mode has its strength and weakness. The decoder combines these modes adaptively to minimize the MSE of each concealed pixel based on the error propagation model.

A. Error Tracking Model

The error tracking model and the general MMSE decoding procedure are reviewed in this section. For more details, readers are referred to [31]. Two specific concealment modes for P-frames will be described in the next section.

In packet video transmission, erroneous packets are detected and discarded by the channel receiver, and only correctly received packets are passed to the video decoder. Consequently, the decoder knows the error locations but has no information about the error magnitudes. Let us define a pixel error as the difference between its decoded value and error-free reconstruction. It is natural to treat each pixel error as a zero mean random variable with a certain variance. Here, we would like to estimate and track the variance of each pixel error. To achieve this goal, we maintain an extra frame buffer called the error variance map. Each element in the error variance map records the error variance σ_p^2 of the corresponding pixel p in the reconstructed video frame P_k.

Suppose that the decoder reconstructs pixel p in P_k by motion-compensating it from a pixel in the previous frame P_{k-1}.

Then, the pixel error of p is affected only by the propagation error. On the other hand, suppose that the value of p is lost so that p is replaced by a pixel q in P_{k-1} using a temporal concealment method. Then, the pixel error of p is given by the sum of the concealment error and the propagation error [31]. The concealment error is caused by the loss of the motion vector and the DCT-encoded residual, and it is defined as the pixel error when the referenced pixel q is not corrupted. The propagation error is caused by the corruption of the referenced pixel q. It is assumed that the concealment error and the propagation error are independent of each other. Thus, we have

$$\sigma_p^2 = \sigma_{\text{conc}}^2 + \sigma_{\text{prop}}^2 \qquad (11)$$

where σ_{conc}^2 and σ_{prop}^2 denote the variances of the concealment error and the propagation error, respectively. Note that $\sigma_{\text{conc}}^2 = 0$ when the data for p are not lost, and $\sigma_{\text{prop}}^2 = 0$ when the referenced pixel q is not corrupted. The concealment error variance σ_{conc}^2 can be obtained from training sequences using various error patterns.

The propagation error variance σ_{prop}^2 in (11) is calculated based on the accuracy of the motion vector of p. Fig. 3 illustrates the interpolation scheme for the half-pixel motion compensation in H.263 or MPEG-4, where ordinary pixels are depicted by black circles and virtual interpolated pixels are depicted by '+' or '×.' Let us consider three cases according to the accuracy of motion vector $\mathbf{v} = (v_x, v_y)$ as discussed below.

- Both v_x and v_y are of integer-pixel accuracy.
  The current pixel p is predicted from an ordinary pixel $q_\bullet$, specified by motion vector $\mathbf{v}$. Then, the error in $q_\bullet$ propagates to p without attenuation, and the propagation error variance is given by

$$\sigma_{\text{prop}}^2 = \sigma_{q_\bullet}^2. \qquad (12)$$

- v_x is of half-pixel accuracy while v_y is of integer-pixel accuracy (and vice versa).
  The motion vector $\mathbf{v}$ specifies a virtual pixel. For instance, suppose that the current pixel p is predicted from the virtual pixel $q_+ = (q_1 + q_2)/2$ in Fig. 3. Let e_1 and e_2 denote errors in q_1 and q_2, respectively. Then, p is corrupted by $(e_1 + e_2)/2$. Consequently, we have

$$\sigma_{\text{prop}}^2 = \sigma_{q_+}^2 = l_+ \cdot \frac{\sigma_{q_1}^2 + \sigma_{q_2}^2}{2} \qquad (13)$$

where

$$l_+ = \frac{1}{2}\left(1 + 2\frac{E\{e_1 e_2\}}{\sigma_{q_1}^2 + \sigma_{q_2}^2}\right)$$

is called a leaky factor with its value in [0,1].

- Both v_x and v_y are of half pixel accuracy.

The current pixel p is predicted from the virtual pixel $q_\times = (q_1 + q_2 + q_3 + q_4)/4$ as shown in Fig. 3. Let e_i denotes the error of q_i for $i = 1,2,3,4$. Then, we have

$$\sigma_{\text{prop}}^2 = \sigma_{q_\times}^2 = l_\times \cdot \frac{\sigma_{q_1}^2 + \sigma_{q_2}^2 + \sigma_{q_3}^2 + \sigma_{q_4}^2}{4} \qquad (14)$$
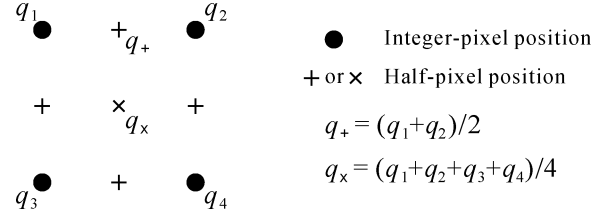

Fig. 3. Interpolation schemes for half-pixel motion compensation.

where

$$l_\times = \frac{1}{4}\left(1 + 2\frac{\sum_{i=1}^{3}\sum_{j=i+1}^{4} E\{e_i e_j\}}{\sum_{i=1}^{4} \sigma_{q_i}^2}\right)$$

is another leaky factor with its value in [0, 1].

Note that due to half-pixel motion compensation, errors attenuate as they propagate. The leaky factors l_+ and $l_\times$ are obtained from training sequences. Typical values of l_+ and $l_\times$ are 0.8 and 0.65, respectively.

To summarize, the propagation error variance σ_{prop}^2 can be calculated from (12)–(14) according to motion vector accuracy. After obtaining σ_{prop}^2, the error variance σ_p^2 of pixel p is updated by $\sigma_{\text{conc}}^2 + \sigma_{\text{prop}}^2$, where σ_{conc}^2 depends on the concealment method for p. As mentioned previously, if the value for p is not lost, we have $\sigma_{\text{conc}}^2 = 0$. In this way, the decoder can estimate and track the error variance of each pixel recursively.

B. MMSE Decoding With Two Concealment Modes

Let us consider multiple concealment methods for a lost pixel simultaneously, where each concealment is called a mode. Based on the error tracking model, we can conceal pixel values by combining several concealment modes. The combination rule is dynamically determined to minimize the MSE of each pixel. Let us describe and analyze the MMSE decoding mechanism in more detail.

Suppose a lost pixel with unknown value p can be concealed by two modes. The first mode replaces the pixel with value $\tilde{p}_1$, and the second mode replaces it with value $\tilde{p}_2$. Then, instead of using one of the two modes directly, p can be concealed by a weighted sum of $\tilde{p}_1$ and $\tilde{p}_2$, given by

$$\tilde{p} = \alpha\tilde{p}_1 + (1 - \alpha)\tilde{p}_2, \qquad (15)$$

where α is a weighting coefficient. Let σ_1^2 and σ_2^2 denote the error variances of $\tilde{p}_1$ and $\tilde{p}_2$, respectively. Then, the error variance σ^2 of $\tilde{p}$ can be written as

$$\begin{aligned}
\sigma^2 &= E\left\{(\tilde{p} - p)^2\right\} \\
&= E\left\{(\alpha(\tilde{p}_1 - p) + (1 - \alpha)(\tilde{p}_2 - p))^2\right\} \\
&= \alpha^2\sigma_1^2 + (1 - \alpha)^2\sigma_2^2 + 2\alpha(1 - \alpha)\rho_{1,2}\sigma_1\sigma_2
\end{aligned}$$

where $\rho_{1,2}$ denotes the correlation coefficient between $(p - \tilde{p}_1)$ and $(p - \tilde{p}_2)$. The optimal α that minimizes σ^2 is given by

$$\alpha = \frac{\sigma_2^2 - \rho_{1,2}\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho_{1,2}\sigma_1\sigma_2} \qquad (16)$$

and the minimum value of σ^2 is given by

$$\sigma_{\min}^2 = \frac{\sigma_1^2\sigma_2^2\left(1 - \rho_{1,2}^2\right)}{\sigma_1^2 + \sigma_2^2 - 2\rho_{1,2}\sigma_1\sigma_2}. \qquad (17)$$

It can be shown that $\sigma_{\min}^2$ has an upper bound, i.e.,

$$\sigma_{\min}^2 \leq \min\left\{\sigma_1^2, \sigma_2^2\right\}.$$

This indicates that the weighted sum $\tilde{p}$ results in a lower error variance than the two modes applied individually, if the decoder selects the optimal weighting coefficient α in (16). However, in real applications, we do not know the accurate values of σ_1^2, σ_2^2 and $\rho_{1,2}$. The weighting coefficient α obtained with inaccurate statistical measurements may result in a huge amount of distortion, especially when $\alpha < 0$ or $\alpha > 1$. Therefore, to be conservative, we impose the following restriction:

$$0 \leq \alpha \leq 1 \qquad (18)$$

such that the absolute error of $\tilde{p}$ is limited by

$$|\tilde{p} - p| \leq \max\left\{|\tilde{p}_1 - p|, |\tilde{p}_2 - p|\right\}.$$

By substituting (16) into (18), we have the following condition:

$$\rho_{1,2} \leq \min\left\{\frac{\sigma_2}{\sigma_1}, \frac{\sigma_1}{\sigma_2}\right\}. \qquad (19)$$

When this condition is satisfied, $\sigma_{\min}^2$ in (17) is an increasing function of $\rho_{1,2}$ since its derivative is nonnegative:

$$\frac{\partial\sigma_{\min}^2}{\partial\rho_{1,2}} = \frac{2\sigma_1^2\sigma_2^2(\sigma_1\rho_{1,2} - \sigma_2)(\sigma_2\rho_{1,2} - \sigma_1)}{(\sigma_1^2 + \sigma_2^2 - 2\rho_{1,2}\sigma_1\sigma_2)^2} \geq 0.$$

This suggests that the smaller the correlation coefficient is, the lower the error variance will be. Note that $\rho_{1,2} = -1$ achieves the minimum value of $\sigma_{\min}^2$ while the maximum occurs when the equality holds in (19). However, $\rho_{1,2}$ is higher than zero in most cases, since any concealment method exploits similar spatial and temporal information. For example, adjacent MBs and previous reconstructed frames are commonly used to conceal the lost MB, even though specific methods may be different. One simple way to lower the correlation coefficient $\rho_{1,2}$ is to select different reference frames in the two concealment modes.

Let us examine the following variance ratio:

$$G_{\text{MMSE}} = \frac{\min\left\{\sigma_1^2, \sigma_2^2\right\}}{\sigma_{\min}^2}. \qquad (20)$$

This can be interpreted as the gain of the weighted MMSE decoding method, compared with the decoding method that chooses the better one between the two concealment modes. By substituting (17) into (20), we have

$$G_{\text{MMSE}} = \frac{1 - \rho_{1,2}^2 + (\delta - \rho_{1,2})^2}{1 - \rho_{1,2}^2} \qquad (21)$$

where

$$\delta = \min\left\{\frac{\sigma_2}{\sigma_1}, \frac{\sigma_1}{\sigma_2}\right\}.$$

It is clear that δ ranges from 0 to 1. Let assume the two concealment modes are selected such that the correlation coefficient $\rho_{1,2}$ is close to 0. Then, the gain in (21) is maximized when $\delta = 1$, that is, when $\sigma_1 = \sigma_2$. This indicates that the error variances of the two concealment modes should be as close as possible to get the best benefit of MMSE decoding.

The MMSE decoding method can be summarized as follows. First, we choose two concealment modes based on the following two criteria.

- They should have a small correlation coefficient $\rho_{1,2}$.
- They should provide similar concealment capabilities, i.e., $\sigma_1 \approx \sigma_2$.

The parameters σ_1, σ_2 and $\rho_{1,2}$ are obtained by training in advance. In the decoder, each pixel is reconstructed via (15) and (16). Then, the corresponding element in the error variance map is updated by (17). During the reconstruction and the map updating, if $\alpha \geq 1$, α is set to 1 and the error variance is updated to σ_1^2 to satisfy the constraint in (18). Similarly, if $\alpha \leq 0$, α is set to 0 and the error variance is updated to σ_2^2.

V. P-FRAME CONCEALMENT MODES

Based on the above discussion, the proposed algorithm employs two temporal concealment modes in the decoder: 1) temporal linear interpolation and 2) motion vector recovery with block matching. Let us describe these two modes in detail below.

A. Temporal Linear Interpolation

Linear interpolation is often used for error concealment. As in (1), four pixel values in spatially adjacent MBs can be linearly interpolated to conceal an erroneous pixel. On the other hand, in this work, four pixel values in the previous frame is linearly interpolated to conceal a pixel temporally. We employ the temporal interpolation rather than the spatial interpolation, since temporal correlation is much higher than spatial correlation in general.
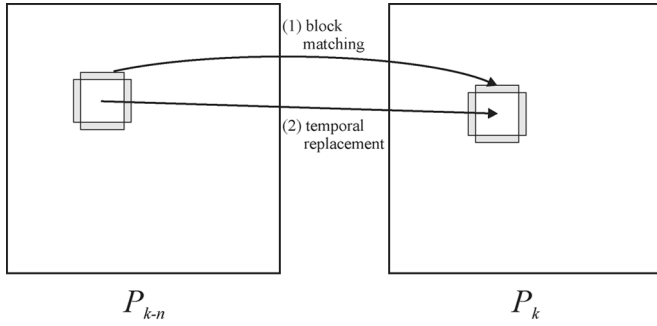
Fig. 4.   Motion vector recovery with block matching.

For each pixel p in a missing MB, four reference pixels are obtained using the motion vectors of the upper, lower, left and right MBs. They are denoted by p_{upper}, p_{lower}, p_{left} and p_{right}. To conceal the pixel p, the four reference pixel values are averaged using the weighting coefficients, which are inversely proportional to the distances between p and the adjacent MBs. Specifically, assume that p is the (x, y)th pixel in the missing MB, where $1 \leq x, y \leq 16$. Then, it is concealed via

$$\tilde{p} = \frac{p_{\text{upper}} \cdot (17 - y) + p_{\text{lower}} \cdot y + p_{\text{left}} \cdot (17 - x) + p_{\text{right}} \cdot x}{34}.$$
(22)

If a neighboring motion vector is not available due to packet loss, the boundary effect or an intra-coded block, only those available motion vectors are used for concealment. If all motion vectors are not available, the erroneous MB is copied from the previous frame with the zero motion vector.

B. Motion Vector Recovery With Block Matching

As mentioned in Section II-B, there are several approaches to recover the motion vector of an erroneous MB. We adopt the block matching approach [21]–[23], which finds the motion vector for the set of surrounding pixels and uses that vector for the erroneous block. Fig. 4 illustrates the idea of motion vector recovery with block matching. First, the decoder estimates the motion vector for the error-free surrounding pixels, which are adjacent to the erroneous block. In this work, the motion vector is searched from the previous frame P_{k-1} or the earlier frame P_{k-2}, and the sum of square differences (SSD) is used as the block matching criterion. Then, the erroneous block is temporally replaced using the retrieved motion vector.

Since MBs are decoded in a raster scan order, when reconstructing a MB, its right and lower adjacent MBs are not decoded yet. Thus, to simplify the decoding procedure, the matching of four sides can be reduced to that of two sides which include only upper and left surrounding pixels. If one side of the surrounding pixels is not error-free, then it is ignored when calculating the SSD. If all surrounding pixels are erroneous, the motion vector is simply set to the zero vector.

To reduce the computational complexity of the block matching, the search area for the motion vector is reduced by exploiting the spatio-temporal correlation between adjacent motion vectors. Let (a_i, b_i), $1 \leq i \leq 4$, denote the motion

vectors of the four adjacent MBs, respectively. Then, the search area from the previous frame P_{k-1} is restricted to

$$\min\{a_i\} \leq a \leq \max\{a_i\}, \quad \min\{b_i\} \leq b \leq \max\{b_i\}$$

where (a, b) denotes the motion vector of the erroneous MB. Also, the search area from the previous previous frame P_{k-2} is restricted to

$$\min\{a_i\} + c \leq a \leq \max\{a_i\} + c$$
$$\min\{b_i\} + d \leq b \leq \max\{b_i\} + d$$

where (c, d) is the motion vector of the MB in P_{k-1}, which is at the same spatial location as the current erroneous MB. In this way, the decoder can reduce the computations for block matching significantly at the cost of slight performance degradation.

VI. SUMMARY OF DECODER IMPLEMENTATION

To reconstruct or conceal frame P_k, the proposed algorithm uses the information from frames P_{k-1} and P_{k-2}. Thus, the decoder should maintain three video frame buffers. Also, the decoder requires additional three frame buffers to record the corresponding error variance maps. Therefore, the decoder needs six frame buffers in total.

Let us first consider the decoding of I-frames. If an MB is error-free, it is reconstructed and the corresponding variances in the error variance map are set to 0. On the other hand, if an MB is erroneous, it is concealed by the directional interpolation in Section III and the error variances are set to the highest value 255.

Next, let us consider the decoding of P-frames. The MMSE weighting method is applied to conceal erroneous MBs using the two concealment modes.
- *Mode 1)* Temporal linear interpolation from frame P_{k-1}.
- *Mode 2)* Motion vector recovery with block matching from frame P_{k-1}.

In rare cases, when a scene contains fast motions or occlusions, it is more efficient to use the spatial concealment than the temporal concealment. Therefore, in our implementation, if an erroneous block is adjacent to more than two intra-coded MBs, it is concealed by the directional interpolation. The MMSE weighting method is used to reconstruct error-free MBs also using the following two modes.
- *Mode 3)* Conventional reconstruction using frame P_{k-1}.
- *Mode 4)* Motion vector recovery with block matching from frame P_{k-2}.

Note that, in a P-frame, even though the pixel values of an MB are received correctly, the MB can still be severely corrupted by the error propagated from frame P_{k-1}. In such a case, mode 4 may provide better reconstruction by concealing the MB using the information in P_{k-2}.

Fig. 5 shows the decoding flowchart for an MB in the P-frame. With the exception of intra-concealed MBs, the proposed algorithm conceals an erroneous MB or reconstructs an error-free MB by combining two modes via (15) and (16), and then updates the error variance map via (17). Table I
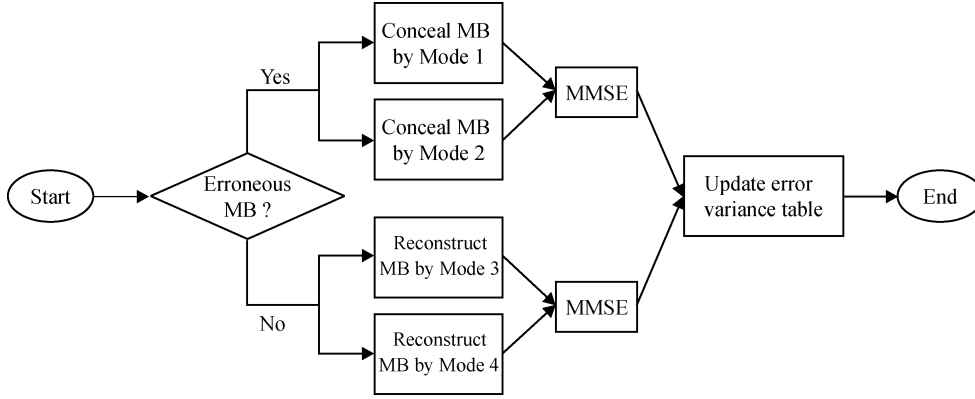
Fig. 5. MMSE decoding of an MB in a P-frame.

TABLE I
PARAMETERS FOR MMSE DECODING, WHERE THE CONCEALMENT ERROR VARIANCES ARE NORMALIZED
WITH RESPECT TO THE INTRA CONCEALMENT ERROR VARIANCE 255

	Parameters	Comments
Erroneous MBs	$\sigma^2_{\text{conc},1} = 38$	Concealment error variance for mode 1
	$\sigma^2_{\text{conc},2} = 45$	Concealment error variance for mode 2
	$\rho_{1,2} = 0.47$	Correlation coefficient between mode 1 and mode 2
Error-free MBs	$\sigma^2_{\text{conc},3} = 0$	No concealment error variance in mode 3
	$\sigma^2_{\text{conc},4} = 80$	Concealment error variance for mode 4
	$\rho_{3,4} = 0.4$	Correlation coefficient between mode 3 and mode 4

summarizes the parameters for P-frame decoding, in which the concealment error variances are normalized with respect to the intra concealment error variance 255. It is worthy to point out that the two concealment modes for erroneous MBs are designed to satisfy the criteria in Section IV-B. They have similar error variances and their correlation coefficient is relatively small.

VII. SIMULATION RESULTS

A. Experimental Setup

The performance of the proposed algorithm is evaluated using the standard H.263 coder [32]. In H.263, the group of blocks (GOB) is defined as a number of MB rows that are dependent on the picture resolution. For example, a GOB consists of a single MB row at the QCIF (176 × 144) resolution. In many cases, each GOB is packetized into one packet. However, in the GOB packetization, if a packet is lost, the information in the left and right MBs cannot be used for the concealment of an MB. To improve the concealment performance, we also implement the interleaving packetization by modifying the syntax of H.263. As shown in Fig. 6, an interleaving packet for a QCIF frame is formed with 11 MBs chosen from every nine consecutive MBs. For instance, the first packet consists of the $(9i + 1)$th MBs, where $0 \leq i < 11$. Thus, as in the GOB packetization, the interleaving packetization also generates nine packets for each frame. However, when one packet is missing, an erroneous MB can be concealed more effectively using the information in the

Fig. 6. Interleaving packetization of a frame at the QCIF resolution.

upper, lower, left and right MBs. As compared with the GOB packetization, the interleaving packetization increases the bit rate slightly. The overhead is less than 5%.

In the following simulations with the proposed algorithm as well as other algorithms for benchmarking, the 16-bit cyclic redundancy check (CRC) [5] code is appended to each packet. Although the CRC requires a small overhead (2 bytes), it can detect most of errors and can be easily implemented. In addition, the 2-byte overhead may be absorbed, when video packets are transmitted using the user datagram protocol (UDP) and the checksum in the UDP header is enabled. The packets, which are declared corrupted by the CRC decoder, are not used in the video decoder.
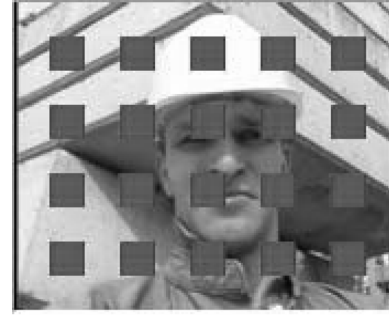
(a)

(b)

(c)

Fig. 7. I-frame concealment results when 3 MBs are lost. (a) Error locations (18.95 dB). (b) Zeng and Liu's algorithm (35.37 dB). (c) Proposed algorithm (39.56 dB).

(a)

(b)

(c)

Fig. 8. I-frame concealment results when 20 MBs are lost. (a) Error locations (10.91 dB). (b) Zeng and Liu's algorithm (30.66 dB). (c) Proposed algorithm (30.75 dB).

B. Error Concealment of I-Frames

Fig. 7 compares the performance of the proposed I-frame concealment method with that of the Zeng and Liu algorithm in [18]. The test image is the first frame of the "Foreman" QCIF sequence, which has 43.20-dB PSNR with error-free reconstruction. The quantization parameter (QP) is set to 2. Three MBs, containing object edges, are lost in this simulation. The top MB contains three parallel edges, and the bottom MB has a single dominant edge. These simple edges are detected and faithfully concealed by both the proposed algorithm and Zeng and Liu's algorithm. The middle MB contains intersecting edges. One of the edges meets another edge within the MB and does not exit the MB. In Zeng and Liu's algorithm, this edge direction is not detected and the resulting interpolation yields a false edge in the concealed MB. On the other hand, the proposed algorithm successfully detects the edge direction and provides a better reconstructed result.

Fig. 8 shows the results when 20 MBs are lost. As in the previous test, QP is set to 2. In Zeng and Liu's algorithm, the directions of a few edges are incorrectly estimated and the concealed image contains blurring artifacts especially around the face. We see that the proposed algorithm reconstructs the edges more accurately and provides better image quality.

The computational complexity of the proposed algorithm is comparable with that of Zeng and Liu's algorithm. The main difference is that the proposed algorithm uses the Sobel operator for the edge detection, while Zeng and Liu's algorithm performs the sorting of pixel values and the two-level quantization. These two approaches require a similar amount of computations.

C. Error Concealment of P-Frames

We evaluate the performance of the proposed MMSE decoding for P-frame concealment. First, three consecutive frames (154th, 155th, and 156th frames) in the "Foreman" CIF (352×288) sequence are encoded with $QP = 11$. Fig. 9(a) shows the error pattern. That is, 99 interleaved MBs are lost from the 156th

Fig. 9.   Performance comparison of P-frame concealment results. (a) Error pattern (9.76 dB). (b) Zhang *et al.*'s algorithm (29.36 dB). (c) Lee *et al.*'s spatial method (29.48 dB). (d) Lee *et al.*'s temporal method (30.20 dB). (e) Lee *et al.*'s combined method (31.61 dB). (f) Proposed algorithm (32.23 dB).

frame, while the preceding 154th and 155th frames are correctly reconstructed without any error. Since the 156th frame contains fast motions, its concealment is relatively difficult. For comparison, we test the performance of Zhang *et al.*'s algorithm in [21], which recovers the motion vector of an erroneous block via block matching of surrounding pixels. Its recovery performance depends on the number of surrounding pixel layers. Table II lists the PSNRs of the reconstructed frames according to the number of layers. The search region for the motion vector is set to $[-16, 16] \times [-16, 16]$. We see that the best PSNR (29.36 dB) is achieved when 11 layers are used for block matching. Fig. 9(b) shows the reconstructed frame in that case. Note that Zhang *et al.*'s algorithm is modified and used as the concealment mode 2 in the proposed MMSE decoding. Specifically, the set of candidate motion vectors is restricted as described in Section V-B. Table II also provides the PSNR values when the restricted search scheme is used. The best PSNR (29.32 dB) is achieved when six layers are used. However, the PSNR performance is less sensitive to the number of layers in the restricted

search, and even the single layer matching provides a relatively high PSNR value. In the following tests, the concealment mode 2 fixes the number of layers to 1 to reduce the computational complexity of block matching.

We also provide the results of Lee *et al.*'s algorithm in [30], which has the spatial average, temporal average and combined average modes. Fig. 9(c)–(e) is concealed by the spatial, temporal and combined modes of Lee *et al.*'s algorithm, respectively. Among them, the combined mode provides the best quality by mixing numerous spatial and temporal concealment candidates. The proposed algorithm provides the reconstruction in Fig. 9(f), which achieves a higher PSNR value and better visual quality than Zhang *et al.*'s algorithm and all three modes of Lee *et al.*'s algorithm.

Next, the 60th–160th frames of the "Foreman" and "News" CIF sequences are encoded with $QP = 11$ and at the frame rate of 10 frames/s. To investigate the error propagation effect, only the first frame is encoded in the I-frame mode and the other frames are encoded in the P-frame mode. The first I-frame is
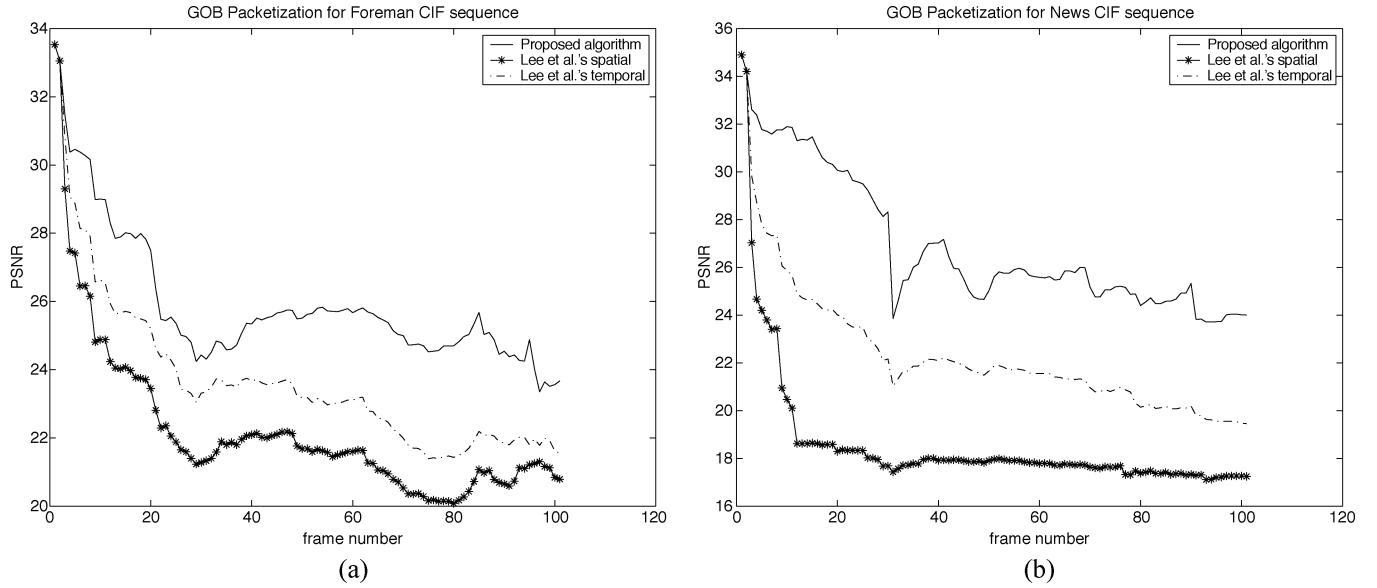
Fig. 10. PSNR performances of P-frame concealment algorithms with the GOB packetization, where the packet loss rate is 10%. (a) Foreman. (b) News.

TABLE II
PSNR PERFORMANCES OF THE MOTION VECTOR RECOVERY SCHEME IN [21] ACCORDING TO THE NUMBER OF SURROUNDING PIXEL LAYERS. THE "FOREMAN" 156TH FRAME IS CORRUPTED BY THE ERROR PATTERN IN FIG. 9(a). IN THE FULL SEARCH, ALL CANDIDATE VECTORS ARE EXAMINED. IN THE RESTRICTED SEARCH, THE CANDIDATE SET IS REDUCED AS DESCRIBED IN SECTION V-B

No. of Layers	Full Search	Restricted Search
1	28.86	29.16
2	28.95	29.28
3	28.88	29.30
4	28.82	29.29
5	28.89	29.28
6	28.87	29.32
7	28.87	29.31
8	28.92	29.30
9	28.94	29.28
10	28.92	29.99
11	29.36	28.99
12	29.08	28.99
13	29.04	28.94
14	29.00	28.96
15	28.98	28.96
16	28.76	29.22

assumed to be error-free, and the packets for the P-frames are dropped randomly with a packet loss rate of 10%.

Fig. 10 compares the PSNR performances, when the GOB packetization is used. The bit rate for the "Foreman" sequence is 111 kbps, and that for the "News" sequence is 65 kbps. In a typical image sequence, the spatial correlation is much lower than the temporal correlation. Thus, in Lee *et al.*'s algorithm, the spatial mode often introduces blurring artifacts, providing significantly worse performance than the temporal mode. Conse-

quently, the combined mode does not provide a meaningful gain by mixing spatial candidates with temporal candidates. Fig. 9(e) is exceptional, since the spatial mode conceals the blocks around the fast moving hand more effectively than the temporal mode. Our simulation results confirmed that the combined mode provides worse performance than the temporal mode on the average, and we show the performance of the spatial and the temporal modes in the following simulations only. From Fig. 10, we see that the proposed algorithm is superior to both the spatial and the temporal modes of Lee *et al.*'s algorithm. Moreover, the performance difference becomes bigger, as the frame number increases. This is because the proposed algorithm adapts the reconstruction of error-free blocks as well as the concealment of erroneous blocks to suppress error propagation.

Fig. 11 shows the PSNR performances, when the interleaving packetization is used. The bit rate for the "Foreman" sequence is 113 kbps, and that for the "News" sequence is 66 kbps. The interleaving packetization allows more information to be exploited for the concealment than the GOB packetization. Thus, each method provides better PSNR performance as compared with its counterpart in Fig. 10.

Let us consider the computational complexity of the proposed P-frame concealment algorithm. In the concealment mode 1, we perform 4 multiplications per pixel (mpp) and 3 additions per pixel (app) to obtain $\tilde{p}$ in (22). In the concealment mode 2, the motion vector of an erroneous block is recovered via block matching of surrounding pixels. We use a single layer of surrounding pixels to compute the absolute sum of differences (SAD). As described in Section V-B, we reduce the search area using neighboring motion vectors. After the reduction, we check about 100 motion vectors on the average. The MMSE decoding then combines the two modes and update error variances via (15)–(17). This requires 6 mpp, 3 app, and 1 square root operation per pixel. Therefore, in total, the proposed algorithm requires 10 mpp, 6 app, 1 square root operation per pixel, and about 100 SAD operations per block.
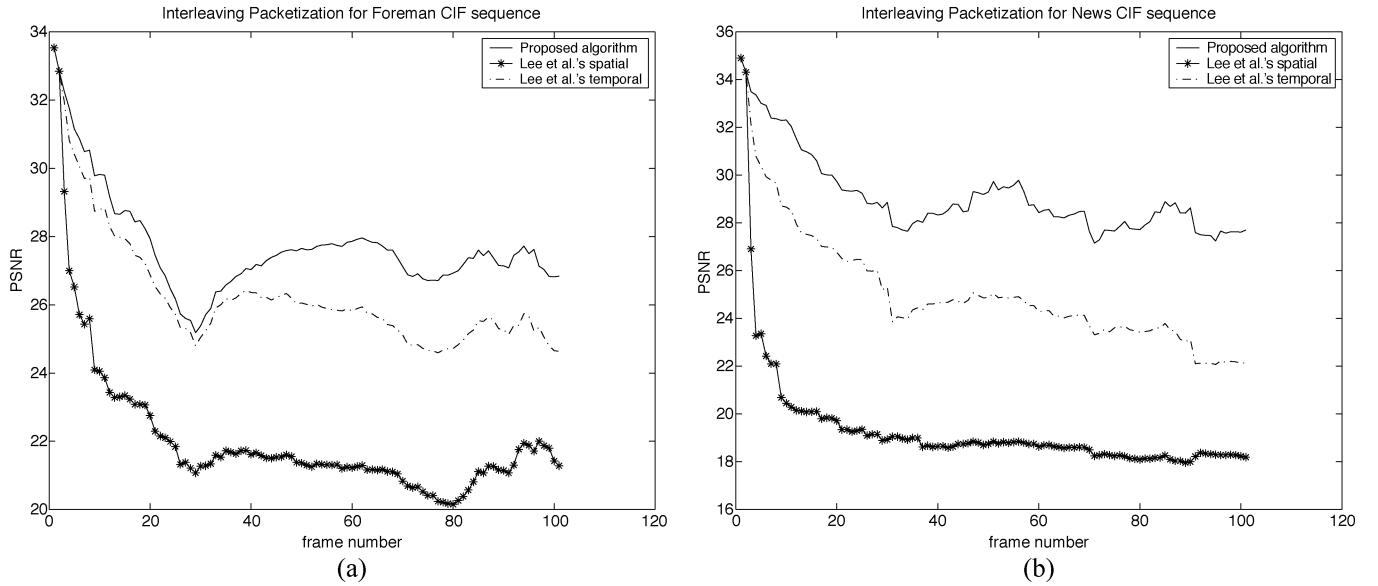
Fig. 11.   PSNR performances of P-frame concealment algorithms with the interleaving packetization, where the packet loss rate is 10%. (a) Foreman. (b) News.
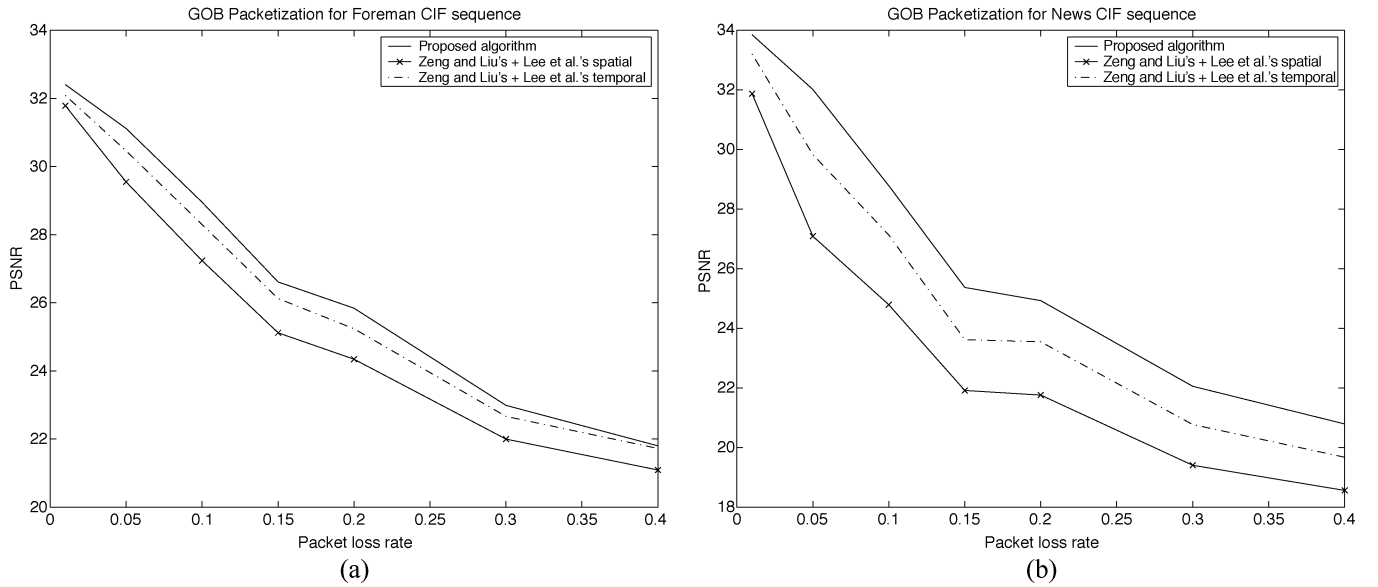

Fig. 12.   PSNR performances in terms of the packet loss rate with the GOB packetization. (a) Foreman. (b) News.

On the other hand, Lee *et al.*'s temporal mode computes the squared sum of differences (SSD) for each motion vector. In their algorithm, the search region contains 225 candidate motion vectors. Then, all 225 prediction blocks are linearly combined with weighting coefficients to recover the erroneous block. The weighting coefficients are computed based on the SSDs using a complex equation. Even though we exclude these computations for weighting coefficients, the temporal mode requires 225 mpp, 224 app, and 225 SSD operations per block. This means that the proposed algorithm requires a lower computational complexity than Lee *et al.*'s temporal mode, while providing better image quality.

D.   Error Concealment of I- and P-Frames

In this test, an I-frame is inserted at the start of every ten frames, and random packet losses occur in both I- and P frames.

The "Foreman" and "News" sequences are encoded with $QP = 11$ and a frame rate of 10 frames/s. For the "Foreman" and "News" sequences, the GOB packetization yields a bit rate of 154 and 108.8 kbps and the interleaving packetization a bit rate of 155 and 108.3 kbps, respectively. Figs. 12 and 13 show the PSNR performance as a function of the packet loss rate with the GOB and the interleaving packetization schemes, respectively. For each packet loss rate, twenty error patterns are simulated and the obtained PSNR's are averaged over all patterns and all frames.

The proposed algorithm uses the directional interpolation and the MMSE decoding to conceal I-frames and P-frames, respectively. For comparison, Zeng and Liu's algorithm and Lee *et al.*'s algorithm are used for the concealment of I-frames and P-frames, respectively. As compared with the better combination of the benchmarking algorithms, the proposed algorithm
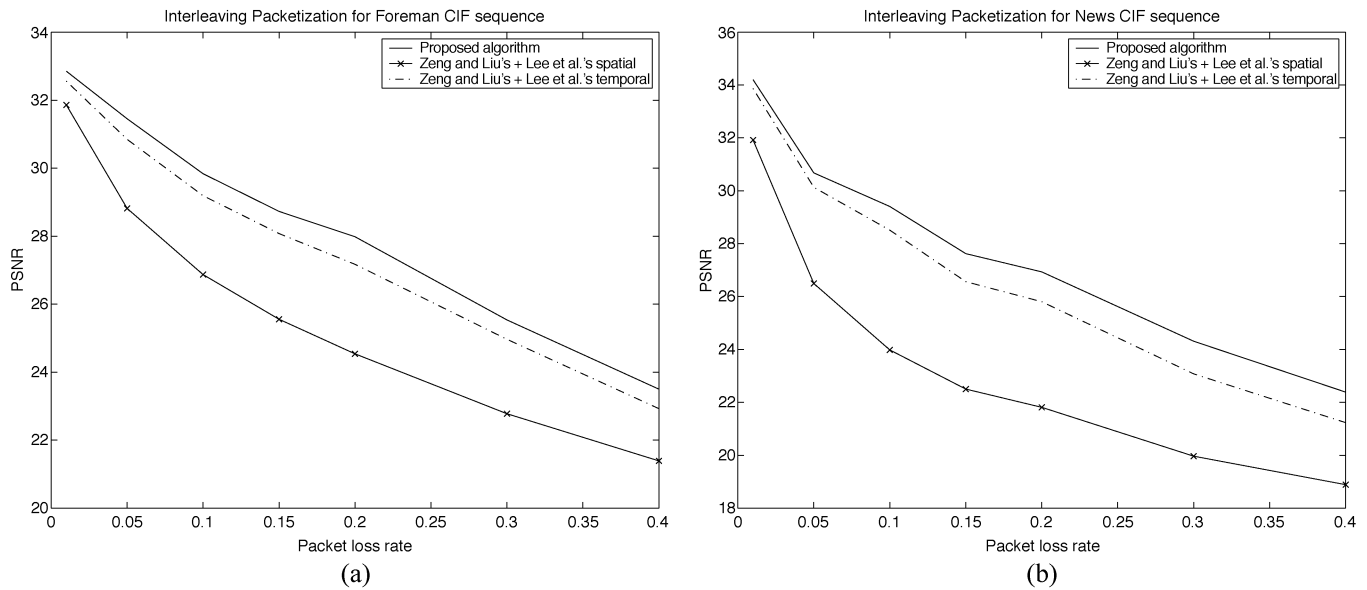
Fig. 13. PSNR performances in terms of the packet loss rate with the interleaving packetization. (a) Foreman. (b) News.

provides up to 1.0 dB PSNR gain. In low bit rate applications, I-frames are inserted less frequently. In such a case, the proposed algorithm provides an even bigger advantage due to the effective suppression of error propagation.

These simulation results indicate that the proposed algorithm offers a promising technique for robust video transmission. Moreover, the proposed algorithm requires neither a feedback channel nor extra delay. Since it only applies to the decoder. It can be easily modified to be compatible with any video coding standards.

VIII. CONCLUSION

In this work, we proposed novel I-frame and P-frame error concealment methods. The I-frame error concealment method employs edge detection and directional interpolation to recover both smooth and edge areas efficiently. The P-frame error concealment method uses error tracking and dynamic mode weighting. It conceals a pixel as a weighted sum of candidate pixels that are reconstructed using different concealment modes. The weighting coefficients are dynamically determined to reduce the propagation error and the concealment error. It was shown with simulation results that the proposed methods provide significantly better performance in error-prone environments than the conventional concealment methods.

REFERENCES

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

[2] V. BeBrunner, L. DeBrunner, L. Wang, and S. Radhakrishnan, "Error control and concealment for image transmission," *IEEE Commun. Surveys & Tuts.*, vol. 3, no. 1, 2000.

[3] B. Girod and N. Farber, "Feedback-based error control for mobile video transmission," *Proc. IEEE*, vol. 87, no. 10, pp. 1707–1723, Oct. 1999.

[4] J. Hagenauer and T. Stockhammer, "Channel coding and transmission aspects for wireless multimedia," *Proc. IEEE*, vol. 87, no. 10, pp. 1764–1777, Oct. 1999.

[5] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[6] V. K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 74–93, Sept. 2001.

[7] C.-S. Kim and S.-U. Lee, "Multiple description coding of motion fields for robust video transmission," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 9, pp. 999–1010, Sep. 2001.

[8] B. W. Wah, X. Su, and D. Lin, "A survey of error-concealment schemes for real-time audio and video transmissions over the internet," in *Proc. Int. Symp. Multimedia Software Engineering*, Dec. 2000, pp. 17–24.

[9] P. Cuenca, L. Orozco-Barbosa, A. Garrido, F. Quiles, and T. Olivares, "A survey of error concealment schemes for MPEG-2 video communications over ATM networks," in *Proc. IEEE 1997 Can. Conf. Elect. and Comput, Eng.*, May 1997, vol. 1, pp. 25–28.

[10] A. Raman and M. Babu, "A low complexity error concealment scheme for MPEG-4 coded video sequences," in *Proc. Tenth Annu. Symp. Multimedia Commun. and Signal Process.*, Bangalore, India, Nov. 2001.

[11] 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. Amsterdam, The Netherlands: Kluwer Academic, Jun. 1998, pp. 212–216.

[12] Y. Wang, Q.-F. Zhu, and L. Shaw, "Maximally smooth image recovery in transform coding," *IEEE Trans. Commun.*, vol. 41, no. 10, pp. 1544–1551, Oct. 1993.

[13] Y. Wang and Q.-F. Zhu, "Signal loss recovery in DCT-based image and video codecs," in *Proc. SPIE VCIP-91*, Boston, MA, Nov. 1991, pp. 667–678.

[14] X. Lee, Y.-Q. Zhang, and A. Leon-Garcia, "Image and video reconstruction using fuzzy logic," in *Proc. IEEE Global Telecommunications Conf.*, Dec. 1993, vol. 2, pp. 975–979.

[15] ——, "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.

[16] 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.

[17] 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.

[18] W. Zeng and B. Liu, "Geometric-structure-based error concealment with novel applications in block-based low-bit-rate coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 4, pp. 648–665, Jun. 1999.

[19] 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.

[20] M.-J. Chen, L.-G. Chen, and R.-M. Weng, "Error concealment of lost motion vectors with overlapped motion compensation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 3, pp. 560–563, Jun. 1997.

[21] J. Zhang, J. F. Arnold, and M. R. Frater, "A cell-loss concealment technique for MPEG-2 coded video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 4, pp. 659–665, Jun. 2000.

[22] S. Tsekeridou and I. Pitas, "MPEG-2 error concealment based on block-matching principles," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 4, pp. 646–658, Jun. 2000.

[23] 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 Process. Syst.*, Sept. 2001, pp. 247–258.

[24] S.-H. Lee, D.-H. Choi, and C.-S. Hwang, "Error concealment using affine transform for H.263 coded video transmissions," *Electron. Lett.*, vol. 37, no. 4, pp. 218–220, Feb. 2001.

[25] Q.-F. Zhu, Y. Wang, and L. Shaw, "Coding and cell-loss recovery in DCT-based packet video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 3, pp. 248–258, Jun. 1993.

[26] P. Salama, N. B. Shroff, and E. J. Delp, "Statistical spatial approach: MAP estimation," in *Signal Recovery Techniques for Image and Video Compression and Transmission*, A. Katsaggelos and N. Galatsanos, Eds. Amsterdam, The Netherlands: Kluwer Academic Publishers, Jun. 1998, pp. 217–219.

[27] S. Shirani, F. Kossentini, and R. Ward, "A concealment method for video communications in an error-prone environment," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 1122–1128, Jun. 2000.

[28] P. Salama, N. B. Shroff, and E. J. Delp, "Error concealment in MPEG video streams over ATM networks," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 1129–1144, Jun. 2000.

[29] D. S. Turaga and T. Chen, "Model-based error concealment for wireless video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 483–495, Jun. 2002.

[30] W. S. Lee, M. R. Frater, M. R. Pickering, and J. E. Arnold, "Spatial temporal concealment of lost blocks in coded video," in *Proc. ICIP*, Oct. 1998, vol. 3, pp. 477–481.

[31] C.-S. Kim, J.-W. Kim, I. Katsavounidis, and C.-C. J. Kuo, "Robust MMSE video decoding: theory and practical implementations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 39–51, Jan. 2005.

[32] *Video Coding for Low Bitrate Communication*, ITU-T Recommendation H.263, 1998.

Wei-Ying Kung received the B.S. degree from the National Taiwan University, Taipei, R.O.C., in 1996, and the M.S., Ph.D. degrees from the University of Southern California, Los Angeles, in 2000, 2004, all in electrical engineering. Her research interests include video/image compression, coding, processing and communication, multimedia communication, wireless communications, and error resilient coding. She is the author or coauthor of more than 20 technical papers. She is currently with Motorola Advanced Technology, San Diego, CA, working on video compression/processing.

Chang-Su Kim (S'95–M'01–SM'05) received the B.S. and M.S. degrees in control and instrumentation engineering from Seoul national University (SNU) in 1994 and 1996, respectively. In 2000, he received the Ph.D. degree in electrical engineering from SNU with a Distinguished Dissertation Award.

From 2000 to 2001, he was a Visiting Scholar with the Signal and Image Processing Institute, University of Southern California, Los Angeles, and a Consultant for InterVideo Inc., Los Angeles. From 2001 to 2003, he coordinated the 3D Data Compression Group in National Research Laboratory for 3D Visual Information Processing in SNU. From 2003 and 2005, he was an Assistant Professor in the Department of Information Engineering, Chinese University of Hong Kong. In Sept. 2005, he joined the Department of Electronics Engineering, Korea University as an Assistant Professor. His research topics include video and 3D graphics processing and multimedia communications. He has published more than 90 technical papers in international conferences and journals.

C.-C. Jay Kuo (S'83–M'86–SM'92–F'99) received the B.S. degree from the National Taiwan University, Taipei, in 1980 and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1985 and 1987, respectively, all in Electrical Engineering. He is Director of the Signal and Image Processing Institute (SIPI) and Professor of Electrical Engineering, Computer Science and Mathematics at the University of Southern California (USC). His research interests are in the areas of digital image/video analysis and modeling, multimedia data compression, communication and networking and multimedia database management. Dr. Kuo has guided about 70 students to their Ph.D. degrees and supervised 15 postdoctoral research fellows. He is a co-author of about 120 journal papers, 650 conference papers and 7 books.

Dr. Kuo is a Fellow of IEEE and SPIE. He is Editor-in-Chief for the Journal of Visual Communication and Image Representation, and Editor for the Journal of Information Science and Engineering, LNCS Transactions on Data Hiding and Multimedia Security and the EURASIP Journal of Applied Signal Processing. Dr. Kuo received the National Science Foundation Young Investigator Award (NYI) and Presidential Faculty Fellow (PFF) Award in 1992 and 1993, respectively.