

A Hybrid Frame Concealment Algorithm for H.264/AVC

Bo Yan, Member, IEEE, and Hamid Gharavi, Fellow, IEEE

Abstract

In packet-based video transmissions, packets loss due to channel errors may result in the loss of the whole video frame. Recently, many error concealment algorithms have been proposed in order to combat channel errors, however most of the existing algorithms can only deal with the loss of macroblocks and are not able to conceal the whole missing frame. In order to resolve this problem, in this paper we have proposed a new hybrid motion vector extrapolation (HMVE) algorithm to recover the whole missing frame, and it is able to provide more accurate estimation for the motion vectors of the missing frame than other conventional methods. Simulation results show that it is highly effective and significantly outperforms other existing frame recovery methods.

Index Terms

Error concealment, error-resilient video transmission, whole-frame losses, H.264/AVC.

I. Introduction

With the rapid development of wireless networks, more and more users are seeking video services over wireless networks. In video transmission, video data must often be compressed to reduce the bit rate for transmission over communication channels. However highly compressed video data is more sensitive to channel errors. Since wireless networks are always error-prone and cannot provide the guaranteed quality of service (QoS), the coded data transmitted over them may be corrupted. Loss of a single bit often results in the loss of a whole block and may cause consecutive block losses. In addition, because of motion compensation, this error may propagate to succeeding frames [2-4].

To alleviate this effect, error concealment techniques are considered to minimize the distortion caused by loss of packets at the decoder. These techniques take advantage of the spatial and/or temporal correlation in the received video to interpolate the missing data. For image coding and INTRA coded pictures (I-pictures), only the spatial correlation will be exploited. A simple and typical spatial error concealment method is bilinear interpolation, which is based on interpolating each pixel in the lost areas from intact neighboring pixels [5]. In [6-10], more advanced concealment techniques were proposed to perform the interpolation adaptively in order to recover the missing areas with a greater accuracy.

In the case of motion compensated interframe coded video (e.g., P-frame), the first step in concealing the missing macroblocks (MBs) is to predict their motion vectors (MVs) using the MVs of the neighboring blocks. The predicted MVs and the reference frame are then used to conceal the missing MBs in the frame [2-4]. If there

¹ B. Yan is with School of Computer Science, Fudan University, Shanghai, China (email: byan@fudan.edu.cn). H. Gharavi is with the National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD USA 20899 (email: gharavi@nist.gov). Parts of this work were presented in [1]. This work is supported in part by NSFC (Grant No.: 60703034), and in part by Shanghai Pujiang Program (Grant No.: 07PJ14017).

are no neighboring MBs, the concealment has to rely entirely on the previously decoded frame. This situation may arise when the entire P-frame is encapsulated into a single packet (i.e., transmitting video at very low bitrates) or the transmission may suffer from a burst of packet loss. Bear in mind that the size of the packet is normally selected to ensure the optimal trade-off between channel conditions and packet overhead. For example, for real-time video streaming the header information is not solely limited to the overhead associated with the video coding layer (i.e., slice header), but also may include RTP, UDP and IP (not to mention MAC/PHY) overheads. Although it is possible to compress the RTP/UDP/IP headers [11], due to the high compression efficiency of recent video coding techniques, it is often possible that the entire P-frame may indeed be encapsulated into a single, but small-sized packet.

H.264/AVC [12, 13] is the latest standard developed by a joint committee of the ISO MPEG and the ITU-T VCEG, which permits a significant reduction in bit rate compared to all previous standards, but at the same quality level [14, 15]. On the one hand, in H.264/AVC a frame can be divided into different slices. Each slice can be decoded independently. Therefore, during video transmission each slice can be packetized into one RTP packet. The loss of one packet due to channel errors will not corrupt the decoding of other correct packets in the same frame. However, multiple packet burst losses are quite common in the wireless networks. Thus all the packets of one frame are very likely to be corrupted by channel errors during transmission. In this case, the losses of multiple RTP packets result in the loss of a whole frame. On the other hand, due to the high compression ratio of H.264/AVC, it is often common that an entire coded picture fits the packet size when transmitting low resolution sequences at low bitrates. Then usually one coded frame data will be packetized into one RTP packet. In such a scenario, the loss of an RTP packet results in the loss of a whole frame. Because of the temporal prediction, the loss of a frame can significantly affect the quality of subsequent frames in H.264/AVC. Conventional block-based spatial and temporal interpolation concealment algorithms [2] are not suitable for frame concealment. Therefore, it is necessary to propose an efficient method in order to recover the lost frames.

Frame concealment is an effective method to combat channel errors, and is aimed at masking the effect of missing frames to create subjectively acceptable images. So far, many frame concealment methods have been proposed [16-22]. Among these methods, they generally generate the MVs of the lost frame with the optical flow estimation [23-28], or the motion vector extrapolation (MVE) [29, 30]. The drawback of these methods is that the quality of the concealed frame is not satisfactory.

In this paper, we propose a new hybrid motion vector extrapolation (HMVE) algorithm to give more accurate estimation for the MVs of the lost frames, than other conventional methods in order to conceal the lost frames. Firstly, it will construct an MV set including all the possible MV values of each pixel in the lost frame. Secondly, wrongly extrapolated MVs in this MV set will be discarded with the proposed algorithm. Finally, the frame will be concealed with the MV of each pixel. Experimental results show that the HMVE algorithm significantly outperforms other existing methods in concealing the lost frame.

This paper is organized as follows. In Section 2, we briefly introduce the existing techniques. In Section 3, we propose our HMVE algorithm, which can improve the decoded video quality significantly after transmission over error-prone channels. Then we evaluate the proposed method by simulations and present the results in Section 4. Finally, in Section 5, we draw the conclusions.

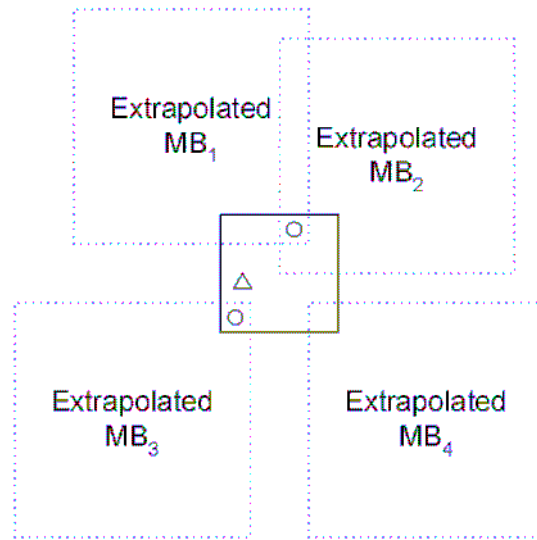


Fig. 1. Pixel-based motion vector extrapolation [30].

II. Existing Techniques

Recently, some frame concealment techniques have been proposed to combat frame loss during video transmission [16-30].

Belfiore *et al.* proposed a method based on optical flow, which is able to conceal the lost frames [25,26]. This method firstly estimates the optical flow for each pixel of the last decoded frame with the MVs of a few past frames. Then the pixels of the last decoded frame are projected onto the lost frame to reconstruct it. For the pixel in the missing frame projected by multiple pixels of the last decoded frame, its luminance or chrominance value is the average of all contributions. The pixel, which is not filled after projection, is interpolated by means of a 9×9 median filter. This method usually provides a relatively good quality, however its performance is not always better than the frame copy (FC). In addition, it's hard to determine the number of the reference frames while calculating the optical flow of each pixel. Wang *et al.* proposed to use both the forward and backward MVs for bi-directional temporal prediction to handle the entire frame losses [22]. However it may not be suitable for delay-sensitive applications.

MV extrapolation (MVE) is a simple but efficient way to achieve the MV of the lost frame [29]. In this method, the MVs of macroblocks (MBs) are first extrapolated from the last decoded frame to the missing frame. Then the overlapped area between the missing block and the motion extrapolation MB is estimated and the best MV is selected. Finally, the corrupted block is concealed with the MV and previous frame. This method is able to overcome the disadvantage of incorrect MB displacement, but the block (8×8) based MV is too rough to cause block artifacts.

Chen *et al.* proposed a pixel-based MVE (PMVE) method to conceal the missing frame, which is able to get the MV by extending the MV extrapolation (MVE) method to the pixel level [30]. As shown in Fig. 1, the pixels in the missing frame can be divided into two parts:

- 1) For a pixel which is covered by at least one extrapolated MB, such as the circle points in Fig. 1, the MV is estimated by averaging the MVs of all overlapped MBs.

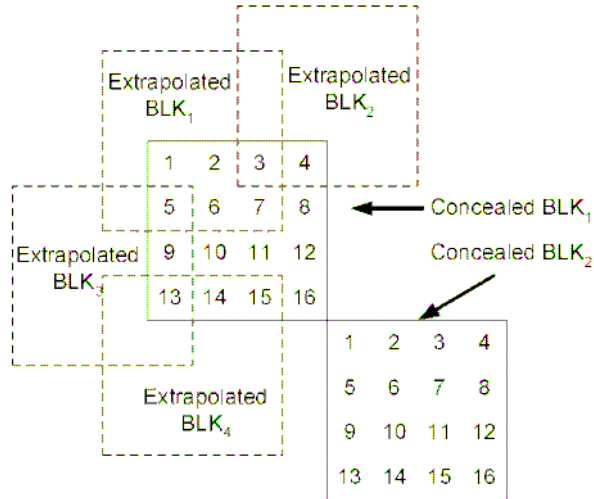


Fig. 2. The proposed hybrid MV extrapolation.

- 2) For a pixel which is not covered by any of the extrapolated MBs, such as the triangle point in Fig. 1, the MV is duplicated from the MV of the same pixel in the previous frame [30].

If the estimated MV is $MV = (MV_x, MV_y)$, each missing pixel $p_m(x, y)$ can be recovered as follows:

$$p_m(x, y) = p_r(x + MV_x, y + MV_y) \quad (1)$$

where $p_r(x, y)$ refers to pixels in the previous frame [30].

This PMVE method provides a similar performance as block-based MVE does in little motions, but it is able to improve the performance greatly in large motion scenes.

III. Our Proposed Algorithm

The PMVE method is able to provide a better performance, however it does have a shortcoming. The MVs of the lost pixels, which are extrapolated from the MVs of the corresponding pixels in the previous frame, may not be accurate. Some MVs are very likely to be extrapolated wrongly, especially in large motion scenes. This shortcoming will damage the accuracy of the MV for the pixel, thus the performance will be degraded. In order to overcome this problem we propose a hybrid MVE (HMVE) method based on PMVE, which uses not only the extrapolated MVs of the pixels, but also the extrapolated MVs of the blocks. This proposed algorithm is able to discard the wrongly extrapolated MVs in order to obtain the accurate MV.

A. Pixels Classification

Different from PMVE, the proposed HMVE method firstly classifies the pixels of the missing frame into three parts:

- 1) Part A: pixels that are covered by at least one extrapolated 4×4 block. For example, in Fig. 2 Part A includes pixels of $\{1, 2, 3, 4, 5, 6, 7, 9, 13, 14, 15\}$ in the concealed block₁.
- 2) Part B: pixels that are not covered by any of the extrapolated 4×4 blocks. But the block which the pixel belongs to, has the overlapped area with the extrapolated block. For example, pixels of $\{8, 10, 11, 12, 16\}$ in the concealed

block₁ belong to Part B.

- 3) Part C: pixels that are not covered by any of the extrapolated 4×4 blocks. And the block which the pixel belongs to, doesn't overlap with the extrapolated block. For example, all the pixels in the concealed block₂ belong to Part C.

Then HMVE will use a new scheme to estimate the MV of each pixel in order to discard the wrongly extrapolated MVs.

In H.264/AVC, the smallest unit for motion estimation and compensation is a 4×4 block [12,13]. Thus in our proposed HMVE algorithm, we use a 4×4 block as the concealment unit as shown in Fig. 2.

B. MV Extrapolation

Two possible MVs of each block in the missing frame will be estimated by the MVE method [29]. As shown in Fig. 2, the MV of the block in the missing frame is estimated according to the extrapolated blocks, which occupy the missing block. The number of pixels in the overlapped areas (as shown in Fig. 2) is used to obtain the weight for the estimation. Let EB_n^j denotes the extrapolated 4×4 block from the j th block in the reference frame to the missing frame n , $MV(EB_n^j)$ denotes the MV of EB_n^j , and B_n^i denotes the i th 4×4 block in frame n . Then the weight is given by [29]:

$$w_n^{i,j} = \sum_{p \in B_n^i} f_j(p), \quad i, j = 1, 2, \dots, M \quad (2)$$

where

$$f_j(p) = \begin{cases} 1, & p \in EB_n^j(4)(5) \\ 0, & p \notin EB_n^j(6)(7) \end{cases} \quad (3)$$

M is the total number of blocks in a video frame.

Then two possible MVs of the missing block B_n^i are obtained, which are denoted by $MV_m(B_n^i)$ and $MV_a(B_n^i)$. $MV_m(B_n^i)$ is obtained by selecting the MV of the extrapolated block with the maximum weight $w_n^{i,j}$. $MV_a(B_n^i)$ is obtained by the weighted mean value of the MVs of all the overlapped extrapolated blocks. They are obtained as follows.

- 1) $MV_m(B_n^i)$:

$$MV_m(B_n^i) = MV(EB_n^{j^*}) \quad (4)$$

where

$$j^* = \operatorname{argmax}\{w_n^{i,j}\}, \quad j = 1, 2, \dots, M$$

- 2) $MV_a(B_n^i)$:

$$MV_a(B_n^i) = \frac{\sum_{j=1}^M MV(EB_n^j) w_n^{i,j}}{\sum_{j=1}^M w_n^{i,j}} \quad (5)$$

If block B_n^i has not been overlapped by any extrapolated blocks, the MV of this block is null.

3.3. MV Selection

After MV extrapolation, an extrapolated MV set $MVS_p(P_n^{x,y})$ for each pixel is obtained. $P_n^{x,y}$ represents the pixel with the coordinate (x,y) in frame n . After extrapolation, pixels in Part A (as shown in Fig. 2) are covered by at least one extrapolated block, thus the $MVS_p(P_n^{x,y})$ of them includes the MVs of all overlapped extrapolated blocks. For pixels in Part B and C, the $MVS_p(P_n^{x,y})$ is null.

Then the new MV set $MVS(P_n^{x,y})$ will be constructed for different types of pixels as follows;

- 1) For pixels in Part A:

$$MVS(P_n^{x,y}) = \{MV_m(B_n^i), MV_a(B_n^i), MVS_p(P_n^{x,y})\} \quad (6)$$

where $P_n^{x,y} \in B_n^i$.

In this MV set, some MVs may be obtained due to the incorrect extrapolation. In order to get more accurate MV, the wrongly extrapolated MVs should be discarded. Let $MV(i)$ denote the i th component in $MVS(P_n^{x,y})$.

- a) Firstly, HMVE calculates the distances between $MV(i)$ and other MVs in $MVS(P_n^{x,y})$.

$$Dis(i, j) = \sqrt{(MV_x(i) - MV_x(j))^2 + (MV_y(i) - MV_y(j))^2} \quad (7)$$

$$1 \leq j \leq N, \quad j \neq i$$

where MV_x and MV_y denote the x and y components of the MV respectively, and N denotes the total number of MV in the set $MVS(P_n^{x,y})$.

- b) Then $num(i)$ is calculated based on $Dis(i, j)$, which denotes the number of $Dis(i, j)$ values that are less than the predefined threshold T for $1 \leq j \leq N$ and $j \neq i$.
- c) Finally, only the $MV(i)$, which satisfies the condition of $num(i) \geq N - 1$, will be considered as the true MV candidate. Other MVs will be discarded from $MVS(P_n^{x,y})$ as the wrongly extrapolated MVs. In this case, the accurate MV set is obtained.

- 2) For pixels in Part B:

$$MVS(P_n^{x,y}) = \{MV_m(B_n^i), MV_a(B_n^i)\} \quad (8)$$

- 3) For pixels in Part C:

$$MVS(P_n^{x,y}) = \{MV(P_{n-1}^{x,y})\} \quad (9)$$

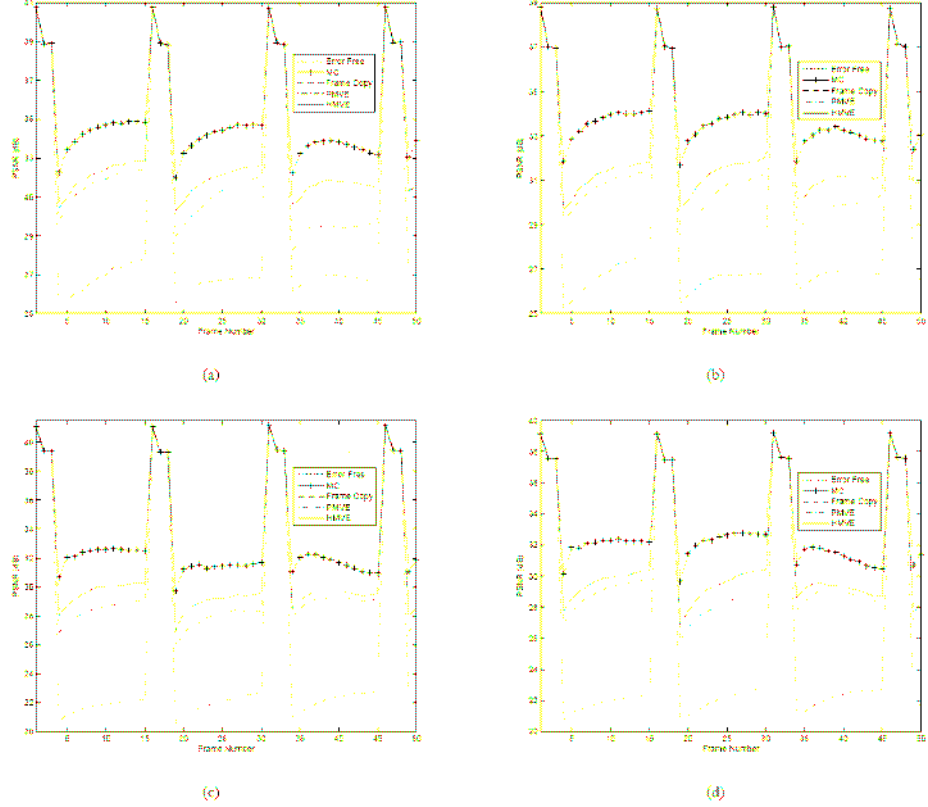


Fig. 3. "Mobile" sequence PSNR comparison vs. frame number. (a)QCIF resolution and QP=22; (b)QCIF resolution and QP=24; (c)CIF resolution and QP=22; (d) CIF resolution and QP=24.

where $MV(P_{n-1}^{x,y})$ is the MV of the same pixel in the previous frame.

Finally, the MV of each pixel $MV = (MV_x, MV_y)$ is estimated by averaging the components of the MV set $MVS(P_n^{x,y})$. With help of the estimated MV, each missing pixel can be recovered as shown in (1).

4. Simulation Results

The proposed HMVE method is simulated using the JM10.2 H.264/AVC codec. Two standard video sequences, "Mobile" and "Bus", are used to evaluate the performance of the proposed algorithm. These sequences are encoded by the H.264/AVC standard. The frame size is both at QCIF (176×144) and CIF (352×288) resolution, and the frame rate is 30 frames/s. The period of I frame reset is 15 and the number of reference frames is 1. A constant quantization parameter (QP) of 22 or 24 is maintained for all frames. A frame only contains one slice. Thus, a slice loss means a frame loss. As for the packetization scheme, all the compressed video streams related to one frame are stuffed into one packet. This is a typical condition for streaming RTP/UDP video over a network in which all packets have the same priority. In this case, packet loss means frame loss.

In order to demonstrate the effectiveness of our proposed method, we compare performances of the proposed HMVE algorithm with PMVE, frame copy (FC) and motion compensation (MC). FC means copying the previous frame directly for concealment. MC means that the original MVs are correctly received, but the residual

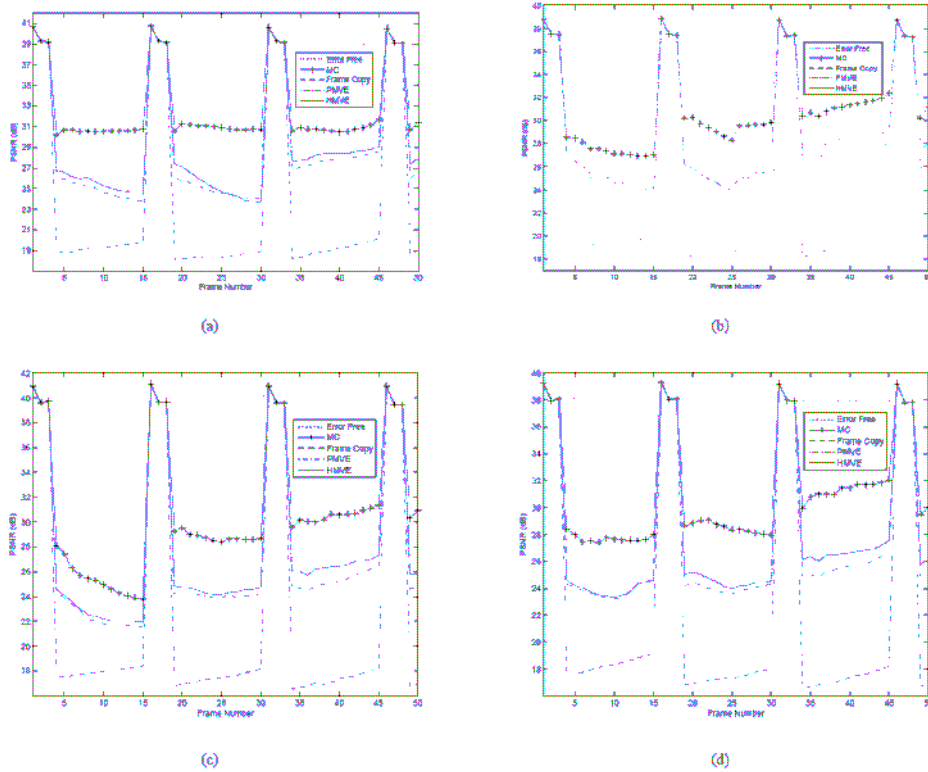


Fig. 4. "Bus" sequence PSNR comparison vs. frame number: (a)QCIF resolution and QP=22; (b)QCIF resolution and QP=24; (c)CIF resolution and QP=22; (d) CIF resolution and QP=24.

information has been lost. Note that this is a very likely scenario if data partitioning is enabled. If the information of one frame is totally lost, its MVs can only be estimated from those of other frames. Different from estimating the MVs for block concealment algorithms, there is no neighboring information for help to estimate the MVs in frame concealment. It's much more difficult for MVs' estimation of frame concealment than that of block concealment. Therefore, MC could be regarded as the "upper bound" while comparing the performances of temporal error concealment methods. In this simulation, a P-frame is dropped in each group of pictures (GOP). The dropped frame is then concealed by MC, FC, PMVE and HMVE respectively. Their corresponding PSNR values are calculated and compared. Fig. 3 and Fig. 4 show the simulation results with PSNR vs. frame number for different test sequences with different resolutions and QP's. Experimental results report that the proposed HMVE algorithm significantly outperforms FC and PMVE for different sequences.

For clearer comparison, Table I presents the average PSNR performances over the erroneous frames only, which are defined as frames corrupted by the frame losses. As shown in this table, the proposed HMVE algorithm yields higher PSNR performance than PMVE and FC, and is able to provide up to 8.88 dB and 1.09 dB better PSNR performances than the FC and PMVE respectively.

In order to evaluate the capability to stop error propagation for the proposed algorithm, Table II provides the average PSNR performances over the erroneous frames and the following P frames, which refer the erroneous frame for encoding. These frames are defined as the frames corrupted not only by the frame losses, but also by the error propagation. From this table, we can observe that the proposed HMVE algorithm yields up to 7.48 dB and 1.15 dB gain over FC and PMVE respectively. Therefore, it is quite effective in stopping the error propagation.

For subjective evaluation, Fig. 5 and Fig. 6 show the results of one frame extracted

TABLE I
COMPARISON OF THE AVERAGE PSNR PERFORMANCE OVER ERRONEOUS FRAMES ONLY.

Sequence	QP	PSNR (dB)					Gain (dB)		
		Error Free	MC	FC	PMVE	HMVE	HMVE-FC	HMVE-PMVE	
QCIF	Mobile	22	38.85	32.38	25.89	29.59	30.67	4.79	1.09
		24	36.91	31.95	25.82	29.36	30.14	4.32	0.78
	Bus	22	39.13	30.54	18.43	26.25	27.30	8.88	1.06
		24	37.29	29.84	18.45	26.21	26.93	8.48	0.72
CIF	Mobile	22	39.35	30.63	20.92	26.93	27.90	6.99	0.98
		24	37.54	30.31	20.92	26.96	27.72	6.80	0.76
	Bus	22	39.63	29.33	16.92	24.41	25.32	8.40	0.91
		24	37.96	29.13	16.93	24.49	25.40	8.47	0.91

TABLE II
COMPARISON OF THE AVERAGE PSNR PERFORMANCE OVER ERRONEOUS AND AFFECTED FRAMES.

Sequence	QP	PSNR (dB)					Gain (dB)		
		Error Free	MC	FC	PMVE	HMVE	HMVE-FC	HMVE-PMVE	
QCIF	Mobile	22	38.80	33.93	26.69	30.68	31.82	5.13	1.15
		24	36.89	33.32	26.63	30.27	31.11	4.48	0.84
	Bus	22	39.08	30.78	18.97	25.84	26.43	7.46	0.59
		24	37.28	29.48	19.11	25.80	26.28	7.17	0.47
CIF	Mobile	22	39.31	31.73	21.91	28.33	29.30	7.39	0.97
		24	37.49	31.76	21.96	28.46	29.26	7.30	0.80
	Bus	22	39.63	28.33	17.52	23.98	24.63	7.11	0.65
		24	37.98	29.20	17.62	24.51	25.10	7.48	0.59

TABLE III
COMPARISON OF THE AVERAGE PSNR PERFORMANCE OVER ALL FRAMES FOR DIFFERENT PLRS

Sequence	QP	PSNR (dB)					Gain (dB)		
		Error Free	MC	FC	PMVE	HMVE	HMVE-FC	HMVE-PMVE	
PLR=3%	Mobile	22	38.97	37.22	34.99	35.85	36.53	1.54	0.68
		24	37.06	35.75	33.65	34.53	34.92	1.27	0.39
	Bus	22	39.23	37.11	34.79	36.25	36.58	1.79	0.33
		24	37.42	35.67	33.37	34.96	35.12	1.75	0.16
PLR=5%	Mobile	22	38.97	36.17	32.1	33.81	34.6	2.5	0.79
		24	37.06	35	31.19	32.75	33.48	2.29	0.73
	Bus	22	39.23	32.45	23.67	28.74	29.35	5.68	0.61
		24	37.42	31.19	23.28	28.18	28.59	5.31	0.41

from different sequences, where (a) is the error-free frame, (b) to (e) are the images reconstructed using MC, FC, PMVE and the proposed HMVE algorithm respectively. Due to discarding the wrongly extrapolated MV, the proposed HMVE algorithm is able to capture local motions more accurately than PMVE and leads to a concealed frame with less block artifacts. In Fig. 5 and Fig. 6, it can be seen that the image quality is consistent with the PSNR measurement and the visual improvements are observed to be more significant than the PSNR improvements. In these figures, (e) is perceptually superior to (c) and (d), especially around the edges of the image objects.

It should be noted that the proposed HMVE algorithm could also be capable of concealing the I-frame in situations where the entire I-frame is lost and the missing I-frame is not representing a scene cut. Bear in mind that in these situations, the receiver, regardless of the frame type, can continue to apply the HMVE algorithm to conceal the lost frame. To examine this, we carried out further tests to assess the performance of HMVE when a missing frame is an I-frame. Fig. 7 shows the PSNR vs. frame number for QCIF "Mobile" and "Bus" when QP is 22. Fig. 7 (a) corresponds to a case when three I-frames (16th, 31st and 46th frames) are dropped for the "Mobile" sequence. The results indicate that HMVE can improve the PSNR performance by up to 7.16 dB (46th frame) and 1.32 dB (16th frame), compared with

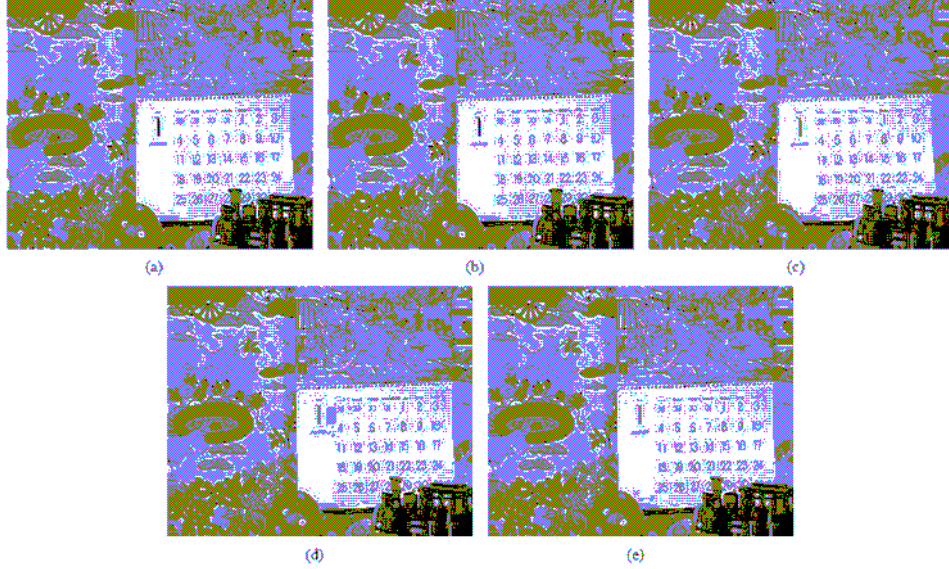


Fig. 5. Restored 28th frame of "Mobile" CIF sequence when QP is 24. (a)error-free frame (37.51 dB); (b)MC frame (32.73 dB); (c)concealed with FC (22.68 dB); (d) concealed with PMVE (28.64 dB); (e)concealed with HMVE (30.25 dB. Gains over FC and PMVE are 7.57 dB and 1.61 dB respectively).

TABLE IV
COMPARISON OF THE AVERAGE PSNR PERFORMANCE OVER ERRONEOUS FRAMES ONLY FOR DIFFERENT PLRS.

Sequence	QP	PSNR (dB)					Gain (dB)		
		Error Free	MC	FC	PMVE	HMVE	HMVE-FC	HMVE-PMVE	
PLR=3%	Mobile	22	38.77	31.87	25.39	27.97	29.73	4.34	1.76
		24	36.91	31.56	25.35	27.88	29.1	3.75	1.22
	Bus	22	39.06	30.03	18.33	25.8	27.12	8.79	1.32
		24	37.23	30.07	18.35	26.16	26.86	8.51	0.7
PLR=5%	Mobile	22	38.86	32.45	25.89	28.01	29.06	3.17	1.05
		24	37.02	32.15	25.82	27.85	28.77	2.95	0.92
	Bus	22	39.16	30.33	18.42	24.55	25.26	6.84	0.71
		24	37.37	29.87	18.44	24.59	25.08	6.64	0.49

FC and PMVE, respectively. Similarly, Fig. 7(b) shows the SNR performance for the "Bus" sequence when the missing I-frames are limited to two frames (31st and 46th frames). Again, we observe that HMVE can significantly outperform FC and PMVE by up to 9.17 dB (31st frame) and 1.23 dB (31st frame), respectively. We should point out that since there are no MVs for the I-frame, the performance of MC is not included in Fig. 7. Nonetheless, these results indicate that the proposed HMVE is capable of providing a much better performance in concealing I-frames than FC and PMVE.

It should also be noted that the main purpose of frame concealment is not only to conceal the effect of missing frames, but also to prevent error propagation. The simulation results in Table I and II clearly show that our algorithm is highly effective in these two aspects. In order to access the performance of the HMVE scheme under "real-world" environments, we used a packet loss model in [31] for QCIF sequences. The simulation results are shown in Table III to V. Table III presents the PSNR performances averaged over all frames for different packet loss ratios (PLRs). These results indicate that the proposed HMVE algorithm can maintain its superiority even under "real-world" packet loss environments (i.e., up to 5.68 dB over FC and 0.79 dB over PMVE). Similarly, Tables IV and V provide the performance comparisons in concealing erroneous frames, as well as stopping error propagation for different PLRs. In particular, Table IV indicates that the proposed algorithm can provide up to 8.79 dB and 1.76 dB PSNR gains over FC and PMVE for erroneous frames.

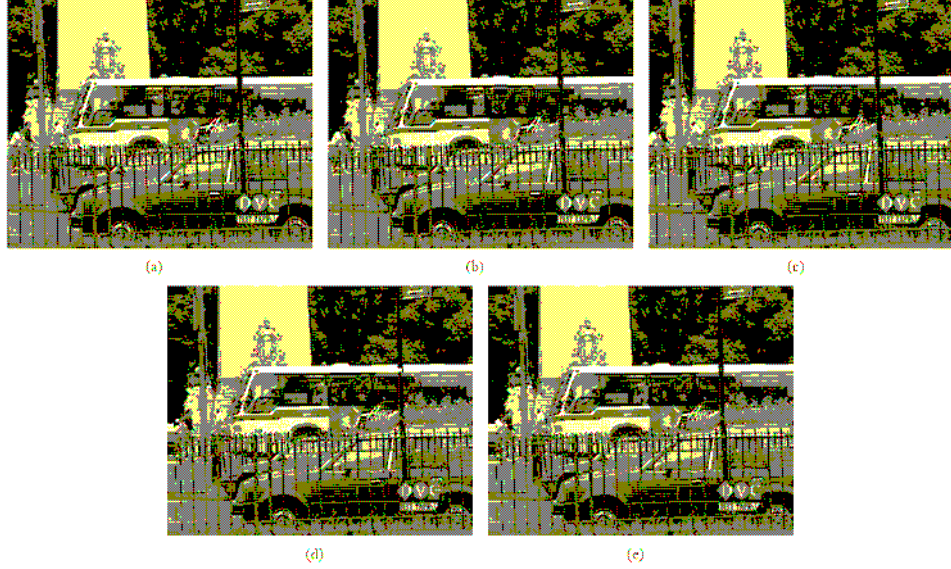


Fig. 6 Restored 34th frame of "Bus" CIF sequence when QP is 24. (a)error-free frame (37.93 dB); (b)MC frame (29.96 dB); (c)concealed with FC (16.57 dB); (d) concealed with PMVE (24.75 dB); (e)concealed with HMVE (26.14 dB). Gains over FC and PMVE are 9.57 dB and 1.39 dB respectively.

TABLE V
COMPARISON OF THE AVERAGE PSNR PERFORMANCE OVER ERRONEOUS AND AFFECTED FRAMES FOR DIFFERENT PLRS.

Sequence	QP	PSNR (dB)					Gain (dB)		
		Error Free	MC	FC	PMVE	HMVE	HMVE-FC	HMVE-PMVE	
PLR=3%	Mobile	22	38.81	33.32	26.34	29.05	31.16	4.82	2.11
		24	36.89	32.8	26.23	28.97	30.22	3.99	1.25
	Bus	22	39.07	29.55	18.84	25.58	27.11	8.27	1.53
		24	37.28	29.47	18.79	26.21	26.94	8.15	0.73
PLR=5%	Mobile	22	38.82	33.78	26.51	29.93	31.52	5.01	1.59
		24	36.91	33.21	26.4	29.52	30.99	4.59	1.47
	Bus	22	39.09	30.37	19.14	26.19	27.03	7.89	0.84
		24	37.29	29.18	19.17	25.97	26.55	7.38	0.58

In our final set of experiments we examine the effect of GOP size on the concealment performance. In these experiments, only the first frame is coded as an I-frame and the rest are coded as P-frames. Three frames are dropped in each sequence. Fig. 8 shows the PSNR vs. frame number for the QCIF "Mobile" and "Bus" sequences when QP is 22. For the "Mobile" sequence (Fig. 8(a)), the HMVE's gain over FC and PMVE are about 8.16 dB (38th frame) and 0.95 dB (34th frame), respectively. For the "Bus" sequence, as shown in Fig. 8(b), the gains are about 7.94 dB (4th frame) and 1.09 dB (18th frame). This clearly indicates that the size of the GOP does not seriously impact the performance of the HMVE algorithm.

5. Conclusions

In this paper, we propose a new frame concealment algorithm, namely HMVE, to combat frame loss during video transmission over error-prone networks. As opposed to conventional methods, HMVE constructs a new MV set and discards the wrongly extrapolated MV. Then the MV of each pixel in the lost frame is estimated more accurately than with other conventional methods. Simulation results have demonstrated the effectiveness of our proposed HMVE algorithm to conceal the lost frame for different sequences. In conclusion, the HMVE algorithm proposed in this paper is highly effective and efficient. It is capable of significantly improving the video quality that has been corrupted by transmission errors.

Finally, we should point out that the application of the proposed method is not

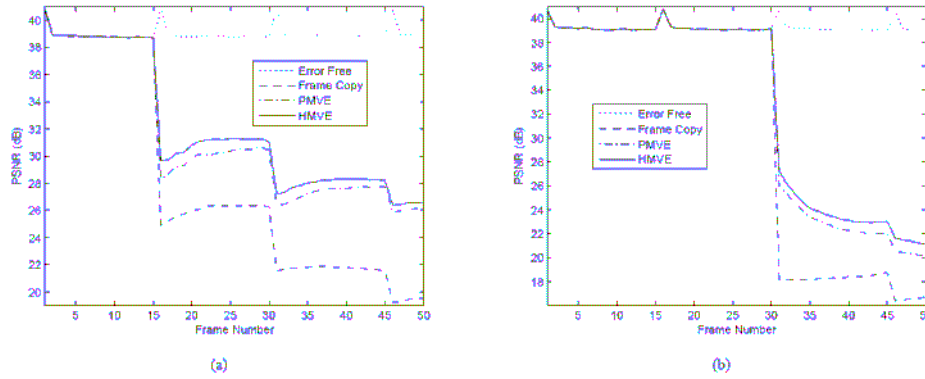


Fig. 7. PSNR comparison vs. frame number when QP is 22 for I-frame loss. (a) "Mobile" QCIF sequence with three I-frame losses (16th, 31st and 46th frames); (b) "Bus" QCIF sequence with two I-frame losses (31st and 46th frames).

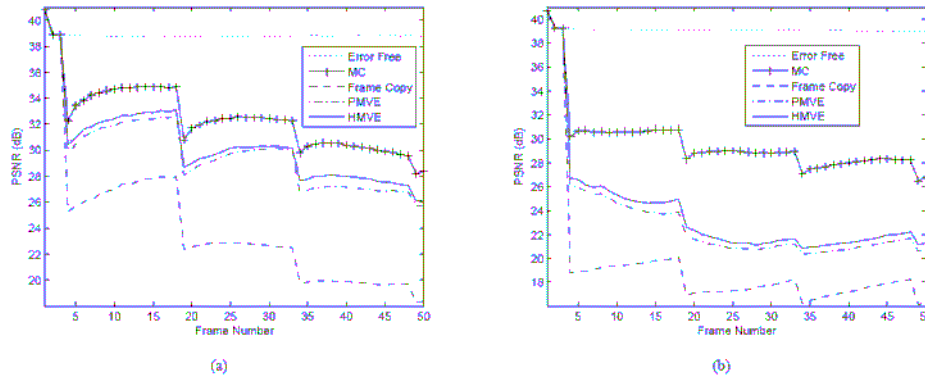


Fig. 8. PSNR comparison vs. frame number when QP is 22 without GOP format. (a) "Mobile" QCIF sequence; (b) "Bus" QCIF sequence.

limited to a single packet per frame transmission. For instance, if a frame is encapsulated into multiple packets, it is possible that none of them may reach the final destination. Under these conditions, the whole frame concealment approach (such as the one proposed in this paper) may still be the only affective approach. At the same time, in situations where not all the packets are dropped, applying a conventional method would be the best solution. Therefore, a combination of both techniques may prove very effective in further enhancing the overall concealment performance, especially for video transmission over wireless fading channels where a burst of packet-loss is expected more frequently.

6. Acknowledgement

The authors would thank the reviewers for their invaluable comments.

References

- [1] B. Yan and H. Gharavi, "Efficient error concealment for the whole-frame loss based on H.264/AVC," in *Proc. of IEEE ICIP'08*, Oct. 2008, pp. 3064–3067.
- [2] Y. Wang and Q.-F. Zhu, "Error control and concealment for video communication: a review," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 974–997, May 1998.
- [3] Z. Wang, Y. Yu, and D. Zhang, "Best neighborhood matching: an information loss restoration technique for block-based image coding systems," *IEEE Transactions on Image Processing*, vol. 7, no. 7, pp. 1056–1061, July 1998.
- [4] B. Yan and K. W. NG, "A novel selective motion vector matching algorithm for error concealment in MPEG-4 video transmission over error-prone channels," *IEEE Transactions on Consumer Electronics*, vol. 49, no. 4, pp. 1416–1423, Nov. 2003.
- [5] P. Salama, N. B. Shroff, E. J. Coyle, and E. J. Delp, "Error concealment techniques for encoded video streams," in *Proc. of IEEE International Conference on Image Processing (ICIP'95)*, 1995, pp. 9–12.
- [6] J. W. Park, J. W. Kim, and S. U. Lee, "DCT coefficients recovery based error concealment technique and its application to the MPEG-2 bit stream error," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, pp. 845–854, Dec. 1997.
- [7] H. Sun and W. Kwok, "Concealment of damaged block transform coded images using projections onto convex sets," *IEEE*

Transactions on Image Processing, vol. 4, pp. 470–477, Apr. 1995.

- [8] J. W. Park and S. U. Lee, "Recovery of corrupted image data based on the NURBS interpolation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, pp. 1003–1008, Oct. 1999.
- [9] T. P.-C. Chen and T. Chen, "Second-generation error concealment for video transport over error prone channels," *Wireless Communications and Mobile Computing, Special Issue on Multimedia over Mobile IP*, vol. 2, pp. 607–624, Oct. 2002.
- [10] H. Gharavi and S. Gao, "Spatial Interpolation Algorithm for Error Concealment," in *Proc. of IEEE ICASSP'08*, Apr. 2008, pp. 1153 – 1156.
- [11] S. Casner and V. Jacobson, "Compressing IP/UDP/RTP Headers for Low Speed Serial Links," *RFC2508*, Feb. 1999.
- [12] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "Draft ITU-T recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 - ISO/IEC 14496/10 AVC)," *IEEE Transactions on Image Processing*, May 2003.
- [13] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, July 2003.
- [14] T. Stockhammer, H. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 657–673, July 2003.
- [15] S. Wenger, "H.264/AVC over IP," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 645–656, July 2003.
- [16] Z. Wu and J. Boyce, "An error concealment scheme for entire frame losses based on H.264/AVC," in *Proc. of IEEE ICASSP'06*, May 2006.
- [17] K. Song, T. Chun, Y. Kim, Y. Oh, and C. Kim, "Error concealment of H.264/AVC video frames for mobile video broadcasting," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 704–711, May 2007.
- [18] X. Chen, Y. Chung, and C. Bae, "A new multi-mode intra-frame error concealment algorithm for H.264/AVC," in *Proc. of IEEE ICME'07*, July 2007, pp. 236–239.
- [19] S. C. Sui, F. Liu, and P. Xue, "Whole-frame error concealment with improved backward motion estimation for H.264 decoders," in *Proc. of Fourth International Conference on Image and Graphics (ICIG'07)*, Aug. 2007, pp. 279–284.
- [20] Y.-C. Lee, Y. Altunbasak, and R. Mersereau, "Multiframe error concealment for MPEG-coded video delivery over error-prone networks," *IEEE Transactions on Image Processing*, vol. 11, no. 11, pp. 1314–1331, Nov. 2002.
- [21] P. Nasiopoulos, L. Coria-Mendozal, H. Mansour, and A. Golikeri, "An improved error concealment algorithm for intra-frames in H.264/AVC," in *Proc. of IEEE ISCAS'05*, vol. 1, May 2005, pp. 320–323.
- [22] Y.-K. Wang, M. M. Hannuksela, V. Varsa, A. Hourunranta, and M. Gabbouj, "The error concealment feature in the H.26L test model," in *Proc. of IEEE International Conference on Image Processing (ICIP'02)*, vol. 2, 2002, pp. II-729–II-732.
- [23] P. Baccichet, D. Bagni, A. Chimienti, L. Pezzoni, and F. Rovati, "Frame concealment for H.264/AVC decoders," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 1, pp. 227–233, Feb. 2005.
- [24] P. Baccichet and A. Chimienti, "A low complexity concealment algorithm for the whole-frame loss in H.264/AVC," in *Proc. of IEEE 6th Workshop on Multimedia Signal Processing*, Sept. 2004, pp. 279 – 282.
- [25] S. Belfiore, M. Granello, E. Magli, and G. Olmo, "An error concealment algorithm for streaming video," in *Proc. of IEEE ICIP'03*, vol. 3, Sept. 2003, pp. 649–652.
- [26] —, "Concealment of whole-frame losses for wireless low bit-rate video based on multiframe optical flow estimation," *IEEE Transactions on Multimedia*, vol. 7, no. 2, pp. 316 – 329, Apr. 2005.
- [27] Z. Zhou and S. Xie, "Error concealment based on robust optical flow," in *Proc. of International Conference on Communications, Circuits and Systems*, vol. 1, May 2005, pp. 547 – 550.
- [28] C. Huang, K. Yang, and J. Wang, "Error resilience supporting bidirectional frame recovery for video streaming," in *Proc. of IEEE ICIP'04*, vol. 1, Oct. 2004, pp. 537–540.
- [29] Q. Peng, T. Yang, and C. Zhu, "Block-based temporal error concealment for video packet using motion vector extrapolation," in *Proc. of IEEE International Conference on Communications, Circuits and Systems and West Sino Expositions*, vol. 1, June 2002, pp. 10–14.
- [30] Y. Chen, K. Yu, J. Li, and S. Li, "An error concealment algorithm for entire frame loss in video transmission," in *Proc. of IEEE Picture Coding Symposium (PCS'04)*, 2004.
- [31] S. Wenger, "Error patterns for Internet experiments," *ITU Telecommunications Standardization Sector, Red Bank, NJ, Doc. Q15-1-16r1*, Oct. 1999.



Bo YAN received his Ph.D. degree in Computer Science and Engineering from the Chinese University of Hong Kong (CUHK) in 2004. Before that, he received his degrees of B.E. and M.E. in Communication Engineering both from Xi'an Jiaotong University (XJTU) in 1998 and 2001 respectively. From 2004 to 2006, he worked in the National Institute of Standards and Technology US (NIST) as a Postdoctoral Guest Researcher.

Since Sept. 2006, he has been the Associate Professor in School of Computer Science at Fudan University, Shanghai, China. His research interests include video processing and multimedia communications.



Hamid Gharavi (Fellow, IEEE) received the Ph.D. degree from Loughborough University, Loughborough, U.K., in 1980. He joined AT&T Bell Laboratories, Holmdel, in 1982. He was then transferred to Bell Communications Research (Bellcore) after the AT&T-Bell divestiture, where he became a Consultant on video technology and a Distinguished Member of Research Staff. In 1993, he joined Loughborough University as Professor and Chair of Communication Engineering. Since September 1998, he has been with the National Institute of Standards and Technology (NIST), US Department of Commerce, Gaithersburg, MD.

Dr Gharavi was a core member of the Study Group XV (Specialist Group on Coding for Visual Telephony) of the International Communications Standardization Body CCITT (ITU-T). He was selected as one of the six university academics to be appointed to the U.K. Government's Technology Foresight Panel in Communications to consider the future through 2015 and make recommendations for allocation of key research funds. His research interests include video/image transmission, wireless multimedia, mobile communications and third generation wireless systems, and mobile ad-hoc networks. He holds eight U.S. patents related to these topics.

Dr Gharavi received the Charles Babbage Premium Award from the Institute of Electronics and Radio Engineering in 1986, and the IEEE CAS Society Darlington Best Paper Award in 1989. He has been a Distinguished Lecturer of the IEEE Communication Society. In 1992 Dr Gharavi was elected a Fellow of IEEE for his contributions to low bit-rate video coding and research in subband coding for image and video applications. He has been a Guest Editor for a number of special issues. Dr Gharavi served as a member of the Editorial Board of the PROCEEDINGS OF THE IEEE from January, 2003 to December, 2008. He is currently a member of the Editorial board, IET Image Processing. Dr Gharavi served as an Associate Editor for the IEEE Transactions on CAS for Video Technology (CSVT) from 1996 to 2006. He then became the Deputy Editor-in-Chief of this IEEE Transactions through December 31, 2009. Dr Gharavi was recently appointed to serve as the new Editor-in-Chief for the IEEE Transactions on CSVT.