A Concealment Method for Video Communications in an Error-Prone Environment

Shahram Shirani, *Member, IEEE*, Faouzi Kossentini, *Senior Member, IEEE*, and Rabab Ward, *Fellow, IEEE*

***Abstract*—In this paper, we propose a two-stage error-concealment method for block-based compressed video which was transmitted in an error-prone environment. In the first stage, we obtain initial estimates of the missing blocks. If the motion vectors associated with the missing blocks are available, motion compensation is used to provide good estimates. Otherwise, a novel algorithm which preserves image continuity is used to estimate the blocks. In the second stage, a maximum *a posteriori* (MAP) estimator, which employs an adaptive Markov random field (MRF) as the image *a priori* model, is used to improve the video reconstruction quality. The adaptive model enables the estimation to incorporate information embedded not only in the immediate neighborhood pixels but also in a wider neighborhood into the reconstruction procedure without increasing the order of the MRF model. The proposed concealment method achieves very good computation–performance tradeoffs, as demonstrated via experimental results.**

***Index Terms*—Error concealment, image reconstruction, Markov random fields.**

I. Introduction

DIGITAL image and video signals require very high bit rates, thus, compression of such signals before their transmission is necessary. Communication channels are not error free and, consequently, the encoded bit streams are vulnerable to transmission errors—usually causing loss of blocks of data and/or loss of synchronization. Despite the various methods that have been proposed to combat or localize the effects of channel errors, the quality of a decoded video sequence may degrade significantly because of the residual errors. Error concealment intends to conceal the effects of such errors by exploiting redundancies in the video signal and limitations of the human visual system, without requiring additional information.

Temporal concealment methods are usually used for error concealment in inter-coded frames. Data of an inter-coded block are composed of the motion vector and the DCT coefficients of the prediction error. If the motion vectors are received without errors, the missing blocks are set to their corresponding motion compensated blocks. Since in many of the video compression algorithms (e.g., H.263) the coded motion vectors and the DCT coefficients are interleaved, the loss of data of a block usually results in the loss of both the motion vectors and the DCT coefficients.[1] Therefore, most of the proposed temporal concealment methods first estimate the motion vector associated with a missing block using the motion vectors of adjacent blocks [1], [2]. The estimated motion vector is then used to find a block in the previous frame which yields the restoration of the missing block. One problem associated with these methods is that if the adjacent blocks are coded in a nonpredictive way (e.g., intra-coded), there would not be any data available to estimate the missing motion vectors. Thus, ad hoc assumptions for estimating the missing motion vector should be made, and the results may be unreliable. Moreover, if the missing block lies on the boundary of two objects moving in opposite directions, such methods will perform quite poorly.

Spatial error-concealment methods restore the missing blocks using the information in the current frame. In [3], to restore the missing data, a measure of variations (e.g., gradient or Laplacian) between adjacent pixels is minimized. The underlying smoothness assumption of this method limits its ability in restoring image details. In [4], each pixel in a damaged block is interpolated from the corresponding pixels in its four neighboring blocks such that the total squared border error is minimized. In [5] and [6], the missing information is interpolated utilizing spatially correlated edge information from a large local neighborhood. Note that although these edge-based methods are generally more accurate than other approaches, they are computationally more intensive. In [7], a computationally simple, spatial directional interpolation scheme has been proposed. The two nearest surrounding layers of pixels of a missing block are converted into a binary pattern to reveal the local geometrical structure. Then, the missing pixels are interpolated in a way to preserve the local geometrical structures. In the statistical error concealment methods, for example the one proposed in [8], it is assumed that the pixel values in an image or video signal are realizations of an underlying statistical model (e.g., MRF model). The statistical approaches of spatial error concealment are expected to outperform the deterministic ones, as they provide a systematic way for incorporating the *a priori* information about the video signal in the restoration procedure.

Previous spatial error-concealment methods employing the MRF model usually yield blurry images with a significant loss of detail in the high frequency or edge portions of the image. This is due to 1) the type of the MRF selected as the image model, and 2) the fact that the amount of information that is

Manuscript received May 4, 1999; revised November 30, 1999. This work was supported by MDSI, BC ASI, and NSERC.

The authors are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada.

Publisher Item Identifier S 0733-8716(00)04337-7.

[1]Recognizing the need to improve the concealment capabilities, in the new generation of video compression standards (e.g., MPEG 4, H.263++), the motion information is separated from other information through partitioning of the coded data.

used in the reconstruction process is often restricted to a single-pixel wide region around the erroneous area. Incorporating more pixels in order to enhance restoration of edges and high-frequency parts would usually require a higher order MRF model. However, this is computationally expensive, as the complexity grows exponentially with the order of the MRF model.

In this paper, we propose a two-stage error-concealment method for compressed video. In the first stage, if applicable, we use the information in the previous frame to obtain initial estimates of the missing blocks. If the motion vectors of the missing blocks are available, motion compensation is used to provide the estimates. Otherwise, an algorithm which preserves image continuity is employed to compute the initial estimates. In the second stage, a MAP estimator is used for refinement of the initial estimates. The MAP estimator employs an adaptive MRF as the image *a priori* model. The proposed adaptive MRF takes into account the local image characteristics embedded not only in the immediate neighborhood pixels of the damaged area but also in a wider neighborhood without a dramatic increase in computational complexity. Our concealment method improves on the existing statistical methods in that it yields good reconstruction performance regardless of the content of the missing blocks.

To make the compressed data more error resilient, most of the standard-compliant video compression systems partitioned a video frame into Groups of Blocks (GOB's) or "slices," which are coded independently. Therefore, the output bit stream usually consists of segments separated with markers, where each segment corresponds to the coded data of the blocks in a GOB or a slice. When channel errors occur, the decoder usually discards the erroneous data between two markers surrounding the erroneous data, effectively discarding the GOB or the slice. Then, loss of data of a slice does not affect the rest of the compressed video sequence. In our work, it is assumed that the frames of the coded video sequence are partitioned into GOB's or slices and, thus, the missing data belong to blocks of a GOB or a slice. Moreover, it is assumed that the decoder knows the locations of the missing blocks. This information (e.g., the checksum information) can be obtained from the network or it can be inferred, for example, by detecting the semantic or syntactic violations as a result of errors [9].

The rest of this paper is organized as follows. In Section II, MAP estimation of missing data is briefly reviewed. Section III presents the proposed error concealment method. Section IV discusses the computational complexity of the proposed method. Sections V and VI present our experimental results and conclusions, respectively.

II. MAP Estimation of Missing Data

Let the corrupted image be represented by the vector $\boldsymbol{Y}$, while the decompressed image without errors can be represented by $\boldsymbol{X}$. Using MAP estimation, the decompressed image estimate is given by

$$\hat{\boldsymbol{X}} = \arg\max_{\boldsymbol{X}} \log \Pr(\boldsymbol{X} \mid \boldsymbol{Y}).$$

Using the Bayes' rule we, obtain

$$\begin{aligned}\hat{\boldsymbol{X}} &= \arg\max_{\boldsymbol{X}} \left\{ \log \frac{\Pr(\boldsymbol{Y} \mid \boldsymbol{X}) \Pr(\boldsymbol{X})}{\Pr(\boldsymbol{Y})} \right\} \\ &= \arg\max_{\boldsymbol{X}} \{ \log \Pr(\boldsymbol{Y} \mid \boldsymbol{X}) + \log \Pr(\boldsymbol{X}) \} \end{aligned} \tag{1}$$

where the term $\log \Pr(\boldsymbol{Y})$ has been dropped because it is independent of $\boldsymbol{X}$. The corrupted image can be expressed as $\boldsymbol{Y} = \mathcal{T}\boldsymbol{X}$, where $\mathcal{T}$ is the transform that converts the decompressed image $\boldsymbol{X}$ into the corrupted one $\boldsymbol{Y}$.[2] The conditional probability of the corrupted image can be then written as follows:

$$\Pr(\boldsymbol{Y} \mid \boldsymbol{X}) = \begin{cases} 1: & \boldsymbol{Y} = \mathcal{T}\boldsymbol{x} \\ 0: & \boldsymbol{Y} \neq \mathcal{T}\boldsymbol{x}. \end{cases} \tag{2}$$

Substituting (2) in (1), the MAP estimation becomes

$$\hat{\boldsymbol{X}} = \arg\min_{\boldsymbol{X} \in \mathcal{S}} \{ -\log \Pr(\boldsymbol{X}) \} \tag{3}$$

where $\mathcal{S} = \{\boldsymbol{z}: \boldsymbol{Y} = \mathcal{T}\boldsymbol{z}\}$.

For image *a priori* model $(\Pr(\boldsymbol{X}))$, we use the MRF. According to the Hammersley–Clifford theorem, every MRF should have a Gibbs distribution [10] in the following form:

$$\Pr(\boldsymbol{X} = x) = \frac{1}{Z} \exp\left\{ -\sum_{c \in C} V_c(x) \right\} \tag{4}$$

where Z is a normalizing constant, $V_c(\cdot)$ is called a potential function and is a function of a local group of pixels c called clique, and C denotes the set of all cliques throughout the image. The potential function characterizes the relationship between a group of pixels by assigning larger costs to configurations of pixels which are less likely to occur. The choice of the potential function impacts substantially the performance of the image model. The function $\sum_{c \in C} V_c(x)$ should be convex in order to have an easily obtainable global minimum. Otherwise, local minima will be present, and the function must then be minimized via a computationally expensive technique such as simulated annealing. Commonly, the potential functions are selected to be of the form

$$\sum_{c \in C} V_c(x) = \sum_{c \in C} \rho\left(\boldsymbol{d}_c^t \boldsymbol{x}\right) \tag{5}$$

where $\boldsymbol{d_c}$ is a coefficient vector, $\boldsymbol{x}$ is the vector of pixels in the clique c, and $\rho(\cdot)$ is the cost function. The convexity property of the cost function insures that the $\sum_{c \in C} V_c(x)$ remains convex. The coefficients $\boldsymbol{d_c}$ are usually selected such that $\boldsymbol{d}_c^t \boldsymbol{x}$ provides an approximation of the first or second derivative of the image at each pixel. For the special case of $\rho(x) = x^2$, the model is called a Gauss–Markov random field (GMRF).

III. Proposed Method

As mentioned earlier, our method consists of two stages. The purpose of the first stage of our method is to obtain an initial estimate of the missing block using information from the previous frame. If the missing block has been intra-coded, the initial estimate is set to zero. If the missing block has been inter-coded, and its motion vector is received correctly (e.g., because the encoded

[2] See [8] for a more detailed discussion of the transform.

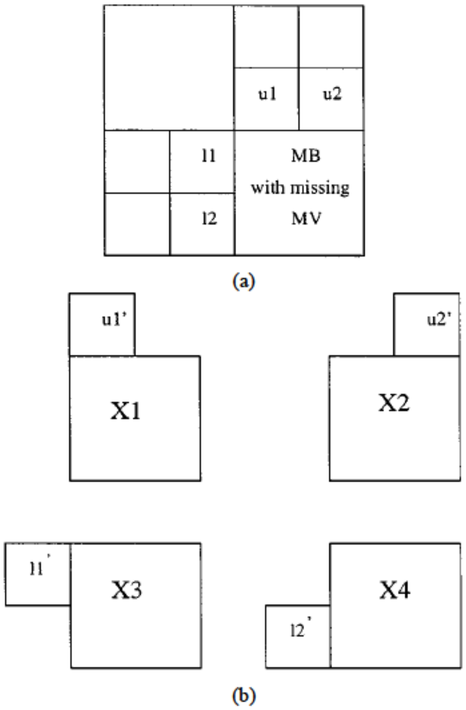

Fig. 1. (a) The four subblocks, (b) their corresponding subblocks in the previous frame, and the blocks connected to them.

data have been partitioned), then the estimate is set to the corresponding motion compensation block. When the motion vector is not available, we find an estimate of the missing block such that image continuity inside the block and across its boundaries is preserved [11]. To do this, subblocks adjacent to the missing block, i.e., $u1, u2, l1$, and $l2$ in Fig. 1(a), are considered. First, for each of these subblocks, a corresponding subblock in the previous frame is determined. The corresponding subblock is found by searching a small area around the point corresponding to the center of each of the subblocks $u1, u2, l1, l2$ in the previous frame. The sum of absolute differences is used as the measure of similarity. The four subblocks $u1, u2, l1, l2$ and their corresponding subblocks $u1', u2', l1', l2'$ in the previous frame are shown in Fig. 1(a) and (b). For example, $u1'$ is the subblock corresponding to $u1$. Then, four blocks, namely $\boldsymbol{X1}, \boldsymbol{X2}, \boldsymbol{X3}$, and $\boldsymbol{X4}$, which are connected to $u1', u2', l1'$, and $l2'$ (respectively), are determined. To obtain an initial estimate of the missing block that smoothly connects to the rest of the image, a block from the above four blocks that minimizes the squared sums of border errors, between the estimated block and its adjacent above and left blocks, is selected. Thus,

$$\hat{\boldsymbol{X}} = \arg \min_{\boldsymbol{X1},\boldsymbol{X2},\boldsymbol{X3},\boldsymbol{X4}} \epsilon^2$$

where

$$\epsilon^2 = \epsilon_T^2 + \epsilon_L^2. \tag{6}$$

Each of the border errors ϵ_T and ϵ_L is defined in terms of pixels by

$$\epsilon_T^2 = \|(\hat{\boldsymbol{x}}_t - \boldsymbol{p}_{\mathrm{Tb}})\|^2 \quad \text{and} \quad \epsilon_L^2 = \|(\hat{\boldsymbol{x}}_l - \boldsymbol{p}_{\mathrm{Lr}})\|^2$$

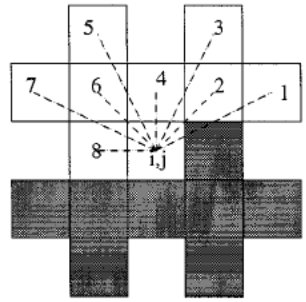

Fig. 2. A pixel, its clique c, and the eight directions. The complement of the clique c' is the dark area.

where the vector $\boldsymbol{p}_{\mathrm{Tb}}$ consists of the bottom line pixels of the block to the top of the missing block, and $\boldsymbol{p}_{\mathrm{Lr}}$ consists of the right column pixels of the block to the left of missing block. The vectors $\hat{\boldsymbol{x}}_t$ and $\hat{\boldsymbol{x}}_l$ are those elements of the estimated block $\hat{\boldsymbol{X}}$ that correspond to the pixels in its top row and left column, respectively. To ensure the online applicability (e.g., concealment while decoding the blocks in raster scan), only subblocks to the left and above the missing block are used. If the condition of online concealment is relaxed, the same method can be extended to include the subblocks of the blocks to the right and below of the missing block.

The estimation of motion of the subblocks has a high computational overhead which can possibly introduce unacceptable requirements on decoder. The computational overhead can be reduced if the search for displacement of each of the subblocks is restricted to a set of candidate motion vectors. This is a decoder option and can be used to trade performance against computational complexity. The set consists of the motion vector of the block corresponding to the missing block in the previous frame, the motion vectors of available neighboring blocks, the median of the motion vectors of available neighboring blocks, the average of the motion vectors of available neighboring blocks, and the zero motion vector [2].

In the second stage of our proposed error-concealment method, the information around each missing block is used to refine the initial estimate. This stage is based on a MAP estimator. We consider a GMRF model with an eight-pixel clique around each pixel as the image *a priori* model (see Fig. 2). The reason for selecting an eight-pixel clique in the manner shown in Fig. 2 will become clear later. The potential function of (5) is selected as

$$\sum_{c \in C} V_c(x) = \sum_i \sum_j \sum_{(k,l) \in c} w_{i,j \to k,l}\, \rho(x_{i,j} - x_{k,l}) \tag{7}$$

where c is the clique, $w_{i,j \to k,l}$ is the weight assigned to the difference between the values of the pixel in position (i,j) and the pixel in its clique at position (k,l) and $\rho(x) = x^2$.

Combining the MAP estimator of (3) with the image model [(4) and (7)], the restoration of missing data eventuates in the following minimization problem:

$$\hat{\boldsymbol{X}} = \min_{x_{i,j}} \sum_{i,j \in \mathcal{M}} \sum_{k,l \in c} w_{i,j \to k,l}\, \rho(x_{i,j} - x_{k,l}) \tag{8}$$

where $\mathcal{M}$ is the set of all missing pixels in the frame. Since $\rho(x) = x^2$ is a convex function, the above minimization problem yields a unique global solution. In fact, the estimated value of a pixel is given by

$$\hat{x}_{i,j} = \sum_{(k,l)\in c\cup c'} w_{i,j\to k,l} \quad x_{k,l} \Big/ \sum_{(k,l)\in c\cup c'} w_{i,j\to k,l} \tag{9}$$

where c is the clique and c' is its complement shown in Fig. 2.

In our adaptive MRF model, the weight corresponding to the difference between a pixel and one of the pixels in its clique [$w_{i,j\to k,l}$ in (9)] is selected adaptively, based on the likelihood of an edge in the direction of the subject pair of pixels. The rationale behind this selection is to weigh more the difference between the pixels in that direction which will cause the values of the pixels in that direction to get closer to each other. The likelihood of edges in each of the eight directions is computed using blocks around the missing block. In this way, the available information in a larger area is exploited without increasing the order of the MRF model (which increases dramatically the computational complexity). To determine the likelihood of edges in each of the eight directions, edges in the blocks surrounding the missing block whose directions imply that they pass through the missing block are determined. That is, we first obtain

$$\begin{aligned} g_x &= x_{i+1,j-1} - x_{i-1,j-1} + 2x_{i+1,j} \\ &\quad - 2x_{i-1,j} + x_{i+1,j+1} - x_{i-1,j+1} \end{aligned}$$

and

$$\begin{aligned} g_y &= x_{i-1,j+1} - x_{i-1,j-1} + 2x_{i,j+1} \\ &\quad - 2x_{i,j-1} + x_{i+1,j+1} - x_{i+1,j-1} \end{aligned}$$

for every pixel in the blocks to the left, right, top, and bottom of the missing block. The magnitude and angular direction of the edge at pixel (i,j) are

$$G = \sqrt{g_x^2 + g_y^2} \quad \text{and} \quad \theta = \arctan\left(\frac{g_y}{g_x}\right)$$

where θ determines if the edge at pixel (i,j) passes through the missing block. Since there are eight pixels in the clique, the value of θ is rounded to one of the eight directions equally spaced in the range from 0° to 180°. There is a counter c_m $(m = 1, 2, \ldots, 8)$ for each of the eight directions. If the extension of an edge at pixel (i,j) belonging to one of the neighboring blocks passes through the missing area, the counter for that particular direction is incremented by the amount of G. Since the employed edge detector is sensitive to the image noise, the values of c_ms, $m = 1, \ldots, 8$ are compared to a threshold. If any of them is less than the threshold, it is set to zero.

There are eight pixels in the clique of each pixel, and eight directions for the detected edges. Each pixel in the clique of a pixel corresponds to a direction. In our proposed method, $w_{i,j\to k,l}$ is selected based on the edge counter of the direction corresponding to (i,j) and (k,l), i.e.,

$$w_{i,j\to k,l} = \beta c_m \tag{10}$$

where β is a constant and c_m is the counter corresponding to direction m, and direction m corresponds to the direction formed by (i,j) and (k,l). Finally, the second stage of the proposed error-concealment method can be summarized as follows. 1) Determine the edges in the neighboring blocks and assign them to eight equally spaced directions. Compute the counter for each direction. 2) Use (10) to find a set of weights for each missing block. 3) Use (9) to obtain an estimate of each missing pixel employing the weights obtained in the previous step. 4) Iteratively reestimate the missing pixels using (9) until convergence. Note that since the cost function is convex, convergence is guaranteed.

For intra-coded blocks, the missing pixel values are initially set to zero and then the MAP estimator, using the adaptive MRF and the data of the neighboring blocks, is applied. For missing inter-coded blocks, an initial estimate are obtained using information from the previous frame. Estimates of the prediction errors are found using the adaptive MRF model along with the prediction error signals of the neighboring blocks. The estimated values are then added to the initial estimates. Since the prediction error signal consists mostly of high-frequency components, the MAP estimation stage (i.e., the second stage) will improve the video reconstruction quality, especially around the edges.

IV. Computational Complexity

The numbers provided below for computational complexity correspond to a block and subblock of sizes 16 × 16 and 8 × 8, respectively. The computational load of the first stage of the proposed method consists of those computations required in 1) estimating the motion, and 2) computing the error of (6). For motion estimation, we used a spiral search method using an area of 16 × 16. This requires approximately 197 000 operations. If the search for the displacement of each of the subblocks is restricted to the set of candidate motion vectors explained in Section III, the required number of operation will reduce to 6100. The number of operations required to find the total error of (6) for four blocks is approximately 380.

For the second stage, the computational load consists of those computations required in adapting the weights of the MRF model, which are approximately 18 400 operations for a missing block, assuming four available neighboring blocks. The estimation of missing pixels using (9) is an iterative procedure which, for each iteration, requires approximately 30 operations for each missing pixel, or 16 × 16 × 30 operations for each missing block. On average, 80 iterations are required for the algorithm to converge for a block.

Clearly, the computational load of our error-concealment method is quite reasonable. Our simulation experiments confirm that the run time of our method is indeed acceptable. In fact, real-time decoding (e.g., 10 frames/s for QCIF video sequences) is still possible on a Pentium 300 MHz PC.

V. Experimental Results

Although the method proposed in this paper is general and can be applied to any block-based video compression method, H.263 is used as our video coding framework. QCIF (176 × 144) video sequences at a temporal resolution of 10 frames/s are

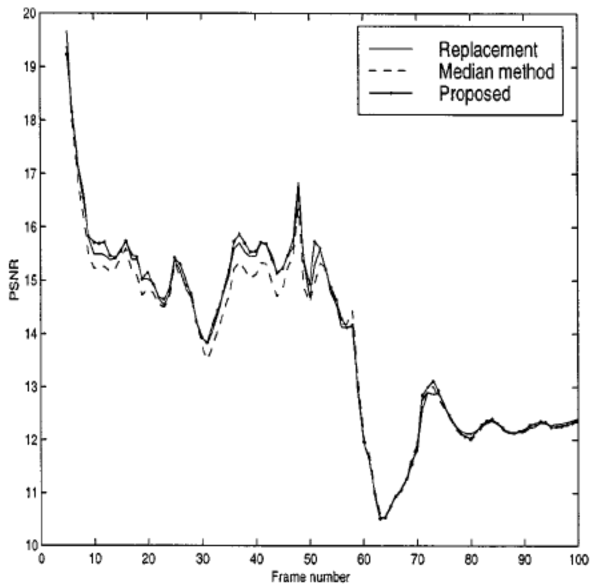

Fig. 3. PSNR values for image sequence Foreman with 10% GOB missing for different concealment methods.

coded at 64 kbps and then decoded in the presence of slice/GOB errors. The size of the blocks is 16×16, that of the subblock is 8×8.

To simulate the channel errors, the following tasks are performed. Coded video information is first grouped into packets, where each video packet consists of the coded data of a GOB or a slice. The video packets are then multiplexed with audio information according to the H.223 standard. We used an H.223 multiplexing simulator which receives video packets, simulates audio traffic, applies errors to the multiplexed bit stream according to an error pattern stored in a file, and outputs the packets [12]. The error pattern that we employed corresponds to a mobile channel [13]. Burst errors will most likely not corrupt two consecutive video packet, since audio packets are inserted between them. The erroneous bit stream is decoded such that the effect of errors will appear as missing slices/GOB's.

For inter-coded blocks, our error concealment consists of obtaining an initial estimate of the missing block using information from the previous frame, and finding an estimate of the prediction error using the adaptive MRF model. We compare the performance of the above method with two other methods: 1) a replacement method, where a missing block is replaced with the same block in the previous frame; and 2) a median method, where the motion vector of a missing block is set to the median of motion vectors of blocks to the left, above and above-right of it, and the estimated motion vector is used to obtain a motion compensation block which serves as the concealment of the missing block. Here, we consider only the case where a GOB is placed in each of the video packets. Shown in Fig. 3 are the concealment PSNR values for the three methods mentioned above for the video sequence Foreman with a 10% packet (GOB) loss rate. As can be seen, for video sequences with a large amount of motion like Foreman, the performance of the proposed method and the median method are close to each other, and both are better than the replacement method. Although the performance

Fig. 4. An inter-coded frame of the sequence Foreman concealed by the (a) replacement, (b) median, and (c) proposed methods.

of the above methods is comparable in terms of PSNR, the video sequences obtained using the proposed method are more visually appealing than the ones obtained using the other two methods. Fig. 4(a), (b), and (c) shows the reconstructed video frame (inter-coded) of the video sequence Foreman. As can be seen in the images, the replacement method does not generate good results, especially around the nose area. Moreover, the concealment result of the proposed method is clearly better than that of the other two methods.

For intra-coded blocks, we compare the performance of our method (which consists of estimating the missing block using the adaptive MRF model) to that of four other efficient methods: 1) a MAP estimator using a nonadaptive GMRF as the *a priori* model where each missing pixel is basically set to the average of the pixels around it, 2) a suboptimal version of the MAP estimator where a missing pixel is set to the median of pixels around it [8], 3) the deterministic method proposed in [4], and 4) the deterministic method proposed in [3]. Although there are other deterministic methods that are likely to outperform the above two methods, for example, the ones proposed in [6] and [7], we restrict our comparison to the above two deterministic methods. In fact, a direct comparison of our method, which is statistical, to deterministic methods is difficult because of different design approaches. Our main aim here is to compare our method to previously proposed statistical methods.

In this experiment, we employ different video sequences and different packetization schemes. In the first set of simulation experiments, we assign a slice which consists of one block, to a packet. Fig. 5(a) shows a frame of the image sequence Foreman encoded and decoded using an H.263 compliant coder. Fig. 5(b)–(g), shows (respectively) the same frame: 1) missing approximately 20% of the packets (blocks), 2) reconstructed using the nonadaptive GMRF model, 3) using the suboptimal MRF model, 4) using the method proposed in [4], 5) using the method proposed in [3], and 6) using our proposed adaptive MRF error concealment method. Clearly, our proposed method performs best in reconstruction quality, particularly in retrieving the edges and in the areas of the frame that correspond to adjacent missing blocks.

In the next set of simulation experiments, we assign the coded data of a GOB, which consists of 11 blocks, to each of the video packets. Fig. 6(a) shows a frame of the video sequence Foreman with two missing packets (GOB's), at approximately 18% loss rate. Fig. 6(b) and (c) shows the error- concealment result obtained using the nonadaptive GMRF model and suboptimal MRF models, respectively. Fig. 6(d) and (e) shows the result obtained using the methods proposed in [4] and [3], respectively. Fig. 6(f) shows the result obtained using our adaptive

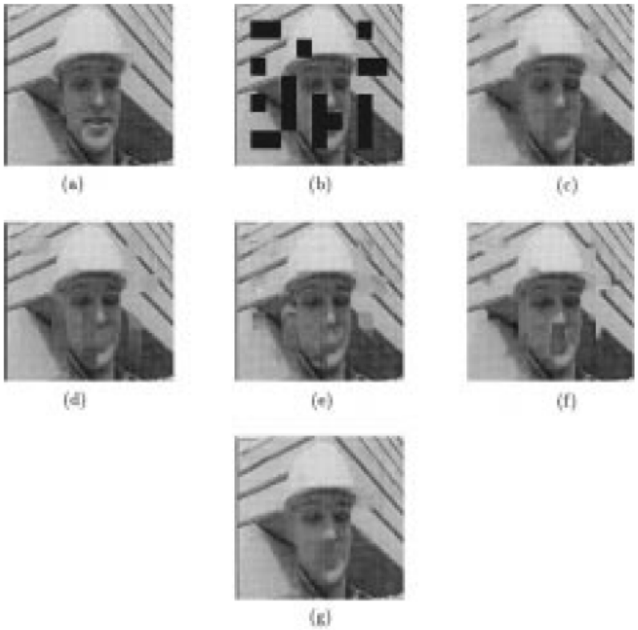

Fig. 5. A frame from the video sequence Foreman: (a) original; (b) missing blocks; reconstructed using (c) a nonadaptive GMRF model; (d) a suboptimal MRF model; (e) the method proposed in [4]; (f) the method proposed in [3]; and (g) our adaptive MRF model.

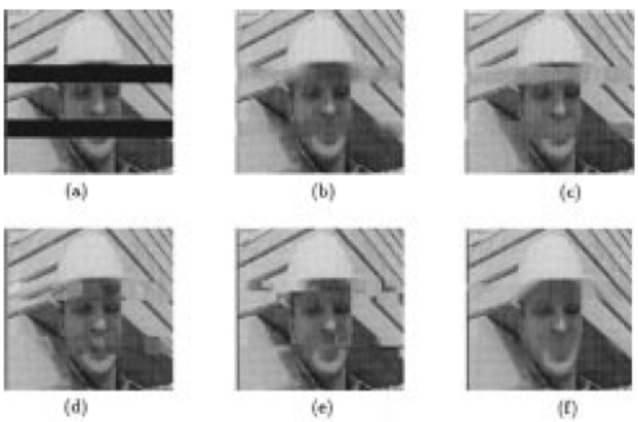

Fig. 6. A frame from the video sequence Foreman: (a) with missing GOB's, reconstructed using (b) a nonadaptive GMRF model; (c) a suboptimal MRF model; (d) the method proposed in [4]; (e) the method proposed in [3]; and (f) our adaptive MRF model.

MRF model. By comparing the figures, the superior performance of our method becomes obvious. This performance advantage is demonstrated in the blocks that contain edges. For a quantitative evaluation, Table I provides the PSNR values of the above concealment methods (for both packetization cases) for the video sequence Foreman. The table demonstrates that our method outperforms the other methods by at least 2 dB.

Finally, we compare the computational complexity of our method in terms of required number of additions and multiplications to the methods proposed in [4] and [3]. The total number of additions and multiplications required for restoring a 16 $\times$ 16 block are approximately 4500 and 20 000 using the method proposed in [4] and our method, respectively. The method proposed in [3] requires approximately 650 000 additions and multiplications to restore a 16 $\times$ 16 block. Therefore, the number of computations of our method is 4.4:1 larger than that of [4] and 1:0.03 less than those of [3]. However, our method outperforms [4] substantially in terms of the quality of the reconstructed images. This can be seen in Figs. 5 and 6. This quality improvement justifies, in most applications, the additional computational complexity.

TABLE I
PSNR (dB) COMPARISON OF DIFFERENT METHODS FOR THE VIDEO SEQUENCE FOREMAN

Method	No loss	20% Block loss	GOB missing
GMRF	32.1	26.3	25.3
Suboptimal MRF	32.1	25.4	23.1
Method proposed in [4]	32.1	24.3	23.8
Method proposed in [3]	32.1	25.2	24.1
Adaptive MRF	32.1	28.6	27.2

VI. CONCLUSION

This paper introduces a new method of error concealment for block-based coded video communications over error-prone channels. The proposed method employs an adaptive MRF as the image *a priori* model in a MAP estimation paradigm. The adaptation enables the estimation procedure to incorporate more information without increasing the order of the MRF. The proposed concealment method is able to restore missing blocks located in smooth and low-frequency areas, as well as in high-frequency and edge portions of a video frame. Our concealment method achieves very good computation–performance tradeoffs, and outperforms previously proposed MRF-based concealment methods.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their insightful comments and constructive suggestions. The authors would also like to gratefully acknowledge the help and support provided by G. Cote, their colleague in the Signal Processing and Multimedia Group at the University of British Columbia.

REFERENCES

[1] M. Ghanbari and V. Seferidis, "Cell loss concealment in ATM video codecs," *IEEE Trans. Circuit System Video Technol.*, vol. 3, pp. 238–247, June 1993.
[2] W. M. Lam, A. R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, vol. V, 1993, pp. 417–420.
[3] 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.
[4] S. S. Hemami and T. H. Y. Meng, "Transform coded image reconstruction exploiting interblock correlations," *IEEE Trans. Image Processing*, vol. 4, pp. 1023–1027, July 1995.
[5] H. Sun and W. Kwok, "Concealment of damaged block transform coded images using projection onto convex sets," *IEEE Trans. Image Processing*, vol. 4, pp. 470–477, Apr. 1995.
[6] W. Kwok and H. Sun, "Multi-directional interpolation for spatial error concealment," *IEEE Trans. Consumer Electron.*, vol. 39, pp. 455–460, Aug. 1993.
[7] 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, pp. 648–665, June 1999.

[8] P. Salama, N. Shroff, E. J. Coyle, and E. J. Delp, "Error concealment in encoded video streams," in *Signal Recovery Techniques for Image and Video Compression and Transmission*, N. P. Galatsanos and A. K. Katsaggelos, Eds. Boston, MA: Kluwer Academic, 1998.
[9] R. Talluri, "Error resilient video coding in the MPEG-4 standard," *IEEE Commun. Mag.*, vol. 26, pp. 112–119, June 1998.
[10] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 689–691, 1984.
[11] S. Shirani, F. Kossentini, and R. Ward, "Reconstruction of motion vector missing macroblocks in H.263 encoded video transmission over lossy networks," in *Proc. Int. Conf. Image Processing*, vol. III, Oct. 1998, pp. 487–491.
[12] G. Sullivan, "A Simple Video Packet Mux Simulator Program for Video Streams in H.263/M Using AL3 Mux of H.223 Annex B," ITU-T Study Group 15, Video Coding Expert Group, Doc. Q15-F-16, Nov. 1996.
[13] Ericsson Sweden, "WCDMA Error Patterns at 64 kb/s," ITU-T Study Group 16, Multimedia Terminals and Systems Expert Group, Cannes, France, June 1998.

Shahram Shirani (S'88–M'89) received the B.Sc. degree in electrical engineering from Esfahan University of Technology, Esfahan, Iran, in 1989 and the M.Sc. degree in biomedical engineering from Amirkabir University of Technology, Tehran, Iran in 1994. Currently, he is pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, B.C., Canada.

From 1994 to 1996, he was with the Department of Electrical Engineering, University of Tehran. His research interests include image and video compression, video communications, signal processing and ultrasonic imaging.

Faouzi Kossentini (S'89–M'89–SM'98) received the B.S., M.S., and Ph.D. degrees from the Georgia Institute of Technology, Atlanta, in 1989, 1990, and 1994, respectively.

During 1995, he was a Research Scientist at Nichols Research Corporation, Huntsville, AL. Since January 1996, he has been an Assistant Professor and then an Associate Professor in the Department of Electrical and Computer Engineering, University of British Columbia, vancouver, B.C., Canada, where he is involved in research in the areas of signal processing, communications, and multimedia. He has coauthored more than 100 journal papers, conference papers, book chapters and patents.

Dr. Kossentini has been active as a Voting Member, and recently as head of delegation, of the Canadian delegate to ISO/IEC JTC1/SC29, which is responsible for the standardization of coded representation of audiovisual, multimedia, and hypermedia information. In particular, he has participated in most current JBIG/JPEG and MPEG-4 standardization activities. He has also participated in most current ITU-T low bit rate video coding standardization activities. He has served as a technical area coordinator and member of the technical program committee of ICIP-1997, and as a Member of the technical program committee of ISCAS-1999. He is also the Vice General Chairman of ICIP-2000. Dr. Kossentini is currently an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING. He is also an Associate Editor for the IEEE TRANSACTIONS ON MULTIMEDIA.

Rabab Ward (F'98) received the B.Sc. degree in electrical engineering from University of Cairo, Egypt, in 1966 and the M.Sc. and the Ph.D. in electrical and computer engineering from University of California, Berkeley, in 1969 and 1972, respectively.

She is a Professor in the Electrical and Computer Engineering Department, University of British Columbia, Vancouver, B.C., Canada, and is the Director of the Center for Integrated Computer Systems Research there. Her research interests are mainly in the areas of signal processing and image processing. She has made contributions in the areas of signal detection, image encoding, compression, recognition, restoration and enhancement, and their applications to infant cry signals, cable TV, HDTV, medical images, and astronomical images. She holds five patents related to cable television picture monitoring, measurement and noise reduction.

Dr. Ward is a fellow of the EIC and the Royal Society of Canada. She is currently serving as the General Chair of ICIP'2000 to be held in Vancouver.