Frame Error Concealment Using Pixel Correlation in Overlapped Motion Compensation Regions

Dinh Trieu Duong, Byeong-Doo Choi, Min-Cheol Hwang, and Sung-Jea Ko

In low bit-rate video transmission, the payload of a single packet can often contain a whole coded frame due to the high compression ratio in both spatial and temporal domains of most modern video coders. Thus, the loss of a single packet not only causes the loss of a whole frame, but also produces error propagation into subsequent frames. In this paper, we propose a novel whole frame error concealment algorithm which reconstructs the first of the subsequent frames instead of the current lost frame to suppress the effects of error propagation. In the proposed algorithm, we impose a constraint which uses side match distortion (SMD) and overlapped region difference (ORD) to estimate motion vectors between the target reconstructed frame and its reference frame. SMD measures the spatial smoothness connection between a block and its neighboring blocks. ORD is defined as the difference between the correlated pixels which are predicted from one reference pixel. Experimental results show that the proposed algorithm effectively suppresses error propagation and significantly outperforms other conventional techniques in terms of both peak signal-to-noise ratio performance and subjective visual quality.

Keywords: Frame error concealment, motion estimation, whole frame loss, video restoration.

Manuscript received June 9, 2008; revised Nov. 5, 2008; accepted Dec. 11, 2008.

This research was supported by Seoul Future Contents Convergence (SFCC) Cluster established by Seoul R&BD Program.

Dinh Trieu Duong (phone: + 82 2 3290 3228, email: duongdt@dali.korea.ac.kr), Min-Cheol Hwang (email: mchwang@dali.korea.ac.kr), and Sung-Jea Ko (email: sjko@korea.ac.kr) are with the Department of Electronic Engineering, Korea University, Seoul, Rep. of Korea.

Byeong-Doo Choi (email: byeong-doo.choi@hhi.fraunhofer.de) is with Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut (HHI), Berlin, Germany.

I. Introduction

For robust video transmission over low bit-rate networks, it is necessary to cope with several problems including transmission error, channel bandwidth variation, and fading effects [1], [2]. Transmission error is the most critical problem because it may cause the loss of synchronization between the encoder and decoder for correct decoding. Compressed video streams are extremely vulnerable to transmission error due to the use of temporal predictive coding and variable length coding [3]. Thus, a single error bit in a video frame can lead to error propagation in subsequent frames. As shown in Fig. 1, when the current frame is lost, the error is temporally propagated in the next and following frames.

Without modifying the source and channel coding schemes, error concealment (EC) at the decoder is an effective approach to cope with transmission error which has been widely investigated [4]. Some EC techniques exploit spatial correlation to conceal a missing block from neighboring pixels in the same frame [5]-[7]. Other EC techniques utilize the temporal correlation of video frames to replace missing blocks by the corresponding blocks in the reference frame [8]-[10]. Most existing EC algorithms assume that either a set of blocks in a video frame or a slice consisting of several consecutive blocks is lost. However, in low bit-rate video transmission, this assumption has been somewhat limited so far [11]. The reason is that, at a low bit-rate, each coded video frame is fairly small and is usually encapsulated in a channel packet for transmission through the network. The loss of the channel packet most likely leads to the loss of a whole video frame. As a consequence, the existing EC algorithms are not applicable.

There have been several algorithms proposed for whole frame loss EC. In conventional whole frame EC methods, the

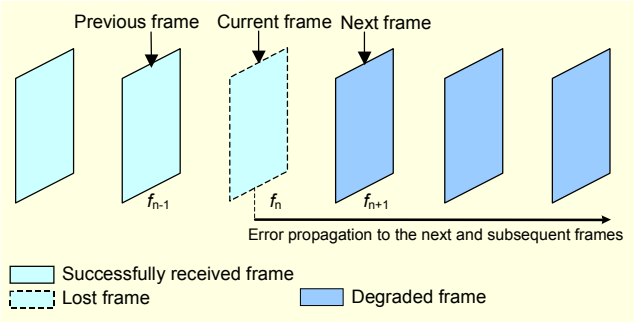

Fig. 1. Effects of error propagation on a video sequence when a frame is lost.

next frame, that is, the first of the subsequent frames, is reconstructed after concealing the current lost frame. Bandyopadhyay and others [12] introduced frame copy (FC) and motion vector copy (MVC) methods for concealing the current lost frame. In the FC method, the lost frame is simply replaced by the previous frame without consideration of the motion between the previous and current frames. In the MVC method, the motion vectors (MVs) and reference indices of the co-located blocks in the previously decoded reference frame are copied to the current lost frame, and then the lost frame is reconstructed using these copied MVs. A more efficient algorithm proposed by Baccichet and others [13] exploits the hypothesis of optical flow estimation to improve the accuracy of estimated motion for the missing frame. This algorithm provides good EC performance in terms of peak signal-to-noise ratio (PSNR) and visual quality while requiring low computational complexity.

In the previously mentioned algorithms, after concealing the current lost frame, the next frame is reconstructed by using the current concealed frame as its reference frame. However, the image quality of the next frame in those algorithms can be severely degraded since the concealed frame is not faithful enough to be used as the reference frame. In this case, errors can be strongly propagated from the next frame to the subsequent frames. To deal with this problem, we propose a novel whole frame EC algorithm that directly performs EC on the next frame instead of the current lost frame. In the proposed algorithm, we impose a constraint that uses side match distortion (SMD) and overlapped region difference (ORD) to re-estimate MVs for the next frame. SMD measures the spatial smoothness connection between a block and its neighboring blocks. ORD is defined as the difference between the correlated pixels which are predicted from one reference pixel. Experimental results show that the proposed algorithm effectively suppresses error propagation and significantly outperforms other conventional techniques in terms of PSNR performance and subjective visual quality.

The rest of the paper is organized as follows. Section II describes the proposed algorithm in detail. Experimental results are discussed in section III. Finally, section IV concludes this paper.

II. Proposed Frame Error Concealment Algorithm

The proposed algorithm consists of two processing steps: backward projection and motion estimation (ME) using ORD and SMD. In the first step, backward projection is performed between the next and current frames. In the second step, the MV of each block in the next frame is re-estimated based on the smoothness measured across the boundaries of neighboring blocks and the difference between overlapped pixels in the overlapped regions which are obtained in the first step. Next, we describe each processing step in detail.

1. Backward Projection

Let f_{n-1}, f_n, and f_{n+1} denote the previous, current, and next frames, respectively, as shown in Fig. 1. Assume that f_{n-1} is the last correctly decoded frame, and f_n is the lost frame caused by the transmission error. At the decoder, MVs of the previous and next frames are correctly received, while those of the current frame are lost. In conventional methods such as FC and MVC [12], f_n is first reconstructed; then, f_{n+1} is reconstructed using the reconstructed frame of f_n as its reference frame. However, in our proposed method, f_n is reconstructed using the conventional methods, and f_{n+1} is directly reconstructed by using f_{n-1} as its reference frame instead of f_n.

Let B_{n+1}^k represent the k-th block in f_{n+1} and $\mathbf{v}_{n+1}^k$ denote the 2-D vector which represents the MV of B_{n+1}^k. Note that $\mathbf{v}_{n+1}^k$ is estimated between f_n and f_{n+1} at the encoder and is correctly received for B_{n+1}^k at the decoder. Then, in the backward projection between f_{n+1} and f_n, every block in f_{n+1} is mapped onto f_n using their received MVs as shown in Fig. 2. In Fig. 2, four blocks in f_{n+1} are mapped onto f_n using their MVs, such as $\mathbf{v}_{n+1}^k$, $\mathbf{v}_{n+1}^p$, $\mathbf{v}_{n+1}^q$, and $\mathbf{v}_{n+1}^h$. After performing the backward projection, each block in f_{n+1} has its own mapped block in f_n and these mapped blocks can overlap one another. Figure 2 shows an example of overlapped regions between $\hat{B}_n^k$ and two blocks $\hat{B}_n^p$ and $\hat{B}_n^q$ in f_n. Here, $\hat{B}_n^k$, $\hat{B}_n^p$, and $\hat{B}_n^q$ are the mapped blocks of B_{n+1}^k, B_{n+1}^p, and B_{n+1}^q, respectively, in the backward projection. It is worth noting that $\hat{B}_n^k$, $\hat{B}_n^p$, and $\hat{B}_n^q$ are also the motion compensated blocks from which the inter-coded blocks B_{n+1}^k, B_{n+1}^p, and B_{n+1}^q are copied in the forward prediction between f_n and f_{n+1}. For each overlapped motion compensation region in $\hat{B}_n^k$, two regions named sibling regions are obtained in f_{n+1}. These sibling regions contain sibling pixels which are predicted from the same reference pixels in the overlapped region. In Fig. 2,

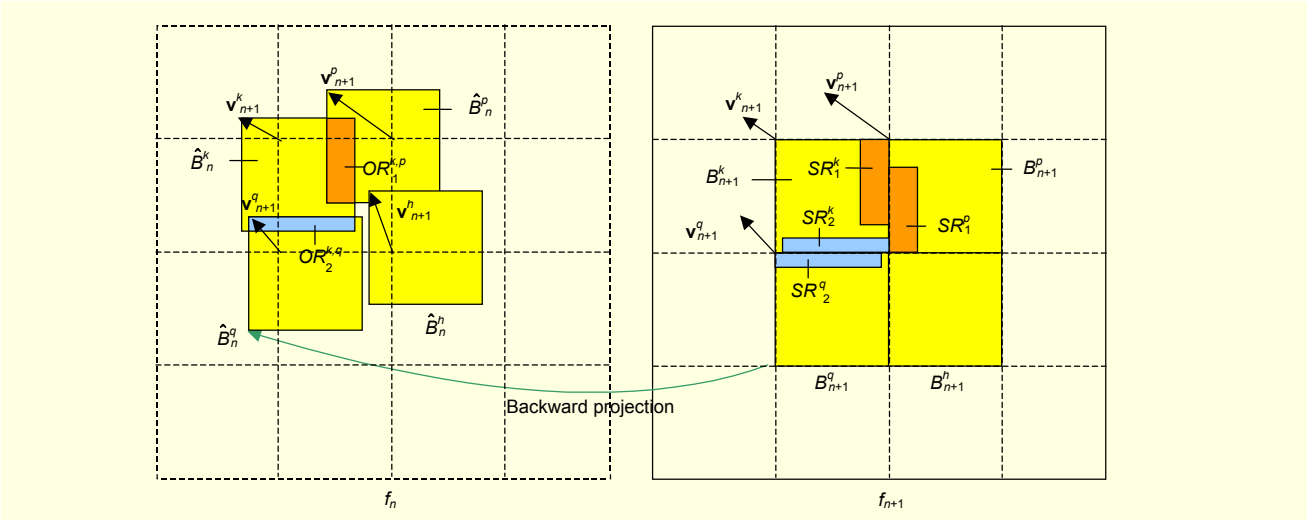

Fig. 2. Backward projection between the next and current frames.

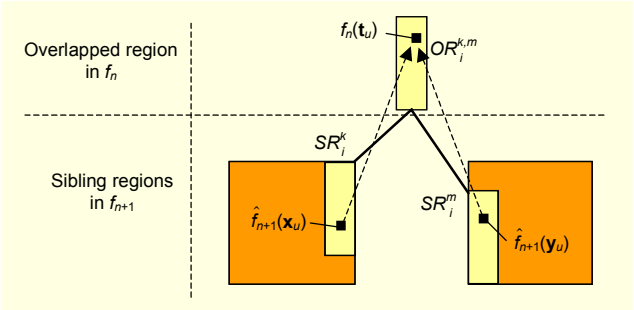

Fig. 3. Tree-structured diagram of an overlapped region and its sibling regions.

$OR_1^{k,p}$ and $OR_2^{k,q}$ are the overlapped motion compensation regions between $\hat{B}_n^k$ and $\hat{B}_n^p$ and between $\hat{B}_n^k$ and $\hat{B}_n^q$, respectively. In f_{n+1}, SR_1^k and SR_1^p are the two sibling regions of $OR_1^{k,p}$. In the backward projection between f_{n+1} and f_n, SR_1^k is overlapped with SR_1^p at the position of $OR_1^{k,p}$ because these sibling regions contain the sibling pixels which are predicted from the same reference pixels in $OR_1^{k,p}$. Similarly, SR_2^k and SR_2^q are the two sibling regions of $OR_2^{k,q}$. Figure 3 shows a tree-structured diagram that explains more details of the relation between an overlapped region and its corresponding sibling regions. In Fig. 3, $OR_i^{k,m}$ is the i-th overlapped region in $\hat{B}_n^k$ of the two sibling regions SR_i^k and SR_i^m. The three regions SR_i^k, SR_i^m, and $OR_i^{k,m}$ are the same size. In the proposed algorithm, for each overlapped region $OR_i^{k,m}$ in $\hat{B}_n^k$, we utilize ORD which measures the absolute differences between sibling pixels in the corresponding sibling regions SR_i^k and SR_i^m as follows:

$$\mathrm{ORD}_i = \frac{1}{L}\sum_{u=1}^{L}\left|\hat{f}_{n+1}(\mathbf{x}_u) - \hat{f}_{n+1}(\mathbf{y}_u)\right|, \quad (1)$$

where $\mathbf{x}_u$ and $\mathbf{y}_u$ represent the positions of the u-th sibling pixels in SR_i^k and SR_i^m, respectively; $\hat{f}_{n+1}(\mathbf{x}_u)$ represents the value of sibling pixel at the position $\mathbf{x}_u$; and L is the number of sibling pixels in SR_i^k or SR_i^m. As shown in Fig. 3, $\hat{f}_{n+1}(\mathbf{x}_u)$ and $\hat{f}_{n+1}(\mathbf{y}_u)$ are predicted from the same reference pixel $f_n(\mathbf{t}_u)$ in $OR_i^{k,m}$ and overlapped at the position of $f_n(\mathbf{t}_u)$ in the backward projection.

Since several overlapped regions can be yielded in $\hat{B}_n^k$, the ORD measured for all of these overlapped regions is given by

$$\mathrm{ORD}(\hat{B}_n^k) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{ORD}_i, \quad (2)$$

where N is the number of overlapped regions in $\hat{B}_n^k$. When there is no overlapped region yielded in $\hat{B}_n^k$, the value of $\mathrm{ORD}(\hat{B}_n^k)$ is assigned to zero. It should be noted that the value of $\mathrm{ORD}(\hat{B}_n^k)$ obtained in the backward projection between f_{n+1} and f_n is always equal to zero since the sibling pixels in (1) are predicted from the same reference pixel in f_n. In the proposed algorithm, this property of ORD has been employed as a constraint to reconstruct f_{n+1}. In the next subsection, we describe in detail the constraint using ORD and the EC algorithm for f_{n+1}.

2. Motion Estimation Using ORD and SMD

At the decoder, although the received MVs of f_{n+1} are utilized to obtain the ORD measurement for each mapped block in the backward projection as described in the previous subsection, these MVs cannot be employed to reconstruct f_{n+1} from f_n because f_n is lost. To deal with this problem, we estimate new MVs for all blocks in f_{n+1} so that f_{n+1} can be directly reconstructed by using the previous frame f_{n-1} as its reference frame instead of f_n.

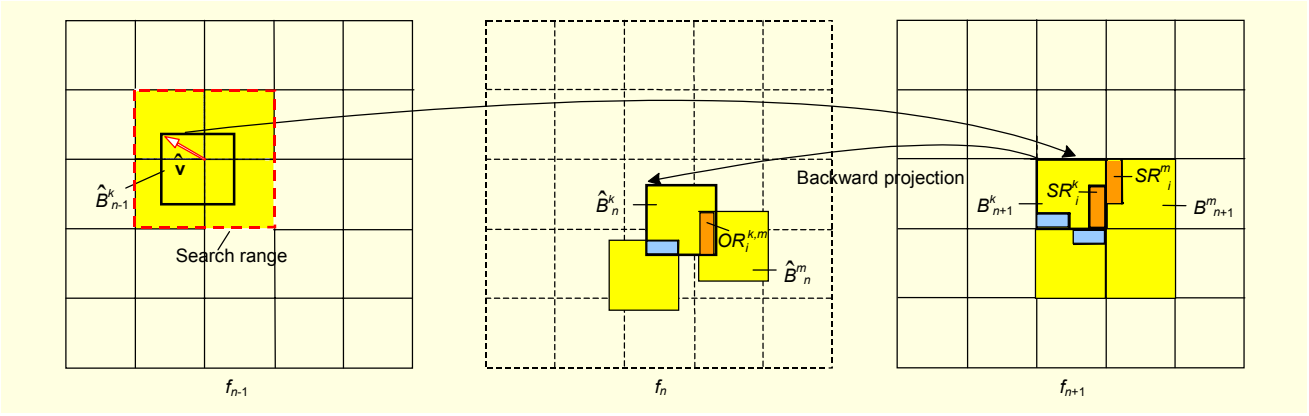

Fig. 4. Motion estimation for the next frame using the previous frame as the reference frame.

Figure 4 shows the ME scheme performed on the k-th block B^k_{n+1} in f_{n+1}. As shown in Fig. 4, $\hat{B}^k_{n-1}$ in f_{n-1} is a predicted block of B^k_{n+1} and the displacement vector $\hat{\mathbf{v}}$ between $\hat{B}^k_{n-1}$ and B^k_{n+1} is one of the candidate MVs estimated for B^k_{n+1}. The goal of the ME scheme is to find an optimal MV among various candidate MVs estimated for B^k_{n+1}.

As described in the previous subsection, after performing the backward projection between f_{n+1} and f_n, each block in f_{n+1} has its own mapped block in f_n, and these mapped blocks can overlap one another. Without loss of generality, we assume that the mapped block $\hat{B}^k_n$ of B^k_{n+1} is overlapped with the mapped block $\hat{B}^m_n$ at the i-th overlapped region $OR^{k,m}_i$ in $\hat{B}^k_n$ as shown in Fig. 4. Two sibling regions of $OR^{k,m}_i$, SR^k_i, and SR^m_i, are located in B^k_{n+1} and B^m_{n+1}, respectively.

For each $\hat{\mathbf{v}}$ in a given search range $\Re$ as shown in Fig. 4, the value of the u-th sibling pixel in the sibling region SR^k_i is predicted from f_{n-1} by

$$\hat{f}_{n+1}(\mathbf{x}_u) = f_{n-1}(\mathbf{x}_u + \hat{\mathbf{v}}), \tag{3}$$

where $f_{n-1}(\mathbf{x}_u)$ denotes the pixel at the position $\mathbf{x}_u$ in f_{n-1}.

Because blocks are decoded in raster scan order in f_{n+1}, the value of sibling pixel $\hat{f}_{n+1}(\mathbf{y}_u)$ in the sibling region SR^m_i is not available if the ME on B^k_{n+1} is performed before it is performed on B^m_{n+1}. In this case, the value of $\hat{f}_{n+1}(\mathbf{y}_u)$ is assigned by

$$\hat{f}_{n+1}(\mathbf{y}_u) = f_{n-1}(\mathbf{y}_u + 2\mathbf{v}^m_{n+1}), \tag{4}$$

where $\mathbf{v}^m_{n+1}$ is the MV which is estimated between f_n and f_{n+1} at the encoder and is correctly received for B^m_{n+1} at the decoder.

Similarly, for the other overlapped regions in $\hat{B}^k_n$, we can obtain the values of sibling pixels in their corresponding sibling regions. Then, in the backward projection between f_{n+1} and f_n, the value of ORD measured for all overlapped regions in $\hat{B}^k_n$ is achieved by modifying (1) and (2) as follows:

$$\mathrm{ORD}_i(\hat{\mathbf{v}}) = \frac{1}{L}\sum_{u=1}^{L}\left|\hat{f}_{n+1}(\mathbf{x}_u) - \hat{f}_{n+1}(\mathbf{y}_u)\right|, \tag{5}$$

and

$$\mathrm{ORD}(\hat{B}^k_n, \hat{\mathbf{v}}) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{ORD}_i(\hat{\mathbf{v}}). \tag{6}$$

For each $\hat{\mathbf{v}}$ in the given search range, we can obtain the corresponding value of $\mathrm{ORD}(\hat{B}^k_n, \hat{\mathbf{v}})$ using (6). However, as mentioned in the previous subsection, the true value of $\mathrm{ORD}(\hat{B}^k_n, \hat{\mathbf{v}})$ obtained for $\hat{B}^k_n$ should be zero or the minimum value among various values of $\mathrm{ORD}(\hat{B}^k_n, \hat{\mathbf{v}})$. Therefore, among different candidate MVs, the MV that minimizes $\mathrm{ORD}(\hat{B}^k_n, \hat{\mathbf{v}})$ can be chosen as the best one estimated for B^k_{n+1} as

$$\mathbf{v} = \arg\min_{\hat{\mathbf{v}}\in\Re}\mathrm{ORD}(\hat{B}^k_n, \hat{\mathbf{v}}). \tag{7}$$

In (7), ORD measurement has been utilized as a constraint to find out the MV estimated for B^k_{n+1}. To further improve the accuracy of $\mathbf{v}$, we employ an additional spatial smoothness constraint using SMD which measures the smoothness connection between B^k_{n+1} and its neighboring blocks. The spatial smoothness constraint using SMD has been successfully applied in a number of EC algorithms [9], [14]. In the proposed algorithm, this constraint is imposed on each $\hat{\mathbf{v}}$ in $\Re$ to obtain an optimal MV estimated for B^k_{n+1}.

Figure 5 shows the boundaries of B^k_{n+1} where SDM is to be imposed. In Fig. 5, B^t_{n+1}, B^b_{n+1}, B^l_{n+1}, and B^r_{n+1} are the neighboring blocks of B^k_{n+1} to the top, bottom, left, and right, respectively. For each $\hat{\mathbf{v}}$, SMD computes the absolute pixel differences between the predicted pixels in B^k_{n+1} and its neighboring blocks along the top, left, bottom, and right boundaries by

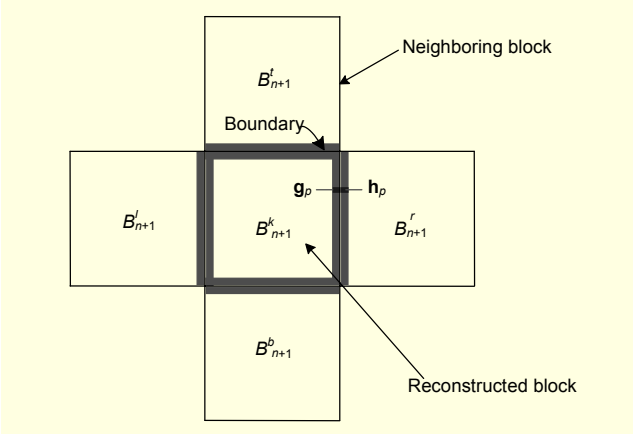

Fig. 5. SMD measurement along the boundary between a block and its neighboring blocks.

$$\mathrm{SMD}(B^k_{n+1}, \hat{\mathbf{v}}) = \frac{1}{P}\sum_{p=1}^{P}\left|\hat{f}_{n+1}(\mathbf{g}_p, \hat{\mathbf{v}}) - \hat{f}_{n+1}(\mathbf{h}_p)\right|, \qquad (8)$$

where $\mathbf{g}_p$ and $\mathbf{h}_p$ represent the positions of the p-th pixel at the boundaries of block B^k_{n+1} and its neighboring blocks, respectively, and P is the number of boundary pixels. Note that the pixel values $\hat{f}_{n+1}(\mathbf{h}_p)$ in the neighboring blocks should be available when the SMD is evaluated. Therefore, the proposed ME scheme is applied iteratively as discussed below.

By minimizing the cost function of $\mathrm{ORD}(\hat{B}^k_n, \hat{\mathbf{v}})$ and $\mathrm{SMD}(B^k_{n+1}, \hat{\mathbf{v}})$, we can obtain the optimal MV estimated for B^k_{n+1} as

$$\mathbf{v}_{\mathrm{opt}} = \arg\min_{\hat{\mathbf{v}}\in\Re}(\mathrm{ORD}(\hat{B}^k_n, \hat{\mathbf{v}}) + \alpha\,\mathrm{SMD}(B^k_{n+1}, \hat{\mathbf{v}})), \qquad (9)$$

where α is a weighting factor of the cost function of the ORD and SMD, $\alpha > 0$.

In (9), the computation of the SMD and ORD requires the pixel values in the neighboring blocks of B^k_{n+1} and in the sibling region SR^m_i in B^m_{n+1}. In addition, the ME of B^k_{n+1} is affected by that of the neighboring blocks and B^m_{n+1}. Therefore, the minimization of the ORD and SMD in (9) is iteratively applied to obtain a reliable MV estimated for B^k_{n+1} as follows:

- MV scaling (MVS): In this stage, the received MVs of f_{n+1} are multiplied by 2. Then, based on these scaled MVs, the pixel values of f_{n+1} are initially assigned using (4).
- In the first ME iteration: Since blocks are decoded in raster scan order in f_{n+1}, the pixel values in the sibling region SR^m_i are not available if the ME on B^k_{n+1} is performed before that on B^m_{n+1}. In this case, the initial pixel values which are obtained in the MVS for B^m_{n+1} are assigned to the pixels in SR^m_i to compute the ORD. Also, the pixel values in the bottom and right blocks of B^k_{n+1} are not available in the first ME. Then, the computation of SMD in (8) for four block boundaries is reduced to that of two boundaries which include only the top and left boundary pixels.
- In the t-th ME iteration $(t \geq 2)$: Since the MVs of B^m_{n+1}, B^r_{n+1}, and B^b_{n+1} are certainly determined in the (t-1)th ME, the reconstructed pixel values of these blocks in the (t-1)th reconstruction are employed to compute the ORD and SMD in the t-th ME.

The minimization of the ORD and SMD in (9) is iteratively applied until a satisfactory result is obtained. Experiment results confirm that the proposed algorithm provides good performance when α is about 0.75, and the performance is not very sensitive to the variation of α. Therefore, in this work, α is fixed to 0.75.

III. Experiment Results

The proposed algorithm was tested on several video sequences including the Foreman, Table Tennis, Mobile, and Stefan sequences. These test sequences are in 4:2:0 format with CIF (352×288) resolution. JM 12.2 which is H.264/AVC reference software [15] was utilized to implement the proposed algorithm. For all experiments, ME was carried out at a quarter-pixel resolution with the search range of ± 16 for both horizontal and vertical directions. We also configured the test video sequences to use the group of pictures (GOP) structure to effectively stop error propagation; specifically, GOPs of 30 frames were formed. The number of reference frames was set to 1 for default settings. This means that only the previous frame is employed for temporal prediction. This setting is necessary when the decoder memory is limited and low encoding complexity is required. The slice mode was not used, so one complete frame could be encapsulated into each network packet. When only one MV per macroblock (MB) is used, the size of a block unit utilized in the proposed algorithm is equal to 16×16. However, in the H.264/AVC standard, since a 16×16 MB can be partitioned into smaller sub-blocks with their own MVs and sizes such as 16×8, 4×8, or 4×4 [16], each MB can be associated with more than one MV. In this case, the size of the block unit is reduced to 4×4 so that MVs of all sub-blocks in an MB can be estimated in the proposed algorithm. In our experiments, quantization parameters (QPs) were set to 25, 30, and 35 for each test sequence. The packets were randomly dropped according to the predefined packet loss rate (PLR). The values of PLR were simulated from 5% to 10% since these are typical values in wireless environments [17].

First, we evaluated the impact of iteration numbers on the PSNR performance of the next frames concealed by the proposed algorithm. The PSNR performance of the next frames obtained in the MVS corresponding to the number of

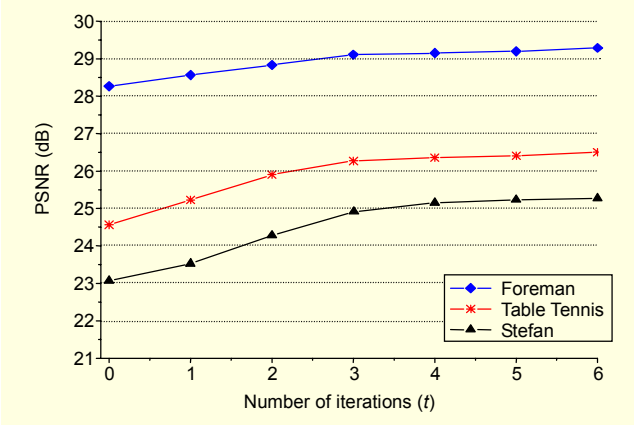

Fig. 6. Impact of the number of iterations on the performance of the proposed algorithm.

iterations t=0 was also evaluated. As described in the previous section, in the MVS, based on the scaled MVs, the next frame is initially reconstructed from the previous frame without using the ORD and SMD constraints. Figure 6 shows the average PSNR of the concealed frames according to different numbers of iterations. For various test sequences, when the number of iterations is equal or higher than 3, the absolute difference between the PSNR performance at the (t+1)th iteration and that at the t-th iteration is very small and generally smaller than 0.1 dB ($\left|PSNR_{t+1} - PSNR_t\right| < \Delta PSNR_{\text{th}} = 0.1$). This means that given a predefined threshold $\Delta PSNR_{\text{th}} = 0.1$ dB, the proposed algorithm is converged when $t \geq 3$. In this work, to reduce the computational complexity, we chose t=3 as the number of iterations after which to stop iterative loops to reconstruct the next frame. The experimental results obtained for the proposed algorithm at the third iteration are utilized in this paper to compare the performance of the proposed algorithm with that of other EC methods.

Next, to evaluate the effectiveness of the proposed algorithm, we compare the performance of the proposed algorithm with that of the MVS and two conventional methods, namely, the FC method supported by JM [15] and the concealment algorithm on blocks (CA_B) method [13]. In the FC method, MVs of all MBs in the lost frame are set to zero. Although the CA_B method uses the optical flow concept for the EC algorithm on pixels (CA_P) in [18], it works at the block level instead of the pixel level to meet the requirement of real-time applications. The current lost frame in the FC and CA_B methods is first concealed, and then the next frame is reconstructed by using the current concealed frame as its reference frame. However, in our proposed algorithm, the current lost frame is concealed using the CA_B method, but the next frame is directly reconstructed by using the previous frame as its reference frame.

For objective evaluation, we compare the PSNR performance of the frames which are concealed by the FC, CA_B, and the proposed algorithm. Tables 1 and 2 summarize the average PSNR with various PLRs and QPs of the conventional and proposed algorithms. When calculating the average PSNR, we consider all frames of the test sequences, including the concealed frames and their subsequent frames deteriorated by error propagation. As shown in Tables 1 and 2, the proposed algorithm consistently provides better performance than the FC and CA_B methods. For example, as shown in Table 1, the proposed algorithm provides up to 4.1 dB and 3.0 dB gains as compared with the FC and CA_B methods, respectively, for the Stefan sequence when QP=25 and PLR= 5%. The effectiveness of the proposed algorithm can be observed by comparing the improvement in the PSNR performance of the proposed algorithm with that of the MVS.

Table 1. Comparison of the average PSNR performance with 5% PLR on the Table Tennis and Stefan sequences. (dB)

Sequence	QP	FC	CA_B	MVS	Proposed
Table Tennis	25	29.02	29.17	31.11	32.08
	30	27.21	27.43	28.75	29.86
	35	24.30	25.14	26.68	27.73
Stefan	25	27.26	28.37	30.25	31.42
	30	26.01	27.03	28.62	29.90
	35	23.25	25.32	26.87	27.80

Table 2. Comparison of the average PSNR performance with 10% PLR on the Foreman and Mobile sequences. (dB)

Sequence	QP	FC	CA_B	MVS	Proposed
Foreman	25	31.73	31.51	33.05	33.93
	30	29.57	29.21	30.32	31.15
	35	26.63	26.82	28.43	29.27
Mobile	25	19.14	22.49	24.31	25.26
	30	18.22	20.84	22.37	23.19
	35	17.40	19.72	21.80	22.51

As reported in Table 1, the proposed algorithm provides effectively higher gains as compared with MVS. The gain is smaller for the Foreman sequence. However, as shown in Table 2, the smallest gain obtained by the proposed algorithm is 0.8 dB compared with that of MVS and 1.9 dB compared with that of the CA_B method.

Figure 7 shows more details of the PSNR performance at each frame in the Table Tennis and Stefan sequences. In Fig. 7, the Table Tennis sequence is corrupted at the 13th, 43rd, and 66th frames, and the Stefan sequence is corrupted at the 12th, 39th, and 68th frames. The PSNR performance at the next frame

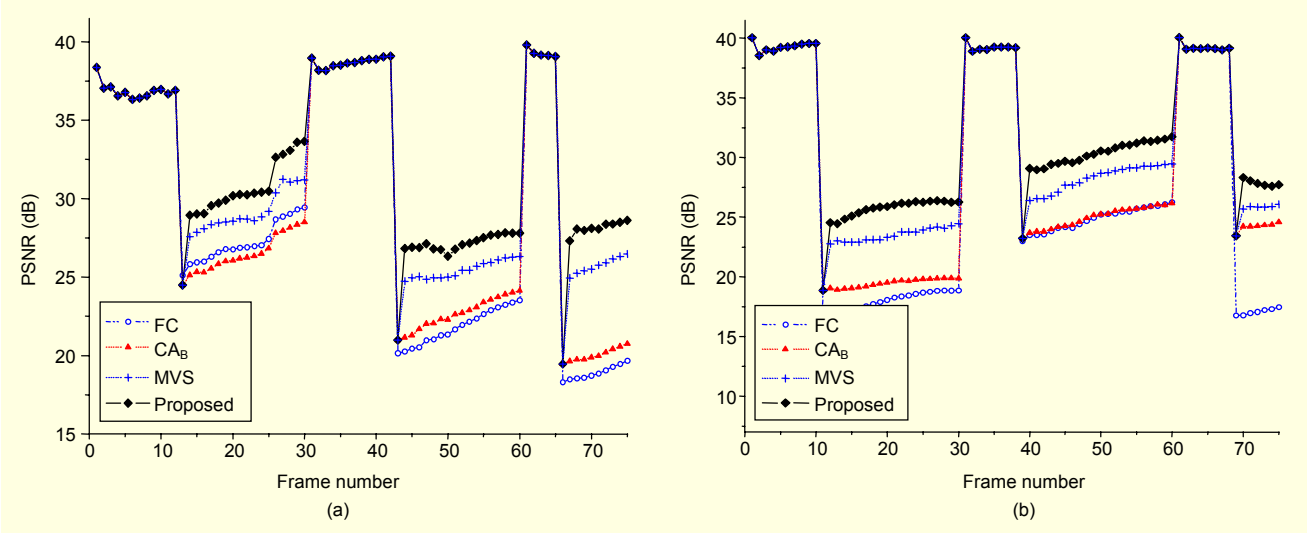

Fig. 7. Comparison of the PSNR performance with QP 25 and 5% PLR on (a) Table Tennis and (b) Stefan sequences.

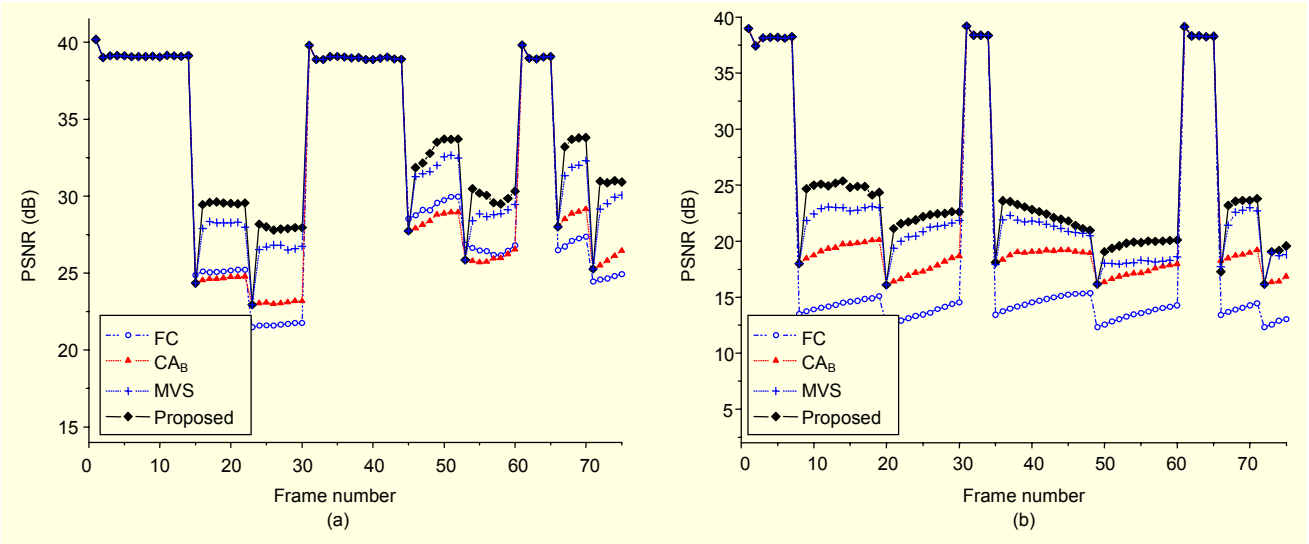

Fig. 8. Comparison of the PSNR performance with QP 25 and 10% PLR on (a) Foreman and (b) Mobile sequences.

represents the error concealment ability of different EC methods, while the PSNRs of the following frames in the same GOP show the effects of the error propagation. For all concealed and error propagated frames, the proposed algorithm provides up to 2 dB to 4 dB better PSNR performance as compared with the conventional methods. Furthermore, as shown in Fig. 7, since the next frame is effectively concealed, the effects of error propagation on the subsequent frames are effectively suppressed in the proposed algorithm as compared with FC, CA_B, and MVS, allowing high PSNR performance to be maintained.

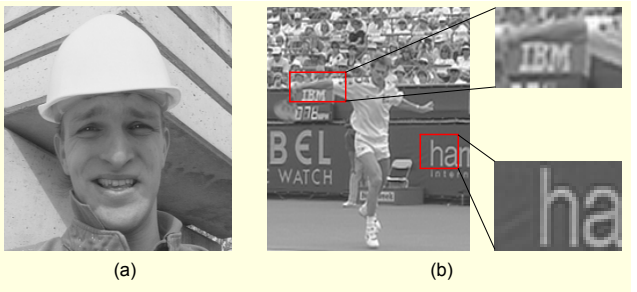

Fig. 9. Original images: (a) Foreman and (b) Stefan.

Similar results are obtained for the Foreman and Mobile sequences, where the proposed algorithm outperforms conventional methods by several decibels in PSNR. This can be observed in the next frame and the subsequent frames of the Foreman and Mobile sequences as shown in Fig. 8. For example, in the 9th reconstructed frame of the Mobile sequence,

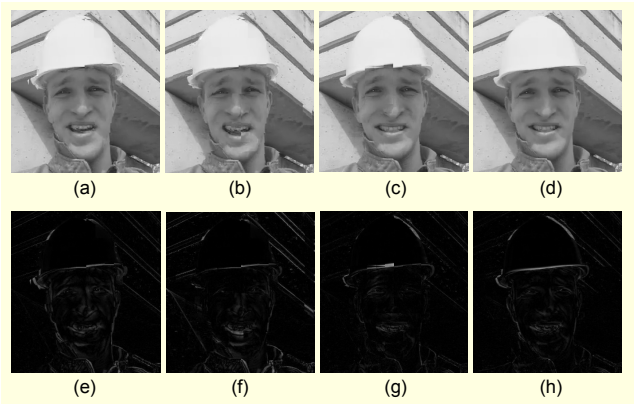

Fig. 10. Comparison of the visual quality for the 23rd concealed frame of the Foreman sequence: (a) FC, (b) CA_B, (c) MVS, and (d) proposed. The bottom row shows the error images between the correctly decoded frame and the concealed frames: (e) FC, (f) CA_B, (g) MVS, and (h) proposed.

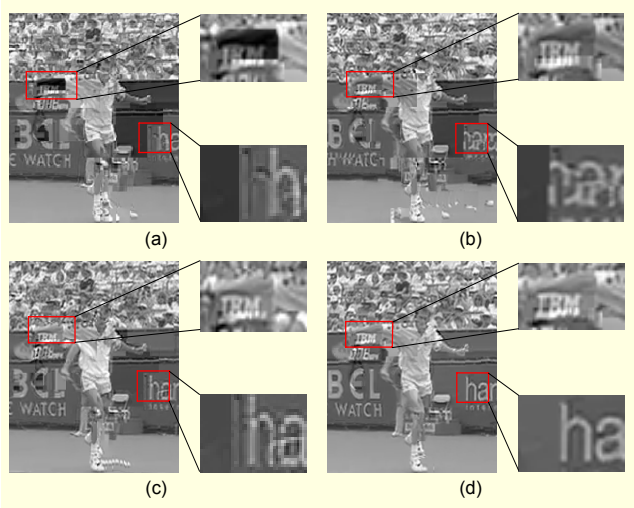

Fig. 11. Comparison of the error propagation effect on the 24th reconstructed frame of the Stefan sequence: (a) FC, (b) CA_B, (c) MVS, and (d) proposed.

the gain is up to 6.2 dB for the proposed algorithm as compared with the CA_B method.

The proposed algorithm also introduces better subjective visual quality than conventional methods. When the 22nd frame is lost, the concealed results of the next frame at the position of the 23rd frame in the Foreman sequence with FC, CA_B, MVS, and the proposed algorithm are shown in Figs. 10(a)-(d). Error images in Figs. 10(e)-(h) show the differences between the correctly decoded frame and the frames concealed by FC, CA_B, MVS, and the proposed algorithm, respectively.

In Figs. 10 (a) and (b), the performance of the FC and CA_B methods are critically degraded due to annoying blocking artifacts seen around the face of the foreman. Therefore, the FC and CA_B methods introduce many error areas in the concealed frames as shown in Figs. 10(e) and (f). In Figs. 10(c) and (g), the scaled MVs obtained in MVS introduce a good image quality for the areas on the wall behind the foreman. However, for the areas around the face and the hat of the foreman, these scaled MVs need to be refined to achieve more satisfactory results. As seen in Figs. 10(d) and (h), with the use of the ORD and SMD constraints, the proposed algorithm yields the most satisfactory image quality with significant reduction in blocking artifacts.

To compare the effects of error propagation after performing different ECs, we show the 24th frame of the Stefan sequence in Fig. 11. When the 12th frame is damaged and concealed by different EC methods, the subsequent frames, including the 24th frame, are degraded by error propagation. As shown in the zoomed images in Figs. 11(a) and (b), the pixel values around the right hand of the tennis player are undesirably inverted from black to white in the FC and CA_B methods. Furthermore, in the conventional methods, the natural smoothness of video frames is severely degraded due to the large area of blocking artifacts in the reconstructed frame. In Fig. 11(c), although the effects of error propagation are reduced in MVS as compared with the conventional methods, the incorrect reconstruction around the head and legs of the tennis player still remains. On the contrary, in the 3rd iteration of the proposed algorithm, since the next frame after the 12th frame is faithfully reconstructed, the concealment error and blocking artifacts propagated from this frame to the subsequent frames are effectively alleviated to provide a smooth and visually pleasant image quality as shown in Fig. 11(d).

IV. Conclusion

In this paper, we have proposed a novel whole frame EC algorithm that directly performs EC on the next frame instead of the current lost frame. In the proposed algorithm, a new concept of ORD which measures the difference between the sibling pixels in sibling regions was introduced and used as the effective constraint in the ME scheme for next frame concealment. To further improve the accuracy of the ME scheme, the spatial smoothness constraint based on the SMD measure is also employed in the proposed algorithm. Since the iterative processes are required to conceal the next frame, the proposed algorithm introduces relatively higher computational complexity than other EC methods. However, experimental results show that the proposed algorithm consistently provides better image quality than conventional methods in terms of both objective and subjective visual quality.

References

[1] J. Huang et al., "Performance of a Mixed-Traffic CDMA2000

Wireless Network with Scalable Streaming Video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 10, Oct. 2003, pp. 973-981.

[2] M. Naghshineh and M. Willebeek-LeMair, "End-to-End QoS Provisioning in Multimedia Wireless/Mobile Networks Using an Adaptive Framework," *IEEE Commun. Mag.*, vol. 35, no. 11, Nov. 1997, pp. 72-81.

[3] Y. Wang, J. Ostermann, and Y.Q. Zhang, *Video Processing and Communication.*, Signal Processing Series, NJ: Prentice Hall, 2002.

[4] Y. Wang and Q.-F. Zhu, "Error Control and Concealment for Video Communication: A Review," *Proc. IEEE*, vol. 86, no. 5, May 1998, pp. 974-997.

[5] W.-Y. Kung, C.-S. Kim, and C.-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, July 2006, pp. 789-802.

[6] W. Zeng and B. Liu, "Geometric-Structure-Based Error Concealment with Novel Applications in Block-Based Low-Bit-Rate Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 4, June 1999, pp. 648-665.

[7] J. Park et al., "Content-Based Adaptive Spatio-Temporal Methods for MPEG-Repair," *IEEE Trans. Image Process.*, vol. 13, no. 8, Aug. 2004, pp. 1066-1077.

[8] S.-H. Lee, D.-H. Choi, and C.-S. Hwang, "Error Concealment Using Affine Transform for H.263 Coded Video Transmissions," *Electron. Lett.*, vol. 37, no. 4, Feb. 2001, pp. 218-220.

[9] M. Friebe and A. Kaup, "Fading Techniques for Error Concealment in Block-Based Video Decoding Systems," *IEEE Trans. Broadcast.*, vol. 53, no. 1, Mar. 2007, pp. 286-296.

[10] L. Tang, "Combined and Iterative Form of Spatial and Temporal Error Concealment for Video Signals," *IEEE Trans. Broadcast.*, vol. 52, no. 2, Sept. 2006, pp. 356-361.

[11] J. Lu, "Signal Processing for Internet Video Streaming: A Review," *Proc. of SPIE Image Video Comm. Process.*, vol. 3974, Jan. 2000, pp. 246-259.

[12] S.K. Bandyopadhyay et al., "Frame Loss Error Concealment for H.264/AVC," 73rd MPEG Meeting and 16th JVT Meeting, Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-P072, July 2005.

[13] P. Baccichet et al., "Frame Concealment for H.264/AVC Decoders," *IEEE Trans. Consumer Electron.*, vol. 51, no. 1, Feb. 2005, pp. 227-233.

[14] Y. Wang, Q.-F. Zhu, and L. Shaw, "Maximally Smooth Image Recovery in Transform Coding," *IEEE Trans. Commun.*, vol. 41, no. 10, Oct. 1993, pp. 1544-1551.

[15] K. Suhring, Ed., JM 12. 2 Reference Software. [Online] Available: http://iphome.hhi.de/suehring/tml/download/

[16] M. Wien, "Variable Block-Size Transforms for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, July 2003, pp. 604-613.

[17] Z. He, J. Cai, and C.W. Chen, "Joint Source Channel Rate-Distortion Analysis for Adaptive Mode Selection and Rate Control in Wireless Video Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, June 2002, pp. 511-523.

[18] S. Belfiore et al., "Concealment of Whole-Frame Losses for Wireless Low Bit-Rate Video Based on Multiframe Optical Flow Estimation," *IEEE Trans. Multimedia*, vol. 7, no. 2, Apr. 2005, pp. 316-329.

Dinh Trieu Duong received the BS degree in electronics engineering from Hanoi Natural Scientific University in 1999. In 2003, he received the MS degree in electronic engineering from Vietnam National University in Hanoi, VNUH. He is working toward the PhD degree in electronic engineering with the Department of Electronic Engineering at Korea University. His current research interests are in the areas of digital signal processing and multimedia communication.

Byeong-Doo Choi received the BS, MS, and PhD degrees, all in electronics engineering, from Korea University, in 2001, 2003, and 2007, respectively. He was a researcher from 2005 to 2007 at the Research Institute for Information and Communication Technology, Korea University. He joined the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut (HHI), Berlin, Germany, as a visiting scholar in 2007. His research interests are in the areas of video communication and image processing.

Min-Cheol Hwang received the BS degree in electronics engineering from Korea University, Seoul, Korea, in 2003, where he is currently working toward the PhD degree in electronics engineering. His research interests are in the areas of image and video signal processing, multimedia communications, DSP systems, and image and video compression including JPEG 2000 and H.264.

Sung-Jea Ko received the BS degree in electronics engineering from Korea University, Seoul, Korea, in 1980, and the MS and PhD degrees in electrical and computer engineering from the State University of New York at Buffalo, in 1986 and 1988, respectively. In 1992, he joined the Department of Electronic Engineering at Korea University, Seoul, Korea, where he is currently a professor. From 1988 to 1992, he was an assistant professor of the Department of Electrical and Computer Engineering at the University of Michigan-Dearborn. He has published more than 300 papers in journals and conference proceedings. He also holds over 30 patents on video signal processing and multimedia communications. He is currently a senior member in the IEEE, a fellow in the IEE (U.K.), and a chairman of the Consumer Electronics Chapter of IEEE Seoul Section. He has been the Special Sessions chair for the IEEE Asia Pacific Conference on Circuits and Systems (1996). He has served as an associate editor for the Journal of the Institute of Electronics Engineers of Korea (IEEK) (1996), the Journal of Broadcast Engineering (1996-1999), and the Journal of the Korean Institute of Communication Sciences (KICS) (1997-2000). He was also an editor of the Journal of Communications and Networks (JCN) (1998-2000). He was the 1999 Recipient of the LG Research Award given to the Outstanding Information and Communication Researcher. He received the Hae-Dong Best Paper Award from the IEEK (1997), the Best Paper Award from the IEEE Asia Pacific Conference on Circuits and Systems (1996), and the Research Excellence Award from Korea University (2004).