

# Packet Video Error Concealment With Auto Regressive Model

Yongbing Zhang, Xinguang Xiang, Debin Zhao, Siwe Ma, *Student Member, IEEE*, and Wen Gao, *Fellow, IEEE*

**Abstract**—In this paper, auto regressive (AR) model is applied to error concealment for block-based packet video coding. In the proposed error concealment scheme, the motion vector for each corrupted block is first derived by any kind of recovery algorithms. Then each pixel within the corrupted block is replenished as the weighted summation of pixels within a square centered at the pixel indicated by the derived motion vector in a regression manner. Two block-dependent AR coefficient derivation algorithms under spatial and temporal continuity constraints are proposed respectively. The first one derives the AR coefficients via minimizing the summation of the weighted square errors within all the available neighboring blocks under the spatial continuity constraint. The confidence weight of each pixel sample within the available neighboring blocks is inversely proportional to the distance between the sample and the corrupted block. The second one derives the AR coefficients by minimizing the summation of the weighted square errors within an extended block in the previous frame along the motion trajectory under the temporal continuity constraint. The confidence weight of each extended sample is inversely proportional to the distance toward the corresponding motion aligned block whereas the confidence weight of each sample within the motion aligned block is set to be one. The regression results generated by the two algorithms are then merged to form the ultimate restorations. Various experimental results demonstrate that the proposed error concealment strategy is able to improve both the objective and subjective quality of the replenished blocks compared to other methods.

**Index Terms**—Auto regressive model, confidence weight, error concealment, spatial continuity constraint, temporal continuity constraint, video coding.

## I. INTRODUCTION

STATE-OF-THE-ART video coding standard H.264/AVC [1] significantly outperforms the previous coding stan-

Manuscript received May 5, 2010; revised August 19, 2010 and November 18, 2010; accepted December 2, 2010. Date of publication March 17, 2011; date of current version January 6, 2012. This work was supported by the National Science Foundation of China, under Grant 60736043, the Joint Funds of National Science Foundation of China, under Grant U0935001, and the Major State Basic Research Development Program of China's 973 Program, under Grant 2009CB320905. This paper was recommended by Associate Editor E. Steinbach.

Y. Zhang was with the Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China. He is now with the Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China (e-mail: ybzhang@mail.tsinghua.edu.cn).

X. Xiang and D. Zhao are with the Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China (e-mail: xgxiang@jdl.ac.cn; dbzhao@jdl.ac.cn).

S. Ma and W. Gao are with the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China (e-mail: swma@jdl.ac.cn; wgao@pku.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2011.2130450

dards, such as MPEG-1 [2], H.262/MPEG-2 [3], and H.263 [4]. Although the highly efficient redundancy removing techniques in spatial and temporal domains leads to the success of H.264/AVC, the highly compressed bit stream is susceptible to transmission errors for error-prone networks. Consequently, packet errors are unavoidable, which will severely degrade the display quality at the decoder side.

Error resilience [5] and error concealment [6] are two major techniques to combat the visual quality degradation caused by noisy channels during transmission. Error resilience is used to combat the transmission errors by adding redundant information at the encoder with the penalty of decreasing the compression efficiency. On the contrary, error concealment is a post-processing technique which conceals the errors utilizing the correctly received information at the decoder side without modifying source and channel coding schemes. In this paper, we mainly study the techniques of error concealment. According to the information utilized, error concealment algorithms can be categorized into spatial approaches, temporal approaches and hybrid approaches that combine the former two ones.

Spatial approaches reconstruct the corrupted macroblock by utilizing the correctly decoded surrounding pixels under smoothness constraint. Wang *et al.* proposed a spatial error concealment method by minimizing the first-order derivative-based smoothness measure [7]. To suppress the induced blurring artifacts, the second-order derivatives were considered in [8]. Although such a smoothness constraint achieves good results for the flat regions, it may not be satisfied in the areas with high frequency edges. To tackle this shortcoming, an edge-preserving algorithm [9] was proposed to interpolate the missing pixels. In [10], smooth and edge areas were efficiently recovered based on selective directional interpolation. In [11], an orientation adaptive interpolation scheme derived from the pixel wise statistical model was proposed. In addition, a spatial error concealment method based on a Markov random field (MRF) model was proposed in [12]. And in [13], a multiframe spatial error concealment considering the error propagation and incorporating the idea of least squares (LS) estimation was proposed.

Spatial approaches may yield better performance than temporal ones in scenes with high motion, or after a scene change [14]. However, they may not restore the detail textures of corrupted blocks [15]. In this case, the information from the past frames (temporal approaches) may improve the quality of corrupted blocks.

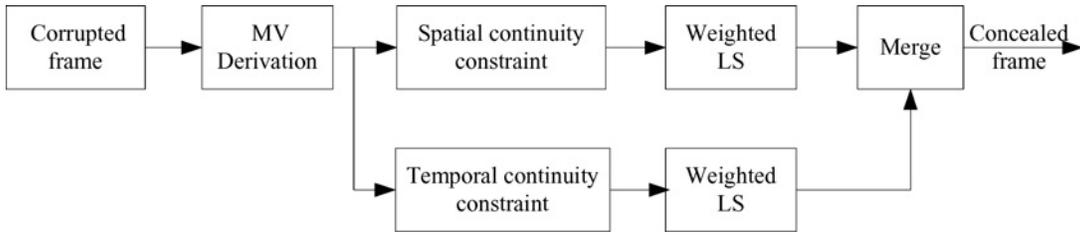


Fig. 1. Proposed AR model based error concealment.

Temporal approaches restore the corrupted blocks by exploiting temporal correlation between successive frames. An important issue in temporal approaches is to find the most suitable substitute blocks from the previous frames, i.e., selecting the optimal motion vectors (MVs) for the corrupted blocks. If the MV of the corrupted block is available at the decoder, it can be utilized directly to motion-compensate the corrupted block. However, when the MV is also lost, it has to be re-estimated. Many pioneering works have been done on recovering the corrupted MVs. Haskell and Messerschmitt [16] took zero MV, the MV of the collocated block in the reference frame, and the average or the median of the MVs from the spatially adjacent blocks as candidate MVs for the lost blocks. Chen *et al.* [17] proposed a side match criterion taking advantage of the spatial contiguity and inter-pixel correlation of image to select the best-fit replacement among the MVs of spatially contiguous candidate blocks. The well known boundary matching algorithm (BMA) proposed in [18] selected the MV that minimizes the total variation between the internal boundary and the external boundary of the reconstructed block as the optimal one to recover the corrupted block. There are also some more sophisticated algorithms [12], [13], [19]–[23] to obtain better replacements for the corrupted blocks. For example, a means of estimating the missing MV based on the use of MRF models [12], an algorithm using the multiframe recovery principle and the boundary smoothness property [13], a vector rational interpolation scheme [19], a bilinear motion field interpolation algorithm [20], a Lagrange interpolation algorithm [21], and a dynamic programming algorithm [22], [23] were proposed for error concealment. In addition, some model aided error concealment algorithms were also proposed. For instance, a projection of convex set (POCS) based error concealment for packet video was proposed in [24]. And in [25], a mixture of principal components was proposed for error concealment.

Besides spatial and temporal approaches, hybrid approaches combining the former two methods have been proposed recently to obtain better replenishment results. For instance, in temporal error concealment, the compensated block can be further improved by spatial smoothing at its edges to make it conform to the neighbors. In [26], the coding mode and block loss patterns are clustered into four groups, and the weighting between spatial and temporal smoothness constraints depends on the group. In [27], a priority-driven region matching algorithm to exploit the spatial and temporal information was proposed. And in [28], a spatio-temporal boundary matching algorithm (STBMA) and partial differential equation (PDE) were proposed.

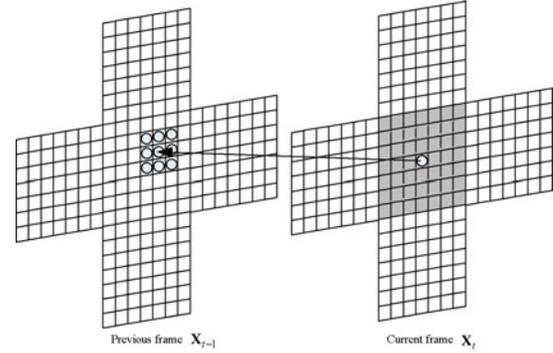


Fig. 2. Auto-regressive model.

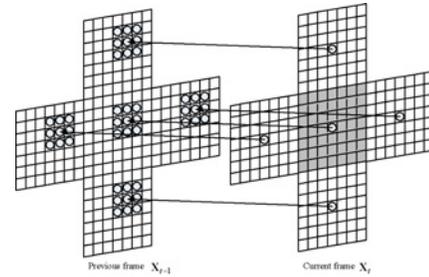


Fig. 3. Spatial continuity constraint.

The most aforementioned error concealment algorithms usually interpolate the previous frames into half or quarter-pel accuracy before deriving the best MVs for the corrupted blocks, due to the fact that the motion of objects between adjacent frames may be of fractional-pel accuracy. The interpolation filters used are usually separable and the coefficients are fixed. Such methods achieve good performance for isotropic regions; however, they may result in poor performance for anisotropic local image structures. To inhibit the inferiority of the separable and fixed interpolation filters, an auto-regressive (AR) model based error concealment is proposed in this paper. It is well known that AR has long been employed to model regular stationary random process [29]. For such a process, its statistical properties have been well studied. For example, Kokaram *et al.* used AR to detect and interpolate “dirt” areas [30], [31]. Efstratiadis and Katsaggelos employed AR to perform motion estimation [32]. Li developed a backward adaptive video encoder exploiting the prediction property of AR model [33].

In our formulation, each pixel within the corrupted block is replenished as the weighted summation of pixels within a square, which is centered at the pixel indicated by the MV

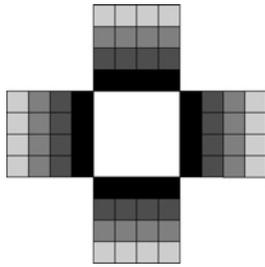


Fig. 4. Probabilistic confidence magnitude within neighboring blocks.

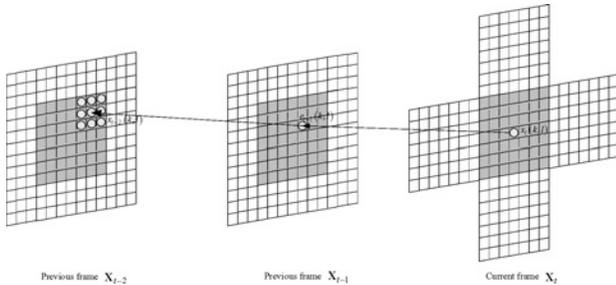


Fig. 5. Temporal continuity constraint.

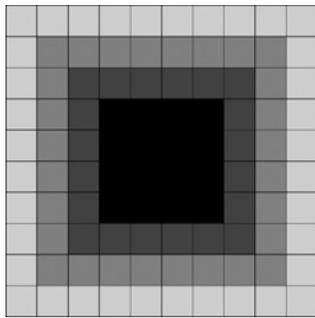


Fig. 6. Probabilistic confidence magnitude within an extended  $4 \times 4$  block.

with integer-pel accuracy in a regression manner. Two block-dependent AR coefficient derivation algorithms are proposed to achieve better performance. The first one is the coefficient derivation algorithm under the spatial continuity constraint, in which the summation of the weighted square errors within the available neighboring blocks is minimized. The confidence weight of each sample within the available neighboring blocks is inversely proportional to the distance between the sample and the corrupted block. The second coefficient derivation algorithm is under the temporal continuity constraint, where the summation of the weighted square errors within an extended block in the previous frame along the motion trajectory is minimized. The confidence weight of the extended sample is inversely proportional to the distance toward the corresponding motion aligned block whereas the confidence weight of each sample within the motion aligned block is set to be one. The interpolations generated by the weights derived under these two constraints are then merged to form the ultimate concealing results.

The proposed AR model based error concealment scheme is the extension of our previous works [34], [35]. In [34],

only the spatial continuity constraint is applied and equal confidence weight is assigned for each training pixel sample. In [35], both spatial and temporal continuity constraints are applied; however, the experimental results and discussions are not enough. For example, in [35], the experimental result is only compared with BMA and our previous work in [34], and the probability confidence effects are not fully discussed. In addition, the merging operation is just simply averaging the results obtained by spatial and temporal continuity constraints in [35], whereas the merging depends on the estimated MV in this paper. Actually, the proposed AR model based error concealment scheme can be considered as a post-processing for any MV recovery scheme (e.g., BMA, the methods in [12] and [13] and STBMA) by adaptively adjusting the AR coefficients according to the local image properties. Our goal is to obtain appropriate AR coefficients, whereas other inter frame error concealments (e.g., BMA, the methods in [12] and [13] and STBMA) are aimed at generating more accurate MVs by certain criterions. Various experimental results demonstrate that the proposed error concealment strategy is able to not only increase the peak signal-to-noise ratio (PSNR) but also improve the visual quality of concealing blocks compared to other methods.

The remainder of this paper is organized as follows. Section II describes the AR model based error concealment scheme. Sections III and IV present the coefficient derivations under the spatial and temporal continuity constraints respectively. Experimental results and analysis conducted on various sequences are given in Section V. Finally, a brief conclusion is provided in Section VI.

## II. AUTO-REGRESSIVE MODEL-BASED ERROR CONCEALMENT

The proposed AR model based error concealment scheme is illustrated in Fig. 1. For each corrupted block, the corresponding MV is first derived by any kind of recovery algorithms (such as BMA and STBMA). The AR model is then applied to the corrupted block along the derived motion trajectory. To improve the quality of concealed frames, two AR coefficient derivation algorithms under the spatial continuity and temporal continuity constraints are performed respectively, utilizing the weighted LS algorithm. The interpolation results generated by the two sets of coefficients are then merged to form the ultimate restorations.

Fig. 2 illustrates the AR model employed by the proposed error concealment. It is noted that the AR model is applied along the motion trajectory. For each corrupted pixel, the corresponding pixel along the motion trajectory with integer-pel accuracy in the previous reconstructed frame is first found, and then all the pixels within a square centered at the corresponding motion aligned pixel are combined in a linear regression form. The linear regression can be expressed as

$$\hat{x}_t(i, j) = \sum_{k=-R}^R \sum_{l=-R}^R \alpha(k, l) x_{t-1}(i + dy + k, j + dx + l) \quad (1)$$

where  $\hat{x}_t(i, j)$  represents the corrupted pixel located at  $(i, j)$  within the current frame  $\mathbf{X}_t$ ,  $R$  represents the range of the AR model,  $(dx, dy)$  represents the estimated MV with the integer-pel accuracy,  $x_{t-1}(i, j)$  represents the pixel within the previous reconstructed frame  $\mathbf{X}_{t-1}$ , and  $\alpha(k, l)$  represents the desired coefficients.

The main merit of the proposed AR model based error concealment, compared with other motion compensated schemes, is that it is able to adapt spatially to local orientation structure. In traditional motion compensated error concealment algorithms, the corrupted block is replaced by the corresponding block indicated by the estimated MV in the previous frames. The best MV is usually found by minimizing the matching errors between the neighboring blocks and the candidate ones in the fractional interpolated version of the previous frames. Such methods achieve good performance for isotropic local regions; however, inferior results may be perceived for anisotropic local image structures, since interpolation filters are separable and fixed along vertical and horizontal directions. In contrast, in the proposed AR model, the interpolation is non-separable and can be along arbitrary direction. Besides the interpolation coefficients can be varied from one local region to the others. This results in strong preservation of details in the restored image and greatly improves the performance of error concealment.

Define  $\Psi_{k,l,R}$  as an operator that extracts a patch of a fixed size (centered at  $(k, l)$  and with  $(2R+1) \times (2R+1)$  pixels) from an image, the expression  $\Psi_{k,l,R}\mathbf{X}_{t-1}$  ( $\mathbf{X}_{t-1}$  is represented as a vector by lexicographic ordering) results with a vector of length  $(2R+1)^2$  being the extracted patch. Consequently, the linear regression in (1) can also be expressed as

$$\hat{x}_t(i, j) = \Psi_{i+dy, j+dx, R}\mathbf{X}_{t-1}\alpha^T \quad (2)$$

where  $\alpha$  represents the coefficient vector of the AR model and  $(dy, dx)$  represents the MV with integer-pel accuracy. The summed square error between the corrupted and the actual pixels is

$$\begin{aligned} \varepsilon^2 &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (x_t(i, j) - \hat{x}_t(i, j))^2 \\ &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (x_t(i, j) - \Psi_{i+dy, j+dx, R}\mathbf{X}_{t-1}\alpha^T)^2 \end{aligned} \quad (3)$$

where  $N$  represents the width and height of the corrupted block. To minimize  $\varepsilon^2$ , the first derivative of  $\varepsilon^2$  to  $\alpha$  should be zero according to the LS algorithm, that is

$$\begin{aligned} \frac{\partial \varepsilon^2}{\partial \alpha} &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left( (\Psi_{i+dy, j+dx, R}\mathbf{X}_{t-1})^T (\Psi_{i+dy, j+dx, R}\mathbf{X}_{t-1}) \right) \alpha^T \\ &\quad - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_t(i, j) (\Psi_{i+dy, j+dx, R}\mathbf{X}_{t-1})^T = 0. \end{aligned} \quad (4)$$

By solving the above equation, we get the optimal coefficients as

$$\alpha^T = \left[ \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left( (\Psi_{i+dy, j+dx, R}\mathbf{X}_{t-1})^T (\Psi_{i+dy, j+dx, R}\mathbf{X}_{t-1}) \right) \right]^{-1} \left[ \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_t(i, j) (\Psi_{i+dy, j+dx, R}\mathbf{X}_{t-1})^T \right]. \quad (5)$$

However, since the actual pixel  $x_t(i, j)$  is not available at the decoder side, we cannot directly obtain the AR coefficients according to (5). Instead, we have to estimate the AR coefficients by exploring the spatial and temporal correlations of the corrupted block with its available spatial and temporal neighboring pixels.

### III. AR COEFFICIENT DERIVATION UNDER SPATIAL CONTINUITY CONSTRAINT

Pixels within adjacent blocks have a high possibility of belonging to the same object, which can be reflected by the phenomenon that adjacent blocks possess similar motion trends. Such a property is termed as spatial continuity constraint in this paper, based on which a set of AR coefficients for the corrupted block can be derived. It is stated that AR coefficients can reflect the MV of each block to some extent [33] and due to the piecewise stationary characteristics of natural image [36], we assume all the pixels within the corrupted block possess the same AR coefficients, just like all the pixels within the corrupted block have the same MV in the traditional motion compensated error concealment method. If we use AR model to represent the motion between successive frames, spatial continuity constraint can be interpreted as that all the pixels within the available neighboring blocks have the same AR coefficients as those within the corrupted block in this paper.

As shown in Fig. 3, under spatial continuity constraint each pixel within the corrupted block and its neighboring blocks can be regressed by the corresponding pixels within the previous reconstructed frame utilizing the same AR coefficients. Let  $\mathbf{B}_t$  be a neighboring block of the current block within the current frame, i.e.,  $\mathbf{B}_t \subset \mathbf{X}_t$ . In addition, let  $b_t(m, n)$  be an arbitrary pixel in  $\mathbf{B}_t$ , i.e.,  $b_t(m, n) \in \mathbf{B}_t$ .  $b_t(m, n)$  can be represented by the regression function of  $\mathbf{X}_{t-1}$  and  $\alpha$  as

$$\hat{b}_t(m, n) = \Psi_{m+dy, n+dx, R}\mathbf{X}_{t-1}\alpha^T \quad (6)$$

where  $\alpha$  represents the AR coefficients. According to (5), the solution of  $\alpha$  can be computed by the LS method.

It is noted that during the coefficient derivation process, different training samples should be assigned different probabilistic confidences so as to achieve better performance. For example, pixels that are closer to the corrupted block or with similar texture should be assigned larger probabilistic confidences. Define the corresponding probabilistic confidence of  $b_t(m, n)$  under the spatial continuity constraint is  $w_\alpha(m, n)$ , with  $0 \leq w_\alpha(m, n) \leq 1$  and  $\sum_{(m,n) \in \mathbf{B}_t} w_\alpha(m, n) = 1$ , the

optimal  $\alpha$  under probabilistic confidences should be

$$\hat{\alpha} = \arg \min_{\alpha} \sum_{(m,n) \in B_t} \|(b_t(m,n) - \hat{b}_t(m,n)) w_{\alpha}(m,n)\|^2. \quad (7)$$

Since the correlation between pixels decreases with the increase of their distance,  $w_{\alpha}(m,n)$  is set to be inversely proportional to the distance between  $b_t(m,n)$  and the corrupted block, that is

$$w_{\alpha}(m,n) = \frac{1}{S} \begin{cases} \frac{1}{N-m}, & \text{if } b_t(m,n) \in \text{upper block} \\ \frac{1}{N-n}, & \text{if } b_t(m,n) \in \text{left block} \\ \frac{1}{m+1}, & \text{if } b_t(m,n) \in \text{lower block} \\ \frac{1}{n+1}, & \text{if } b_t(m,n) \in \text{right block} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

with  $S = S_L + S_R + S_A + S_B$ , where

$$S_L = \begin{cases} \sum_{n=0}^{N-1} \frac{1}{N-n}, & \text{if left block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$S_R = \begin{cases} \sum_{n=0}^{N-1} \frac{1}{n+1}, & \text{if right block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$S_A = \begin{cases} \sum_{m=0}^{N-1} \frac{1}{N-m}, & \text{if upper block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$S_B = \begin{cases} \sum_{m=0}^{N-1} \frac{1}{m+1}, & \text{if lower block is available} \\ 0, & \text{otherwise.} \end{cases}$$

Here  $N$  represents the width and height of the corrupted block.

Fig. 4 graphically shows the probabilistic confidence magnitudes within a  $4 \times 4$  block given by (8) as an example. The white block represents the corrupted block, which is surrounded by its four neighboring blocks. Each neighboring block is composed of 15 pixels whose gray value is inverse proportional to the magnitude  $w_{\alpha}(i,j)$  of the sixteen samples. It can be observed that much larger probabilistic confidence values are assigned for the pixels closer to the corrupted block than those for the pixels farther toward the corrupted block.

It is noted that Figs. 3 and 4 exhibit a universal case, where the four neighboring blocks are all available to train AR coefficients. Actually, there are two cases. In the first case, if any of the neighboring blocks are correctly received, the correctly received neighboring blocks are utilized to train AR coefficients of the corrupted block. In the second case, if all the neighboring blocks are lost, the already concealed neighboring blocks are utilized to train AR coefficients of the corrupted block.

By setting the first derivative of the weighted errors in (7) to zero, the AR coefficients under spatial continuity constraint are computed as

$$\alpha^T = [\mathbf{C}_P^L + \mathbf{C}_P^R + \mathbf{C}_P^A + \mathbf{C}_P^B]^{-1} [\mathbf{D}_P^L + \mathbf{D}_P^R + \mathbf{D}_P^A + \mathbf{D}_P^B] \quad (9)$$

where

$$\mathbf{C}_P^L = \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} (\vec{P}_{\alpha}(m,n) * C)^T C, & \text{if left block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{C}_P^R = \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} (\vec{P}_{\alpha}(m,n) * C)^T C, & \text{if right block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{C}_P^A = \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} (\vec{P}_{\alpha}(m,n) * C)^T C, & \text{if upper block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{C}_P^B = \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} (\vec{P}_{\alpha}(m,n) * C)^T C, & \text{if lower block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{D}_P^L = \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} w_{\alpha}(m,n) x_t(m,n) C^T, & \text{if left block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{D}_P^R = \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} w_{\alpha}(m,n) x_t(m,n) C^T, & \text{if right block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{D}_P^A = \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} w_{\alpha}(m,n) x_t(m,n) C^T, & \text{if upper block is available} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{D}_P^B = \begin{cases} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} w_{\alpha}(m,n) x_t(m,n) C^T, & \text{if lower block is available} \\ 0, & \text{otherwise} \end{cases}$$

with  $\vec{P}_{\alpha}(m,n) = \underbrace{[w_{\alpha}(m,n), w_{\alpha}(m,n), \dots, w_{\alpha}(m,n)]}_{(2R+1)^2}$  and  $C =$

$$(\Psi_{m+dy, n+dx, R} \mathbf{X}_{t-1}).$$

The operator “\*” represents element by element multiplication of two vectors. With the obtained AR coefficient  $\alpha$ , the corrupted block is restored according to (2).

#### IV. AR COEFFICIENT DERIVATION UNDER TEMPORAL CONTINUITY

Besides spatial continuity constraint, video sequence also has temporal continuity constraint, which can be proved by the observation that the same object among adjacent frames is usually threaded by the same motion trajectory. Similar to spatial continuity constraint, we assume all the pixels within the corrupted block possess the same AR coefficients. The temporal continuity constraint in this paper can be interpreted as that all the pixels within the corrupted block have the same AR coefficients as those within the corresponding motion aligned block in the previous frame. Utilizing temporal continuity constraint, we can derive another set of AR coefficients, which is shown as Fig. 5. It is noted that we extend the motion aligned block, shown as the gray pixels as well as the pixels surrounding them in the previous frame  $X_{t-1}$  in Fig. 5, to find sufficient training samples for the derivation of AR coefficients. Define  $\mathbf{E}_{t-1}$  to be the extended motion aligned

TABLE I  
AVERAGE PSNR RESULTS OF EACH TEST SEQUENCE WITH AND WITHOUT THE PROPOSED PROBABILISTIC CONFIDENCE (PLR = 10%)

Sequence		QP	PSNR (dB)					
			BMA+AR					
			Spatial		Temporal		Combined	
		Uniform Weight	Proposed Weight	Uniform Weight	Proposed Weight	Uniform Weight	Proposed Weight	
QCIF	Mobile	24	30.86	30.96	31.17	31.46	31.55	31.62
		28	29.54	29.56	30.04	29.99	30.04	29.95
	Paris	24	29.84	30.00	30.59	30.72	30.85	31.04
		28	29.56	29.74	30.21	30.27	30.54	30.72
	Suzie	24	35.29	35.65	35.48	35.51	35.83	36.01
		28	34.06	34.19	34.03	34.07	34.38	34.51
<b>Average</b>			<b>31.53</b>	<b>31.68</b>	<b>31.92</b>	<b>32.00</b>	<b>32.20</b>	<b>32.31</b>
CIF	Foreman	24	31.49	31.63	31.66	31.70	32.24	32.33
		28	30.86	30.99	30.86	30.86	31.47	31.58
	Mobile	24	28.37	28.48	28.51	28.63	28.86	29.05
		28	27.74	27.76	27.98	28.02	28.26	28.28
	Flower	24	28.37	28.43	27.97	28.08	28.57	28.69
		28	27.59	27.60	27.33	27.47	27.84	27.94
<b>Average</b>			<b>29.07</b>	<b>29.15</b>	<b>29.05</b>	<b>29.13</b>	<b>29.54</b>	<b>29.65</b>
Sequence		QP	STBMA+AR					
			Spatial		Temporal		Combined	
			Uniform Weight	Proposed Weight	Uniform Weight	Proposed Weight	Uniform Weight	Proposed Weight
QCIF	Mobile	24	30.62	30.78	31.53	31.56	31.25	31.39
		28	29.36	29.44	29.67	29.64	29.67	29.75
	Paris	24	30.77	31.00	31.73	31.89	31.68	31.90
		28	29.90	30.05	30.80	30.98	30.71	30.83
	Suzie	24	35.34	35.53	35.39	35.44	35.83	36.04
		28	34.04	34.19	34.11	34.10	34.34	34.47
<b>Average</b>			<b>31.67</b>	<b>31.83</b>	<b>32.21</b>	<b>32.27</b>	<b>32.25</b>	<b>32.40</b>
CIF	Foreman	24	31.12	31.30	31.37	31.39	31.73	31.83
		28	30.63	30.89	30.90	30.91	31.19	31.30
	Mobile	24	28.48	28.59	28.97	29.06	28.88	29.07
		28	27.85	27.91	28.35	28.46	28.28	28.41
	Flower	24	29.13	29.25	28.56	28.73	29.21	29.30
		28	28.09	28.18	27.57	27.76	28.07	28.24
<b>Average</b>			<b>29.22</b>	<b>29.35</b>	<b>29.29</b>	<b>29.39</b>	<b>29.56</b>	<b>29.69</b>

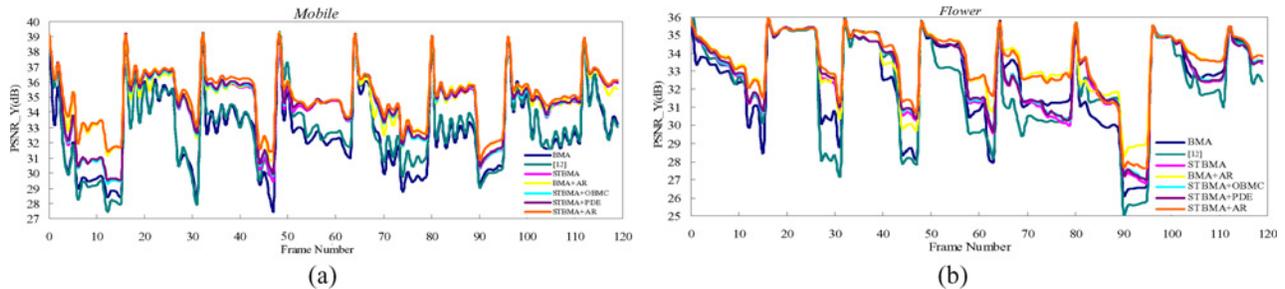


Fig. 7. PSNR performance comparison versus the frame number while the PLR is 10% (each slice contains one row of MBs). (a) Mobile (CIF). (b) Flower (CIF).

TABLE II  
AVERAGE PSNR RESULTS OF EACH QCIF TEST SEQUENCE USING DIFFERENT METHODS

Sequence	QP	PLR = 5%, PSNR (dB)										
		BMA	[12]	STBMA	STBMA +OBMC	STBMA +PDE	BMA+AR			STBMA+AR		
							Spatial	Temporal	Combined	Spatial	Temporal	Combined
Football	16	35.85	34.92	36.06	36.15	36.08	36.01	36.52	36.42	36.13	36.48	36.51
	24	33.67	33.45	33.82	33.88	33.89	33.96	33.85	33.94	34.00	33.85	34.02
	28	32.27	32.35	32.57	32.64	32.60	32.32	32.59	32.42	32.54	32.64	32.69
	40	27.06	26.72	26.92	26.95	26.93	27.11	27.07	27.14	26.99	26.97	26.99
Mobile	16	38.33	38.52	39.51	39.54	39.53	39.96	39.59	40.12	40.35	40.54	40.57
	24	34.92	35.05	35.46	35.48	35.49	35.69	35.67	35.85	35.67	35.85	35.84
	28	32.41	32.37	32.68	32.70	32.69	32.74	32.78	32.79	32.76	32.83	32.82
	40	24.26	24.23	24.27	24.27	24.27	24.28	24.28	24.28	24.28	24.28	24.29
Paris	16	38.29	38.39	38.95	38.99	38.95	39.01	39.86	40.72	39.48	41.46	41.42
	24	35.27	35.74	36.05	36.10	36.06	36.17	36.65	36.63	36.25	36.66	36.75
	28	33.31	32.84	33.33	33.36	33.33	33.64	34.00	34.16	33.62	34.04	34.12
	40	25.36	25.45	25.36	25.37	25.36	25.38	25.60	25.59	25.39	25.60	25.60
Suzie	16	43.26	43.28	44.07	44.11	44.08	44.20	44.03	44.17	44.24	44.10	44.26
	24	38.99	38.81	39.16	39.18	39.17	39.23	39.15	39.27	39.36	39.25	39.28
	28	36.72	36.81	36.94	36.96	36.95	37.00	37.04	37.07	37.02	37.07	37.06
	40	30.58	30.53	30.57	30.58	30.59	30.58	30.60	30.61	30.59	30.61	30.61
<b>Average</b>		<b>33.78</b>	<b>33.72</b>	<b>34.11</b>	<b>34.14</b>	<b>34.12</b>	<b>34.21</b>	<b>34.33</b>	<b>34.45</b>	<b>34.29</b>	<b>34.51</b>	<b>34.55</b>
Sequence	QP	PLR = 10%, PSNR (dB)										
		BMA	[12]	STBMA	STBMA +OBMC	STBMA +PDE	BMA+AR			STBMA+AR		
							Spatial	Temporal	Combined	Spatial	Temporal	Combined
Football	16	25.45	24.95	25.62	25.78	25.65	26.03	25.74	26.15	26.02	25.57	26.14
	24	24.88	24.66	25.20	25.16	25.20	25.46	24.80	25.47	25.43	24.66	25.53
	28	24.28	23.82	23.97	24.40	23.97	24.94	24.36	25.04	24.49	24.23	24.54
	40	22.91	22.59	23.37	23.53	23.36	23.60	23.06	23.74	23.59	22.62	23.62
Mobile	16	30.25	30.82	32.14	32.21	32.18	32.48	33.00	33.32	32.23	33.03	33.05
	24	29.18	29.69	30.65	30.67	30.69	30.96	31.46	31.62	30.78	31.56	31.39
	28	28.44	28.72	29.46	29.47	29.47	29.56	29.99	29.95	29.44	29.64	29.75
	40	23.27	23.33	23.47	23.50	23.48	23.55	23.56	23.57	23.53	23.53	23.54
Paris	16	30.43	31.16	32.31	32.36	32.32	31.77	32.87	32.93	32.17	33.53	33.56
	24	28.95	29.71	30.76	30.85	30.78	30.00	30.72	31.04	31.00	31.89	31.90
	28	28.53	29.16	29.98	30.03	29.99	29.74	30.27	30.72	30.05	30.98	30.83
	40	23.91	24.17	24.55	24.62	24.57	24.44	24.67	24.71	24.52	24.80	24.78
Suzie	16	36.68	36.30	37.51	37.62	37.54	37.66	37.28	37.95	37.91	37.56	37.99
	24	35.25	34.57	35.53	35.56	35.54	35.65	35.51	36.01	35.53	35.44	36.04
	28	33.75	33.38	34.16	34.21	34.18	34.19	34.07	34.51	34.19	34.10	34.47
	40	29.66	29.49	29.62	29.64	29.63	29.53	29.45	29.63	29.50	29.42	29.59
<b>Average</b>		<b>28.49</b>	<b>28.53</b>	<b>29.27</b>	<b>29.35</b>	<b>29.28</b>	<b>29.35</b>	<b>29.43</b>	<b>29.78</b>	<b>29.40</b>	<b>29.54</b>	<b>29.80</b>
Sequence	QP	PLR = 20%, PSNR (dB)										
		BMA	[12]	STBMA	STBMA +OBMC	STBMA +PDE	BMA+AR			STBMA+AR		
							Spatial	Temporal	Combined	Spatial	Temporal	Combined
Football	16	23.22	22.90	23.07	23.11	23.10	23.96	23.46	23.90	23.90	23.70	23.92
	24	23.13	22.81	23.00	23.13	23.03	23.71	23.27	23.73	23.45	23.17	23.61
	28	22.60	21.92	22.35	22.46	22.36	23.14	22.91	23.14	22.75	22.75	22.81
	40	22.16	21.73	21.98	22.12	21.99	22.24	22.25	22.63	22.00	22.10	22.32
Mobile	16	27.19	27.72	29.27	29.27	29.28	29.58	29.61	30.34	29.80	30.06	30.28
	24	26.54	27.24	28.28	28.40	28.35	28.66	28.58	29.06	28.83	28.91	29.23
	28	26.09	26.63	27.35	27.52	27.46	27.76	27.60	28.21	27.73	27.79	28.01
	40	22.37	22.37	22.82	22.88	22.83	22.98	22.81	22.91	22.99	22.82	22.94
Paris	16	29.36	29.71	30.76	30.91	30.77	30.78	31.81	32.26	31.18	32.15	32.31
	24	28.11	28.30	29.53	29.61	29.47	29.50	30.40	30.60	29.91	30.98	31.04
	28	27.49	27.83	28.51	28.61	28.52	28.58	29.30	29.72	28.80	29.77	29.81
	40	23.37	23.52	23.83	23.92	23.85	23.31	24.19	23.92	23.37	24.36	23.92
Suzie	16	34.99	34.98	36.14	36.17	36.16	36.57	34.86	36.31	36.64	35.14	36.41
	24	33.78	33.70	34.53	34.74	34.54	34.82	33.73	34.80	34.86	33.72	34.79
	28	33.05	32.81	33.42	33.48	33.44	33.67	32.61	33.59	33.71	32.60	33.57
	40	29.05	28.74	29.02	29.09	29.04	29.00	28.73	28.89	28.98	28.64	28.80
<b>Average</b>		<b>27.03</b>	<b>27.06</b>	<b>27.74</b>	<b>27.84</b>	<b>27.76</b>	<b>28.02</b>	<b>27.88</b>	<b>28.38</b>	<b>28.06</b>	<b>28.04</b>	<b>28.36</b>

TABLE III  
AVERAGE PSNR RESULTS OF EACH CIF TEST SEQUENCE USING DIFFERENT METHODS

Sequence	QP	PLR = 5%, PSNR (dB)										
		BMA	[12]	STBMA	STBMA +OBMC	STBMA+PDE	BMA+AR			STBMA+AR		
							Spatial	Temporal	Combined	Spatial	Temporal	Combined
Foreman	16	38.40	38.59	39.89	39.99	40.11	40.21	39.97	40.24	40.24	40.25	40.29
	24	36.40	36.33	37.05	37.04	37.10	37.17	37.13	37.15	37.17	37.26	37.30
	28	34.95	34.79	35.38	35.39	35.40	35.50	35.31	35.56	35.50	35.25	35.45
	40	29.74	29.73	29.93	29.93	29.95	29.95	29.99	30.07	29.95	30.00	30.08
Mobile	16	34.93	34.70	38.04	38.10	38.11	38.39	38.04	38.79	38.65	38.57	38.97
	24	32.69	32.93	34.48	34.52	34.59	34.81	34.78	34.99	35.04	35.01	35.25
	28	31.30	31.35	32.62	32.58	32.65	32.81	32.78	32.98	32.92	32.96	32.97
	40	24.71	24.72	25.06	25.06	25.06	25.07	25.04	25.10	25.09	25.11	25.13
Flower	16	36.46	35.88	37.46	37.61	37.49	38.71	38.24	38.69	39.11	38.75	39.11
	24	34.34	33.95	35.10	35.21	25.15	35.84	35.43	25.86	36.13	35.82	36.13
	28	32.58	32.12	32.98	33.06	33.01	33.57	33.37	33.62	33.55	33.49	33.60
	40	24.88	24.82	24.92	24.94	24.93	25.00	25.01	25.03	24.97	25.00	24.98
<b>Average</b>		<b>32.62</b>	<b>32.49</b>	<b>33.58</b>	<b>33.62</b>	<b>32.80</b>	<b>33.92</b>	<b>33.76</b>	<b>33.18</b>	<b>34.03</b>	<b>33.96</b>	<b>34.11</b>
Sequence	QP	PLR = 10%, PSNR (dB)										
		BMA	[12]	STBMA	STBMA +OBMC	STBMA+PDE	BMA+AR			STBMA+AR		
							Spatial	Temporal	Combined	Spatial	Temporal	Combined
Foreman	16	31.33	30.94	31.80	31.82	31.90	32.70	32.53	33.44	32.44	31.51	32.45
	24	30.45	29.90	31.03	31.05	31.14	31.63	31.70	32.33	31.30	31.39	31.83
	28	29.70	29.47	30.56	30.59	30.67	30.99	30.86	31.58	30.89	30.91	31.30
	40	26.64	26.57	27.36	27.38	27.42	27.44	27.56	27.89	27.44	27.55	27.83
Mobile	16	26.59	26.34	28.84	28.98	28.96	29.22	29.45	29.85	29.50	29.93	29.89
	24	25.95	25.63	28.03	28.15	28.13	28.48	28.63	29.05	28.59	29.06	29.07
	28	25.46	25.19	27.53	27.65	27.59	27.76	28.02	28.28	27.91	28.46	28.41
	40	22.11	22.28	23.29	23.36	23.31	23.48	23.47	23.63	23.47	23.52	23.58
Flower	16	26.66	26.14	28.10	28.35	28.21	29.04	28.73	29.22	29.83	29.12	29.90
	24	26.33	25.85	28.02	28.06	28.05	28.43	28.08	28.69	29.25	28.73	29.30
	28	25.78	25.27	26.89	27.13	27.04	27.60	27.47	27.94	28.18	27.76	28.24
	40	22.86	22.19	23.14	23.22	23.24	23.54	23.47	23.62	23.54	23.38	23.53
<b>Average</b>		<b>26.66</b>	<b>26.31</b>	<b>27.88</b>	<b>27.98</b>	<b>27.97</b>	<b>28.36</b>	<b>28.33</b>	<b>28.79</b>	<b>28.53</b>	<b>28.44</b>	<b>28.78</b>
Sequence	QP	PLR = 20%, PSNR (dB)										
		BMA	[12]	STBMA	STBMA +OBMC	STBMA+PDE	BMA+AR			STBMA+AR		
							Spatial	Temporal	Combined	Spatial	Temporal	Combined
Foreman	16	29.13	28.90	30.21	30.22	30.25	30.60	29.63	30.91	30.48	29.64	30.66
	24	28.79	28.38	29.57	29.60	29.63	29.83	29.28	30.27	29.91	29.04	29.92
	28	28.44	27.87	29.27	29.26	29.32	29.25	29.14	29.76	29.43	29.02	29.79
	40	25.83	25.72	26.55	26.62	26.64	26.61	26.39	26.96	26.63	26.18	26.86
Mobile	16	24.39	24.22	26.84	26.95	26.87	27.13	27.11	27.72	27.51	27.65	27.93
	24	23.88	23.67	26.28	26.35	26.33	26.34	26.32	26.81	26.78	26.91	27.18
	28	23.44	23.28	25.65	25.73	25.68	25.70	25.70	26.32	26.01	26.30	26.50
	40	20.93	21.24	22.40	22.49	22.44	22.24	22.25	22.55	22.35	22.42	22.52
Flower	16	24.66	23.88	26.03	26.28	26.13	26.76	26.26	26.81	27.40	26.70	27.46
	24	24.30	23.64	25.70	25.91	25.79	26.19	25.71	26.30	26.97	26.19	27.05
	28	23.84	23.25	24.91	25.18	24.94	25.64	25.28	25.85	26.12	25.65	26.27
	40	21.82	21.14	22.27	22.57	22.29	22.64	22.46	22.65	22.92	22.59	22.97
<b>Average</b>		<b>24.95</b>	<b>24.60</b>	<b>26.31</b>	<b>26.43</b>	<b>26.36</b>	<b>26.58</b>	<b>26.29</b>	<b>26.91</b>	<b>26.88</b>	<b>26.52</b>	<b>27.09</b>

block in the closest previous frame  $\mathbf{X}_{t-1}$ , i.e.,  $\mathbf{E}_{t-1} \subset X_{t-1}$ . And define  $e_{t-1}(k, l)$  to be an arbitrary pixel within  $\mathbf{E}_{t-1}$ , i.e.,  $e_{t-1}(k, l) \in \mathbf{E}_{t-1}$ . As shown in Fig. 5, for each corrupted pixel  $x_t(k, l)$ , the corresponding motion aligned pixel  $e_{t-1}(k, l)$  in the extended block is first found, and then the corresponding pixel  $x_{t-2}(k, l)$  in the second closest reconstructed frame is also found by the same MV. Apparently,  $e_{t-1}(k, l)$  can be regressed by the pixels within a square neighborhood which is centered at  $x_{t-2}(k, l)$  as

$$\hat{e}_{t-1}(k, l) = \Psi_{k+dy, l+dx, R} \mathbf{X}_{t-2} \beta^T \quad (10)$$

where  $\beta$  represents the AR coefficients derived under the temporal continuity constraint. The derived coefficient  $\beta$  is then utilized to restore the corrupted pixel  $x_t(k, l)$ . Apparently, the solution of  $\beta$  should be the one that satisfy

$$\hat{\beta} = \arg \min_{\beta} \sum_{(k, l) \in \mathbf{E}_{t-1}} \|(e_{t-1}(k, l) - \hat{e}_{t-1}(k, l))\|^2. \quad (11)$$

However, this imposes the same probabilistic confidence on each training sample, which will limit the accuracy of the derived AR coefficients. To tackle such a problem, we assigned



Fig. 8. Error concealment results of the eighth frame over *Foreman* (CIF) at the PLR of 10%. (a) Original image. (b) Corrupted image. (c) Concealed image using BMA (34.221 dB). (d) Concealed image using STBMA (35.569 dB). (e) Concealed image using STBMA+PDE (35.568 dB). (f) Concealed image using STBMA+OBMC (35.929 dB). (g) Concealed image using the proposed AR model under temporal continuity constraint (36.292 dB). (h) Concealed image using the proposed AR model under temporal continuity constraint (36.292 dB). (i) Concealed image using the proposed AR model by combining spatial and temporal continuity constraints (36.308 dB).

appropriate probabilistic confidence for each sample within the extended block. That is to say, the optimal  $\beta$  should be

$$\hat{\beta} = \arg \min_{\beta} \sum_{(k,l) \in \mathbf{E}_{t-1}} \|(e_{t-1}(k,l) - \hat{e}_{t-1}(k,l)) w_{\beta}(k,l)\|^2 \quad (12)$$

where  $w_{\beta}(k,l)$  represents the probabilistic confidence of each training sample  $e_{t-1}(k,l)$ . For the samples located within the corresponding motion aligned block, the probabilistic confidence is set to be one; and for the samples located at the extended regions, the probabilistic confidence is defined to be inversely proportional to the distance toward the center of the extended block. To be more specific, the probabilistic confidence of each sample can be formulated as (13) at the bottom of the next page.

Here  $M$  represents the extended range and  $N$  represents the width and height of the corrupted block, respectively.

Fig. 6 depicts the probabilistic confidence magnitudes within an extended  $4 \times 4$  block as an example. It is noted that  $M = 3$ , and  $N = 4$  in Fig. 6. The sixteen black pixels

correspond to the motion aligned block, and all the remaining gray pixels correspond to the extended region. The gray value is inverse proportional to the probabilistic confidence of the corresponding sample.

According to the weighted LS, the closed-form solution of  $\beta$  should be

$$\beta^T = \left[ \sum_{(k,l) \in \mathbf{E}_{t-1}} \left( \left( \vec{P}_{\beta}(k,l) * (\Psi_{k+dy, l+dx, R} \mathbf{X}_{t-2}) \right)^T (\Psi_{k+dy, l+dx, R} \mathbf{X}_{t-2}) \right)^{-1} \right. \\ \left. \sum_{(k,l) \in \mathbf{E}_{t-1}} e_{t-1}(k,l) (\Psi_{k+dy, l+dx, R} \mathbf{X}_{t-2})^T w_{\beta}(k,l) \right] \quad (14)$$

where  $\vec{P}_{\beta}(k,l) = \underbrace{[w_{\beta}(k,l), w_{\beta}(k,l), \dots, w_{\beta}(k,l)]}_{(2R+1)^2}$ , and the operator “\*” represents element by element multiplication of two vectors.



Fig. 9. Error concealment results of the 50th frame over *Mobile* (CIF) at the PLR of 20%. (a) Original image. (b) Corrupted image. (c) Concealed image using BMA (29.627 dB). (d) Concealed image using STBMA (29.985 dB). (e) Concealed image using STBMA+PDE (29.997 dB). (f) Concealed image using STBMA+OBMC (30.066 dB). (g) Concealed image using the proposed AR model under spatial continuity constraint (31.475 dB). (h) Concealed image using the proposed AR model under temporal continuity constraint (32.516 dB). (i) Concealed image using the proposed AR model by combining spatial and temporal continuity constraints (32.610 dB).

$$w_{\beta}(k, l) = \frac{1}{S} \begin{cases} 1, & M \leq k, l < M+N \\ 1/\max(|k - (M+N/2)| - N/2+1, |l - (M+N/2)| - N/2+1), & 0 \leq k < M+N, 0 \leq l < M+N \\ 1/\max(|k - (M+N/2)| - N/2+1, |l - (M+N/2-1)| - N/2+1), & 0 \leq k < M+N, M+N/2 \leq l < M+N \\ 1/\max(|k - (M+N/2-1)| - N/2+1, |l - (M+N/2)| - N/2+1), & M+N/2 \leq k < M+N, 0 \leq l < M+N \\ 1/\max(|k - (M+N/2-1)| - N/2+1, |l - (M+N/2-1)| - N/2+1), & M+N/2 \leq k < M+N, M+N/2 \leq l < M+N \end{cases} \quad (13)$$

with

$$S = N^2 + \sum_{\substack{0 \leq k < M+N \\ 0 \leq l < M+N}} 1/\max(|k - (M+N/2)| - N/2+1, |l - (M+N/2)| - N/2+1) + \sum_{\substack{0 \leq k < M+N \\ M+N/2 \leq l < M+N}} 1/\max(|k - (M+N/2)| - N/2+1, |l - (M+N/2-1)| - N/2+1) \\ + \sum_{\substack{M+N/2 \leq k < M+N \\ 0 \leq l < M+N}} 1/\max(|k - (M+N/2-1)| - N/2+1, |l - (M+N/2)| - N/2+1) + \sum_{\substack{M+N/2 \leq k < M+N \\ M+N/2 \leq l < M+N}} 1/\max(|k - (M+N/2-1)| - N/2+1, |l - (M+N/2-1)| - N/2+1)$$

TABLE IV  
AVERAGE PSNR RESULTS OF EACH TEST SEQUENCE EXCLUDING ERROR PROPAGATION

Sequence		QP	PSNR (dB)											
			BMA	[12]	STBMA	STBMA+OBMC	STBMA+PDE	BMA+AR			STBMA+AR			
								Spatial	Temporal	Combined	Spatial	Temporal	Combined	
QCIF	Football	16	38.15	37.50	38.29	38.39	38.33	38.49	38.20	38.54	38.49	38.25	38.56	
		24	34.49	34.02	34.64	34.72	34.71	34.69	34.59	34.76	34.80	34.67	34.89	
		28	32.59	32.25	32.78	32.84	32.81	32.82	32.76	32.88	32.93	32.87	33.00	
		40	27.06	26.94	27.11	27.13	27.13	27.12	27.11	27.14	27.16	27.15	27.18	
	Mobile	16	40.29	40.08	40.94	40.98	40.97	41.08	41.37	41.33	41.30	41.60	41.61	
		24	35.27	34.99	35.67	35.71	35.70	35.73	35.88	35.87	35.81	35.97	35.96	
		28	32.58	32.41	32.78	32.79	32.79	32.84	32.91	32.91	32.87	32.96	32.95	
		40	24.50	24.50	24.53	24.53	24.53	24.54	24.54	24.55	24.54	24.54	24.55	
	Paris	16	41.01	41.32	41.85	41.92	41.87	41.70	42.29	42.20	42.12	42.66	42.58	
		24	36.40	36.50	36.95	36.99	36.98	36.76	37.02	36.98	37.05	37.32	37.29	
		28	33.88	33.99	34.19	34.22	34.20	34.07	34.29	34.26	34.27	34.46	34.43	
		40	25.69	25.70	25.74	25.75	25.75	25.74	25.79	25.77	25.74	25.79	25.79	
	Suzie	16	43.68	43.60	44.11	44.15	44.13	44.22	44.27	44.42	44.38	44.28	44.46	
		24	39.21	39.22	39.52	39.55	39.53	39.49	39.46	39.57	39.62	39.54	39.66	
		28	37.10	37.07	37.28	37.29	37.28	37.30	37.28	37.37	37.35	37.31	37.40	
		40	31.00	30.93	31.02	31.02	31.03	31.00	31.01	31.02	31.01	31.01	31.02	
<b>Average</b>			<b>34.56</b>	<b>34.44</b>	<b>34.84</b>	<b>34.87</b>	<b>34.86</b>	<b>34.85</b>	<b>34.92</b>	<b>34.97</b>	<b>34.97</b>	<b>35.02</b>	<b>35.08</b>	
CIF	Foreman	16	43.53	43.38	43.92	43.92	43.92	44.01	43.86	44.03	44.03	43.86	44.05	
		24	38.69	38.51	38.90	38.91	38.96	38.96	38.90	38.99	38.95	38.89	38.95	
		28	36.61	36.52	36.76	36.77	36.77	36.78	36.78	36.83	36.80	36.76	36.83	
		40	30.31	30.29	30.33	30.33	30.33	30.34	30.36	30.37	30.34	30.36	30.37	
	Mobile	16	41.70	41.56	42.61	42.63	42.63	42.74	42.72	42.73	42.77	42.98	42.99	
		24	36.32	36.12	36.91	36.93	36.92	36.89	37.05	37.06	37.02	37.15	37.16	
		28	33.59	33.46	33.99	34.00	34.00	34.02	34.12	34.14	34.07	34.16	34.17	
		40	25.22	25.22	25.29	25.30	25.30	25.30	25.32	25.34	25.30	25.32	25.32	
	Flower	16	42.86	42.73	43.54	43.59	43.57	43.56	43.60	43.64	43.86	43.85	43.91	
		24	37.34	37.22	37.86	37.89	37.88	37.86	37.89	37.92	38.06	38.05	38.10	
		28	34.54	34.38	34.89	34.91	34.90	34.92	34.92	34.96	35.02	35.02	35.05	
		40	25.64	25.61	25.72	25.72	25.72	25.72	25.73	25.74	25.74	25.74	25.75	
	<b>Average</b>			<b>35.53</b>	<b>35.42</b>	<b>35.89</b>	<b>35.91</b>	<b>35.91</b>	<b>35.93</b>	<b>35.94</b>	<b>35.98</b>	<b>36.00</b>	<b>36.01</b>	<b>36.05</b>

After having obtained the AR coefficients  $\alpha$  and  $\beta$ , we merge the two regression results as

$$\hat{x}_t(i, j) = \tau \bullet \Psi_{i+dy, j+dx, R} \mathbf{X}_{t-1} \alpha^T + (1 - \tau) \bullet \Psi_{i+dy, j+dx, R} \mathbf{X}_{t-1} \beta^T \quad (15)$$

where  $\tau$  is the merging factor, and it is computed as

$$\tau = \begin{cases} 1, & \text{if } \max(\text{abs}(mv[0]), \text{abs}(mv[1])) \geq 16 \\ 0.5, & \text{if } \max(\text{abs}(mv[0]), \text{abs}(mv[1])) = 0 \\ \max(\text{abs}(mv[0]), \text{abs}(mv[1]))/16, & \text{otherwise.} \end{cases} \quad (16)$$

Here  $mv[0]$  and  $mv[1]$  represent the horizontal and vertical components of the MV for the corrupted block selected by BMA or STBMA with quarter-pel accuracy.

It is noted that we will pad the corresponding pixels outside the boundary when the training area (in the current and/or reference frame) is close to the boundary of the frame. The padded pixels (if it is necessary) as well as those pixels close to the boundary can be simultaneously utilized during the coefficient derivation of the AR model. If there are no solutions in (9) or (14), we will use the traditional methods (BMA, method in [12], or STBMA) to restore the missing blocks accordingly.

## V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, various experiments were conducted to verify the performance of the proposed AR model based error concealment scheme. H.264/AVC reference software JM 10.0 is utilized to evaluate the proposed algorithm; however, it should be noted that the proposed algorithm can be extended to any block-based video compression scheme. We compare the performance of the proposed algorithm with the inter-frame error concealment schemes implemented in the reference software, which are based on the classical BMA [18], the method in [12], and STBMA [28]. We test the performance on seven video sequences: QCIF: *Mobile*, *Paris*, *Suzie*, *Football* and CIF: *Foreman*, *Mobile*, *Flower*. All the test sequences are encoded at 30Hz. The first 120 frames of each test sequence are encoded, where no B frames are utilized. Slice mode is enabled and no intra mode is used in P frames. Each row of MBs composes a slice and is transmitted in a separate packet. The packet loss rates (PLR) at 5%, 10% and 20% [37] are tested in the experiments. Quantization parameters (QP) are set to be 16, 24, 28 and 40, respectively.

In all the following experiments, parameter R is set to be 1, and parameter M is set to be 4 and 8 for QCIF and CIF sequences respectively.

In the experiments, we will show the effect of the probabilistic confidence first, and then we will give the comparisons of the proposed algorithm in terms of objective and subjective criteria, and finally we will present the computational complexity analysis.

#### A. Probability Confidence Effects

In this subsection, we provide comparisons of regression results under spatial and temporal continuity constraints as well as the merged results with and without probabilistic confidence, respectively. The encoding group of picture (GOP) is set to be IPPP..., where I frames are encoded every 16 frames. The transmission errors are assumed to only occur in P frames and the PLR is 10%. The MV is estimated by BMA and STBMA. The average PSNR results of the first 120 restored frames within each test sequence are provided in Table I. It is noted that BMA+AR and STBMA+AR represent that BMA and STBMA are utilized to obtain the MV of the corrupted block before AR model is applied, respectively. “Spatial” and “temporal” represent that AR coefficients are derived under spatial continuity constraint and temporal continuity constraint, respectively. And “combined” represents the combined results of “spatial” and “temporal.” “Uniform weight” represents that equal probabilistic confidence is assigned to all the training samples. “Proposed weight” represents that the proposed probabilistic confidence scheme is applied to all the training samples. It can be observed that for all the test sequences, the PSNR results get improved when the proposed probabilistic confidence scheme is applied except that there is a about 0.09 dB loss for *Mobile* (QCIF) when QP is 28. Especially for *Mobile* (CIF), when QP is set to be 24 and under BMA+AR, the PSNR gains are 0.11 dB, 0.12 dB and 0.19 dB for the spatial, temporal and combined methods, respectively. And for *Paris* (QCIF), when QP is set to be 24 and under BMA+AR, the PSNR gains are 0.16 dB, 0.13 dB and 0.19 dB for the spatial, temporal and combined methods, respectively. This mainly benefits from the fact that by assigning proper probabilistic confidence to different training samples, the accuracy of the derived AR coefficients can be improved.

#### B. Objective and Subjective Evaluation

In this subsection, we will give the subjective and objective comparison results. BMA, the passive error concealment in [12], STBMA+PDE, STBMA+ overlapped block motion compensation (OBMC) [38] and the proposed AR model with BMA and STBMA are utilized to restore the corrupted frames within each test sequence.

Tables II and III present the average PSNR results of each test sequence using different methods. The encoding GOP is set to be IPPP..., where I frames are encoded every 16 frames. The transmission errors are assumed to only occur in P frames. It is observed that [12] and BMA achieve similar performance for QCIF sequences, but [12] has some degradation for CIF sequences. STBMA outperforms BMA in most cases. This is because more spatial and temporal information is utilized in STBMA during block matching. Applying PDE and OBMC to

STBMA, the performance can be further improved; however the performance gain is not too much.

When applying the proposed AR model (BMA+AR or STBMA+AR), the error concealment performance is able to get a significant improvement. This is due to the fact that the AR model is able to adaptively adjust the coefficients according to the spatio-temporal coherence information. Another observation is that under BMA+AR combination, the proposed AR model is able to outperform STBMA. It strongly confirms that the proposed AR model is able to remedy the inference of BMA results brought by the inaccurate MVs. This is because although STBMA is able to achieve more accurate MVs, it does not guarantee better replenishing results for anisotropic regions, due to the fixed interpolation taps along horizontal and vertical directions. In contrast, the proposed AR model can perform the interpolation along arbitrary directions by properly tuning the coefficients. In addition, we also found that the combined results achieve better performance than those just under spatial or temporal continuity constraint. This is mainly attributed to the fact that combination operation is of higher ability to capture the variation properties of local image structure.

Fig. 7 shows the PSNR performance of *Mobile* (CIF) and *Flower* (CIF) versus the frame number. It is noted that both BMA+AR and STBMA+AR represent the combined results. We can see that STBMA has better performance than BMA and [12], while with the AR model (both BMA+AR and STBMA+AR), the performance can be further improved. Especially for the frames around the tenth frame in Fig. 7(a), the PSNR gains achieved by the proposed AR model (BMA+AR and STBMA+AR) are more than 2 dB compared with other competing methods (e.g., the BMA, [12], STBMA, STBMA+OBMC and STBMA+PDE). In addition, for the frames around the 105th frame in Fig. 7(b), the PSNR gains achieved by the proposed AR model (BMA+AR and STBMA+AR) are more than 1 dB compared with other competing methods.

To better represent the superior performance of the proposed AR model, we give the subjective quality comparisons for *Foreman* (CIF) and *Mobile* (CIF) in Figs. 8 and 9, respectively. Here the AR model is applied after the MV is found via STBMA. It is noted that there are consecutive slice errors in Fig. 8. For the corrupted MBs in the upper row, only the MVs of their upper neighboring MBs and the zero MVs are utilized to generate the optimal MVs in BMA or STBMA. Similarly, for the corrupted MBs in the lower row in the consecutive slice errors in Fig. 8, only the MVs of their lower neighboring MBs and the zero MVs are utilized to generate the optimal MVs in BMA or STBMA. In Fig. 8, for BMA, STBMA, STBMA+PDE, and STBMA+OBMC methods, we can easily observe the blocking artifacts caused by motion, as shown the regions surrounded by the red ellipse. In the replenished image using the proposed AR model under spatial continuity constraint, the blocking artifact is weakened, although it can still be observed. And in the replenished images using the proposed AR model under temporal continuity constraint and combining spatial and temporal continuity constraints, the blocking artifacts are completely removed. In Fig. 9, for the BMA, STBMA, STBMA+PDE, and STBMA+OBMC

TABLE V  
AVERAGE PSNR RESULTS OF EACH TEST SEQUENCE WHEN BOTH I AND P FRAMES SUFFER LOSS DURING TRANSMISSION

Sequence		QP	PSNR (dB)											
			BMA	[12]	STBMA	STBMA+OBMC	STBMA+PDE	BMA+AR			STBMA+AR			
							Spatial	Temporal	Combined	Spatial	Temporal	Combined		
QCIF	Football	16	23.40	22.17	23.45	23.58	23.54	23.21	22.96	23.55	23.24	23.09	23.60	
		24	23.30	22.09	23.40	23.42	23.41	22.79	22.70	23.43	22.84	22.85	23.45	
		28	22.89	21.76	22.37	22.72	22.73	22.67	22.57	22.84	22.63	22.61	22.88	
		40	21.93	21.84	22.67	22.71	22.79	22.38	22.31	22.81	22.63	22.63	22.84	
	Mobile	16	22.31	22.40	22.99	22.97	22.99	22.94	23.15	23.18	23.03	23.20	23.25	
		24	22.10	22.37	22.72	22.74	22.75	22.66	22.80	22.81	22.64	22.73	22.83	
		28	22.01	22.22	22.51	22.52	22.51	22.44	22.52	22.57	22.45	22.49	22.53	
		40	19.67	19.70	19.81	19.83	19.85	19.77	19.74	19.91	19.78	19.75	19.98	
	Paris	16	23.01	23.17	23.39	23.43	23.41	23.44	23.44	23.68	23.46	23.46	23.68	
		24	22.35	22.76	23.09	23.13	23.15	22.89	23.04	23.19	23.14	23.41	23.43	
		28	22.44	22.49	22.82	22.85	22.84	22.85	23.02	23.07	22.88	23.02	23.07	
		40	20.48	20.47	20.67	20.66	20.66	20.62	20.52	20.67	20.63	20.56	20.69	
	Suzie	16	30.04	30.20	30.43	30.50	30.47	30.21	30.27	30.62	30.27	30.35	30.71	
		24	29.41	29.31	29.45	29.49	29.50	29.35	29.19	29.52	29.42	29.16	29.56	
		28	28.78	29.04	28.96	29.00	28.99	28.96	28.91	29.14	28.98	28.90	29.19	
		40	26.22	26.09	26.24	26.25	26.25	26.11	26.06	26.24	26.14	26.09	26.28	
<b>Average</b>			<b>23.77</b>	<b>23.63</b>	<b>24.06</b>	<b>24.11</b>	<b>24.12</b>	<b>23.96</b>	<b>23.95</b>	<b>24.20</b>	<b>24.01</b>	<b>24.02</b>	<b>24.25</b>	
CIF	Foreman	16	27.74	27.84	27.70	27.71	27.71	27.74	27.80	27.98	27.70	27.56	27.77	
		24	27.37	27.46	27.01	27.01	27.01	27.37	27.58	27.69	27.01	27.34	27.28	
		28	26.89	26.90	26.95	26.95	26.96	26.89	27.10	27.17	26.95	27.09	27.08	
		40	25.35	25.36	25.27	25.28	25.28	25.35	25.50	25.60	25.27	25.50	25.53	
	Mobile	16	23.69	23.64	23.82	23.82	23.83	23.69	23.95	23.95	23.82	24.06	24.00	
		24	23.42	23.23	23.43	23.43	23.44	23.42	23.63	23.66	23.43	23.65	23.62	
		28	23.21	23.13	23.30	23.30	23.30	23.21	23.49	23.48	23.30	23.53	23.48	
		40	21.14	21.13	21.10	21.11	21.11	21.14	21.24	21.26	21.10	21.17	21.19	
	Flower	16	29.04	28.76	29.83	29.84	29.84	29.04	28.98	29.09	29.83	29.39	29.88	
		24	28.43	28.25	29.25	29.26	29.15	28.43	28.42	28.69	29.25	28.91	29.30	
		28	27.60	27.52	28.18	28.19	28.19	27.60	27.67	27.94	28.18	27.88	28.24	
		40	23.54	23.21	23.54	23.55	23.54	23.54	23.53	23.67	23.54	23.37	23.53	
	<b>Average</b>			<b>25.62</b>	<b>25.54</b>	<b>25.78</b>	<b>25.79</b>	<b>25.78</b>	<b>25.62</b>	<b>25.74</b>	<b>25.85</b>	<b>25.78</b>	<b>25.79</b>	<b>25.91</b>

methods, we cannot observe the figures “1” in “31,” as shown the regions surrounded by the red circle. However, in the replenished images by the proposed AR model, figures “1” in “31” can be clearly observed.

We also conducted another experiment excluding error propagation effects. The encoding GOP is set to be IPPIPPI..., and we assume the error only occurs at the second P of each GOP with the PLR being 10%. The experimental result is provided in Table IV, from which we can observe that the proposed AR model still outperforms other comparing methods in average.

Table V further tabulates the PSNR results of each test sequence when both I and P frames suffer loss during transmission. The PLR is 10% and the encoding GOP is IPPP..., where I frames are inserted every 16 frames. The first I frame is assumed to be error free. To give a fair comparison of different inter frame concealment methods, the errors in I frames are replenished using the weighted pixel averaging method in the reference software JM10.0, and the errors in P frames are restored utilizing the proposed AR models and other competing methods. From the experimental results we can observe that the performance of all the inter frame error concealment methods drop dramatically. This is due to the reason that the badly concealed MBs in I frames would greatly degrade the quality of the following P frames. However, the

proposed AR model still has a little better performance than other competing methods, although the performance gain is rather small compared to the case when I frames are error free.

### C. Computational Complexity Analysis

Most computational complexity of the proposed AR model based error concealment scheme is concentrated on the calculation of AR coefficients. Take (9) for example, the AR coefficient derivation involves matrix multiplication and inverse matrix operations. It should be noted that the dimension of matrix in (9) depends on the range of the AR model. The smaller the range of the AR model, the lower computational complexity it would be of.

Besides, there are many fast algorithms, e.g., [29] to speed up the calculation of AR coefficients. In Table VI, we examine the consumed time of decoding the first 120 frames of each test sequence using different error concealment methods (the encoding GOP is IPPP..., with I frames being inserted every sixteen frames) on a typical computer (2.5 GHz Intel Dual Core, 2 GB Memory). Except for BMA, which owns the lowest computational complexity, [12] consumes fewer time than other comparing methods. STBMA and STBMA+OBMC have similar computational complexity. STBMA+PDE takes longer time than the former four due to the iteration

TABLE VI  
AVERAGE SYSTEM TIME USING DIFFERENT METHODS

Sequence		QP	PLR	Time (s)											
				BMA	[12]	STBMA	STBMA+PDE	STBMA+OBMC	BMA+AR			STBMA+AR			
									Spatial	Temporal	Combined	Spatial	Temporal	Combined	
QCIF	Football	24	5%	0.953	1.028	0.982	1.077	0.984	0.999	1.029	1.105	1.045	1.076	1.087	
			10%	0.967	1.040	1.122	1.764	1.171	1.270	1.273	1.514	1.404	1.450	1.748	
			20%	0.983	1.084	1.326	2.480	1.388	1.435	1.544	1.997	1.779	1.905	2.355	
		28	5%	0.874	0.942	0.904	1.045	0.915	0.921	0.983	1.014	0.999	0.998	1.028	
			10%	0.890	0.981	1.107	1.686	1.124	1.170	1.179	1.388	1.356	1.389	1.622	
			20%	0.843	1.032	1.279	2.401	1.288	1.420	1.481	1.902	1.764	1.842	2.278	
	Mobile	24	5%	1.029	1.146	1.061	1.170	1.092	1.114	1.193	1.255	1.119	1.078	1.155	
			10%	1.013	1.093	1.216	1.763	1.223	1.289	1.310	1.468	1.498	1.452	1.717	
			20%	1.013	1.130	1.357	2.404	1.419	1.476	1.529	1.933	1.857	1.826	2.232	
		28	5%	0.889	0.994	0.982	1.108	0.984	1.015	0.998	1.088	1.046	0.998	1.062	
			10%	0.920	1.036	1.138	1.670	1.154	1.184	1.217	1.467	1.419	1.388	1.621	
			20%	0.967	1.070	1.296	2.310	1.357	1.487	1.489	1.841	1.794	1.749	2.201	
	Paris	24	5%	0.764	0.817	0.764	0.874	0.780	0.785	0.795	0.874	0.795	0.827	0.890	
			10%	0.654	0.816	0.843	1.372	0.847	0.998	1.031	1.311	1.092	1.169	1.405	
			20%	0.733	0.836	0.921	1.825	0.951	1.141	1.279	1.778	1.356	1.482	1.951	
		28	5%	0.671	0.756	0.702	0.827	0.716	0.774	0.699	0.843	0.778	0.796	0.857	
			10%	0.655	0.779	0.795	1.279	0.810	0.921	1.014	1.249	0.998	1.077	1.372	
			20%	0.686	0.798	0.797	1.731	0.905	1.269	1.279	1.652	1.297	1.467	1.918	
	Suzie	24	5%	0.704	0.800	0.764	0.842	0.766	0.783	0.781	0.798	0.794	0.827	0.859	
			10%	0.749	0.827	0.889	1.311	0.894	0.999	1.046	1.279	1.092	1.202	1.436	
			20%	0.749	0.848	0.936	1.793	1.014	1.198	1.310	1.778	1.435	1.530	1.981	
		28	5%	0.655	0.739	0.703	0.796	0.704	0.781	0.779	0.794	0.758	0.765	0.842	
			10%	0.670	0.764	0.733	1.232	0.795	0.889	0.999	1.201	1.061	1.077	1.373	
			20%	0.671	0.787	0.874	1.653	0.920	1.155	1.201	1.687	1.357	1.435	1.902	
<b>Average</b>				<b>0.821</b>	<b>0.923</b>	<b>0.979</b>	<b>1.517</b>	<b>1.008</b>	<b>1.103</b>	<b>1.143</b>	<b>1.384</b>	<b>1.246</b>	<b>1.284</b>	<b>1.537</b>	
CIF	Foreman	24	5%	2.838	3.211	3.043	3.448	3.012	3.151	3.292	3.510	3.308	3.400	3.558	
			10%	2.979	3.306	3.697	5.695	3.681	4.167	4.991	5.990	4.804	5.600	6.520	
			20%	2.978	3.457	4.133	8.004	4.306	4.976	6.677	8.394	6.243	7.597	9.281	
		28	5%	2.714	2.960	2.731	3.212	2.855	2.917	3.056	3.275	3.027	3.261	3.400	
			10%	2.745	3.190	3.369	5.273	3.400	3.933	4.883	5.804	4.462	5.211	6.396	
			20%	2.776	3.200	3.729	7.346	3.947	4.741	6.490	8.220	5.708	7.286	9.031	
	Mobile	24	5%	3.821	4.654	3.916	4.415	3.916	4.055	4.071	4.276	4.181	4.228	4.461	
			10%	3.808	4.233	4.632	6.973	4.647	4.977	5.631	6.724	5.757	6.350	7.362	
			20%	3.807	4.203	5.228	9.296	5.398	5.711	7.082	8.939	7.224	8.330	10.063	
		28	5%	3.480	3.667	3.681	4.121	3.618	3.728	3.840	3.977	3.821	3.964	4.103	
			10%	3.541	3.832	4.305	6.645	4.371	4.742	5.368	6.443	5.476	5.975	7.051	
			20%	3.634	3.950	4.993	9.019	5.102	4.554	6.863	8.675	6.865	7.925	9.812	
	Flower	24	5%	3.244	3.525	3.353	3.681	3.339	3.448	3.558	3.759	3.541	3.651	3.899	
			10%	3.229	3.498	3.743	5.522	3.821	4.338	5.272	6.209	4.881	5.773	6.833	
			20%	3.215	3.570	4.212	7.503	4.307	5.148	6.897	8.566	6.038	7.862	9.610	
		28	5%	3.025	3.221	3.043	3.448	3.089	3.260	3.336	3.542	3.290	3.541	3.667	
			10%	3.057	3.308	3.590	5.319	3.603	4.180	5.148	6.039	4.711	5.615	6.647	
			20%	3.104	3.397	3.993	7.036	4.072	4.945	6.726	8.456	5.804	7.706	9.392	
	<b>Average</b>				<b>3.222</b>	<b>3.577</b>	<b>3.855</b>	<b>5.886</b>	<b>3.916</b>	<b>4.276</b>	<b>5.177</b>	<b>6.155</b>	<b>4.952</b>	<b>5.738</b>	<b>6.727</b>

operation in PDE. The computational complexity of the proposed AR model is higher than the comparing methods, but still acceptable, especially when only spatial continuity constraint or temporal continuity constraint is applied.

## VI. CONCLUSION

In this paper, we developed an AR model based error concealment scheme for block-based packet video coding. For each corrupted block, we first derived the motion vector and then replenished each corrupted pixel as the weighted summation of pixels within a square centered at the pixel indicated by the derived motion vector with integer-PEL accuracy in

a regression manner. To obtain better concealment results, we proposed two block-dependent AR coefficient derivation algorithms under spatial and temporal continuity constraints. We then combined the regression results generated by the two algorithms to form the ultimate concealment results. The simulation results demonstrate the superiority of the proposed scheme over other inter-frame concealments with acceptable computational complexity.

## REFERENCES

- [1] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

- [2] *Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 mbit/s—Part 2: Video*, ISO/IEC 11 172-2 (MPEG-1), Int. Standards Organization/Int. Electrotechnical Commission (ISO/IEC) JTC 1, Mar. 1993.
- [3] *Generic Coding of Moving Pictures and Associated Audio Information—Part 2: Video*, Rec. H.262 and ISO/IEC 13 818-2 (MPEG-2 Video), Int. Telecommunication Union-Telecommunication (ITU-T) and Int. Standards Organization/Int. Electrotechnical Commission (ISO/IEC) JTC 1, Nov. 1994.
- [4] *Video Coding for Low Bit Rate Communication*, Int. Telecommunication Union-Telecommunication (ITU-T) Rec. H.263, 1995.
- [5] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Process. Mag.*, vol. 17, no. 7, pp. 61–82, Jul. 2000.
- [6] Y. Wang and Q.-F. Zhu, "Error control and concealment for video communication: A review," *Proc. IEEE*, vol. 86, no. 5, pp. 974–997, May 1998.
- [7] 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.
- [8] W. Zhu, Y. Wang, and Q.-F. Zhu, "Second-order derivative-based smoothness measure for error concealment in DCT-based codecs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 6, pp. 713–718, Oct. 1998.
- [9] S. D. Rane, G. Sapiro, and M. Bertalmio, "Structure and texture filling-in of missing image blocks in wireless transmission and compression," *IEEE Trans. Image Process.*, vol. 12, no. 3, pp. 296–303, Mar. 2003.
- [10] W. Y. Kung, C. S. Kim, and C. J. Kuo, "Spatial and temporal error concealment techniques for video transmission over noisy channels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 7, pp. 789–802, Jul. 2006.
- [11] X. Li and M. Orchard, "Novel sequential error concealment techniques using orientation adaptive interpolation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 857–864, Oct. 2002.
- [12] 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.
- [13] Y. C. Lee, Y. Altunbasak, and R. M. Mersereau, "Multiframe error concealment for MPEG-coded video delivery over error-prone networks," *IEEE Trans. Image Process.*, vol. 11, no. 11, pp. 1314–1331, Nov. 2002.
- [14] D. Persson, T. Eriksson, and P. Hedelin, "Packet video error concealment with Gaussian mixture models," *IEEE Trans. Image Process.*, vol. 17, no. 2, pp. 145–154, Feb. 2008.
- [15] D. Persson and T. Eriksson, "Mixture model and least squares based packet video error concealment," *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 1048–1054, May 2009.
- [16] P. Haskell and D. Messerschmitt, "Resynchronization of motion compensated video affected by ATM cell loss," in *Proc. IEEE ICASSP*, vol. 3, May 1992, pp. 545–548.
- [17] 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.
- [18] W. M. Lam, A. R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 3, Apr. 1993, pp. 417–420.
- [19] S. Tsekeridou, F. A. Cheikh, M. Gabbouj, and I. Pitas, "Motion field estimation by vector rational interpolation for error concealment purposes," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 6, Mar. 1999, pp. 3397–3400.
- [20] M. Al-Mualla, N. Canagarajah, and D. R. Bull, "Error concealment using motion field interpolation," in *Proc. IEEE Int. Conf. Image Process.*, vol. 3, Oct. 1998, pp. 512–516.
- [21] J. H. Zheng and L. P. Chau, "A temporal error concealment algorithm for H.264 using Lagrange interpolation," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2004, pp. 133–136.
- [22] Z. W. Gao and W. N. Lie, "Video error concealment by using Kalman filtering technique," in *Proc. IEEE Int. Symp.-Circuits Syst.*, May 2004, pp. 69–72.
- [23] W. N. Lie and Z. W. Gao, "Video error concealment by integrating greedily suboptimization and Kalman filtering techniques," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 8, pp. 982–992, Aug. 2006.
- [24] G. S. Yu, M. K. Liu, and M. Marcellin, "POCS-based error concealment for packet video using multiframe overlap information," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 4, pp. 422–434, Aug. 1998.
- [25] D. 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.
- [26] 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. 13, no. 3, pp. 248–258, Jun. 1993.
- [27] Y. Chen, X. Sun, F. Wu, Z. Liu, and S. Li, "Spatio-temporal video error concealment using priority-ranked region-matching," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2005, pp. 1050–1053.
- [28] Y. Chen, Y. Hu, O. Au, H. Li, and C. Chen, "Video error concealment using spatio-temporal boundary matching and partial differential equation," *IEEE Trans. Multimedia*, vol. 10, no. 1, pp. 2–15, Jan. 2008.
- [29] X. Wu, K. U. Barthel, and W. Zhang, "Piecewise 2-D autoregression for predictive image coding," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 1998, pp. 901–904.
- [30] A. C. Kokaram, R. D. Morris, W. J. Fitzgerald, and P. J. W. Rayner, "Detection of missing data in image sequences," *IEEE Trans. Image Process.*, vol. 4, no. 11, pp. 1496–1508, Nov. 1995.
- [31] A. C. Kokaram, R. D. Morris, W. J. Fitzgerald, and P. J. W. Rayner, "Interpolation of missing data in image sequences," *IEEE Trans. Image Process.*, vol. 4, no. 11, pp. 1509–1519, Nov. 1995.
- [32] S. N. Efstratiadis and A. K. Katsaggelos, "A model-based PEL-recursive motion estimation algorithm," in *Proc. IEEE ICASSP*, Apr. 1990, pp. 1973–1976.
- [33] X. Li, "Least-square prediction for backward adaptive video coding," *EURASIP J. Appl. Signal Process. (Special Issue on H.264 and Beyond)*, vol. 2006, no. 18, pp. 1–13, Mar. 2007.
- [34] X. Xiang, Y. Zhang, D. Zhao, S. Ma, and W. Gao, "A high efficient error concealment scheme based on auto-regressive model for video coding," in *Proc. PCS*, May 2009.
- [35] Y. Zhang, X. Xiang, S. Ma, D. Zhao, and W. Gao, "Auto regressive model and weighted least squares based packet video error concealment," in *Proc. IEEE Data Compression Conf.*, Mar. 2010, pp. 455–464.
- [36] Y. Zhang, D. Zhao, X. Ji, R. Wang, and W. Gao, "A spatio-temporal auto regressive model for frame rate up-conversion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 9, pp. 1289–1301, Sep. 2009.
- [37] S. Wenger, *Error Patterns for Internet Experiments*, ITU-T SG16 document Q15-I-16r1, 1999.
- [38] M. T. Orchard and G. J. Sullivan, "Overlapped block motion compensation: An estimation theoretic approach," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 693–699, Sep. 1994.



**Yongbing Zhang** received the B.A. degree in English, and the M.S. and Ph.D degrees in computer science from the Department of Computer Science, Harbin Institute of Technology, Harbin, China, in 2004, 2006, and 2010, respectively.

He is currently with the Graduate School at Shenzhen, Tsinghua University, Shenzhen, China. His current research interests include video processing, image and video coding, video streaming, and transmission.



**Xinguang Xiang** received the B.S. and M.S. degrees in computer science from the Harbin Institute of Technology, Harbin, China, in 2005 and 2007, respectively. Since 2007, he has been pursuing the Ph.D. degree from the Department of Computer Science, School of Computer Science and Technology, Harbin Institute of Technology.

His current research interests include video compression, multi-view/stereoscopic video coding, and robust video transmission.



**Debin Zhao** received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology, Harbin, China, in 1985, 1988, and 1998, respectively.

He is currently a Professor with the Department of Computer Science, Harbin Institute of Technology. He has published over 200 technical articles in refereed journals and conference proceedings in the areas of image and video coding, video processing, video streaming and transmission, and pattern recognition.



**Siwei Ma** (S'03) received the B.S. degree from Shandong Normal University, Jinan, China, in 1999, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005.

From 2005 to 2007, he was a Post-Doctorate with the University of Southern California, Los Angeles. Then, he joined the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing, where he is currently an Associate Professor. He has published

over 70 technical articles in refereed journals and proceedings in the areas of image and video coding, video processing, video streaming, and transmission.



**Wen Gao** (M'92–SM'05–F'09) received the M.S. degree in computer science from the Harbin Institute of Technology, Harbin, China, in 1985, and the Ph.D. degree in electronics engineering from the University of Tokyo, Tokyo, Japan, in 1991.

He is currently a Professor of computer science with the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing, China. Before joining Peking University, he was a Full Professor of computer science with the Harbin Institute of Technology from

1991 to 1995, and with the Chinese Academy of Sciences, Beijing, from 1996 to 2005. He has published extensively, including four books and over 500 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics.

Dr. Gao is the Editor-in-Chief of the *Journal of Computer* (a journal of the China Computer Federation), an Associate Editor of the *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, *IEEE TRANSACTIONS ON MULTIMEDIA*, *IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT*, an Area Editor of the *EURASIP Journal of Image Communications*, and an Editor of the *Journal of Visual Communication and Image Representation*. He chaired a number of prestigious international conferences on multimedia and video signal processing, and also served on the advisory and technical committees of numerous professional organizations.