Content-Based Adaptive Spatio-Temporal Methods for MPEG Repair

Jiho Park, *Member, IEEE*, Dong-Chul Park, *Senior Member, IEEE*, Robert J. Marks, II, *Fellow, IEEE*, and Mohamed A. El-Sharkawi, *Fellow, IEEE*

Abstract—Block loss and propagation error due to cell loss or missing packet information during the transmission over lossy networks can cause severe degradation of block and predictive-based video coding. Herein, new fast spatial and temporal methods are presented for block loss recovery. In the spatial algorithm, missing block recovery and edge extention are performed by pixel replacement based on range constraints imposed by surrounding neighborhood edge information and structure. In the temporal algorithm, an adaptive temporal correlation method is proposed for motion vector (MV) recovery. Parameters for the temporal correlation measurement are adaptively changed in accordance to surrounding edge information of a missing macroblock (MB). The temporal technique utilizes pixels in the reference frame as well as surrounding pixels of the lost block. Spatial motion compensation is applied after MV recovery when the reference frame does not have sufficient information for lost MB restoration. Simulations demonstrate that the proposed algorithms recover image information reliably using both spatial and temporal restoration. We compare the proposed algorithm with other procedures with consistently favorable results.

Index Terms—Block loss recovery, H.263, motion vector (MV) recovery, MPEG, spatial and temporal error concealment, video communication.

I. INTRODUCTION

MANY VIDEO coding standards are based on block and prediction coding techniques [1]–[3]. MPEG and H.263 are examples. When coded data are transmitted and lost due to channel error or other reasons, corresponding coding blocks and the frames following are degraded. Several methods have been proposed to reduce the degradation including *automatic retransmission request* (ARQ) and block recovery/error concealment. Herein, we consider the latter. Block recovery refers to any technique wherein missing blocks in the decoder are recovered without retransmission. Such restoration techniques can be categorized into two types: temporal and spatial. Temporal block recovery techniques use information in temporally adjacent frames while spatial error concealment procedures exploit the surrounding spatial neighborhood of a missing *mac-*

roblock (MB). Some techniques utilize both temporal and spatial information [4].

The stream in standard video coding includes motion vectors (MVs) and DCT coefficients of coding blocks. Image degradation and error propagation resulting in lost MVs require MV recovery procedures. Several approaches have been proposed for MV recovery. Haskell and Messerschmitt [5] suggest setting lost MVs to zero or utilizing MV information of surrounding blocks. Al-Mualla *et al.* [6] suggest a motion field interpolation technique. In this approach, the MV of each pixel in a missing MB is computed using bilinear interpolation of MVs of surrounding blocks. Lam *et al.* [7] propose a *boundary matching algorithm* (BMA) which finds the MV minimizing the difference between boundary pixels of the lost MB in the current frame and inner pixels of an MB in the previous frame. A recovery technique of lost MVs with overlapped motion compensation is suggested by Chen *et al.* [8]. Here, a lost MV is first restored by a side match criterion. The restored blocks are then subdivided into four subblocks, and, finally, pixels in each subblock are computed by a weighted average pixel value of the originally restored and neighboring blocks. Other methods exploiting the temporal correlation between frames also have been reported. Zhang *et al.* [9] propose a decoder motion-vector estimation algorithm incorporating temporal correlation between surrounding pixels of a lost MB and candidate blocks in the reference frame. Tsekeridou and Pitas [10] suggest MV recovery by a block matching method. Distances are computed between the above and/or below MBs of a lost MB and those of candidate MBs in the reference frame. The best matching candidate is then selected.

In spatial error concealment, numerous methods have been developed to restore missing blocks. A maximally smooth recovery approach is presented by Wang *et al.* [11]. It produces a maximally smooth recovery image with boundary pixels by imposing smooth constraints on the surrounding and restored pixels. To get reliable results in diagonal edges, Shirani *et al.* [12] suggest an image reconstruction algorithm for restoring a missing block by minimizing spatial differences among four adjacent blocks using linear least squares and linear interpolation. A recovery method wherein high frequency DCT coefficients of a missing block and adjacent pixels are set to zero is reported by Alkachouh and Bellanger [13]. Some alternating projections-based algorithms have been proposed to sustain edge structure in surrounding neighborhoods by exploiting neighboring edge information and corresponding convex sets. Sun and Kwok [14] suggest use of a spatial interpolation algorithm using *projections onto convex sets* (POCS) [15]. Park *et al.* [16] propose a spatial

block loss recovery algorithm using the method of alternating projections.

Temporal MV recovery and fast spatial block loss restoration algorithms using edge information is presented in this paper. In spatial restoration, pixels of a missing block are recovered using the pixel values in neighborhoods. The algorithm morphs surrounding edge structure into missing blocks. In the temporal case, we suggest use of adaptive temporal correlation for MV recovery. Adaptive parameter alteration as a function of surrounding information is introduced. The parameters are determined adaptively by the edge information of neighborhood blocks. The adaptive matching parameters enable the MV recovery to be more robust and to utilize surrounding information. Pixels both in the reference frame and current frame are used to find the MB having the minimum difference in the reference frame. A simple range constraint between restored pixels is applied after MV recovery when the reference frame does not have sufficient information for lost MB recovery.

This paper is organized as follows. In Section II, the proposed spatial block loss recovery algorithm is explained. A content-based adaptive temporal block recovery technique is suggested in Section III. The performance and experimental results of the proposed algorithms with other existing procedures are presented in Section IV. Conclusions are drawn in Section V.

II. SPATIAL BLOCK LOSS RECOVERY

A. Glossary for Section II

$\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$	$N \times N$ surrounding blocks of a missing block.
$\mathbf{M}$	$N \times N$ missing block of pixels.
$\{\mathbf{r}_i \mid i = 1, 2\}$	Recovery vectors. $N \times N$.
$\{\mathbf{s}_j \mid 1 \leq j \leq 8N\}$	Surrouding vectors. $N \times N$.
$\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$	Best-matching surrounding block.
$\mathbf{T}$	Two-dimensional (2-D) DCT operator.
$\mathbf{T}^{-1}$	2-D IDCT operator.
$\mathbf{R}_i$	$\mathbf{T} \cdot \mathbf{r}_i$.
$\mathbf{S}_j$	$\mathbf{T} \cdot \mathbf{s}_j$.
$\hat{\mathbf{S}}_i$	$\mathbf{T} \cdot \hat{\mathbf{s}}_i$.
$\hat{\mathbf{R}}_i$	$\langle \mathbf{R}_i, \hat{\mathbf{S}}_i \rangle / \|\hat{\mathbf{S}}_i\|^2 \cdot \hat{\mathbf{S}}_i$.
$\hat{\mathbf{r}}_i$	$\mathbf{T}^{-1} \cdot \hat{\mathbf{R}}_i$.
$\mathbf{f}$	Line of missing pixels in recovery vector, $\mathbf{r}$. $N \times 1$ or $1 \times N$.
$\mathbf{g}$	Adjacent line to $\mathbf{f}$ in recovery vector, $\mathbf{r}$. $N \times 1$ or $1 \times N$.
$\mathbf{h}$	$\mathbf{f} - \mathbf{g}$. $N \times 1$ or $1 \times N$.

B. Preliminary

A fast spatial *content-based adaptive block loss recovery* (CABLR) algorithm based on a spatial recovery method using alternating projections [16] is presented in this section. In order to find proper edge orientation by using the line mask operators [18], the four neighbor blocks, $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$, are first defined for the missing block $\mathbf{M}$ with the size of $N \times N$ as shown in Fig. 1 [16]. According to the edge orientation found by the line mask operator, the recovery vectors, $\{\mathbf{r}_i \mid i = 1, 2\}$, are determined either as (a) or as (b) in Fig. 2. Case (a) is for the horizontal line dominating area while case (b) is for the
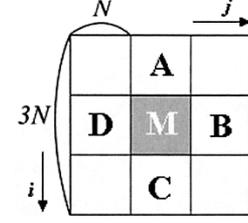

Fig. 1. Missing block with surrounding neighborhood blocks of correctly received data: missing block $\mathbf{M}$ (grey color), surrounding neighborhood, and four connected blocks, $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$.
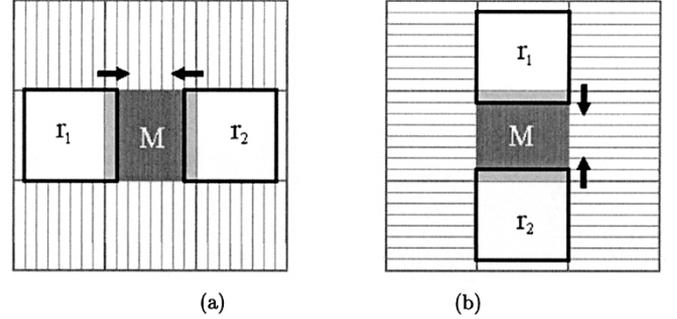

Fig. 2. Missing block with surrounding neighborhood and two $N \times N$ recovery vectors $\mathbf{r}_i$. (a) Recovery vectors $\mathbf{r}_i$ for horizontal line dominating area and (b) recovery vectors $\mathbf{r}_i$ for vertical line dominating area. [16].
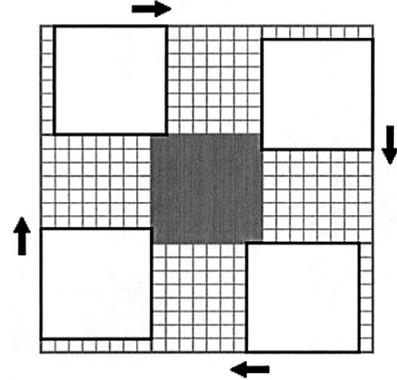

Fig. 3. Missing block with its surrounding neighborhood and an $N \times N$ window to make the surrounding vector $\mathbf{s}$. [16].

vertical line dominating area. For notational simplicity, only the horizontal line dominating case (a), as shown in Fig. 2, is considered throughout the paper. A similar and straightforward explanation of the proposed algorithm can be applied to the vertical line dominating case. After extracting recovery vectors, the best matching vector among surrounding vectors, $\{\mathbf{s}_j \mid 1 \leq j \leq 8N\}$, is found. The surrounding vectors are defined as shown in Fig. 3. The best matching surrounding vectors are $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$. The recovery vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ are projected onto a line in Hilbert space defined by the vector $\hat{\mathbf{s}}_i$

$$\hat{\mathbf{R}}_i = \frac{\langle \mathbf{R}_i, \hat{\mathbf{S}}_i \rangle}{\|\hat{\mathbf{S}}_i\|^2} \cdot \hat{\mathbf{S}}_i \qquad (1)$$

where $i = 1$ and 2, $\mathbf{R}_i = \mathbf{T} \cdot \mathbf{r}_i$, $\hat{\mathbf{S}}_i = \mathbf{T} \cdot \hat{\mathbf{s}}_i$, $\mathbf{T}$ is 2-D DCT kernel, $\langle \cdot, \cdot \rangle$ is the inner product of two vectors, $\|\cdot\|$ is the ℓ_2 vector norm, and $\hat{\mathbf{s}}_i = \mathbf{s}_{J_i}$ when $J_i = \arg\min_j \|\mathbf{r}_i - \mathbf{s}_j\| (1 \leq j \leq 8N)$.
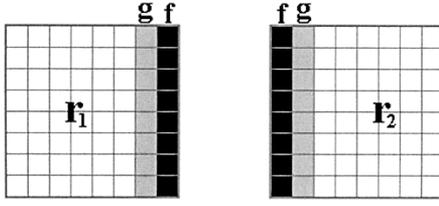
Fig. 4. $\mathbf{f}$ and $\mathbf{g}$ in a recovery vector $\mathbf{r}_1$ and $\mathbf{r}_2$.

Missing pixels in recovery vectors, $\mathbf{r}_i$, are then replaced by the pixels in the same relative position of new recovery vectors in the pixel domain, with $\hat{\mathbf{r}}_i$, where $\hat{\mathbf{r}}_i = \mathbf{T}^{-1} \cdot \hat{\mathbf{R}}_i$ and $\mathbf{T}^{-1}$ is 2-D IDCT operator, $i = 1, 2$. We define $\mathbf{f}$ as the missing pixels (line) in the recovery vectors $\mathbf{r}_i$ and $\mathbf{g}$ as the adjacent line to $\mathbf{f}$ as shown in Fig. 4. A range constraint is applied between replaced pixels, $\mathbf{f}$, and adjacent the original pixels, $\mathbf{g}$. That is, after the missing pixels in the recovery vectors are replaced, $|h_i| \leq \alpha$, $i = 1, \ldots, N$ where $\mathbf{h} = [h_1, h_2, \ldots, h_N] = \mathbf{f} - \mathbf{g}$ and α is a constant. The corresponding projection operator P_r is

$$[P_r \cdot f]_i = \begin{cases} g_i - \alpha, & \text{for } h_i < -\alpha \\ g_i + \alpha, & \text{for } h_i > \alpha \\ f_i, & \text{otherwise} \end{cases} \quad (2)$$

where $1 \leq i \leq N$. The projections defined in (1) and (2) are performed iteratively to restore missing pixels in each recovery vector. After the missing pixels in the recovery vectors are restored, recovery windows move toward each other to extract new recovery vectors and to restore the next line of missing pixels as indicated by the arrows in Fig. 2(a). Using these projections, all missing pixels in a lost block can be restored.

C. Spatial CABLR Algorithm

Although the iteration algorithm based on alternating projections [16] gives better results than simple interpolation [13], [17], the approach is significantly more computationally complex. In order to relieve the computational burden required by the iterative algorithm [16], several operations used by Park *et al.* [16] are modified for the fast spatial CABLR algorithm. For detecting edge orientation with less computational effort, Sobel operator masks are adopted instead of line masks used by Park *et al.* [16]. Edge gradients G_x and G_y by Sobel operator masks at (i, j) [18] are defined by

$$G_x = 2 \cdot [f(i+1, j) - f(i-1, j)] + f(i+1, j-1) + f(i+1, j+1) - f(i-1, j-1) - f(i-1, j+1) \quad (3)$$

and

$$G_y = 2 \cdot [f(i, j+1) - f(i, j-1)] + f(i-1, j+1) + f(i+1, j+1) - f(i-1, j-1) - f(i+1, j-1) \quad (4)$$

where $f(i, j)$ is a pixel value at the point (i, j). Total responses T_x and T_y of the gradient Sobel masks are calculated in four connected blocks $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ in Fig. 1 as

$$T_x = \sum_{B,D} |G_x|, \qquad T_y = \sum_{A,C} |G_y|. \quad (5)$$

After detecting the edge orientation, the recovery blocks are determined either by horizontal line dominating area or by vertical line dominating area as shown in Fig. 2. Formulation of $\mathbf{r}_i$, $\mathbf{s}_j$, $\mathbf{f}$, and $\mathbf{g}$ are the same as described in Section II-B. For finding missing pixels $\mathbf{f}$ in $\mathbf{s}_i$, two projections defined in (1) and (2) are used in Section II-B [16]. However, here, for computational simplicity, missing pixels $\mathbf{f}$ of $\mathbf{r}_i$ are replaced by the pixels of the same position from the best matching surrounding block $\hat{\mathbf{s}}_i$. For $\mathbf{r}_1$ and $\hat{\mathbf{s}}_1$, this pixel replacement process can be expressed as

$$\mathbf{r}_1 = \{\mathbf{r}_{1,1}, \ldots, \mathbf{r}_{1,N-1}, \mathbf{f}\} \quad (6)$$
$$\hat{\mathbf{s}}_1 = \{\hat{\mathbf{s}}_{1,1}, \ldots, \hat{\mathbf{s}}_{1,N}\}, \quad (7)$$
$$\mathbf{f} = \hat{\mathbf{s}}_{1,N}. \quad (8)$$

By adopting this pixel replacement process, a 0.3–0.6-dB loss in *peak-signal noise ratio* (PSNR) is typically experienced when compared with the projection operation approach in (1).

In order to reduce the computational overhead for optimal vector matching search, we propose another approach. Specifically, only half of the surrounding vectors $\{\mathbf{s}_j | 1 \leq j \leq 8N\}$ are used. Only the even numbered (or odd numbered) vectors are first searched. Once the minimum distanced surrounding vector is found, the two adjacent vectors are also searched for the pseudo-best matching surrounding vector. For example, assume that we have $8 \times N$ surrounding vectors and we choose only the vectors with even indices for candidates. Among the $4 \times N$ candidate vectors, assume the kth vector is the best matched. Then the $(k-1)$th, kth, and $(k+1)$th vectors are also tested for the best matched surrounding vectors. When calculating distance between two image blocks the distance can be computed based on *mean absolute difference* (MAD), *mean squared difference* (MSD), or the *cross correlation function* (CCF). When compared with the full search method [16], this fast spatial CABLR gives approximately 0.1–0.2 dB-loss in recovered image quality. However, this reduces the search space by approximately half and results in savings of the same amount in search time.

Next, for removing reconstruction artifacts, a range constraint between replaced pixels and adjacent known pixels used in (2) is adopted. After the restoration of missing pixels in the recovery vectors by pixel replacement and range constraint imposition, another N missing pixels in the next column of row of $\mathbf{M}$ are restored by moving the recovery window toward the center of the missing block $\mathbf{M}$, as shown in Fig. 2. Similar direct matching methods are presented by Wang *et al.* [19]. The proposed algorithm is different in that only a small portion of the surrounding pixels are used and a smaller area of the adjacent surrounding blocks is searched. Specifically, when an 8×8-size block is missing, the size of search space in spatial CABLR is 24×24, while an 80×80 block is used in [19]. This smaller matching block size in a smaller search space enables the spatial CABLR to be implemented more efficiently.

III. Spatio-temporal Block Loss Recovery

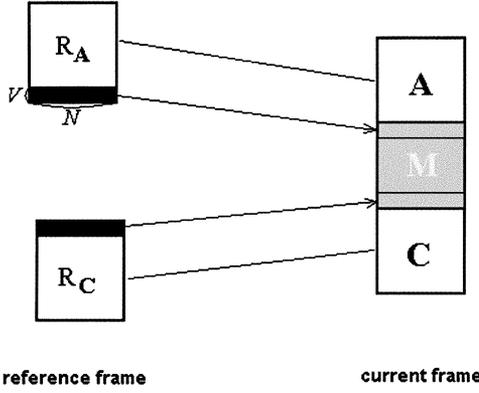In this section, we present the spatio-temporal CABLR algorithm.

Fig. 5. Edge extention from the reference frame to the missing block.

A. Background for MV Recovery

Among the proposed techniques, approaches [5]–[10], [20] that adopt spatial and/or temporal correlation have shown relatively promising recovery performance. However, two problems remain unaddressed. First, the performance may not hold for different local image structures. Second, propagated errors from a previous frame on the considered matching area are not used. Consequently, spatial or temporal correlation methods can fail to provide correct solutions when the surrounding pixels of a missing MB are damaged due to error in the pixel itself or the propagation error from the reference frame. These deficiencies motivate the CABLR algorithm.

B. Content-Based Adaptive Spatio-Temporal Method

When an MV is missing, the lost MV is recovered using surrounding information of the corresponding missing MB. The surrounding information includes known MVs or pixels in surrounding neighborhoods.

In this section, an adaptive distortion metric based on the temporal correlation method is introduced. Weights are imposed on the metric according to the contents of the video sequence. The structure of local image and propagation errors within a *group of pictures* (GOP) are considered as the contents of sequences.

1) Surrounding Image Structure Estimation: In order to get image structure of a missing block, the edge structure is examined to determine information contained in the surrounding neighborhoods of a missing block. Let the vertically connected surrounding blocks of a missing block $\mathbf{M}$ with the size of $N \times N$, be $\mathbf{A}$ and $\mathbf{C}$, as shown in Fig. 5. The magnitudes of gradients T_{vA} and T_{vC} at all (i,j) in the blocks, $\mathbf{A}$ and $\mathbf{C}$, respectively, are computed as

$$T_{vA} = \sum_A |G_x|, \qquad T_{vC} = \sum_C |G_x|, \qquad T_v = T_{vA} + T_{vC}.$$
$$(9)$$

Only the vertical edge is considered in the proposed algorithm since blocks in MPEG/H.26x tend to be corrupted horizontally. Surrounding conditions can be classified into three cases according to the computed edge information. T_G is a predetermined threshold.

1) $T_v < T_G$: Since no vertically oriented image structure or weak edge exists in connected blocks, the surrounding pixels in $\mathbf{A}$ and $\mathbf{C}$ are assumed not to have sufficient information
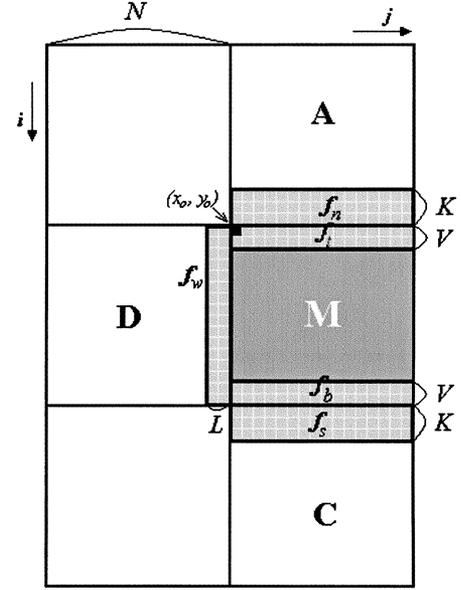
Fig. 6. Missing block $\mathbf{M}$ and the parts of surrounding neighborhoods.

for the matching criterion. This assumption suggests use of pixels in the left block, $\mathbf{D}$, with less of a charge on the pixels in $\mathbf{A}$ and $\mathbf{C}$ for MV recovery. Stronger weights are imposed on the pixels in the left block with the assumption that the pixel information in the left block of the lost MB is more important than pixels in vertically connected blocks for the recovery. The pixels in the left block can be considered in both cases of being correctly received.

2) $T_v \geq T_G$: In this case, the pixels in $\mathbf{A}$ and $\mathbf{C}$ have sufficient information for recovery. With this assumption, stronger pixel weights in $\mathbf{A}$ and $\mathbf{C}$ are required than in $\mathbf{D}$ and pixels in $\mathbf{A}$ and $\mathbf{C}$ should be used more to measure the distortion between the damaged blocks and candidates blocks in the reference frame.

3) $T_v \geq T_G$ and $d_e < T_d$: This is a special case of 2). Let d_e be the sum of transmitted prediction errors between $\mathbf{A}$ and $\mathbf{C}$ and its reference blocks, R_A and R_C, respectively, as shown in Fig. 5

$$d_e = \sum_{i=1}^{N} \sum_{j=1}^{N} (|f_A(i,j) - f_{R_A}(i,j)| + |f_C(i,j) - f_{R_C}(i,j)|) \quad (10)$$

where $f(\cdot)$ is the pixel value and $N \times N$ is the MB size. When $d_e < T_d$, where T_d is a threshold, the missing block and surrounding neighborhoods are assumed to be similar to the reference blocks, R_A and R_C. It is also assumed that the edge in $\mathbf{A}$ and $\mathbf{C}$ extends to the missing block and, consequently, pixels adjacent to R_A and R_C in the reference frame also have information of the lost block. Under this assumption, pixels in the missing block, $\mathbf{M}$, as well as those in $\mathbf{A}$ and $\mathbf{C}$, are utilized for the distortion measurement. $V \times N$ top and bottom pixels in the missing block $\mathbf{M}$ are replaced by pixels below and above the reference block of $\mathbf{A}$ and $\mathbf{C}$, as shown in Fig. 5. Edge structure in $\mathbf{A}$ and $\mathbf{C}$ should be extended to the missing MB by this pixel copy. Since the process is simple pixel replacement using the MVs of two vertically connected blocks, no computation except that of the the error d_e in (10)
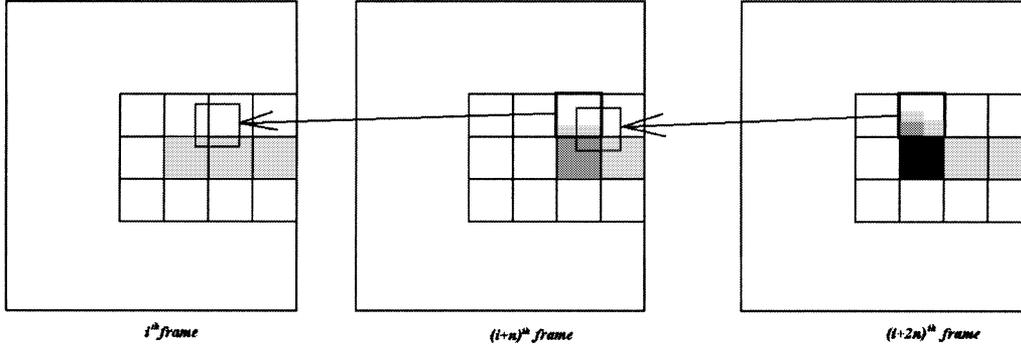
Fig. 7.   Error propagation within a GOP.

is required. In order to avoid additional computation, the approach is applied only to blocks having uncoded or skipped surrounding MBs. When $T_v < T_G$ or $d_e \geq T_d$, there is no vertical edge or similarity in the surrounding neighborhoods and the pixel copy process is not performed.

In order to make the distortion metric adaptive to the local image structure after edge extension, the pixels in missing and surrounding blocks are divided into five areas: f_n, f_s, f_t, f_b, and f_w, as shown in Fig. 6. The size of f_n and f_s is $K \times N$, the size of f_t and f_b is $V \times N$, and the size of f_w is $N \times L$. f_t and f_b are pixels in the missing MB. $N \times N$ is the MB size in all cases. The distance measures corresponding to the pixel (x_0, y_0) of the missing block, as shown in Fig. 6, consist of four components:

1) the spatial difference between the neighbor of missing block and pixels from the previous frame;
2) the spatial distance related coefficient from the missing block η_j;
3) the image structure related coefficient $w_{\alpha,i,j}$;
4) the propagated error related coefficient $\gamma_{i,j}$.

The distance measures, $D_t(x,y)$, $D_b(x,y)$, $D_m(x,y)$, and $D_w(x,y)$ between (x_0, y_0), a pixel position in the missing block of nth frame, and (x,y), a relative pixel position to (x_0, y_0) in the $(n-1)$th frame, is defined as

$$D_t(x,y) = \sum_{i=1}^{K} \sum_{j=0}^{N-1} \eta_j \cdot w_{t,i,j} \cdot \gamma_{i,j} \cdot |f_n(x_0 - i, y_0 + j, n)$$
$$- f_n(x_0 + x - i, y_0 + y + j, n - 1))| \quad (11)$$

$$D_b(x,y) = \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} \eta_j \cdot w_{b,i,j} \cdot \gamma_{i,j} \cdot$$
$$|f_s(x_0 + N + i, y_0 + j, n) -$$
$$f_s(x_0 + N + x + i, y_0 + y + j, n - 1)| \quad (12)$$

$$D_m(x,y) = \sum_{i=0}^{V-1} \sum_{j=0}^{N-1} (\eta_j \cdot \gamma_{i,j} \cdot |f_t(x_0 + i, y_0 + j, n)$$
$$- f_t(x_0 + x + i, y_0 + y + j, n - 1)|)$$
$$+ \sum_{i=1}^{V} \sum_{j=0}^{N-1} (\eta_j \cdot \gamma_{i,j} \cdot |f_b(x_0 + N - i, y_0 + j, n)$$
$$- f_b(x_0 + N + x - i, y_0 + y + j, n - 1)|) \quad (13)$$
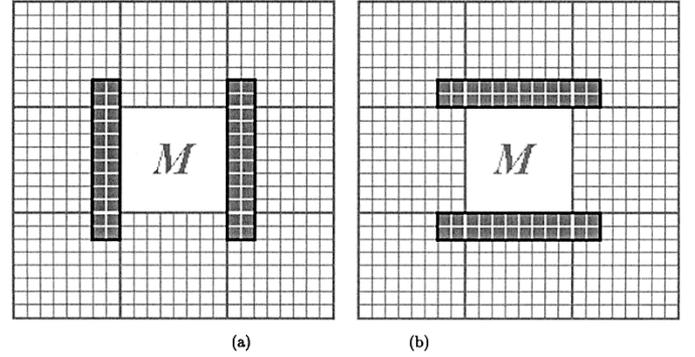
Fig. 8.   Areas for setting parameters α_i. (a) Horizontal edge dominating blocks. (b) Vertical edge dominating blocks. [16].

$$D_w(x,y) = \sum_{i=0}^{N-1} \sum_{j=1}^{L} \cdot \eta_i \cdot w_{w,i,j} \cdot \gamma_{i,j} \cdot$$
$$|f_w(x_0 + i, y_0 - j, n) -$$
$$f_w(x_0 + x + i, y_0 + y - j, n - 1)| \quad (14)$$

where η_i, $w_{\alpha,i,j}$, and $\gamma_{i,j}$ are weights, $f_a(i,j,n)$ is the pixel value at (i,j) of the area a in the nth frame, (x_0, y_0) is the top-left coordinate of the missing MB, $|\cdot|$ is the absolute value, and $x \in S_x$ and $y \in S_y$ when S_x and S_y are the search space in the $(n-1)$th frame. $D_m(x,y)$ of (13) is used only when surrounding edges are extended to the missing block and f_t and f_b are available.

The image structure-related coefficient, $w_{\alpha,i,j}$ and $\alpha \in \{t,b,m,w\}$ in (11)–(14), are adaptively changed according to the edge structure in the surrounding blocks. For example, $w_{t,i,j}$ and $w_{b,i,j}$ are increased while $w_{w,i,j}$ is decreased when a strong edge exists in A and C. The weights $w_{\alpha,i,j}$ enable the distortion metric to adapt to the surrounding characteristic of a missing block. The weights, $w_{\alpha,i,j}$ in (11)–(14) are denoted as

$$w_{t,i,j}, \ w_{b,i,j} = \varepsilon \cdot round \left[\frac{T_v}{\Delta_v} \right]$$
$$w_{w,i,j} = 1 - \frac{(w_{t,i,j} + w_{b,i,j})}{2}$$
$$0 < w_{t,i,j}, \ w_{b,i,j} < 1 \quad (15)$$

where Δ_v is a threshold, ε is a constant, and $round$ is the rounding operation.
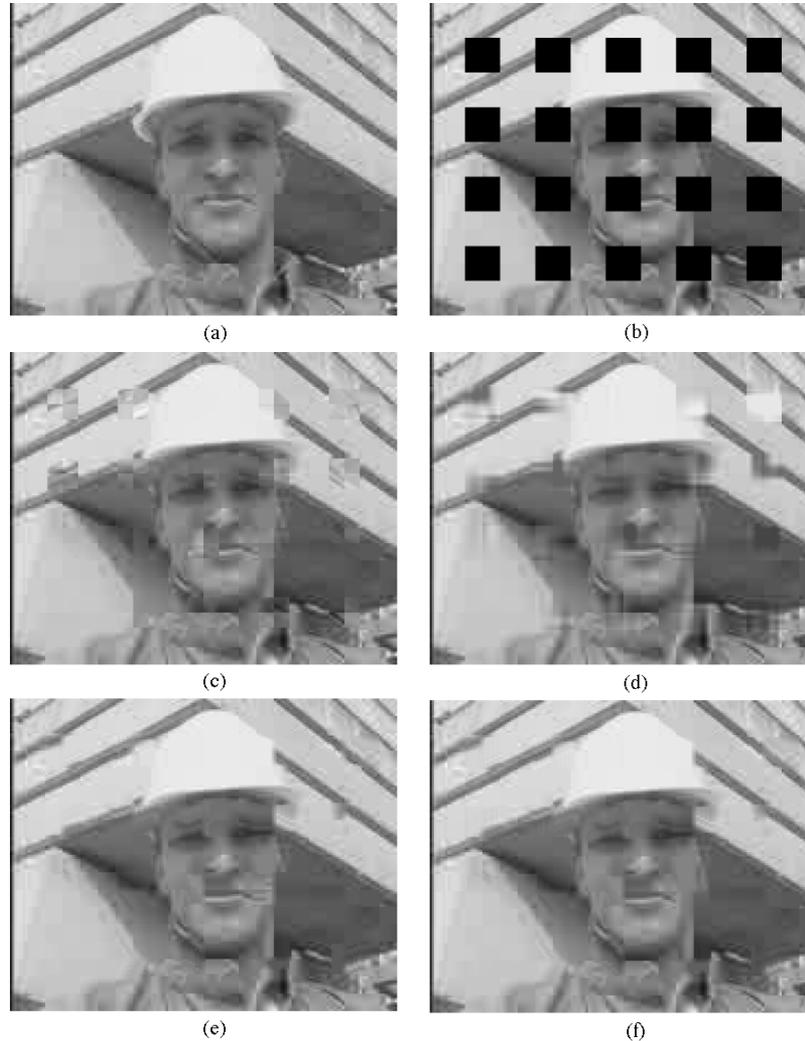
Fig. 9. Experiment on a lost block size of 16×16 pixels of the "foreman" sequence. (a) Original. (b) Damaged image. (c) Image recovered by the Shirani *et al.* (PSNR = 25.66 dB). (d) Image recovered by Alkachouh (PSNR = 24.53 dB). (e) Image recovered by CABLR (PSNR = 27.75 dB). (f) Image recovered by Park *et al.* (PSNR = 27.94 dB).

Spatial distance related coefficient from the missing block η_i are also considered in (11)–(14) for a distance between the considered surrounding pixel and a missing MB. From the assumption that closer pixels in the surrounding block to the missing MB are more correlated to the lost pixels, higher weights are imposed on closer pixels to measure distortion. The weights η can be calculated by

$$\eta = \rho \cdot e^{-(d-1)/\sigma} \qquad (16)$$

where d is the distance between the pixel and the missing block, and ρ and σ are predetermined constants. The constant σ can be used to reduce or increase the effect of the distance d.

2) Error Propagation Estimation: Damaged pixels in an MPEG block cause error propagation to the following blocks within a GOP as shown in Fig. 7. In Fig. 7, MBs with a shaded region in the ith frame are damaged MBs. During the decoding and the post recovery process, damaged MBs in the ith frame

are restored using the surrounding blocks. Since an MB in the $(i + n)$th frame references the ith frame, error in the ith frame is propagated to the $(i + n)$th frame. When the MB below the error propagated MB in the $(i + n)$th frame is also damaged, the missing MB is restored using the surrounding error propagated pixels. This causes another recovery error since damaged blocks in the $(i + n)$th frame can be replaced by restored pixels in the ith frame determined by surrounding pixels in the $(i + n)$th frame. Restored errors are propagated to the $(i + 2 \cdot n)$th and all frames in a GOP until a new intra frame is encountered. To reduce the restoration error caused by the errors in the surrounding blocks, error weights are imposed on the distortion metric.

Let $P_{i,j}$ be the propagation error state related to the propagation and restoration error at a pixel (i, j). The number P is enumerated when the pixel is corrupted and is added to the reference pixel's error state $P_{i,j}^r$. For example, a pixel's error state $P_{i,j}$ is set to zero when the pixel does not have errors, and the reference pixel's error state $P_{i,j}^r$ is added to the number $P_{i,j}$. A
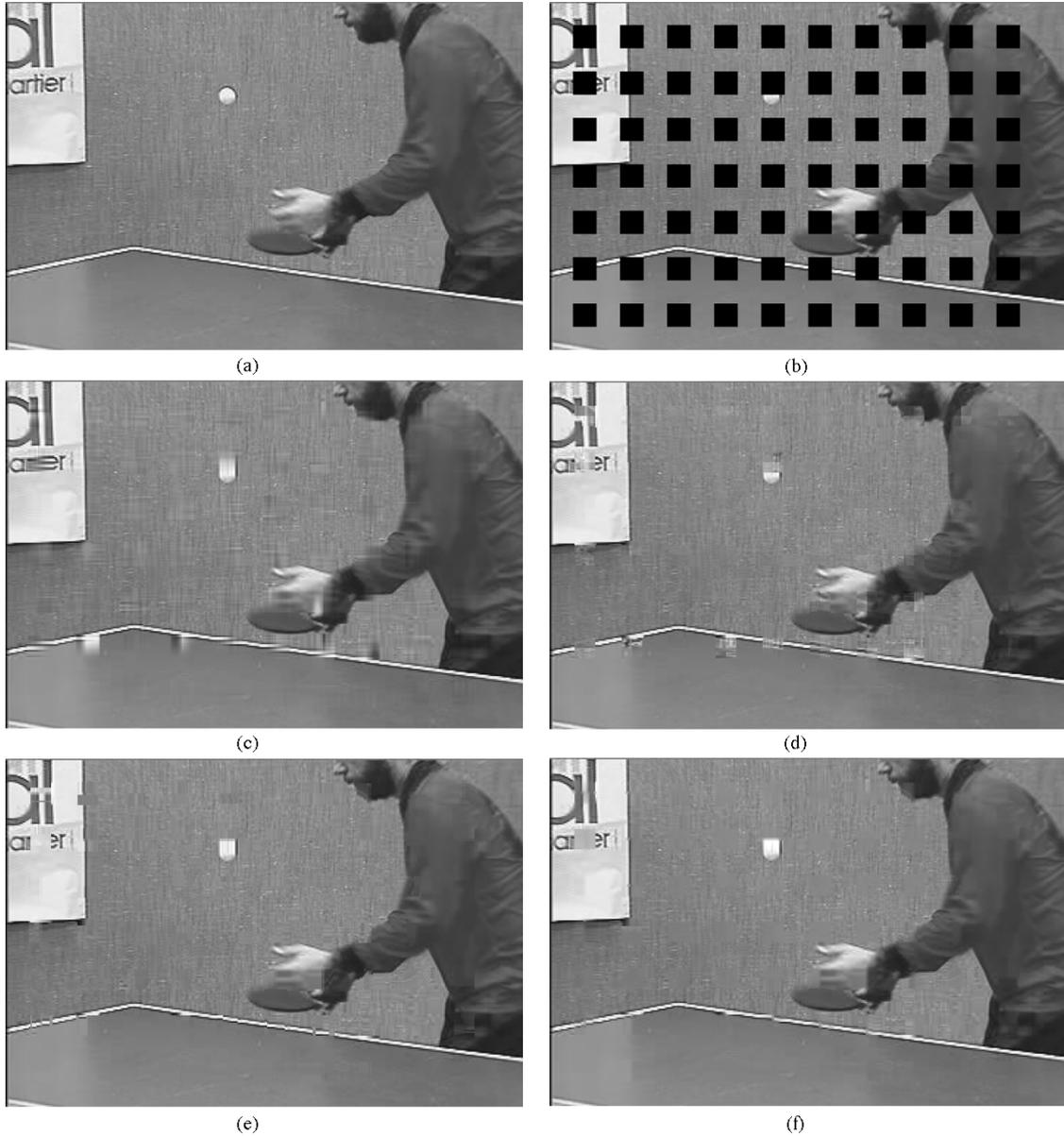
Fig. 10. Experiment on a lost block size of 16×16 pixels of the "table tennis" sequence. (a) Original. (b) Damaged image. (c) Image recovered by the Shirani *et al.* (PSNR = 27.54 dB). (d) Image recovered by Alkachouh (PSNR = 27.58 dB). (e) Image recovered by CABLR (PSNR = 28.86 dB). (f) Image recovered by Park *et al.* (PSNR = 29.29 dB). Figures show table lines are better recovered by CABLR and Park *et al.*'s and the result of CABLR is similar as Park *et al.*'s.

pixel's error state $P_{i,j}$ is set to one when the pixel is not correctly received and restored. With this process, propagation and recovery error occurrence numbers are known. The weights $\gamma_{i,j}$ are imposed in (11)–(14) according to the count P of each pixel

$$P_{i,j} = \begin{cases} 0, & \text{when the pixel is correctly received} \\ P_{i,j}^r + 1, & \text{otherwise} \end{cases}$$

(17)

$$\gamma_{i,j} = 1 - \mu \cdot P_{i,j}, \quad 0 < \gamma_{i,j} \leq 1 \tag{18}$$

where μ is a predetermined constant. The weights $\gamma_{i,j}$ enable the distortion metric to adapt to the propagation error in the video sequences.

3) MV Recovery With CABLR: With the calculation of difference measures as shown in (11)–(14) for the missing block, the MV recovery metric is defined as

$$(\hat{x}, \hat{y}) = \arg \min_{(x,y)} (D_t(x,y) + D_b(x,y) + D_m(x,y) + D_w(x,y))$$

(19)

where $(\hat{x}, \hat{y})$ is a recovered MV of the missing block.

4) Spatial Constraint: MV recovery techniques work well to protect the degradation and error caused by block loss. However, when there are no similar pixels in the reference frame with respect to the missing block, MV recovery does not provide an acceptable solution. This happens especially when the damaged block is intra coded. Spatial restoration gives an alternate solution in this case. Sun *et al.* [21] suggest an adaptive recovery technique
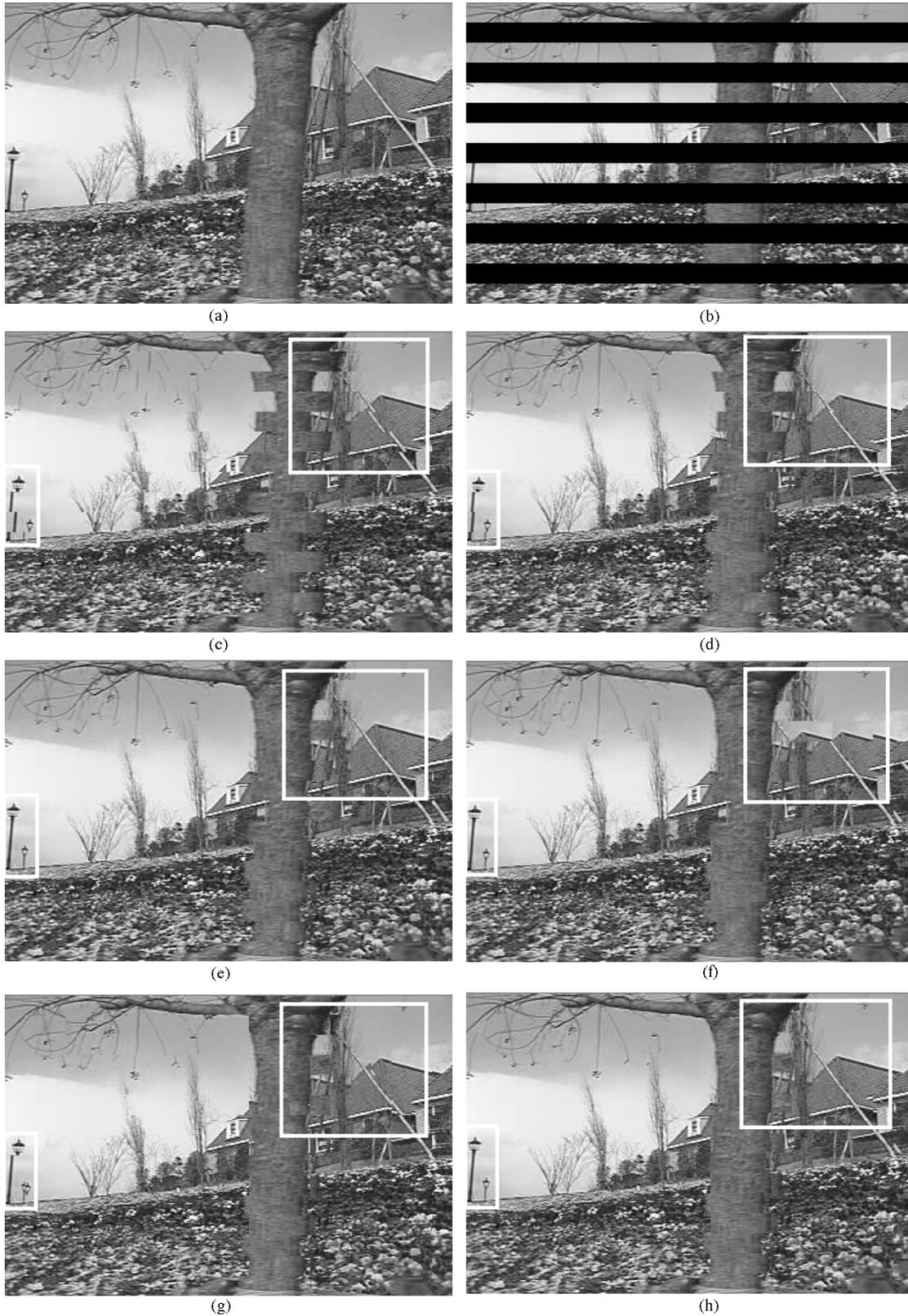
Fig. 11. Experiment on slice loss of a P-frame in the "flower garden" sequence. (a) Original. (b) Damaged. (c) ZM (PSNR = 16.96 dB). (d) AV (PSNR = 19.06 dB). (e) BMA (PSNR = 21.65 dB). (f) DMVE (PSNR = 21.63 dB). (g) FB-BM (PSNR = 20.40 dB). (h) CABLR (PSNR = 22.73 dB).

in which temporal recovery or spatial interpolation is determined according to the spatial and temporal activity of frames. Chen *et al.* [8] report overlapped motion compensation. This may, however, degrade the restored image quality when the reference frame has accurately restored lost pixel values. CABLR proposes a simple error adaptive spatial compensation approach.

TABLE I
PSNR OF SPATIAL RECOVERY IN I-FRAME OF "FLOWER GARDEN," "TABLE TENNIS," "FOOTBALL," "MOBILE," AND "FOREMAN"

	Flower garden	Table tennis	Football	Mobile	Foreman
Shirani *et al.* [12]	20.87	27.54	27.18	19.86	25.66
Alkachouh and Bellanger [13]	19.67	27.58	26.77	19.11	24.53
CABLR	21.01	28.86	28.03	19.54	27.75
Park *et al.* [16]	21.48	29.29	28.49	20.30	27.94

The corresponding error at the recovery MV $(\hat{x}, \hat{y})$ in (19) is obtained by

$$e = D_t(\hat{x}, \hat{y}) + D_b(\hat{x}, \hat{y}) + D_m(\hat{x}, \hat{y}) + D_w(\hat{x}, \hat{y}). \quad (20)$$

Assume the pixels from the recovered MV by (19) are not perfectly matched to the original values when the error e in (20) is larger than a threshold T_e. That is, in the case of $e > T_e$, after recovering the MV, the pixel values in the recovered block are spatially constrained by (2).

The constraint is applied from the top-most and bottom-most pixels to center pixels in the same manner as the spatial case. This range constraint limits the difference between missing and adjacent known pixel values. This method is applied only after the lost MV is recovered.

5) Pseudocode of CABLR Algorithm: The steps of the CABLR are 1) determination of vertical edge existence in two connected MBs and setting the weights w, 2) extension of the vertical edge to the missing MB when there is a strong vertical edge and small prediction error in the connected blocks, 3) estimation of restoration and propagation error, and setting the weights, γ, and 4) spatial compensation when $e > T_e$. Here is pseudocode.

```
Algorithm
 Compute compute gradients, TvA and TvC,
in A and C, respectively
 if TvA and TvC ≥ and de < Td
  copy V  ×  N pixels from the reference
frame
 endif
 set weights, w, according to TvA and TvC
 set weights, η, according to the distance
between pixel and the missing block
 estimate propagation and recovery error
within the GOP
 set weights, γ, according to the error
number, P
 if e > Te
  apply a spatial constraint
 endif
End.
```

IV. EXPERIMENTAL RESULTS

The proposed algorithm is tested on 352×240 pixel size *flower garden, football, calendar and mobile, table tennis*, and 176×144-pixel size *foreman* sequences. The *foreman* sequence is encoded at the rate of 128 Kbit/s and other sequences are encoded at the rate of 1.15 Mbit/s by MPEG-2 encoder. The frame rate is 30 fps in all cases. The frame number of the GOP is 12, and the B frame number between I and P frames is 3. In the I-frame, we consider a missing MB surrounded by known MBs on the assumption that MBs are interleaved on packing [22]. We examine a missing MB from every four MBs for maximum error experiment and apply spatial error concealment to the I-frames. The proposed spatial algorithm is compared with other procedures such as Shirani *et al.*'s technique [12] and Alkachouh and Bellanger's scheme [13].

In the P- and B-frames, we consider a missing *slice* with the assumption that a synchronization marker is inserted in the first MB of every row of the MB. Temporal error concealment is applied in this case. We compare the proposed method with existing algorithms such as zero MV (ZM), average MV (AV) [5], BMA [7], Zhang *et al.*'s algorithm [*decoder motion-vector algorithm* (DMVE)] [9], and Tsekeridou and Pitas' algorithm [*forward-backward block-matching* (FB-BM)] [10].

Spatial block loss recovery is tested in the I-frames of the sequences. Each initial missing pixels of recovery vector **r** is set to the adjacent known pixel in the same vector. Missing block size is 16×16 $(N = 16)$ and MSD is used. Parameter α in (2) are set to the maximum value of the differences between two adjacent pixels in the same column in the blocks with shaded region of Fig. 8(a) and (b) as in [16]. The cases of (a) and (b) in Fig. 8 are for the horizontal edge and vertical edge dominating area, respectively. In the case of Fig. 8(a), α is

$$\alpha = \max |f(x, y) - f(x, y-1)|$$
$$x_0 - 2 < x < x_0 + N + 1,$$
$$y = y_0 - 1 \text{ or } y_0 + N + 1 \quad (21)$$

where (x_0, y_0) is the top-left pixel of a missing block, **M**. The peak signal-to-noise ratio (PSNR) [23] is a measure of the restored image quality, and is given by

$$PSNR = 10 \cdot \log \left(\frac{I \cdot J \cdot 255^2}{\sum_{i=1}^{I} \sum_{j=1}^{J} | f(i,j) - \hat{f}(i,j) |^2} \right) \quad (22)$$

where f and $\hat{f}$ are the value of original and restored image of $I \times J$ pixels, respectively. Figs. 9 and 10 shows the spatial block loss recovery in the intra frame of *foreman* and *table tennis*. Images (a) and (b) of Figs. 9 and 10 are the original and the damaged intra frame, respectively. Images (c) and (d) of Figs. 9 and 10 are the restored by Shirani *et al.*'s and Alkachouh and Bellanger's

TABLE II
AVERAGE PSNR OF TEMPORAL RECOVERY IN P- AND B-FRAMES OF "FLOWER GARDEN," "TABLE TENNIS," "FOOTBALL," "MOBILE," AND "FOREMAN"

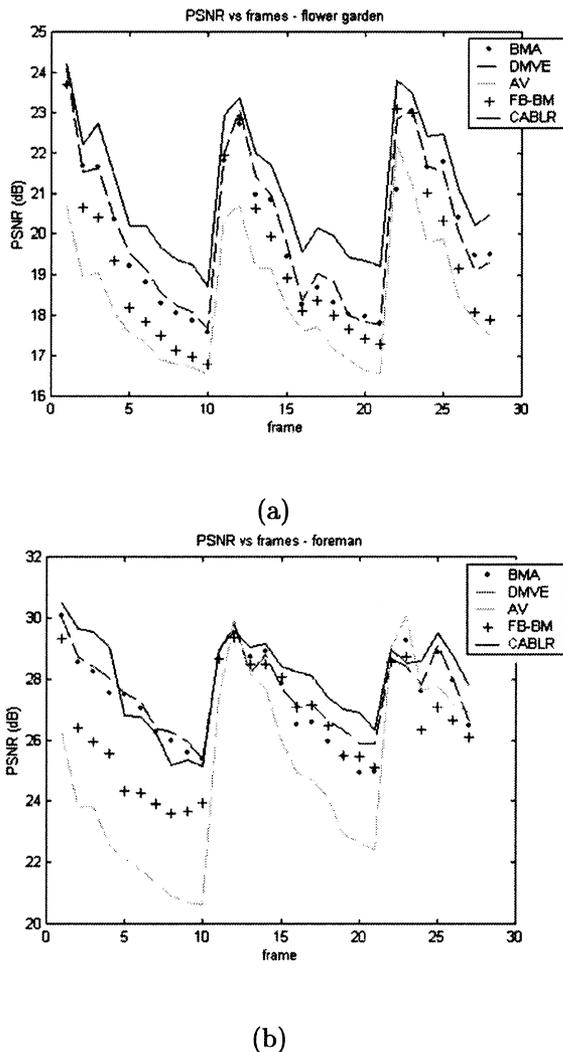|	Flower garden (115)	Table tennis (112)	Football (125)	Mobile (140)	Foreman (150)
ZM	16.15	22.40	18.06	17.49	23.86
AV	18.64	21.98	18.72	19.03	25.15
BMA	19.83	23.55	19.41	19.75	26.83
DMVE	19.88	24.04	19.64	19.99	26.95
FM-BM	19.21	22.49	19.05	19.59	25.91
CABLR	20.70	24.52	20.32	20.66	27.41

(a)

(b)

Fig. 12.   PSNR versus frames. (a) Flower garden. (b) Foreman.

algorithm, respectively. Images (e) and (f) show restored images by the proposed method and [16]. The result of the proposed method is approximately 0.2–0.5 dB lower than that of [16], and typically requires 40% less computation compared to [16] when the number of iteration is 1 for [16]. This is because Park *et al.*'s algorithm requires full search and projections while the proposed algorithm does not. Table I summarizes the PSNR of restored intra frames in all sequences.

Temporal recovery is tested in damaged blocks of the P and B frames. Parameters in the proposed algorithm are set to $K = 3$, $V = 1$, $L = 1$, and $T_d = 1800$. Parameters for weights are set as $\varepsilon = 0.25$, $\Delta_v = 3500$, $\lambda_i = 1.0$, and $\mu = 0.1$. $\eta_i = 1.1$ and 1.0 when $i = 1$ and $i \neq 1$, respectively, since altering the values of parameters concerning η are observed as having not much an affect on the reconstruction results. For the surrounding edge extension, T_G is set to 6000 and it is applied only when a missing block is uncoded or skipped. The whole missing block, including edge extended pixels, are replaced with the reference block after MV recovery. T_e is set to 1700 for the spatial constraint. $x_0 - N < x < X_0 + N, y_0 - N < y < y_0 + N$ is used for the search space in (19), where N is the MB size ($N = 16$). Tested parameters are determined by experiments and the same parameter values described are used for the test of all sequences. In BMA, a full search method is applied instead of the five candidates search. In DMVE, two pixel lines the left of, above, and below a missing MB are used for MV search. MAD is used for the distance measure in all cases.

Fig. 11 shows the test on slice error in a P-frame of *flower garden*. Images of (a) and (b) in Fig. 11 are the original and the damaged image, respectively. Images restored by ZM, AV, BMA, DMVE, and FB-BM are shown in (c), (d), (e), (f), and (g), respectively. Images restored by CABLR are shown in (h) of Fig. 11.

Table II summarizes the average PSNR of the restored frames in all sequences. Tests are performed under the same conditions as previously mentioned, and the same slice loss pattern is used as shown in Fig. 11. Fig. 12(a) and (b) show PSNR versus the first 30 reconstructed P- and B-frames of *flower garden* and *foreman*, respectively, by different methods. Frame numbers for tests are given with sequence names. The table shows that CABLR results in higher PSNR than other algorithms in all cases.

V. CONCLUSION

In this paper, we present new spatial and temporal block loss recovery algorithms. In the spatial algorithm, two $N \times N$ pixel recovery vectors are extracted using a Sobel mask operation. Extracted vectors, including known and unknown pixels, are compared with surrounding $N \times N$ pixels. The best matching block of each vector is found using a fast search method. The values of unknown pixels in the pixel vectors are replaced with the pixels

in the same position of the best matching block. For pixel continuity, a range constraint is applied between the newly replaced pixels and the adjacent known pixels. After missing pixels in the recovery vectors are restored, the new $N \times N$ pixel recovery vectors are extracted and, consecutively, all missing pixels are thereby restored.

In the temporal CABLR algorithm, vertical edge detection is applied in the vertically connected surrounding blocks of a missing MB. Weights to the distortion metric are adaptively determined according to the surrounding information. When there is a strong vertical edge in connected blocks, vertically positioned pixels are more heavily weighted for MV recovery. In the converse case, pixels in the left block are more important. When surrounding pixels have a high error propagation state, lower weights are also imposed on these pixels. When the previous frame does not have sufficient information for the missing MB, simple spatial compensation is applied after MV recovery.

The proposed algorithms are tested on several standard MPEG sequences. Spatial CABLR is applied on I-frames and temporal-spatial CABLR is tested on P- and B-frames of MPEG sequences. CABLR utilizes local image structure and error propagation characteristic of video coding standards for error concealment. The reconstruction quality by the proposed algorithms is consistently higher than other techniques tested.

ACKNOWLEDGMENT

REFERENCES

[1] *ISO/IEC 13 818–2, Generic Coding of Moving Pictures and Associated Audio Information-Part 2: Video*, ISO/IEC 13 818–2 MPEG-2 Video Coding Std., 1995.

[2] *CCIR Recommendation H.261:Codec for Audiovisual Services at $p \times 64$ kbits/sec*, ITU-T, 1990.

[3] *Draft ITU-T Recommendation H.263 Version 2:Video Coding for Low Bitrate Communication*, ITU-T, 1997.

[4] Y. Wang and Q. F. Zhu, "Error control and concealment for video communication: A review," in *Proc. IEEE Multimedia Signal Processing*, May 1998, pp. 974–997.

[5] P. Haskell and D. Messerschmitt, "Resynchronization of motion compensated video affected by ATM cell loss," in *Proc. ICASSP*, San Francisco, CA, Mar. 1992, pp. III545–548.

[6] M. E. Al-Mualla, C. N. Canagarajah, and D. R. Bull, "Motion field interpolation for temporal error concealment," *Proc. Inst Elect. Eng. Vis. Image Signal Process.*, vol. 147, no. 5, Oct. 2000.

[7] W.-M. Lam, A. R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors," in *Proc. ICASSP*, Minneapolis, MN, Apr. 1993, pp. V417–V420.

[8] 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, pp. 560–563, June 1997.

[9] 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, pp. 659–665, June 2000.

[10] S. Tsekeridou and I. Pitas, "MPEG-2 error concealment based on block-matching principles," in *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, June 2000, pp. 646–658.

[11] Y. Wang, Q. F. Zhu, and L. Shaw, "Maximally smooth image recovery in transform coding," *IEEE Trans. Commun.*, vol. 41, pp. 1544–1551, Oct. 1993.

[12] S. Shirani, F. Kossentini, and R. Ward, "Reconstruction of baseline JPEG coded images in error prone environments," *IEEE Trans. Image Processing*, vol. 9, pp. 1292–1299, July 2000.

[13] Z. Alkachouh and M. G. Bellanger, "Fast DCT-based spatial domain interpolation of blocks in images," *IEEE Trans. Image Processing*, vol. 9, pp. 729–732, Apr. 2000.

[14] H. Sun and W. Kwok, "Concealment of damaged blocks transform coded images using projections onto convex sets," *IEEE Trans. Image Processing*, vol. 4, pp. 470–477, Apr. 1995.

[15] R. J. Marks II, "Alternating projections onto convex sets," in *Deconvolution of Images and Spectra*, P. A. Jansson, Ed. San Diego, CA: Academic, 1997, pp. 476–501.

[16] J. Park, D.-C Park, R. J. Marks, and M. A. El-Sharkawi, "Block loss recovery in DCT image encoding using POCS," in *Proc. IEEE Int. Symp. Circuits and Systems*, May 2002, pp. V245–248.

[17] M. Ancis and D. D. Giusto, "Reconstruction of missing blocks in JPEG picture transmission," in *IEEE Pacific Rim Conf.*, 1999, pp. 288–291.

[18] R. Gonzalez and R. Woods, *Digital Image Processing*. Reading, MA: Addison-Wesley, 1992.

[19] Z. Wang, Y. Yu, and D. Zhang, "Best neighborhood matching: An information loss restoration technique for block-based image coding systems," *IEEE Trans. Image Processing*, vol. 7, pp. 1056–1061, July 1998.

[20] K.-W. Kang, S.-H. Lee, and T. Kim, "Recovery of coded video sequences from channel errors," *SPIE*, vol. 2501, pp. 19–27, 1995.

[21] H. Sun, J. W. Zdpski, W. Kwok, and D. Raychaudhuri, "Error concealment algorithms for robust decoding of MPEG compressed video," *Signal Process.: Image Commun.*, vol. 10, no. 4, pp. 249–268, 1997.

[22] A. S. Tom, C. L. Yeh, and F. Chu, "Packet video for cell loss protection using deinterleaving and scrambling," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Toronto, ON, Canada, May 1991, pp. 2857–2860.

[23] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.

Jiho Park (M'93) received the Ph.D. degree in electrical engineering from the University of Washington, Seattle, in 2002. His Ph.D. research involved error-resilient image/video communications and post-processing error concealment.

He joined Samsung Electronics Communication Research Laboratories in 2002. He is currently working on video and multimedia signal precessing and communications. His current interests include video and signal processing, joint source and channel coding, pattern recognition, and communications.

Dong-Chul Park (S'80–M'90–SM'99) received the B.S. degree in electronics engineering from Sogang University, Seoul, Korea, in 1980, the M.S. degree in electrical and electronics engineering from the Korea Advanced Institute of Science and Technology, Seoul, in 1982, and the Ph.D. degree in electrical engineering, with a dissertation on system identification using neural networks, from the University of Washington (UW), Seattle, in 1990.

From 1990 to 1994, he was with the Department of Electrical and Computer Engineering, Florida International University, The State University of Florida, Miami. Since 1994, he has been with the Department of Information Engineering, MyongJi University, Korea, where he holds the title of Professor. From 2000 to 2001, he was a Visiting Professor at UW. He is a pioneer in the area of electric load forecasting using neural networks. He has published more than 50 papers, including 20 archival journals in the area of neural network (NN) algorithms and their applications. His current interests include the development of NN algorithms and their applications to various engineering problems, including financial engineering, image compression, speech recognition, and pattern recognition. He is the founder of NeuroSolutions, Co., Korea, where he develops IT solutions, including fraud detection system for credit card business.

Dr. Park was a member of the Editorial Board for the IEEE TRANSACTIONS ON NEURAL NETWORKS from 2000 to 2002.

Robert J. Marks, II (F'94) holds the position of Distinguished Professor of Electrical and Computer Engineering, Department of Engineering, Baylor University, Waco, TX. He has over 250 publications. He is the Author and Co-Author of the books *Introduction to Shannon Sampling and Interpolation Theory* (New York: Springer-Verlag, 1991), and *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks* (Cambridge, MA: MIT Press, 1999). He is a Co-Editor of five other volumes. He was also the Topical Editor for Optical Signal Processing and Image Science for the *Journal of the Optical Society of America A*.

Prof. Marks is a Fellow of the Optical Society of America. He was awarded the Outstanding Branch Counselor award by the IEEE and the IEEE Centennial Medal. He was named a Distinguished Young Alumnus of the Rose-Hulman Institute of Technology and is an inductee into the Texas Tech Electrical Engineering Academy. In 2000, he was awarded the Golden Jubilee Award by the IEEE Circuits and Systems Society. He is also the first recipient of the IEEE Neural Networks Society Meritorious Service Award. He was given the honorary title of Charter President of the IEEE Neural Networks Council. He served as an IEEE Distinguished Lecturer. He served a six-year stint as the Editor-in-Chief of the IEEE TRANSACTIONS ON NEURAL NETWORKS and as an Associate Editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS. He serves as the Administrative Vice President of the IEEE Circuits and Systems Society.

Mohamed A. El-Sharkawi (F'95) received the B.Sc. degree in electrical engineering from the Cairo High Institute of Technology, Cairo, Egypt, in 1971 and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of British Columbia, Vancouver, BC, Canada, in 1977 and 1980, respectively.

In 1980, he joined the University of Washington, Seattle, where he is presently a Professor of electrical engineering. He is the Founder of the International Conference on the Application of Neural Networks to Power Systems (ANNPS) and a Co-Founder of the International Conference on Intelligent Systems Applications to Power (ISAP). He is a current or past member of the editorial boards or Associate Editor of several journals. He is the Author of a textbook on *Fundamentals of Electric Drives* (Pacific Grove, CA: Brooks Cole, 2000). He has organized and taught several international tutorials on intelligent systems applications, power quality, and power systems, and has organized and chaired numerous panel and special sessions in international conferences. He has published over 160 papers and book chapters in these areas and holds five licensed patents.

Dr. El-Sharkawi is the Editor and Co-Editor of several IEEE tutorial books on the applications of intelligent systems. He is the Vice President for technical activities of the IEEE Neural Networks Society. He is a founding Chairman of several IEEE task forces, working groups, and subcommittees.