Motion-Compensated Frame Interpolation Using Bilateral Motion Estimation and Adaptive Overlapped Block Motion Compensation

Byeong-Doo Choi, *Student Member, IEEE*, Jong-Woo Han, *Student Member, IEEE*, Chang-Su Kim, *Senior Member, IEEE*, and Sung-Jea Ko, *Senior Member, IEEE*

***Abstract*—In this work, we develop a new motion-compe (MC) interpolation algorithm to enhance the temporal resolution of video sequences. First, we propose the bilateral motion estimation scheme to obtain the motion field of an interpolated frame without yielding the hole and overlapping problems. Then, we partition a frame into several object regions by clustering motion vectors. We apply the variable-size block MC (VS-BMC) algorithm to object boundaries in order to reconstruct edge information with a higher quality. Finally, we use the adaptive overlapped block MC (OBMC), which adjusts the coefficients of overlapped windows based on the reliabilities of neighboring motion vectors. The adaptive OBMC (AOBMC) can overcome the limitations of the conventional OBMC, such as over-smoothing and poor de-blocking. Experimental results show that the proposed algorithm provides a better image quality than conventional methods both objectively and subjectively.**

***Index Terms*—Bilateral motion estimation (ME), frame rate up-conversion, MC interpolation (MCI), overlapped block motion compensation (OBMC).**

I. Introduction

To TRANSMIT a huge amount of video data over bandwidth-limited channels, the spatiotemporal resolution of a video signal is often reduced to achieve the target bit rate. In the temporal domain, the bit rate can be reduced by skipping frames, but the quality of the reconstructed video is inevitably degraded in such a case. To restore the loss of temporal resolution, frame rate up-conversion (FRUC) is required at the decoder side as a post-processing tool.

Various FRUC algorithms have been developed. A simple approach is to combine the pixel values at the same spatial location without considering object motions, e.g., frame repetition or frame averaging. Although these algorithms provide acceptable visual quality in the absence of motions, they can cause motion jerkiness and blurring of moving objects. Motion compensation techniques can be applied to reduce these artifacts. Such methods are called MC interpolation (MCI).

There are three issues to be addressed in MCI. First, the motion information should be accurately estimated to reconstruct interpolated frames faithfully. Most MCI algorithms utilize the block-matching algorithm (BMA) for motion estimation (ME). BMA is simple and easy to implement [1]–[3]. It also generates a compactly represented motion field. However, unlike video compression, it is more important to find true motion trajectories in MCI. Note that the objective of MC in MCI is not to minimize the energy of MC residual signals, but to reconstruct interpolated frames with better visual quality. Since motion vectors from the conventional BMA are often not faithful to true object motions, several approaches for more accurate ME have been proposed in recent works [4]–[6]. In [7], de Haan *et al.* proposed a 3-D recursive ME, which has been applied to several MCI schemes due to its fast convergence and high accuracy.

Second, when BMA is used for MCI, a pixel in an interpolated frame can be passed through by multiple motion trajectories or no motion trajectory. The former case is called the overlapping problem, and the latter the hole problem. Several approaches have been proposed to handle overlapped and hole regions, for example the median filtering [8], the spatial interpolation [9], and the MC using neighboring motion fields [10], [11]. However, these methods demand complicated operations and may produce undesirable artifacts.

Third, blocking artifacts can occur in the conventional blockwise processing. They arise when a block contains multiple objects with different motions. We can reduce blocking artifacts using the overlapped block MC (OBMC) technique [12], [13]. However, when OBMC is applied to all blocks uniformly, the quality of the interpolated frame may be degraded due to over-smoothing of edges. Another approach to avoid blocking artifacts is to use spatial transform techniques, such as control grid interpolation (CGI) [14]–[16]. As compared with the translational block model, the spatial transformation techniques enable superior motion tracking and rendition, especially in the presence of rotation and zooming. However, these techniques require much higher computational complexity than the translational block model. Moreover, a globally smooth motion field model is assumed in the spatial transform techniques, which is not true for most image sequences.

In this paper, we propose a novel MCI method that employs the bilateral ME and the adaptive OBMC (AOBMC). The proposed bilateral ME can prevent the hole problem as well as the overlapping problem by estimating the motion vectors of an interpolated frame directly. The proposed algorithm uses AOBMC, which reconstructs the interpolated frame faithfully by controlling the weights of overlapping windows according

Manuscript received May 26, 2006; revised September 16, 2006. This research was supported by Seoul Future Contents Convergence (SFCC) Cluster established by Seoul Industry-Academy-Research Cooperation Project. This paper was recommended by Associate Editor E. Steinbach.

The authors are with Department of Electronics Engineering, Korea University, Seoul 136–701, Korea (e-mail: bottle02@dali.korea.ac.kr; jwhan@dali.korea.ac.kr; cskim@ieee.org; sjko@dali.korea.ac.kr).

Digital Object Identifier 10.1109/TCSVT.2007.893835

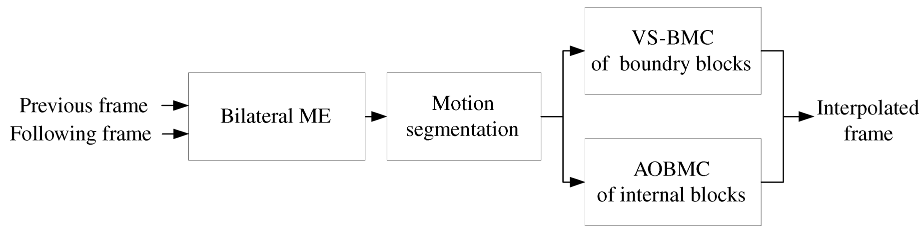


Fig. 1. Architecture of the proposed MCI algorithm.

to the reliabilities of motion vectors. Moreover, the proposed algorithm applies a motion segmentation scheme to divide a frame into several object regions and uses the variable-size block MC (VS-BMC) scheme to interpolate the object boundaries accurately.

The rest of the paper is organized as follows. Section II surveys the previous work on the motion vector prediction and the MC for interpolated frames. Section III describes the proposed algorithm, and Section IV discusses experimental results. Finally, Section V concludes this paper.

II. Previous Work

In MCI, the quality of an interpolated frame is mainly dependent on how to estimate accurate motion trajectories through the interpolated frame and how to merge MC blocks without unnatural defects. We briefly review the previous work on these two issues subsequently.

A. Motion Estimation for Interpolated Frames

In the blockwise MCI, the skipped (or intermediate) frame is divided into blocks, and then each block is interpolated using the information in the previous frame and the following frame. Let f_{n-1}, f_n, and f_{n+1} denote the previous, the intermediate, and the following frames, respectively. Also, let $\mathbf{s}$ be a 2-D vector which represents a pixel location. Then, the interpolated intensity of the $\mathbf{s}$th pixel in f_n is given by

$$\widehat{f}_n[\mathbf{s}, \mathbf{v}] = \frac{1}{2}\left(f_{n-1}[\mathbf{s}-\mathbf{v}] + f_{n+1}[\mathbf{s}+\mathbf{v}]\right) \tag{1}$$

where $\mathbf{v}$ denotes the estimated motion vector of the interpolated pixel, and $f_n[\mathbf{s}]$ denotes the original intensity of the $\mathbf{s}$th pixel in the nth frame f_n.

In MCI, we should estimate the motion vector of each interpolated pixel. If the conventional BMA is used to find a blockwise motion vector field between the previous frame and the following frame, the motion trajectories may not cover all pixels in the interpolated frame, consequently yielding hole regions. In addition, multiple trajectories may pass through the same pixel, causing overlapping regions. Therefore, we should estimate the motion vectors for the blocks in the interpolated frame, instead of using the motion vectors between the previous frame and the following frame. Next, we summarize conventional algorithms to obtain the motion vectors for the interpolated frame.

1) All motion vectors for the interpolated frame are set to zero. It is known as the frame hold (or repetition) scheme [1].
2) The motion vectors for the interpolated frame are copied from those for the previous frame [17].
3) The motion vector for an interpolated block is obtained by linearly combining the candidate motion vectors between the previous and the following frames [18], [19].
4) The pixelwise motion vectors for hole regions are predicted from the neighboring block motion vectors using the CGI technique [11].

In this paper, we propose the bilateral ME, which finds the motion vectors for the interpolated frame directly by comparing a block at the shifted position in the previous frame and another block at the opposite position in the following frame. The bilateral ME will be discussed in more detail in Section III-A.

B. Advanced Motion Models

In the conventional BMA, an input image is decomposed into blocks, and each block is assumed to experience a translational motion. BMA is the most widely adopted approach in video compression standards such as MPEG-4 and H.264. However, BMA cannot represent complex motions, such as zooming, rotation and local deformation. Moreover, it results in serious blocking artifacts since object boundaries may not agree with block boundaries and adjacent blocks can have significantly different motion vectors. To overcome these problems, several approaches have been proposed to express more complicated motions and reduce blocking artifacts.

1) *Spatial transform or CGI*: It first estimates the motion vectors for control grid points and then interpolates them to obtain a smooth motion vector field [14].
2) *VS-BMC*: It partitions a frame into blocks of different sizes. Regions with homogeneous translational motions are divided into larger blocks, whereas those with complicated motions are partitioned into smaller blocks [20].
3) *OBMC*: It predicts a frame by positioning overlapped blocks from a reference frame using a weighting window [12], [21].
4) *Hybrid approach*: It integrates the advantages of both CGI and OBMC using dynamic programming [5], [22]. But, its computational complexity is relatively high.

In the proposed algorithm, we employ the OBMC scheme, which can substantially reduce MC errors, as well as blocking artifacts, at the cost of modest computational overhead.

III. Proposed Algorithm

Fig. 1 shows the overall architecture of the proposed MCI algorithm, which consists of four processing units. First, the bilateral ME method predicts the motion vectors for the intermediate frame using the information in the previous and the following frames. Second, the frame is partitioned into boundary and internal blocks using a motion segmentation method. Third,

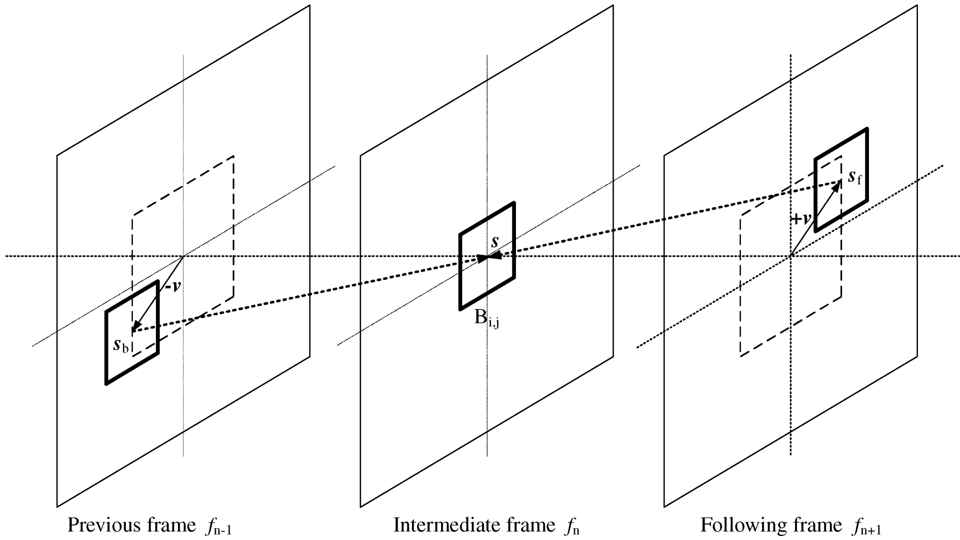


Fig. 2. Bilateral ME.

VS-BMC is applied to the boundary blocks to reduce undesirable blurring artifacts. Finally, the intermediate frame is interpolated by AOBMC with adaptive window coefficients.

Compared with our previous work [23], the proposed algorithm employs more advanced ME and compensation techniques. First, the proposed bilateral ME improves the reliability of the ME method in [23] by applying a spatial smoothness constraint. Second, the AOBMC method is proposed to reconstruct the intermediate frame without blocking defects and blurring artifacts. Third, the motion segmentation with VS-BMC improves the performance of the frame interpolation, especially around object boundaries.

A. Bilateral Motion Estimation

Since we do not have the original pixel values of an intermediate frame, it is not straightforward to estimate the motion vectors of the blocks in the intermediate frame. In [23], we proposed the bilateral criterion for ME to find the motion vectors using the information in the previous and the following frames. Specifically, we obtain the motion trajectory passing through a block in the intermediate frame by comparing a block at a shifted position in the previous frame and another block at the opposite position in the following frame.

Fig. 2 illustrates the bilateral ME. Let $\mathbf{s}$ denote a pixel in the intermediate frame f_n and $\mathbf{v}$ denote its candidate motion vector. Then, the pixel is mapped backwardly to a pixel in the previous frame f_{n-1}, which is given by

$$\mathbf{s}_b = \mathbf{s} - \mathbf{v}. \tag{2}$$

The forward motion vector is assumed to be the same as the backward motion vector. Therefore, the intermediate pixel is mapped forwardly to the following frame f_{n+1} via

$$\mathbf{s}_f = \mathbf{s} + \mathbf{v}. \tag{3}$$

In this way, $\mathbf{s}_b$ and $\mathbf{s}_f$ are at the opposite positions with respect to $\mathbf{s}$. The bilateral ME uses a block motion model. Let $B_{i,j}$ denote a block in the intermediate frame. Then, for each candidate motion vector $\mathbf{v}$, we compute the sum of bilateral absolute differences (SBAD) by

$$\mathrm{SBAD}[B_{i,j}, \mathbf{v}] = \sum_{\mathbf{s} \in B_{i,j}} |f_{n-1}[\mathbf{s} - \mathbf{v}] - f_{n+1}[\mathbf{s} + \mathbf{v}]|. \tag{4}$$

By searching the vector that minimizes the SBAD, we can find out the motion vector for the intermediate block using the information in the previous and the following frames only.

It was demonstrated that bilateral ME provides acceptable motion vectors if the video sequence contains simple translational motions [23]. However, it does not yield reliable motion trajectories when objects have rotational or zooming motions. This is because many trajectories through the current block can be strong candidates with small SBADs. In this work, we impose a spatial smoothness constraint additionally to improve the robustness of the bilateral ME. More specifically, we adopt the side match distortion (SMD), which measures the spatial smoothness across block boundaries as shown in Fig. 3. It computes the average of the absolute pixel differences between the predicted block and the neighboring blocks along the top, bottom, left and right boundaries by

$$\mathrm{SMD}[B_{i,j}, \mathbf{v}] = \frac{1}{N} \sum_{k=0}^{N-1} |\widehat{f}_n[\mathbf{g}_k, \mathbf{v}] - \widehat{f}_n[\mathbf{h}_k]| \tag{5}$$

where $\mathbf{g}_k$ and $\mathbf{h}_k$ indicate the positions of the kth pixels at the boundaries in the predicted block and the neighboring blocks, respectively, and N is the number of boundary pixels. As in (1), $\widehat{f}_n[\mathbf{g}_k, \mathbf{v}]$ denotes the interpolated pixel value in the predicted block when its motion vector is set to $\mathbf{v}$. Note that the interpolated pixel values, $\widehat{f}_n[\mathbf{h}_k]$, in the neighboring blocks should be available when the SMD is evaluated. Therefore, the proposed ME method is applied iteratively as discussed below.

We find the best motion vector $\mathbf{v}_{i,j}$ for the intermediate block $B_{i,j}$ by minimizing the weighted sum of the SBAD and the SMD.

$$\mathbf{v}_{i,j} = \arg\min_{\mathbf{v}} \{\mu \cdot \mathrm{SBAD}[B_{i,j}, \mathbf{v}] + (1 - \mu) \cdot \mathrm{SMD}[B_{i,j}, \mathbf{v}]\} \tag{6}$$

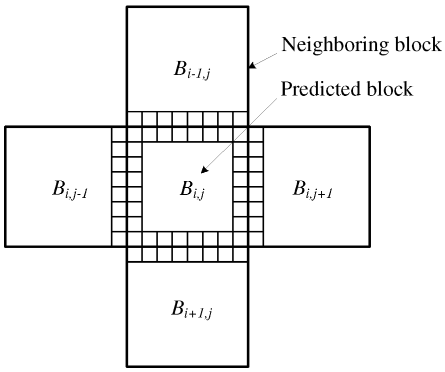


Fig. 3. Side match criterion.

where μ is a weighting coefficient. Since the computation of the SMD requires the pixel values in the right and the bottom blocks, as well as the left and the top blocks, the estimation of a motion vector is affected by those of later blocks. Therefore, the minimization of the cost function in (6) is iteratively applied for all blocks in the intermediate frame, until the estimated motion vectors do not change.

Experimental results confirmed that the proposed algorithm provides a good performance, when the weighting coefficient μ in (6) is about 0.4, and the performance is not very sensitive to the variation of μ. Therefore, in this work, μ is fixed to 0.4.

B. Segmentation of Moving Objects

After the bilateral ME, we classify the blocks in the intermediate frame into several partitions according to the motion vectors. Each partition contains blocks with similar motion vectors and represents a coarsely segmented moving object. For the object segmentation, similar motion vectors are grouped using the k-means clustering algorithm [24] as follows.

Step 1) All blocks in the frame are considered as a single object, and the cluster center is set to the average motion vector of the blocks.

Step 2) If the difference between the motion vector of each block and the cluster center is larger than a predefined threshold τ, the block is declared to belong to a new object. The average motion vector of the blocks in the new object is set as a new cluster center. In this work, the threshold τ is set experimentally to 8.

Step 3) Each motion vector is assigned to the nearest cluster center.

Step 4) Each cluster center is updated to the average of the motion vectors in the cluster.

Step 5) Steps 2–4 are iteratively repeated until there is no change in the cluster centers.

As shown in Fig. 4, a block is called a boundary block if it is at the boundary between two objects, or an internal block otherwise. Then, we use VS-BMC and AOBMC to motion-compensate boundary blocks and internal blocks, respectively.

C. Variable-Size Block Motion Compensation (VS-BMC)

Since the proposed segmentation scheme uses an 8×8 block as the smallest region unit, the blockwise boundary may not match perfectly with the true object boundary. For example,

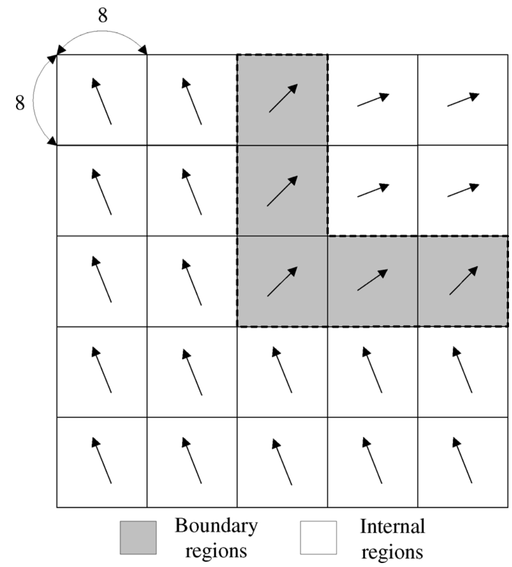


Fig. 4. Block classification based on the motion segmentation.

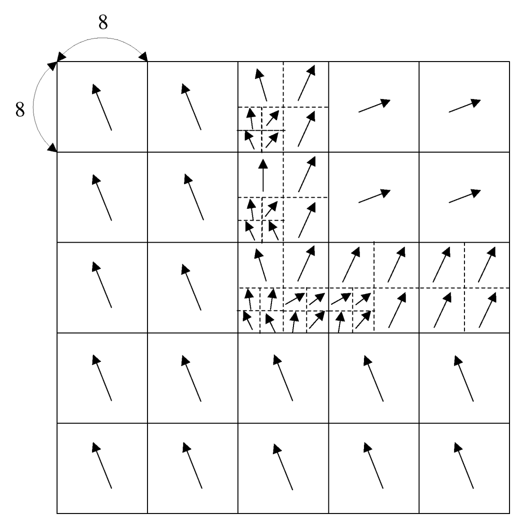


Fig. 5. VS-BMC for boundary blocks.

a boundary block in Fig. 4 can actually contain multiple objects with different motions. To express such complex motions, we adopt VS-BMC to reconstruct boundary blocks. VS-BMC is known to be more effective than the fixed-size block MC in representing irregular motions. Smaller blocks are used to describe complex motions near object boundaries, while larger blocks are used for a homogeneous region.

As shown in Fig. 5, we adopt a quadtree-based VS-BMC, which divides an 8×8 boundary block into 4×4 or 2×2 subblocks. Let $B_l^{(n)}$ denote the lth subblock of size $2^n \times 2^n$ and $\mathbf{v}_l^{(n)}$ denote its motion vector.

Step 1) Initialize $n = 3$.

Step 2) Divide the current $2^n \times 2^n$ block into four $2^{n-1} \times 2^{n-1}$ subblocks

$$B_l^{(n)} \rightarrow \{B_0^{(n-1)}, B_1^{(n-1)}, B_2^{(n-1)}, B_3^{(n-1)}\}. \tag{7}$$

Then, find the motion vector $\mathbf{v}_m^{(n-1)}$ of each subblock $B_m^{(n-1)}$ using the SBAD measure in (4).

Step 3) If the SBAD for every subblock is less than a quarter of the SBAD for the original $2^n \times 2^n$ block, accept the subdivision. More specifically, if

$$\mathrm{SBAD}[B_m^{(n-1)}, \mathbf{v}_m^{(n-1)}] < \frac{1}{4}\mathrm{SBAD}[B_l^{(n)}, \mathbf{v}_l^{(n)}], \quad \text{for all } 0 \leq m \leq 3 \quad (8)$$

accept the subdivision and decrease n to $n-1$. Otherwise, maintain the original block and terminate the procedure.

Step 4) If $n = 1$, terminate the procedure. Otherwise, go to Step 2.

In this way, we obtain finer motion fields at object boundaries and alleviate the shortcomings of the fixed-size block MC. In VS-BMC, the interpolated signal is obtained by

$$\widehat{f}_n[\mathbf{s}, \mathbf{v}_l^{(n)}] = \frac{1}{2}\left(f_{n-1}[\mathbf{s} - \mathbf{v}_l^{(n)}] + f_{n+1}[\mathbf{s} + \mathbf{v}_l^{(n)}]\right), \quad \text{for } \mathbf{s} \in B_l^{(n)}. \quad (9)$$

D. Overlapped Block Motion Compensation Using Adaptive Window

In the conventional MC, each block is associated with a motion vector, which is used for that block only. In comparison, in OBMC, the motion vector of a block is applied to a larger set of pixels using a weighting window.

For the sake of simpler notations, in this section, we assume that $\mathbf{s}$ is located on the lower right quadrant of the (i, j)th block $B_{i,j}$. Then, each pixel $f_n[\mathbf{s}]$ in $B_{i,j}$ is predicted using the motion vectors of the block itself, the lower block $B_{i+1,j}$, the right block $B_{i,j+1}$, and the lower right block $B_{i+1,j+1}$. Let $\mathbf{v}_{i+p,j+q}$ denote the motion vector of $B_{i+p,j+q}$ and $w_{p,q}[\mathbf{s}]$ denote the corresponding weighting coefficient. Then, the prediction $\widehat{f}_n[\mathbf{s}]$ can be expressed as

$$\widehat{f}_n[\mathbf{s}] = \frac{1}{2}\sum_{p=0}^{1}\sum_{q=0}^{1} w_{p,q}[\mathbf{s}]\{f_{n-1}[\mathbf{s} - \mathbf{v}_{i+p,j+q}] + f_{n+1}[\mathbf{s} + \mathbf{v}_{i+p,j+q}]\} \quad (10)$$

where

$$\sum_{p=0}^{1}\sum_{q=0}^{1} w_{p,q}[\mathbf{s}] = 1. \quad (11)$$

The coefficient $w_{p,q}[\mathbf{s}]$ is determined by the relative position of $\mathbf{s}$ within $B_{i,j}$. In practice, a smooth function, such as the raised cosine window [12], is often used for the weighting window $w_{p,q}[\cdot]$. The prediction rules for pixels in the other quadrants can be derived in a similar way. For example, for the upper left quadrant of a block, we use the motion vectors of the block itself, the upper block, the left block, and the upper left block.

In general, when motion activities are low, OBMC can effectively reduce blocking artifacts and provide good visual quality. On the contrary, if adjacent blocks have substantially different motions, OBMC can yield blurring or over-smoothing artifacts. Suppose that two blocks are adjacent to each other. One block is a part of a fast moving object, while the other belongs to the background. The motion vector of the moving block should not affect the pixels in the background block. Unfortunately, in the conventional OBMC, the weighting coefficients are determined only by the relative distances of the pixels within the block, and the boundary pixels in the background block are significantly influenced by the motion vector of the moving block.

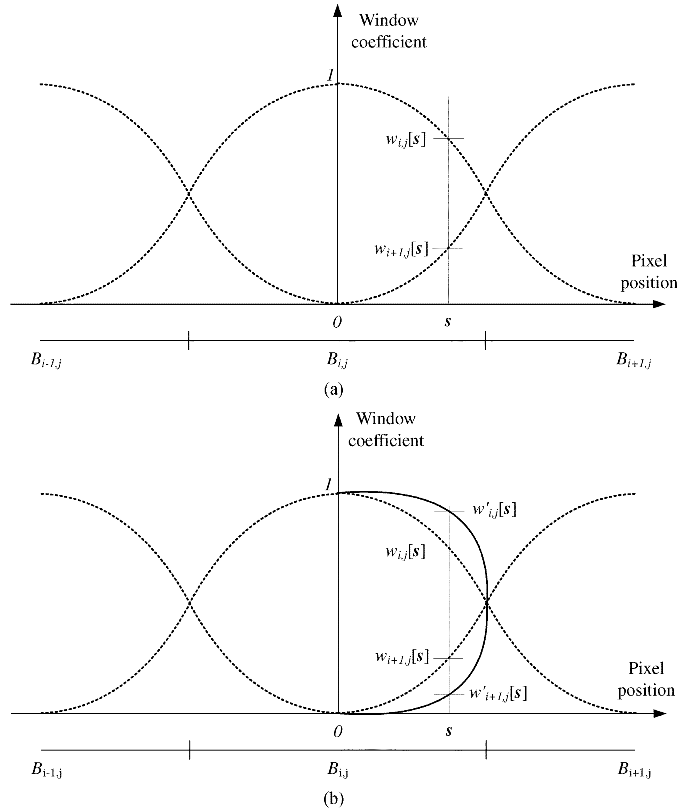


Fig. 6. Comparison of (a) the conventional OBMC and (b) the proposed AOBMC.

To overcome this shortcoming, we propose AOBMC that controls the weighting coefficients adaptively according to the reliabilities of the neighboring motion vectors. Fig. 6 compares the proposed AOBMC with the conventional OBMC. For the sake of simple illustrations, Fig. 6 shows 1-D weighting windows, instead of 2-D ones. Each pixel $\mathbf{s}$ in the right half of $B_{i,j}$ is MC by the weighted sum of two pixel values, which are specified by the motion vectors of $B_{i,j}$ and $B_{i+1,j}$. In the conventional OBMC, the weight coefficients $w_{i,j}[\mathbf{s}]$ and $w_{i+1,j}[\mathbf{s}]$ are determined by the relative position of $\mathbf{s}$ within $B_{i,j}$. On the other hand, AOBMC modifies those weighting coefficients using the reliability of the neighboring motion vector. For instance, suppose that the motion vector of $B_{i+1,j}$ is not reliable enough to be used for the prediction of $B_{i,j}$. In such a case, the corresponding weights $w_{i+1,j}[\mathbf{s}]$ are reduced to $w'_{i+1,j}[\mathbf{s}]$, while the other weights $w_{i,j}[\mathbf{s}]$ are increased to $w'_{i,j}[\mathbf{s}]$ as shown in Fig. 6(b).

In this work, the reliability of the neighboring motion vector $\mathbf{v}_{i+p,j+q}$ for the prediction of the current block $B_{i,j}$ is defined as

$$\Phi_{B_{i,j}}[\mathbf{v}_{i+p,j+q}] = \frac{\mathrm{SBAD}[B_{i,j}, \mathbf{v}_{i,j}]}{\mathrm{SBAD}[B_{i,j}, \mathbf{v}_{i+p,j+q}]}. \quad (12)$$

Note that $\mathbf{v}_{i,j}$ is selected to minimize the cost function in (6), which includes the SBAD measure. Therefore, in most cases, $\mathrm{SBAD}[B_{i,j}, \mathbf{v}_{i,j}]$ is smaller than or equal to $\mathrm{SBAD}[B_{i,j}, \mathbf{v}_{i+p,j+q}]$, and $0 \leq \Phi_{B_{i,j}}[\mathbf{v}_{i+p,j+q}] \leq 1$. As $\Phi_{B_{i,j}}[\mathbf{v}_{i+p,j+q}]$ becomes larger, the neighboring motion vector $\mathbf{v}_{i+p,j+q}$ can be more reliably used to motion-compensate the current block $B_{i,j}$.

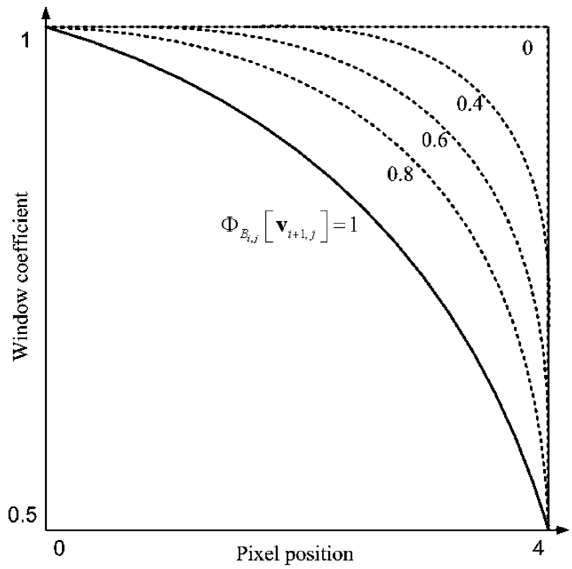


Fig. 7. Reshaping of the raised cosine window in Fig. 6(a) according to the reliability of the neighboring motion vector $\Phi_{B_{i,j}}[\mathbf{v}_{i+1,j}]$.

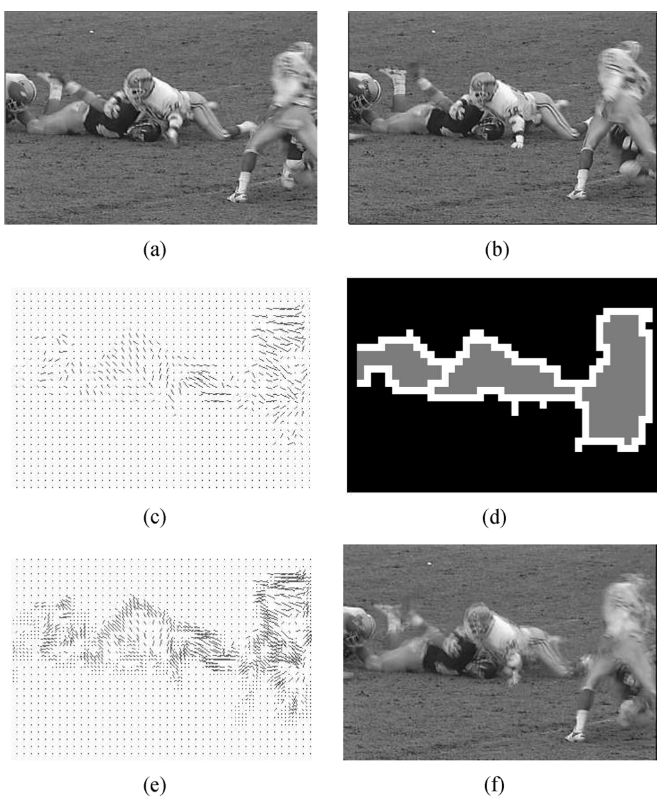


Fig. 8. Step-by-step processing results: (a) the previous frame, (b) the following frame, (c) the motion vector field, (d) the motion segmentation, (e) the motion vector field after VS-BMC, and (f) the interpolated frame with AOBMC.

Consequently, using the reliability information, the weighting coefficients $w_{p,q}[\mathbf{s}]$ in (10) are modified to

$$w'_{p,q}[\mathbf{s}] = \frac{\Phi_{B_{i,j}}[\mathbf{v}_{i+p,j+q}] \cdot w_{p,q}[\mathbf{s}]}{\sum_{a=0}^{1}\sum_{b=0}^{1}\Phi_{B_{i,j}}[\mathbf{v}_{i+a,j+b}] \cdot w_{a,b}[\mathbf{s}]}. \tag{13}$$

The modified weighting coefficient is proportional to the reliability $\Phi_{B_{i,j}}[\mathbf{v}_{i+p,j+q}]$. The denominator in (13) is necessary to normalize the weighting coefficients such that $\sum_{p=0}^{1}\sum_{q=0}^{1} w'_{p,q}[\mathbf{s}] = 1$.

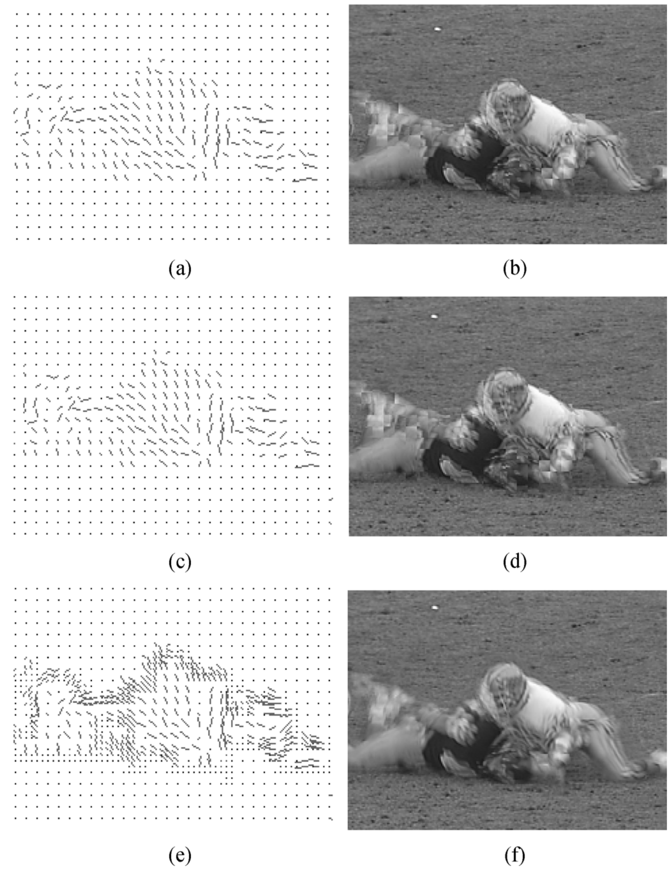


Fig. 9. Comparison of motion vector fields. (a) Motion vector field. (b) Interpolated frame by the bilateral ME without the side match criterion. (c) Motion vector field. (d) Interpolated frame by the bilateral ME with the side match criterion. (e) Motion vector field. (f) Interpolated frame by the proposed bilateral ME with VS-BMC.

Fig. 7 shows how the raised cosine window in Fig. 6(a) is modified according to the reliability of the neighboring motion vector. If the neighboring motion vector has reliability 1, the raised cosine window is not modified. In the other extreme case, if the neighboring motion vector has reliability 0, it is not used for the current block and the window becomes rectangular as in the conventional block MC. Thus, the proposed algorithm can overcome the disadvantage of the conventional OBMC by controlling the window shapes adaptively.

IV. Experimental Results

We demonstrate the performance of the proposed MCI algorithm using several test sequences, which are in the standard CIF (352×288) format. In all experiments, the block size is set to 8×8 and the search range is set to 16 pixels for both horizontal and vertical directions. However, a boundary block can be divided into multiple 4×4 or 2×2 subblocks in VS-BMC. The following tests double the original frame rate 30–60 fps.

A. Step by Step Processing Results

The proposed algorithm consists of four steps: the bilateral ME, the motion segmentation based on the K-means clustering, VS-BMC for boundary blocks, and AOBMC. Fig. 8 shows

(a) (b) (c)

Fig. 10. Comparison of the MC methods. (a) No-OBMC. (b) Conventional OBMC. (c) Proposed AOBMC with VS-BMC.

the processing result after each step on the *Football* sequence. Using the previous and the following frames in Fig. 8(a) and (b), the bilateral ME estimates the motion vector field in Fig. 8(c). From these motion vectors, we extract object regions using the K-means clustering scheme, as shown in Fig. 8(d). Gray and white blocks, respectively, represent the interiors and boundaries of moving objects, while black blocks depict the background. Then, we apply VS-BMC to the boundary blocks to obtain the detailed motion vectors in Fig. 8(e). Finally, AOBMC is used to interpolate the intermediate frame in Fig. 8(f).

(a) (b) (c)

Fig. 11. Interpolated images of the *Football* sequence. (a) 3-D recursive MCI. (b) Zhai *et al.* MCI. (c) Proposed algorithm.

(a) (b) (c)

Fig. 12. Interpolated images of the *Foreman* sequence. (a) 3-D recursive MCI. (b) Zhai *et al.* MCI. (c) Proposed algorithm.

(a) (b) (c)

Fig. 13. Interpolated images of the *Table Tennis* sequence. (a) 3-D recursive MCI. (b) Zhai *et al.* MCI. (c) Proposed algorithm.

(a) (b) (c)

Fig. 14. Interpolated images of the *Mobile* sequence. (a) 3-D recursive MCI. (b) Zhai *et al.* MCI. (c) Proposed algorithm.

B. Subjective Evaluation

Next, we compare the subjective image qualities of interpolated frames. In many applications, subject image quality is as important as objective image quality, since the mean square error (MSE) or peak signal-to-noise ratio (PSNR) measures cannot accurately represent the image quality perceived by the human visual system.

Fig. 9(a) and (c) show the motion vector fields of the *Football* sequence, which are predicted by the bilateral ME without and with the side match criterion, respectively. We see that the proposed bilateral ME with the side match criterion provides a more regularized and accurate motion vector field by imposing the smoothness constraint across block boundaries. Moreover, if VS-BMC is combined with the proposed bilateral ME, the detailed motions around the object boundaries can be more effectively represented as shown in Fig. 9(e). The corresponding predicted images are shown in Fig. 9(b), (d), and (e). It is observed that the proposed bilateral ME with VS-BMC yields the most faithful image quality.

Fig. 10 compares the performance of the MC methods: the conventional block compensation (No-OBMC), the conventional OBMC [12], and the proposed AOBMC with VS-BMC. In each scheme, all motion vectors are obtained by the proposed

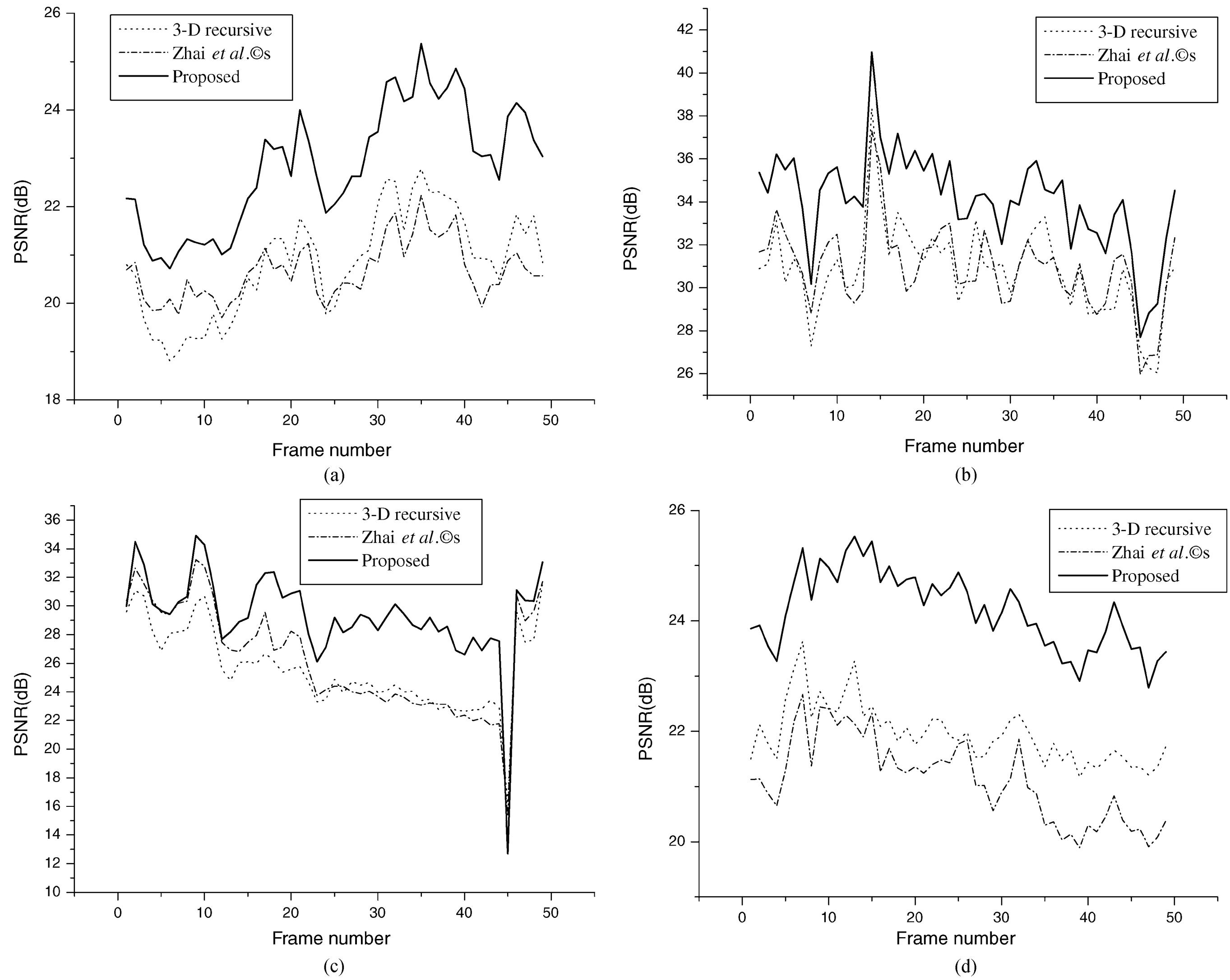


Fig. 15. Comparison of the PSNR performances on the (a) *Football*. (b) *Foreman*. (c) *Table Tennis*. (d) *Mobile* sequences.

bilateral ME. Notice that the conventional block compensation yields severe blocking artifacts. The conventional OBMC alleviates these blocking artifacts, but it causes blurring artifacts especially around the back numbers. On the other hand, the proposed algorithm reconstructs crisp edges efficiently, providing the best image quality, as shown in Fig. 10(c).

Fig. 11–14 show the interpolated images of the *Football*, *Foreman*, *Table Tennis*, and *Mobile* sequences, respectively. The performance of the proposed algorithm is compared with those of the 3-D recursive MCI and the Zhai *et al.* method [21]. In the 3-D recursive MCI, motion vectors are estimated by the 3-D recursive block matching [7], and holes and overlapping regions are processed as described in [25] and [26]. In the Zhai *et al.* method, the motion information in the compressed video signal is refined with the SBAD measure in [23], and each block is compensated with the conventional OBMC. On the other hand, the proposed algorithm consists of the bilateral ME, VS-BMC, and AOBMC. The *Football* sequence has fast and complicated motions, the *Foreman* sequence contains modest motions, the *Table Tennis* sequence has fast moving objects and scene changes, and the *Mobile* sequence contains multiple objects with diverse motions. For all these four sequences, the proposed algorithm reduces blocking artifacts effectively, preserves the sharp edges of objects, and provides significantly better image quality. The quality improvement can be easily observed around the football players' bodies, the foreman's face, the table tennis player's arms, and the rolling ball.

C. Objective Evaluation

Next, we compare the PSNR performance of the proposed algorithm with those of the 3-D recursive MCI and the Zhai *et al.* method. To measure the objective qualities of interpolated frames, the nth frame is interpolated from the original $(n-1)$th and $(n+1)$th frames, and PSNR is computed between the original nth frame and the interpolated nth frame. Let f_n and $\widehat{f}_n$ denote the original frame and the interpolated frame, respectively. Then, PSNR is defined as

$$\mathrm{PSNR} = 10\log_{10} \frac{255^2}{\frac{1}{W\times H}\sum_{j=0}^{H-1}\sum_{i=0}^{W-1}\left(f_n[i,j] - \widehat{f}_n[i,j]\right)^2} \tag{14}$$

where $W \times H$ is the spatial resolution of the sequence.

Fig. 15 shows the PSNRs of the first 100 interpolated frames of the *Football*, *Foreman*, *Table Tennis*, and *Mobile* sequences.

TABLE I
COMPARISON OF THE AVERAGE PROCESSING TIMES (S/FRAME)

	Motion Estimation	Motion Compensation	Total
3-D recursive	0.057	0.004	0.062
Zhai *et al.*'s	0.122	0.025	0.147
Proposed	0.287	0.028	0.315

It is seen that the proposed algorithm provides up to 2 dB better PSNR performance than the conventional MCI schemes. The performance improvement of the proposed algorithm is achieved especially for the *Football* and *Mobile* sequences which have fast and complex motions.

Table I compares the average processing times. In all three methods, ME takes a significantly longer time than the MC. The 3-D recursive scheme employs the simplest ME and requires the shortest processing time. The average processing time of the proposed algorithm is about twice longer than that of the Zhai *et al.* method. However, at the cost of higher computational complexity, the proposed algorithm provides much better image quality.

V. CONCLUSION

In this paper, we proposed a new MCI method based on the bilateral ME and AOBMC. The bilateral ME can predict motion vectors effectively without any hole and overlapping artifacts. Moreover, we improved the reliability of the ME by employing the side match criterion. Also, for the MC, we proposed AOBMC that controls weighting coefficients adaptively based on the reliability of neighboring motion vectors.

We compared the performance of the proposed algorithm with those of the 3-D recursive MCI and the Zhai *et al.* method. Experimental results demonstrated that the proposed algorithm provides better image quality than the conventional schemes both objectively and subjectively. Specifically, the proposed algorithm provides up to 2 dB better PSNR performance than the conventional schemes. Although the experiments were mainly for doubling the input frame rate, the proposed algorithm can be easily extended for various applications, such as the reconstruction of lost frames in wireless video communications or the conversion of 60 fps videos to 90 or 120 fps for the elimination of motion blurring in LCD TVs.

In this work, only two frames (the previous and the following one) are employed to interpolate an intermediate frame. The quality of the interpolated frame can be improved further by using more than three neighboring frames. However, in such a case, the complexity of the algorithm will be also increased. It is one of our future research topics to develop a low complexity MCI method using multiple reference frames.

REFERENCES

[1] P. Haavisto, J. Juhola, and Y. Neuvo, "Fractional frame rate up-conversion using weighted median filters," *IEEE Trans. Consum. Electron.*, vol. 35, no. 3, pp. 272–278, Aug. 1989.

[2] S.-H. Lee, Y.-C. Shin, S.-J. Yang, H.-H. Moon, and R.-H. Park, "Adaptive MC interpolation for frame rate up-conversion," *IEEE Trans. Consum. Electron.*, vol. 48, no. 3, pp. 444–450, Aug. 2002.

[3] R. Castagno, P. Haavisto, and G. Ramponi, "A method for motion adaptive frame rate up-conversion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 5, pp. 436–446, Oct. 1996.

[4] Y.-K. Chen and S. Y. Kung, "Rate optimization by true motion estimation," in *Proc. IEEE Workshop Multimedia Signal Process.*, Jun. 1997, pp. 187–194.

[5] C.-C. Cheng, W.-L. Hwang, Z. Shen, and T. Xia, "Advanced motion compensation techniques for blocking artifacts reduction in 3-D video coding systems," in *Proc. ICIP*, Sep. 2005, vol. 3, pp. 89–92.

[6] Y.-K. Chen, A. Vetro, H. Sun, and S. Y. Kung, "Frame-rate up-conversion using transmitted true motion vectors," in *Proc. IEEE Workshop Multimedia Signal Process.*, Dec. 1998, vol. 2, pp. 622–627.

[7] G. de Haan, P. W. Biezen, H. Huijgen, and O. A. Ojo, "True-motion estimation with 3-D recursive search block matching," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 5, pp. 368–379, Oct. 1993.

[8] T. Y. Kuo and C.-C. J. Kuo, "Motion-compensated interpolation for low-bit-rate video quality enhancement," *Proc. SPIE Vis. Commun. Image Process.*, vol. 3460, pp. 277–288, Jul. 1998.

[9] A. Kaup and T. Aach, "Efficient prediction of uncovered background in interframe coding spatial extraprediction," in *Proc. ICASSP*, Apr. 1994, vol. 5, pp. 501–504.

[10] R. J. Schutten and G. D. Haan, "Real-time 2–3 pull-down elimination applying motion estimation/compensation in a programmable device," *IEEE Trans. Consum.Electron.*, vol. 44, no. 3, pp. 501–504, Aug. 1998.

[11] B.-W. Jeon, G.-I. Lee, S.-H. Lee, and R.-H. Park, "Coarse-to-fine frame interpolation for frame rate up-conversion using pyramid structure," *IEEE Trans. Consum. Electron.*, vol. 49, no. 3, pp. 499–508, Aug. 2003.

[12] M. T. Orchard and G. J. Sullivan, "Overlapped block motion compensation: An estimation-theoretic approach," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 693–699, Sep. 1994.

[13] H. Wantanabe and S. Singhal, "Windowed motion compensation," in *Proc. SPIE VCIP*, Nov. 1991, vol. 1605, pp. 582–589.

[14] Y. Nakaya and H. Harashima, "Motion compensation based on spatial transformations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, no. 3, pp. 339–356, Jun. 1994.

[15] C.-L. Huang and C.-Y. Hsu, "A new motion compensation method for image sequence coding using hierarchical grid interpolation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, no. 1, pp. 42–52, Feb. 1994.

[16] Y. Wang and O. Lee, "Active mesh – A feature seeking and tracking image sequence representation scheme," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 610–624, Sep. 1994.

[17] K. Hilman, H.-W. Park, and Y.-M. Kim, "Using motion compensated frame-rate conversion for the correction of 3:2 pulldown artifacts in video sequences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 896–877, Sep. 2000.

[18] G. Shen, B. Zeng, Y.-Q. Zhang, and M. Liou, "Transcoder with arbitrarily resizing capability," in *Proc. ISCAS*, May 2001, vol. 5, pp. 25–28.

[19] Y. Liang, L.-P. Chau, and Y.-P. Tan, "Arbitrary downsizing video transcoding using fast motion vector reestimation," *IEEE Signal Process. Lett.*, vol. 9, no. 11, pp. 352–355, Nov. 2002.

[20] M. H. Chan, Y. B. Yu, and A. G. Constantinides, "Variable size block matching motion compensation with application to video coding," *Proc. Inst. Elect. Eng.* , vol. 137, no. 4, pp. 205–212, Aug. 1990.

[21] J. Zhai, K. Yu, J. Li, and S. Li, "A low complexity motion compensated frame interpolation method," in *Proc. ISCAS*, May 2005, vol. 5, pp. 4927–4930.

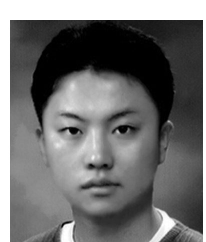
[22] M. C. Chen and J. A. N. Wilson, "Motion-vector optimization of control grid interpolation and overlapped block motion compensation using iterated dynamic programming," *IEEE Trans. Image Process.*, vol. 9, no. 7, pp. 1145–1157, Jul. 2000.

[23] B.-T. Choi, S.-H. Lee, and S.-J. Ko, "New frame rate up-conversion using bi-directional motion estimation," *IEEE Trans. Consum. Electron.*, vol. 46, no. 3, pp. 603–609, Aug. 2000.

[24] J. Y. A. Wang and E. H. Adelson, "Representing moving images with layers," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 625–638, Sep. 1994.

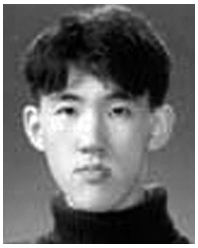
[25] P. Csillag and L. Borocszky, "Enhancement of video data using MC postprocessing techniques," in *Proc. ICASSP*, Apr. 1997, vol. 4, pp. 2897–2900.

[26] D. W. Kim, J. T. Kim, and I. H. Ra, "A new video interpolation technique based on motion-adaptive subsampling," *IEEE Trans. Consum. Electron.*, vol. 45, no. 3, pp. 782–786, Aug. 1999.



Byeong-Doo Choi (S'06) received the M.S. and B.S. degrees in electronics engineering from the Department of Electronics Engineering, Korea University, Seoul, Korea, in 2003, and 2001, respectively, where he is currently working toward the Ph.D. degree in multimedia processing and communication.

His research interests include video compression and image processing.



Jong-Woo Han (S'06) received the B.S. degree in electronics engineering from Korea University, Seoul, Korea, in 2005, where he is currently working toward the M.S. degree in electronics engineering.

His research interests are image and video processing and multimedia communications.

Chang-Su Kim (S'95–M'01–SM'05) received the B.S. and M.S. degrees in control and instrumentation engineering and the Ph.D degree in electrical engineering with a Distinguished Dissertation Award from Seoul national University (SNU), Seoul, Korea, in 1994, 1996, and 2000, respectively.

From 2000 to 2001, he was a Visiting Scholar with the Signal and Image Processing Institute, University of Southern California, Los Angeles, and a Consultant for InterVideo Inc., Los Angeles. From 2001 to 2003, he coordinated the 3-D Data Compression Group in National Research Laboratory for 3-D Visual Information Processing, SNU. From 2003 to 2005, he was an Assistant Professor in the Department of Information Engineering, Chinese University of Hong Kong. In September 2005, he joined the School of Electrical Engineering, Korea University, Seoul, Korea, as an Assistant Professor. His research topics include video and 3-D graphics processing and multimedia communications. He has published more than 90 technical papers in international journals and conferences.

Dr. Kim is an Editorial Board Member of *Journal of Visual Communication and Image Representation*.

Sung-Jea Ko (M'88–SM'97) received the B.S. degree in electronic engineering from Korea University, Seoul, Korea, in 1980. and the M.S. and Ph.D. 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, 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, University of Michigan-Dearborn. From 1981 to 1983, he was with Daewoo Telecom where he was involved in research and development on data communication systems. He has published more than 200 papers in journals and conference proceedings. He also holds over ten patents on data communication and video signal processing.

He is currently a Fellow of 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 *Journal of the Institute of Electronics Engineers of Korea* (IEEK) (1996), *Journal of Broadcast Engineering* (1996–1999), and the *Journal of the Korean Institute of Communication Sciences* (1997–2000). He has been an Editor of the *Journal of Communications and Networks* (1998–2000). He is 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) and the Best Paper Award from the IEEE Asia Pacific Conference on Circuits and Systems (1996).