

Improved error concealment of region of interest based on the H.264/AVC standard

Zhengyi Luo

Li Song

Shibao Zheng

Yi Xu

Xiaokang Yang

Shanghai Jiao Tong University

Institute of Image Communication and Information Processing

800 Dong Chuan Road

Min Hang, Shanghai 200240

China

E-mail: song_li@sjtu.edu.cn

Abstract. Video transmission over error-prone channels often suffers from inevitable transmission errors, which necessitates proper error concealment (EC) for acceptable image quality. Furthermore, the region of interest (ROI) in images usually draws much attention, and so the EC of the ROI receives special treatment during encoding and decoding. We explore a data hiding–based scheme to effectively improve the EC of the ROI in the case of erasures of large continuous regions, which becomes impractical for conventional EC methods. At the encoder side, motion vectors of the ROI are adaptively embedded in the background based on original quantized coefficients of background macroblocks. Considering the limited embedding capacity of the background, we further propose to assign priorities to each ROI macroblock based on a predefined metric of error propagation. Our scheme is applied with the state-of-the-art H.264/AVC standard in a packet loss scenario, and better video quality can be obtained. Experimental results show that the scheme can improve the EC of the ROI significantly without much loss of coding efficiency. © 2010 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3381178]

Subject terms: error concealment; region of interest; flexible macroblock ordering; data hiding; H.264/AVC.

Paper 090571RR received Jul. 28, 2009; revised manuscript received Jan. 28, 2010; accepted for publication Feb. 8, 2010; published online Apr. 15, 2010.

1 Introduction

With the continuing trend toward the provision of multimedia services, video transmission over unreliable channels, such as the Internet or wireless networks, has become quite common nowadays. In order to reduce transmission errors and achieve better image quality, many technologies have been proposed up to now. For example, when a feedback channel is available and decent delay is permitted, automatic repeat request (ARQ) can be utilized. However, when there are no feedback channels or little delay is allowed, forward error correction (FEC) may be used at the cost of bandwidth resources. Still, probably perfect transmission cannot be guaranteed, as the underlying channels provide only best-effort services, and video data might suffer from inevitable loss. Therefore, suitable error concealment (EC)¹ is always desired for better image quality at the decoder side.

Conventionally, as a non-normative feature, EC is performed only by decoders. Recently, EC methods based on data hiding techniques² have been developed, which require some side information to be embedded into bit streams by encoders. Then with extracted ancillary information, decoders can achieve better EC performance. The embedded side information helpful for concealment can be edge directions and key point values to assist interpolation,^{3,4} the mean value of blocks to serve as substitutes upon errors,⁵ prediction residuals,⁶ part of transform coefficients,⁷ or half-toned images.⁸ For video data with highly temporal coherence, motion vector (MV)–related cues are considered as good side information for decoders

to find suitable substitutes from reference frames.^{9–14} It is noted that data hiding–based EC methods obtain good image quality at the cost of coding performance. Especially, these methods may produce many extra nonzero residual blocks when applied to the state-of-the-art H.264/AVC standard,¹⁵ ultimately resulting in significant increase of bitrate.

Research on the human visual system (HVS) reveals that people generally pay more attention to the region-of-interest (ROI) areas. Thus, an acceptable compromise between coding efficiency and EC performance can be achieved by performing data hiding only for ROI regions. In recent research works, some efforts have been taken to develop data hiding–based EC methods for ROIs. Lin et al. embedded low-frequency discrete cosine transform (DCT) coefficients of ROI in the background,¹⁶ and Jue and Liang embedded part of wavelet coefficients of ROI.¹⁷ When it comes to video coding (H.264, for example), however, both schemes lead to too much degradation in coding efficiency. The data hiding schemes suitable for ROIs of video images still deserve exploration.

Inspired by the work in Ref. 10, where the authors proposed to embed an additional MV for each macroblock only in intraframes, in this paper, we develop an enhanced scheme to further improve EC performance in ROI regions. At the encoder side, a separate MV is searched for each ROI macroblock and embedded in the background both in intra- and interframes. In the case of foreground and background slices coded and transmitted independently, a corrupted ROI can be properly restored by EC methods as long as its background is correctly received. Instead of performing data embedding as in the aforementioned schemes, we

adapt the amount of data to be embedded according to the analysis of the original quantized coefficients of background macroblocks. Considering the limited embedding capacity of the background, we further propose to assign priorities to each ROI macroblock based on a predefined metric of error propagation. Therefore, neighboring erroneous ROI areas are ensured to be successively restored in priority-ordered concealment instead of one-off concealment, so better concealment can be expected because more neighboring blocks are previously restored. It is observed that the proposed scheme can be applied jointly with many existing EC methods and only minor loss of coding performance is incurred when applied to the H.264/AVC standard. Moreover, experimental results show that decoders can use extracted information from our scheme to greatly improve EC performance in ROI regions.

The remainder of this paper is organized as follows. Section 2 introduces ROI coding and transmission briefly and states the addressed EC problem. In Sec. 3, the proposed scheme and its implementation details are described. Experimental results validating the effectiveness of our scheme are shown in Sec. 4. Last, Sec. 5 draws conclusions.

2 Preliminaries

2.1 Flexible Macroblock Ordering and the Mechanism of Packetization

In H.264, a picture can be split into one or several slices, and a slice consists of a certain number of macroblocks. Given the necessary parameters and reference pictures, a slice can be correctly decoded without the other slices in the same picture. Generally, slices are composed of macroblocks in the order of the raster scan. Specifically, flexible macroblock ordering (FMO) of H.264/AVC advanced the new concept of slice groups. The pictures are partitioned into slices and macroblocks optionally in different patterns, where each macroblock is assigned to a slice group statically according to the macroblock allocation map. FMO was introduced into H.264 mainly as an error robustness feature.¹⁸ If a slice group gets lost during transmission, there exists great probability for other correct slice groups to conceal it. In addition, FMO can be used for other purposes such as the ROI coding. We can encode all the ROI macroblocks in the same slice using FMO and treat them specifically for higher visual quality during encoding, transmission, and decoding.

The network abstraction layer (NAL) in H.264 adapts bit streams in a network-friendly way during packetization and transportation. An NAL unit can carry a coded slice, a data partition, or a parameter set, etc. As far as ROI coding is concerned, the ROI and background can be coded in different slices, as stated earlier. Accordingly, NAL units will be obtained containing coded ROI slices or background slices, which may be further packetized into data packets [such as real-time transport protocol (RTP) packets¹⁹] for transmission and enable ROI and background slices to suffer from packet loss independently.

2.2 ROI EC Problem

For simplicity, suppose that a picture contains one ROI and that each slice group contains one slice. Pictures are then divided into ROI and background slices, which may be

transmitted independently in separate packets. Given a channel packet-loss rate p_{roi} and a distortion D_{roi} of the ROI at the decoder side, we can get

$$E(D_{roi}) = (1 - p_{roi})E(D_{dec}) + p_{roi}E(D_{ec}), \quad (1)$$

where D_{dec} denotes the distortion of ROI in the case of correct decoding (such as the quantization distortion), and D_{ec} denotes the distortion of ROI due to EC. With proper EC methods, we can mitigate the distortion factor D_{ec} , which can be further derived as

$$E(D_{ec}) = (1 - p_b)E(D_{ec_b}) + p_bE(D_{ec_bu}), \quad (2)$$

where p_b depicts the packet-loss rate of the background, D_{ec_b} denotes the distortion of the ROI due to EC with the background available, and D_{ec_bu} denotes the distortion of the ROI due to the EC with the background unavailable. Combining Eqs. (1) and (2), we obtain

$$\begin{aligned} E(D_{roi}) &= (1 - p_{roi})E(D_{dec}) + p_{roi}[(1 - p_b)E(D_{ec_b}) \\ &\quad + p_bE(D_{ec_bu})] \\ &= (1 - p_{roi})E(D_{dec}) + p_{roi}(1 - p_b)E(D_{ec_b}) \\ &\quad + p_{roi}p_bE(D_{ec_bu}). \end{aligned} \quad (3)$$

From investigation, we find that the quality of the ROI is more sensitive to distortion due to the EC. Considering p_{roi} and p_b usually taking a small value, we can infer that the error term related to D_{ec_b} usually plays a dominant role over the one related to D_{ec_bu} . Thus, the visual quality of the ROI can be expected to improve if the proper EC is applied using the available background, which is the very case where FMO works.

As compared with simple copying or extrapolation, the EC of the lost ROI can be improved with reference information about the correctly received background. However, common EC methods usually cannot produce pleasing results for lost macroblocks unless there are enough correct neighboring macroblocks. When the whole ROI slice gets lost, those inner ROI macroblocks receiving increased error propagation still need methods for better EC. In the following section, we propose to solve this problem using data hiding of motion cues, so that the EC of the lost ROI can be significantly improved as long as the background is correctly received.

3 Error Concealment of ROI Based on Adaptive Motion Vector Embedding

As video often shows high correlation between frames, accurate MVs are always desired by the EC for better performance. On this point, temporal error concealment is developed to repair lost regions, which generally consists of two steps: (1) estimate MVs for missing macroblocks, and (2) find suitable substitutes based on estimated MVs.

Data hiding is originally proposed to resolve information security issues. Since it can convey useful information to decoders, data hiding-based EC methods have been proposed recently to achieve better image quality, where MVs are commonly used as the embedded information.^{9,10} However, blind embedding would lead to distinct degradation in coding efficiency. As a compromise, we suggest only MVs

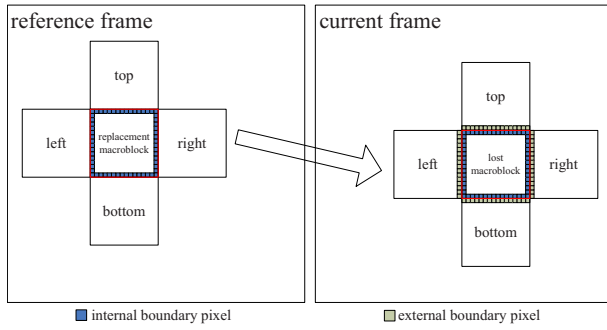


Fig. 1 Illustration of the smoothness EC criterion.

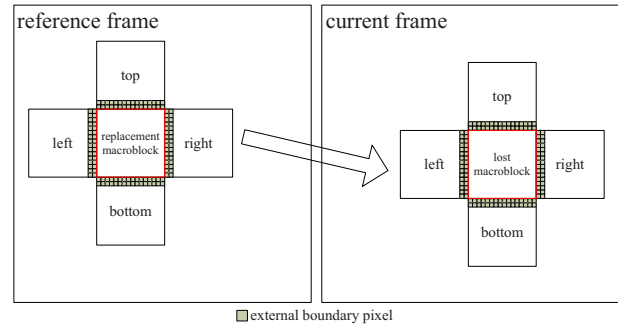


Fig. 2 Illustration of the motion similarity EC criterion.

of ROI macroblocks be embedded in the background. Instead of adopting one-off error concealment for a large ROI region as in the current methods, we divide a large ROI region into neighboring smaller ROI elements, which are ensured to be successively restored in priority-ordered concealment, so that better concealment for each element can be expected because more neighboring blocks are previously restored. In the following subsections, the details of our ROI EC scheme are provided.

3.1 Motion Vectors to Be Embedded

To achieve the best coding efficiency, encoders generally choose the encoding modes and motion vectors for each macroblock that minimize the rate-distortion Lagrangian cost—namely,

$$J = \min(D + \lambda R), \quad (4)$$

where R and D denote rate and distortion, respectively, and λ is a constant. However, in terms of EC performance, there is no rate involved, so the embedded MV should be searched separately, which accounts for minimum distortion only. In our scheme, we embed only one MV for each ROI macroblock and simply use mean square error (MSE) as the matching criterion for embedded MVs. The search range of embedded MVs is limited within the range of ± 31 pixels, and the search precision is set at a half pixel. For each ROI macroblock, the total bits of its MV to be embedded can be computed as

$$L = 2 \times (\lceil \log_2(2 \times 31 + 1) \rceil + 1) = 14. \quad (5)$$

Although ROI macroblocks in intraframes need no MVs during encoding, we still search the corresponding MVs relative to the previous frame. Then, restoration of lost ROIs in both intra- and interframes can benefit from correctly received backgrounds during EC.

3.2 Embedding Procedure of Motion Vectors

3.2.1 Embedding methods

We adopt the odd-even method²⁰ to embed MV bits of ROI in certain quantized coefficients of background macroblocks. Suppose that $\bar{z}(x, y)$ denotes the original quantized coefficient and that b denotes the bit to be hidden. Then, after embedding, the quantized coefficient will become¹⁰

$$\bar{z}'(x, y) = \begin{cases} \bar{z}(x, y) + 1 & b = 1 \ \& \ \text{mod}[\bar{z}(x, y), 2] = 0 \\ \bar{z}(x, y) & b = 1 \ \& \ \text{mod}[\bar{z}(x, y), 2] = 1 \\ \bar{z}(x, y) & b = 0 \ \& \ \text{mod}[\bar{z}(x, y), 2] = 0 \\ \bar{z}(x, y) - 1 & b = 0 \ \& \ \text{mod}[\bar{z}(x, y), 2] = 1 \end{cases} \quad (6)$$

In our scheme, 4×4 luma and chroma blocks of background macroblocks are examined one by one. Considering visual artifact, bits are embedded only from the fourth to the seventh quantized coefficients^{3,10} in zigzag order, with the other coefficients left unchanged. At the decoder side, MV bits of ROI macroblocks can be extracted from the background by

$$b = \begin{cases} 1 & \text{mod}[\bar{z}'(x, y), 2] = 1 \\ 0 & \text{mod}[\bar{z}'(x, y), 2] = 0 \end{cases} \quad (7)$$

3.2.2 Adaptive embedding

Although MV embedding can improve the EC of the ROI, we still have to take into account its impact on the background. If additional nonzero 4×4 blocks come into being due to data hiding, it is found that coded block pattern (CBP) may change and the bit rate will increase greatly. In addition, if data are embedded into blocks that have few prediction residuals (i.e., exhibit high correlation to neighbors), the impaired correlation is likely to lead to perceptible distortion. Considering coding efficiency and embedding distortion jointly, we design an adaptive embedding scheme as follows.

In the H.264/AVC standard it is known that for the 4×4 luma blocks in I16 \times 16 mode and all the 4×4 chroma blocks, the quantized DC coefficients are scanned and encoded separately from the AC coefficients. To ensure that no additional nonzero 4×4 blocks will appear as a result of data embedding, we perform embedding in these blocks only when the first or the second AC coefficients are nonzero. As for the other 4×4 blocks, as the DC coefficients are scanned and encoded together with the AC coefficients, we perform embedding only when there exist nonzero values in the DC or the first or second AC coefficients. Thus, the bit rate increase can be restrained. As we perform embedding only in blocks that have nonzero prediction residuals, embedding will not cause many perceptible artifacts. The image quality is maintained as well.

In the proposed embedding scheme, a different number of bits would be embedded for each background macroblock

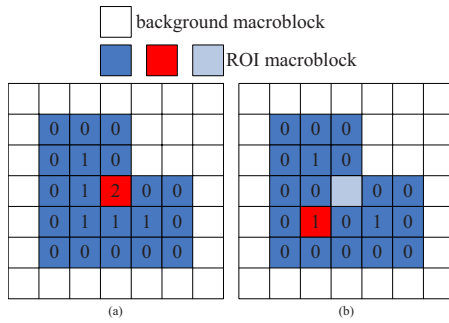


Fig. 3 Illustration of priority calculation.

lock according to different candidate encoding modes. Suppose that some bits are available to be embedded, and then the distortion D' of each valid encoding mode results from both quantization distortion D and possible distortion ΔD incurred by data hiding:

$$D' = \begin{cases} D + \Delta D & \text{if some bits are embedded} \\ D & \text{if no bits are embedded.} \end{cases} \quad (8)$$

We incorporate the distortion due to data hiding into the mode decision process of the background macroblocks. With each mode yielding a new rate, the selected encoding mode will be the one minimizing the new rate-distortion cost.

3.3 Priorities of ROI Macroblocks

It is obvious that the embedding capacity of background varies with the image contents. This poses the problem that in some cases, not all the MVs of the ROI macroblocks can be embedded in the background. When the ROI is quite large, or when the background has few prediction residuals, we have to select out some ROI macroblocks among all the candidates, so that better EC performance can be expected if their MVs are embedded rather than the others. To tackle this problem, we propose to order the ROI macroblocks with assigned priorities that are determined by a predefined metric of error propagation, where the “divide and conquer” methodology is used.

Before explanation of priority computation, two classical EC criteria used by traditional methods are reviewed first. One criterion relies on the smooth intensity variations among neighboring macroblocks. The estimated MVs for lost macroblocks minimize the difference of boundary pixels between the recovered macroblock and its neighbors, which can be roughly illustrated in Fig. 1 (Ref. 21) and formulated as

$$\arg \min_{mv} \sum_{i=1}^N \|F_{cur_i}^{out} - F_{ref}(mv)_i^{in}\|, \quad (9)$$

where $F_{cur_i}^{out}$ is the i 'th pixel of the external boundary in the current frame, $F_{ref}(mv)_i^{in}$ is the i 'th pixel of the internal boundary using the candidate MV mv in the reference frame, and N is the total number of calculated boundary pixels. The frequently used boundary matching algorithm (BMA) algorithm²² belongs to this category. Another criterion exploits the smoothness of the motion field among

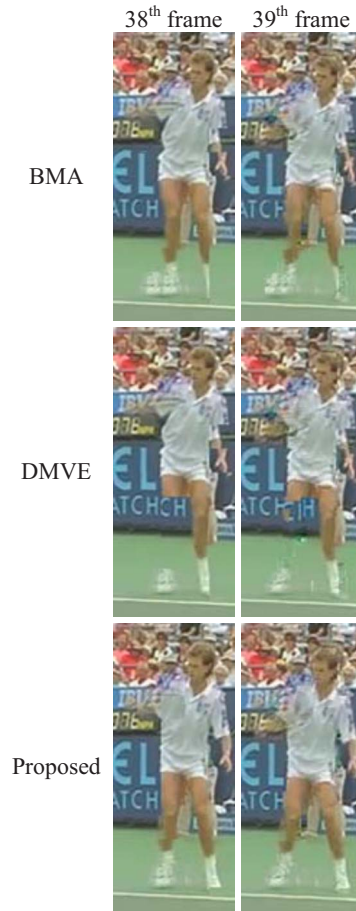


Fig. 4 Comparison of subjective quality when the ROI from the 38th frame of Stefan is lost (QP=28). Left: concealed ROI in the 38th frame; right: error propagation in the 39th frame.

neighboring macroblocks. The estimated MVs for missing macroblocks minimize the difference of the multiple-pixel external boundary between the lost macroblock's neighbors in the current frame and the replacement macroblock's neighbors in the reference frame, which can be roughly illustrated in Fig. 2 (Ref. 21) and formulated as

$$\arg \min_{mv} \sum_{i=1}^N \|F_{cur_i}^{out} - F_{ref}(mv)_i^{out}\|, \quad (10)$$

where $F_{ref}(mv)_i^{out}$ denotes the i 'th pixel of the external boundary using candidate mv . The popularly used decoder motion-vector estimation (DMVE) algorithm²³ belongs to this category. From both criteria, we can see that the performance of EC methods highly depends on the quality of the missing macroblock's neighbors along the top, left, bottom, and right directions.

With regard to the previously mentioned EC criteria, it is observed that the closer to correct macroblocks a lost macroblock is, or the more correct neighbors a lost macroblock has along four directions, the higher the concealment quality that can be achieved at this lost macroblock. Therefore, the distance of error propagation between a lost macroblock and its closest correct neighbors is a suitable indicator of the quality after EC. In order to obtain better EC perfor-

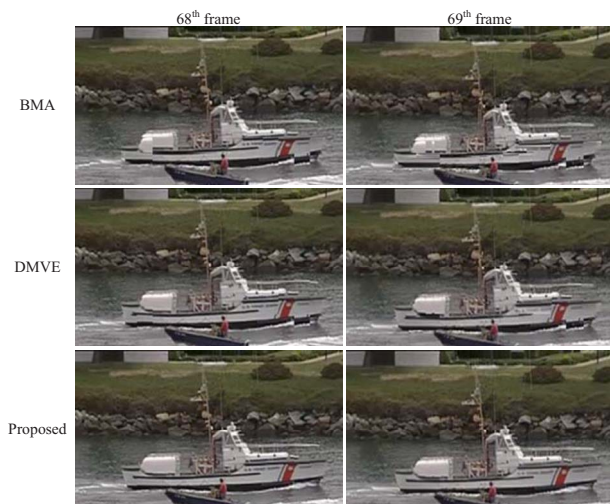


Fig. 5 Comparison of subjective quality when the ROI from the 68th frame of Coastguard is lost (QP=28). Left: concealed ROI in the 68th frame; right: error propagation in the 69th frame.

mance, we should emphasize EC performance for those lost macroblocks far from correct ones. After computing the top distance d_T , left distance d_L , bottom distance d_B , and right distance d_R for each ROI macroblock, which respectively denote the distance to the closest correct macroblocks along each direction if only the background can be correctly decoded, we define the priority metric for each ROI macroblock as

$$Pri \propto \min\{d_T, d_L, d_B, d_R\}, \quad (11)$$

where \propto denotes the direct proportion operator. That is, the MV of the ROI macroblock with the largest Pri value is embedded first, so that error propagation can be blocked earlier for the most inner parts of ROI regions. If several ROI macroblocks share the same Pri value, we assign the one with the highest priority if around it there are the fewest neighbors ranked as the top ones along four directions.

Since extracted MVs are usually quite accurate, EC methods using these MVs can be supposed to conceal ROI macroblocks reliably. Each time the MV of an ROI macroblock is embedded, the remnant ROI macroblocks should have their priorities updated accordingly. In this way, the lost ROI may be separated into smaller regions by reliable macroblocks during EC. As a result, better EC performance can be expected in our scheme due to greatly reduced error propagation.

For clarity, Fig. 3 demonstrates an example of priority calculation for ROI macroblocks. In Fig. 3(a), the priorities have been marked assuming that the MVs of the ROI macroblocks need to be embedded. It is noted that the red macroblock, which enjoys the maximum priority value, gets its MV embedded first. If this macroblock is concealed reliably, the priorities of the remnant ROI macroblocks are updated accordingly, as shown in Fig. 3(b). Now, three macroblocks share the priority value 1, among which the red macroblock has fewer reliable neighbors one grid away from it along four directions than the other two. So the red macroblock enjoys the highest priority according to our metric. As this process continues, ROI macroblocks will be

ordered in terms of decent priorities. Subsequently, macroblocks with higher priorities will have MVs embedded first.

Now, with operations for each frame at the encoder and decoder side included, our scheme is concluded as follows.

Encoder side:

For every ROI macroblock

Encode the ROI macroblock;

Search an additional MV to be embedded and convert it to binary bits as in Sec. 3.1;

End for

Order ROI macroblocks in decreasing order of priorities as in Sec. 3.3, and arrange their MV bit series to be embedded correspondingly;

For every background macroblock

If the coefficient condition in Sec. 3.2.2 is satisfied

Encode the background macroblock and embed bits as in Sec. 3.2.1;

Else

Encode the background macroblock without bits embedded;

End If

End For

Decoder side:

If ROI is lost but background is received correctly

For every background macroblock

If the coefficient condition in Sec. 3.2.2 is satisfied

Decode the background macroblock and extract some bits as in Sec. 3.2.1;

Else

Decode the background macroblock without bits extracted;

End If

End For

Convert the extracted bits to MV series as in Sec. 3.1;

Order ROI macroblocks in decreasing order of priorities as in Sec. 3.3;

Assign each extracted MV to a ROI macroblock sequentially;

Conceal ROI macroblocks with MVs assigned;

Conceal remnant ROI macroblocks normally;

Else

Decode or conceal the frame normally;

End If

Definition of the coefficient condition: There exist 4×4 luma blocks in I16 \times 16 mode having nonzero first or second AC coefficients, or 4×4 luma blocks in other modes have nonzero DC or first or second AC coefficients, or 4×4 chroma blocks having nonzero first or second AC coefficients.

4 Experimental Results

The proposed scheme is evaluated for the first 100 frames of two standard CIF sequences Stefan and Coastguard. The ROIs are the player and the large ship, respectively, and always follow the standard MPEG-4 segmentation aligned with macroblock boundaries.

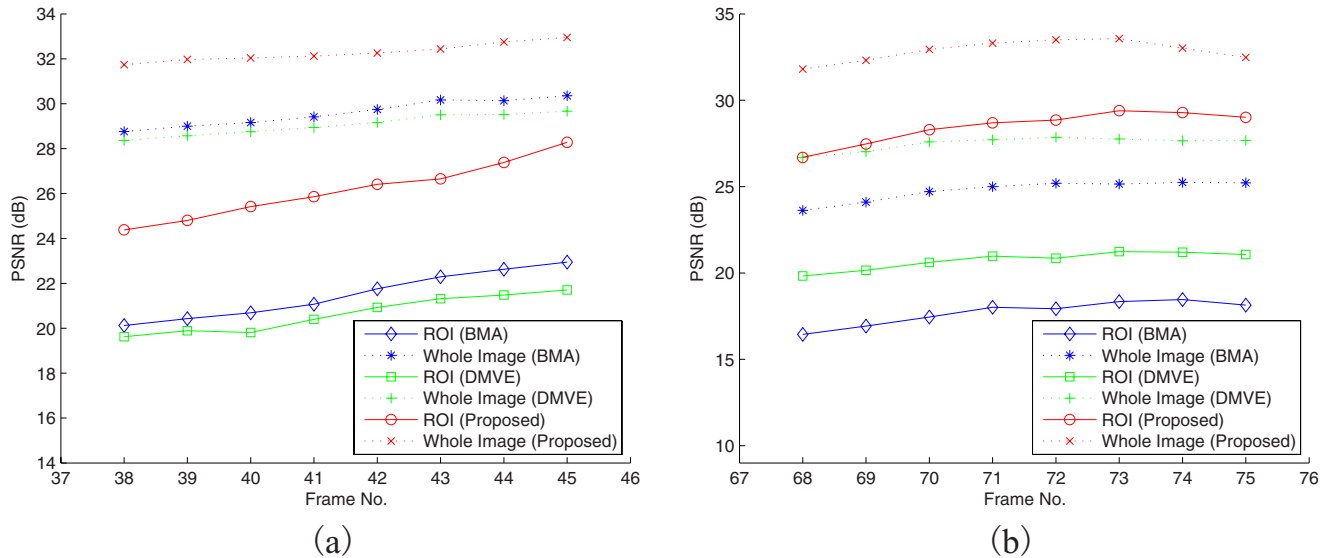


Fig. 6 EC results when the ROIs from (a) the 38th frame of Stefan and (b) the 68th frame of Coastguard are lost (QP=28).

Table 1 Impact on the coding efficiency of Stefan.

QP	18	28	38
(No embedding) PSNR _{ROI} (dB)	42.627	31.591	27.217
(With embedding) PSNR _{ROI} (dB)	42.640	34.584	27.222
ΔPSNR _{ROI} (dB)	0.013	-0.007	0.005
(No embedding) PSNR _{Frame} (dB)	43.233	35.298	27.806
(With embedding) PSNR _{Frame} (dB)	43.179	35.236	27.554
ΔPSNR _{Frame} (dB)	-0.054	-0.062	-0.252
(No embedding) bit rate (kbps)	4331.56	1391.64	288.19
(With embedding) bit rate (kbps)	4342.61	1408.44	304.87
ΔBit rate (%)	0.255	1.207	5.788

Table 2 Impact on the coding efficiency of Coastguard.

QP	18	28	38
(No embedding) PSNR _{ROI} (dB)	42.501	34.331	26.819
(With embedding) PSNR _{ROI} (dB)	42.502	34.332	26.824
ΔPSNR _{ROI} (dB)	0.001	0.001	0.005
(No embedding) PSNR _{Frame} (dB)	42.632	34.581	27.899
(With embedding) PSNR _{Frame} (dB)	42.578	34.457	27.649
ΔPSNR _{Frame} (dB)	-0.054	-0.124	-0.250
(No embedding) bit rate (kbps)	4528.28	1334.98	217.42
(With embedding) bit rate (kbps)	4542.78	1350.95	223.20
ΔBit rate (%)	0.320	1.196	2.658

JM 14.2 (Ref. 24) is modified to support our experiments. Images are divided into ROI and background slices, with FMO type 6 of H.264 applied during encoding. Groups of pictures (GOPs) of IPPP structure with one I frame inserted every 15 frames are considered. One reference frame is used for prediction, and the search range of motion estimation is set to 32 pixels. Except for the first instantaneous decoding refresh (IDR) frame, the MVs of the ROI in all the remaining frames are embedded in the background by means of our proposed scheme.

4.1 Impact on Coding Efficiency

In experiments, images are encoded at 30 fps. When quantization parameter (QP) is set to 18, 28, and 38, respectively, the coding results are shown in Table 1 and Table 2. From the results, we can see that the peak signal to noise ratio (PSNR), especially that of the ROI, does not change much. Therefore, with our scheme, image quality is well maintained. In addition the bit rate basically does not increase much as well. In a word, only minor loss of coding efficiency might be incurred by the adaptive MV embedding of our scheme.

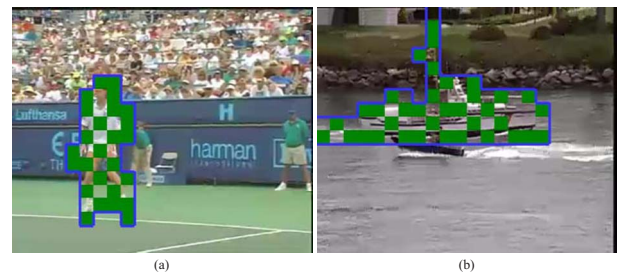


Fig. 7 Partly restored ROIs from (a) the 51st frame of Stefan and (b) the 70th frame of Coastguard by extracted MVs (QP=38).

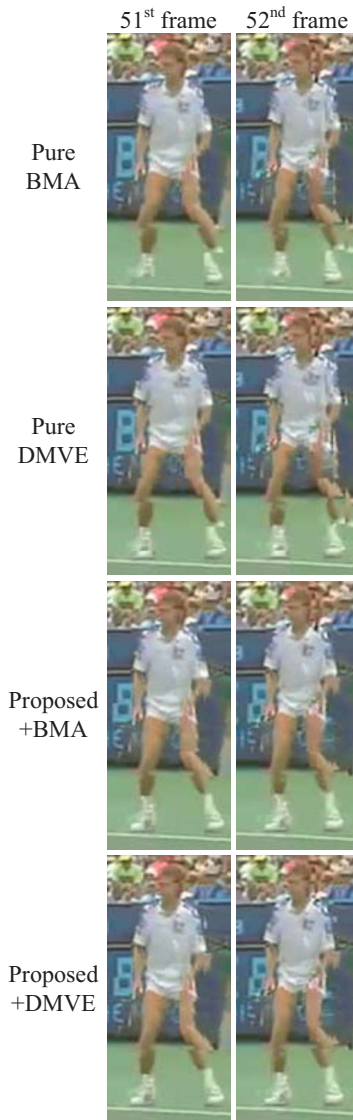


Fig. 8 Comparison of subjective quality when the ROI from the 51st frame of Stefan is lost (QP=38). Left: concealed ROI in the 51st frame; right: error propagation in the 52nd frame.

4.2 Performance of Error Concealment

Two frequently used EC algorithms, DMVE²³ with 2-pixel-wide borders and improved BMA,²⁵ are compared with our proposed scheme. We first present the results when only a single ROI is lost and then show the performance in the case of random packet loss.

4.2.1 Error concealment in the case of a single ROI loss

As described in Sec. 3, when only the background is available, our EC process of ROI can be divided into two steps. First, ROI macroblocks with MVs embedded in the background are concealed via extracted MVs. Then, conventional EC methods are resorted to for the remnant ROI macroblocks. In experiments, conventional methods perform concealment for the lost ROI from outside to inside similar to the order in Ref. 24.

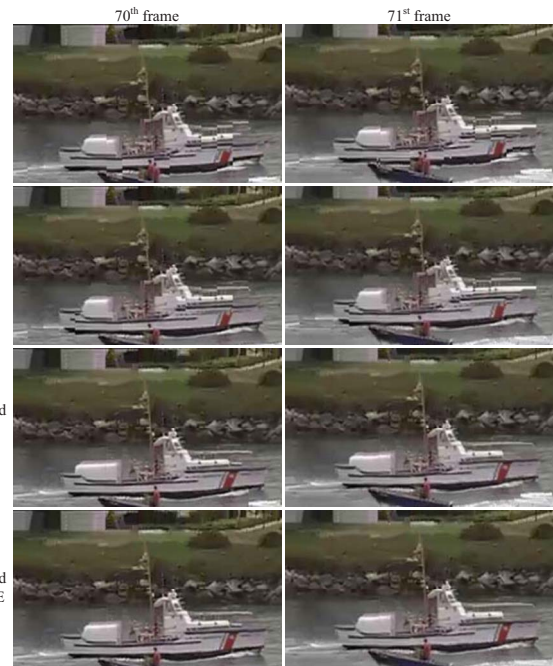


Fig. 9 Comparison of subjective quality when the ROI from the 70th frame of Coastguard is lost (QP=38). Left: concealed ROI in the 70th frame; right: error propagation in the 71st frame.

First, we present the EC performance when all MVs of the ROI macroblocks are embedded in the background. Experiments are carried out for two sequences when QP is 28. Assume that the ROIs from the 38th frame of Stefan and the 68th frame of Coastguard get lost during transmission. The concealed ROIs and the error propagation in the next frame are shown in Fig. 4 and Fig. 5. Qualities of the subsequent frames are shown in Fig. 6. It is obvious that our scheme performs much better than conventional methods. The gain is mainly attributed to better guidance of the embedded MVs during the EC of the ROI.

Next, let us see the results when only part of ROI macroblocks have MVs embedded in the background. We carry out the experiments for two sequences when QP is 38. Assume that the ROIs from the 51st frame of Stefan and the 70th frame of Coastguard get lost during transmission. After the first step of the EC with extracted MVs, the ROIs are partly restored, as shown in Fig. 7. For the other ROI macroblocks, we resort to the BMA and DMVE algorithms, respectively. The concealed ROIs and the error propagation in the next frame are shown in Fig. 8 and Fig. 9. The quality of every part of concealed ROIs is shown in Table 3 and Table 4, where ROI macroblocks concealed by extracted MVs are named *MV part*, and the other ROI macroblocks are named *remnant part*. Obviously, the proposed scheme presents higher quality and less error propagation. In addition, from both Table 3 and Table 4, we can see that not only the MV part, which is concealed with extracted MVs, exhibits higher quality, but also that the remnant part benefits from our scheme. This is because the remnant part has more reliable neighbors, which help to provide accurate motion information and evaluate candidate MVs. Therefore, both the MV embedding and the priority metric are effective for the improvement of the EC of the ROI.

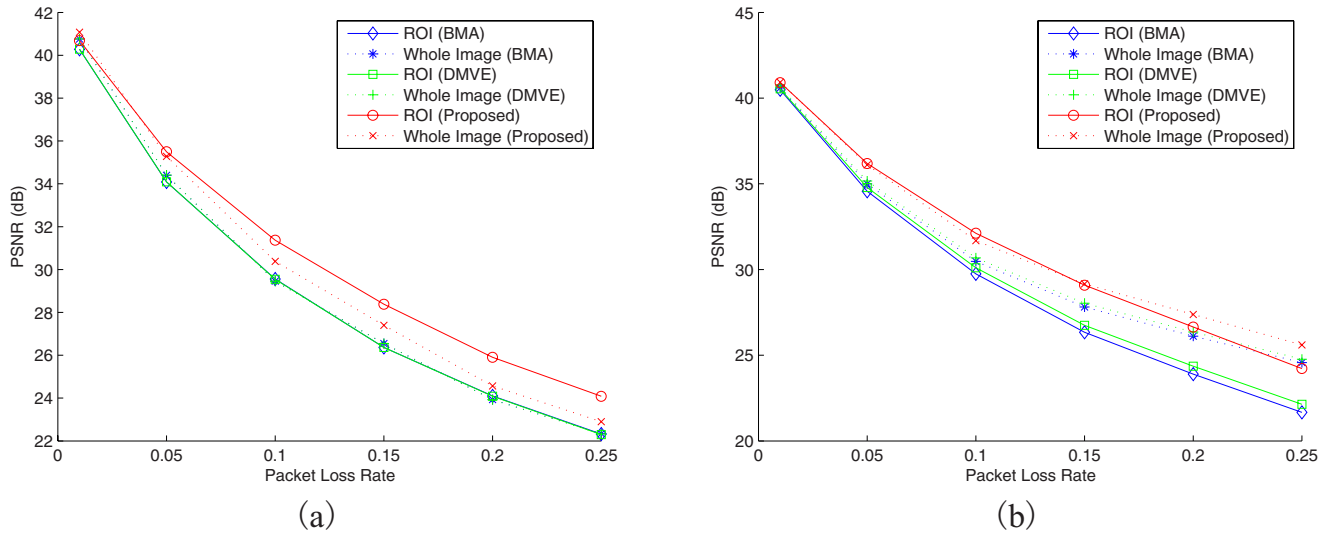


Fig. 10 EC results for (a) Stefan and (b) Coastguard in the case of random packet loss (QP=18).

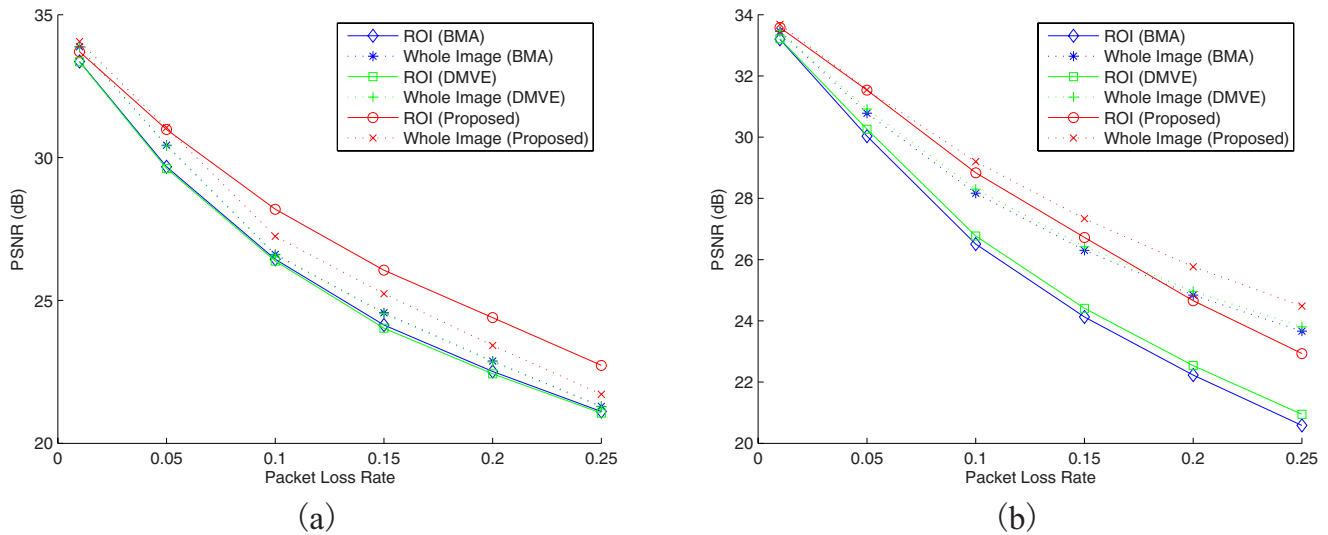


Fig. 11 EC results for (a) Stefan and (b) Coastguard in the case of random packet loss (QP=28).

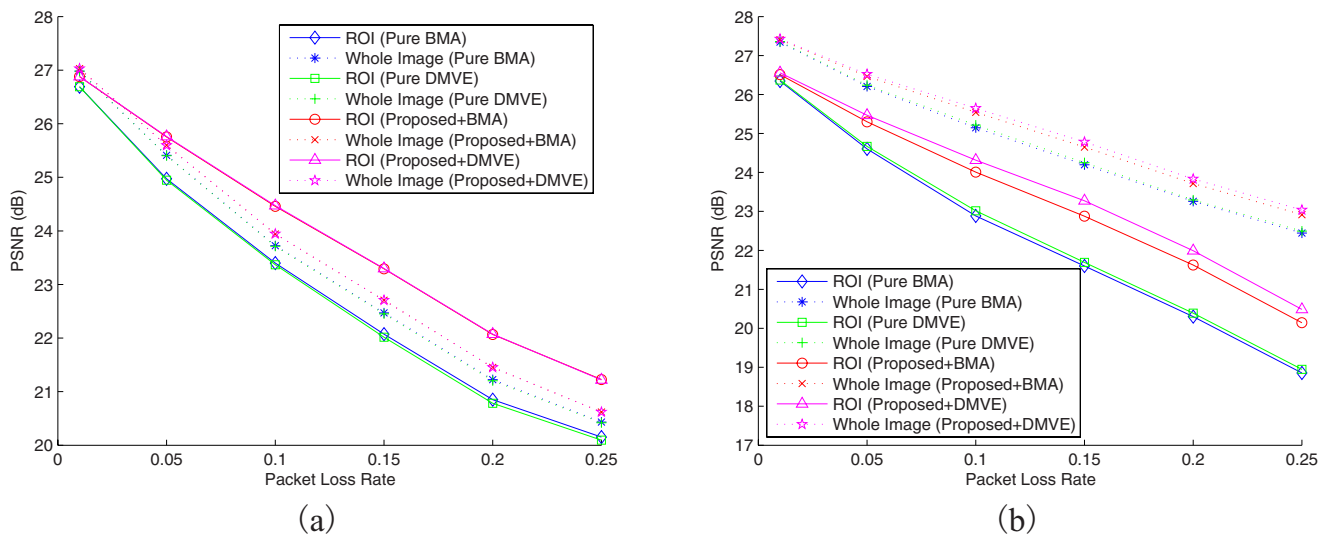


Fig. 12 EC results for (a) Stefan and (b) Coastguard in the case of random packet loss (QP=38).

Table 3 EC results for the ROI from the 51st frame of Stefan (QP=38).

EC method	MV part (dB)	Remnant part (dB)	Whole ROI (dB)
Pure BMA	18.712	20.183	19.635
Proposed+BMA	23.745	21.624	22.223
Pure DMVE	18.291	19.421	19.011
Proposed+DMVE	23.745	21.613	22.215

4.2.2 Error concealment in the case of random packet loss

To test the performance of our scheme, when the whole frame is lost, the EC is simply performed by copying from the previous frame. When only the background slice is lost, background macroblocks having collocated background macroblocks in the previous frame are concealed by direct copying, while the others are concealed by distance-weighted intra-interpolation. In the experiments, the ROI and background slices are both lost at the predefined packet-loss rate. The results of the EC are shown in Figs. 10–12, where all of the reported PSNR is averaged over 100 simulations. From the results, we can see that when MVs of ROI macroblocks can all be embedded (i.e., when QP is 18 or 28), our scheme performs much better than BMA and DMVE. When QP is 38, although only part of ROI macroblocks can have MVs embedded, the EC performance is still improved significantly. Thus, in the case of random packet-loss, the EC of the ROI is improved obviously by our scheme.

5 Conclusion

In this paper, a simple yet effective EC scheme of ROI based on data hiding is proposed. At the encoder side, MVs of the ROI are embedded in the background adaptively based on the original quantized coefficients of background macroblocks. Considering the limited embedding capacity of the background, we further propose to assign priorities to each ROI macroblock based on a predefined metric of error propagation. When an ROI gets lost but its background is available, previously embedded MVs can be extracted from the background to facilitate the EC of the ROI at the de-

Table 4 EC results for the ROI from the 70th frame of Coastguard (QP=38).

EC method	MV part (dB)	Remnant part (dB)	Whole ROI (dB)
Pure BMA	18.997	19.596	19.456
Proposed+BMA	24.452	21.163	21.708
Pure DMVE	19.799	19.831	19.824
Proposed+DMVE	24.452	22.231	22.636

coder side. Even if not all ROI macroblocks can be concealed by extracted MVs, remnant ROI macroblocks can still benefit from previously concealed ones.

When applied to the H.264/AVC standard, our scheme incurs only minor loss of coding efficiency. But experimental results show that in the case of foreground–background independent slice coding and transmission, our scheme has performance advantages across a range of packet-loss rates. Especially, the EC performance of the ROI is improved significantly, which is always desired for higher image quality.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (60702044, 60902073, 60932006, and 60625103).

References

1. Y. Wang and Z. Qin-Fan, "Error control and concealment for video communication: a review," *Proc. IEEE* **86**, 974–997 (1998).
2. F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding—a survey," *Proc. IEEE* **87**, 1062–1078 (1999).
3. Y. Peng, L. Bede, and H. H. Yu, "Error concealment using data hiding," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1453–1456 (2001).
4. A. Piva, R. Caldelli, and F. Filippini, "Data hiding for error concealment in H.264/AVC," in *Proc. IEEE 6th Workshop on Multimedia Signal Processing*, pp. 199–202 (2004).
5. M. Yang and N. G. Bourbakis, "An efficient packet loss recovery methodology for video streaming over IP networks," *IEEE Trans. Broadcast.* **55**, 190–201 (2009).
6. A. K. Anhari, S. Sodagari, and A. N. Avanaki, "Hybrid error concealment in image communication using data hiding and spatial redundancy," presented at *Int. Conf. Telecommunications*, Lyon (2008).
7. G. Gur, Y. Altug, E. Anarim, and F. Alagoz, "Image error concealment using watermarking with subbands for wireless channels," *IEEE Commun. Lett.* **11**, 179–181 (2007).
8. C. B. Adsumilli, M. C. Q. Farias, S. K. Mitra, and M. Carli, "A robust error concealment technique using data hiding for image and video transmission over lossy channels," *IEEE Trans. Circuits Syst. Video Technol.* **15**, 1394–1406 (2005).
9. M. Kurosaki and H. Kiya, "Error concealment using a data hiding technique for MPEG video," presented at *European Conf. Circuit Theory and Design*, Espoo, Finland (2001).
10. S. Chen and H. Leung, "A temporal approach for improving intra-frame concealment performance in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.* **19**, 422–426 (2009).
11. K. Li-Wei and L. Jin-Jang, "An error resilient coding scheme for H.264 video transmission based on data embedding," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, pp. iii, 257–260 (2004).
12. H. Wang, S. A. Tsiftaris, and A. K. Katsaggelos, "Joint source-channel coding for wireless object-based video communications utilizing data hiding," *IEEE Trans. Image Process.* **15**, 2158–2169 (2006).
13. P. Yin, M. Wu, and B. Liu, "A robust error resilient approach for MPEG video transmission over internet," in *Visual Communication and Image Processing*, *Proc. SPIE* **4671**, 103–111 (2002).
14. J. Song and K. J. R. Liu, "A data embedded video coding scheme for error-prone channels," *IEEE Trans. Multimedia* **3**, 415–423 (2001).
15. ISO/IEC, "Advanced video coding for generic audiovisual services," Rec. H.264 and ISO 14496–10 (March 2005).
16. S. D. Lin, S. C. Shie, and J. W. Chen, "Image error concealment based on watermarking," in *Proc. VIIth Digital Image Computing: Techniques and Applications*, Sydney, Australia, pp. 137–143 (2003).
17. J. Wang and J. Liang, "A region and data hiding based error concealment scheme for images," *IEEE Trans. Consum. Electron.* **47**, 257–262 (2001).
18. S. Wenger and M. Horowitz, "Scattered slices: a new error resilience tool for H.26L," JVT-B027, NVT of ISO/IEC MPEG & ITU-T VCEG Meeting (2002).
19. S. Wenger, M. M. Hannuksela, T. Stockhammer, M. Westerlund, and D. Singer, "RFC 3984: RTP payload format for H.264 video," IETF (2005).
20. M. Wu, H. H. Yu, and A. Gelman, "Multi-level data hiding for digital image and video," in *SPIE Photonics East Conf. on Multimedia Systems and Applications*, SPIE Press, Bellingham, WA (1999).
21. D. Agrafiotis, D. R. Bull, and C. N. Canagarajah, "Enhanced error

concealment with mode selection," *IEEE Trans. Circuits Syst. Video Technol.* **16**, 960–973 (2006).

22. W. M. Lam, A. R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 417–420 (1993).
23. J. Zhang, J. F. Arnold, and M. R. Frater, "A cell-loss concealment technique for MPEG-2 coded video," *IEEE Trans. Circuits Syst. Video Technol.* **10**, 659–665 (2000).
24. "JM 14.2 Reference Software," available at: <http://iphome.hhi.de/suehring/tml/download>.
25. W. Ye-Kui, M. M. Hannuksela, V. Varsa, A. Hourunranta, and M. Gabbouj, "The error concealment feature in the H.26L test model," presented at *Int. Conf. Image Processing* (2002).

Zhengyi Luo received a BS degree in information engineering from Nanjing University of Posts and Telecommunications in 2004 and an MS degree in electronic engineering from Shanghai Jiao Tong University in 2007. He is currently working toward a PhD degree at the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University. His research interest is video coding.



Li Song received BS and MS degrees in electronic engineering from Nanjing University of Science and Technology in 1997 and 2000 and a PhD degree from Shanghai Jiao Tong University, in 2005. He is currently an associate professor at the Institute of Image Communication and Information Processing, Department of Electronic Engineering, Shanghai Jiao Tong University. He has published over 60 research papers and filed 18 patents. His research interests are in the areas of signal processing, image and video coding, computer vision, and machine learning.

areas of signal processing, image and video coding, computer vision, and machine learning.



Shibao Zheng is a professor in the Department of Electronic Engineering, Shanghai Jiao Tong University, and is an IEEE member. He received BS and MS degrees from Xidian University, Xi'an, China in 1983 and 1986, respectively. From 1986 to 1999, he was an expert for the National Project in HDTV. He has made great achievements in the field of image communication, DTV, and IC design. In recent years, he has done a lot of research work and made great progress

in the field of intelligent video analysis and video surveillance systems. His research interests include DTV, intelligent video surveillance, and network multimedia.



Yi Xu received BS and MS degrees in electronic engineering from Nanjing University of Science and Technology in 1996 and 1999 and a PhD degree from Shanghai Jiao Tong University in 2005. She is currently a lecturer at the Institute of Image Communication and Information Processing, Department of Electronic Engineering, Shanghai Jiao Tong University. Her research interests include quaternion wavelet theory and application, computer vision, and artificial intelligence.



Xiaokang Yang received a BSc degree from Xiamen University, China, in 1994, an MEng degree from the Chinese Academy of Sciences, Beijing, in 1997, and a PhD degree from Shanghai Jiao Tong University in 2000. He is currently a professor at the Institute of Image Communication and Information Processing, Department of Electronic Engineering, Shanghai Jiao Tong University. From 2002 to 2004, he was a research scientist at the Institute for Info-comm Research, Singapore. His current research interests include scalable video coding, video transmission over networks, video quality assessment, digital television, and pattern recognition.