BLOCK-BASED MOTION ESTIMATION USING THE PIXELWISE CLASSIFICATION OF THE MOTION COMPENSATION ERROR

Jun-Yong Kim[1], Rae-Hong Park[1] and Seungjoon Yang[2]

[1]Department of Electronic Engineering, Sogang University, Seoul, Korea
{jykimfv, rhpark}@sogang.ac.kr
[2]School of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology, Ulsan, Korea
syang@unist.ac.kr

ABSTRACT

In this paper, we propose block-based motion estimation (ME) algorithms based on the pixelwise classification of two different motion compensation (MC) errors: 1) displaced frame difference (DFD) and 2) brightness constraint constancy term (BCCT). Block-based ME has drawbacks such as unreliable motion vectors (MVs) and blocking artifacts, especially in object boundaries. The proposed block matching algorithm (BMA)-based methods attempt to reduce artifacts in object-boundary blocks caused by incorrect assumption of a single rigid (translational) motion. They yield more appropriate MVs in boundary blocks under the assumption that there exist up to three nonoverlapping regions with different motions. The proposed algorithms also reduce the blocking artifact in the conventional BMA, in which the overlapped block motion compensation (OBMC) is employed especially to the selected regions to prevent the degradation of details. Experimental results with several test sequences show the effectiveness of the proposed algorithms.

KEYWORDS

Block Matching Algorithm, Motion Estimation, Brightness Constancy Constraint, Pixel Classification, Overlapped Block Motion Compensation

1. INTRODUCTION

Motion estimation (ME) is one of the well-known methods for various video processing applications. Among a large number of ME approaches, block-based ME such as the block matching algorithm (BMA) [1,2] has been adopted in a number of international video coding standards including motion picture experts group (MPEG)-2/4 and H.26x [3-6]. Block-based ME is tractable and simple to implement with a lower complexity than pixel-based ME methods, thus has a large number of applications such as interlaced-to-progressive conversion (IPC) [7], and frame rate-up conversion (FRC) [8-10]. Block-based ME reduces the redundancy of the video sequence in the time domain, whereas the discrete cosine transform (DCT) reduces the redundancy in the spatial domain.

Generally, ME by the BMA with two successive video frames can be classified into two types: global and local. The global motion is occurred by camera motions such as translation, scale, and rotation, whereas the local motion is due to motions of individual objects contained in the video sequence. More than one object motion can be possible in some blocks, and thus the BMA

usually has difficulty in accurately finding these local (multiple) object motions in video sequences.

In block-based ME, an image is divided into a number of blocks of pixels with an assumption that each block has a single motion. The optimal motion vector (MV) of each block is found with the given ME criterion. MV in the block-based ME represents the displacement of the block in the current frame with respect to the corresponding block in the previous frame that has the smallest matching criterion, e.g., mean absolute difference (MAD) or mean square error (MSE).

Though the BMA is simple and thus applicable to various applications, it has drawbacks such as unreliable MVs and blocking artifacts which degrade visual quality of the processed video [11]. In detecting MVs using the BMA, the assumption that a block has a single (translational) motion is not likely to hold, especially in boundary blocks containing multiple objects with different motions. To reduce these problems, MPEG-4 visual considers object-based image processing. A video sequence is considered as a collection of a single or multiple video object planes (VOPs). A video object (VO) that constructs a video scene is segmented by shape, motion, and so on. However, it is not easy to accurately extract VOs from a video sequence.

Overlapped block MC (OBMC) recently has provided an effective extension of the conventional block MC (BMC) [12-20], in which blocks are overlapped with each other to reduce the blocking artifacts and residual errors in MC video. The complete estimate of the pixel value in the target block is decided as a linear combination of the previous estimate given by the MVs of the target block and the pixel values of neighboring blocks. The noncausal spatial dependency between the blocks leads to the iterative search for the optimal MV. To reduce the estimation complexity, modified noniterative OBMC schemes [16-17] have been proposed with the reasonable coding results.

In this paper, we propose ME algorithms that have the simplicity of the BMA by considering up to three objects in a block. They consider the motion compensation (MC) errors, and attempt to reduce blocking artifacts, especially in object-boundary blocks caused by incorrect assumption of a single (translational) object motion, providing better image quality with a more appropriate representation of MVs. Allowing more than one object in a block, the proposed BMA-based ME algorithms can obtain good results especially in boundary blocks with multiple motions. Also, the proposed algorithms use the OBMC for the selected region to reduce the blocking artifacts without the degradation of details.

The rest of the paper is organized as follows. In Section 2, we show the block diagram of the proposed BMA-based ME algorithms, followed by their detailed description. Experimental results and discussions are shown in Section 3. Finally, conclusions are given in Section 4.

2. PROPOSED BMA-BASED ME ALGORITHMS

In this section, we illustrate the block diagram of the proposed BMA-based ME algorithms and then describe the algorithmic procedures in detail. The first step of the proposed algorithms corresponds to the conventional BMA. Next, our algorithms are further refined by region-based processing of the MC error. The proposed algorithms reduce the MC error, especially blocking artifacts in boundary blocks caused by unreliable MVs. They also reflect the characteristics of a moving object such as covered and uncovered regions, and thus the refined motion vectors are more accurate and consistent to object's motions.

2.1. Block Diagram of the Proposed Algorithms

Figure 1 shows the block diagram of the proposed algorithms. They consist of four steps: blockwise ME/MC, pixelwise classification, region-based ME/MC, and MC confidence map. In Figure 1(a), the first step represents the blockwise ME/MC. Blockwise ME means the conventional BMA, in which a single translational object motion is assumed in a block. The MV detected by blockwise ME may be incorrect, in which the MC error is large, especially in boundary blocks that contain more than a single object or motion, yielding degraded reconstruction images.

In Figure 1(a), the image compensated by the blockwise MV is passed through the second step (pixelwise classification step). Generally, the intensity difference between two successive frames is small for the stationary background whereas large for moving objects. The large difference occurs near boundary pixels of a moving object. The region with large positive (negative) intensity difference values corresponds to the covered (uncovered) part of a moving object or vice versa. Thus, the proposed algorithm partitions a block into three types of non-overlapping regions (region with small frame differences, region with large positive frame differences, and region with large negative frame differences) based on the aspect of the frame difference. The accuracy of the motion estimation (ME) process is reduced if a block consists of more than two types of regions, for example, the uncovered region often has no information in the forward ME. The objective of the proposed algorithm is to have more reliable ME by considering differently each of these regions in the ME process. The second step divides each block into up to three nonoverlapping regions, which will be explained in detail in Section 2.2. The pixelwise classification is based on the MC error obtained in the first step. Considering this MC error, we can get better motion-compensated images. The output of the second step gives a sequence, in which each block contains up to three different regions.

The third step performs the region-based ME/MC. In Figure 1(b), the details of the third step in Figure 1(a) are shown. With up to three nonoverlapping regions in a block, the third step uses different processes for region sequences R_1, R_2, and R_3. For region sequences R_1 and R_2, the regionwise ME/MC, which represents the conventional BMA, is performed. Since region sequences R_1 and R_2 almost correspond to the object-boundary regions, another ME process separating the different motion regions finds more accurate MVs close to the true motion. For region sequence R_3, the overlapped MC is performed to reduce the blocking artifact.
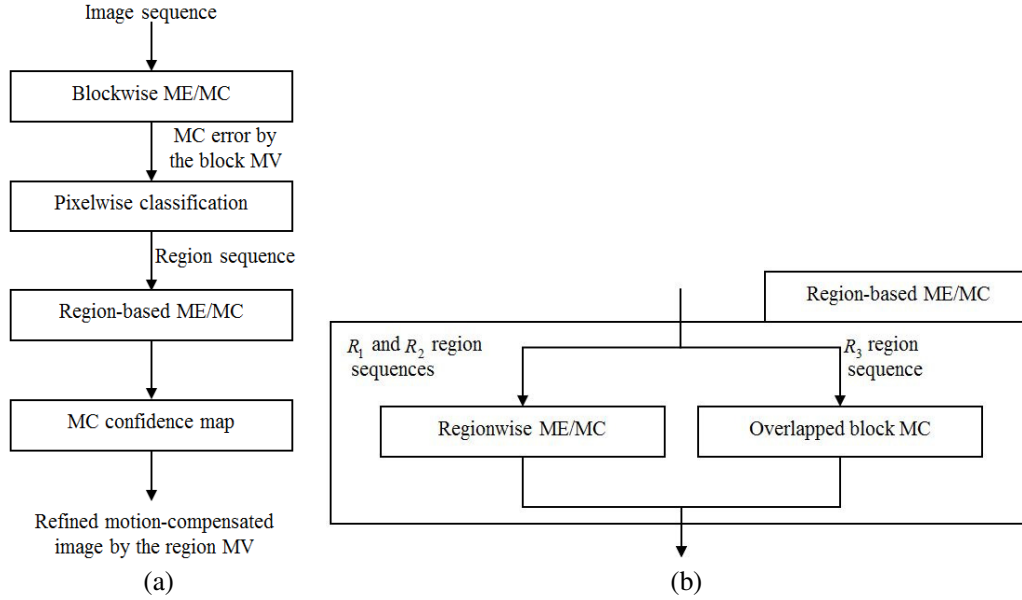
Figure 1. Block diagram of the proposed BMA-based ME algorithms: (a) Overall block diagram, and (b) Detailed block diagram of the region-based ME/MC.

Finally, the MC confidence map (MCCM) shows the confidence between the two MC processes mentioned above. The MC process among the overall process is repeated twice in the first and third steps, as illustrated in Figure 1(a). We consider the confidence between these two MC processes. Also, since the second step, i.e., the pixelwise classification, affects the third step, the MCCM considers the accuracy of the pixelwise classification. In consequence, we can obtain a more accurate and natural reconstructed image sequence.

Figure 2 shows the absolute MC error of the 5th frame of the Salesman sequence, in which darker pixels signify the pixels with large absolute ME errors. The MC error is large in boundaries of objects as expected. Also, regions having large motions produce large MC errors.

2.2. Description of the Proposed Algorithms

We propose two blockwise ME algorithms depending on the expression of the MC error. Let $f(x, y, t)$ indicate the intensity at pixel (x, y) in the t-th frame, with t representing the time (temporal) axis. The first proposed algorithm based on the displaced frame difference (DFD) error (hereafter called "the proposed algorithm (DFD)") is described as follows. In the first step of Figure 1, we find the approximate MV (u^*, v^*) of each block by the BMA based on the DFD defined by $f(x, y, t) - f(x - u, y - v, t - 1)$:

$$
\begin{aligned}
(u^*, v^*) &= \arg \min_{(u,v) \in S} \{ \xi(u, v, t) \} \\
&= \arg \min_{(u,v) \in S} \left\{ \sum_x \sum_y \left| f(x, y, t) - f(x - u, y - v, t - 1) \right| \right\},
\end{aligned}
\tag{1}
$$

where ξ represents the sum of absolute differences (SAD) as a matching measure and S denotes a set of candidate MVs in the search range.

The BMA assumes that all the pixels in each block have a single rigid (translational) motion. Thus a block containing more than one motion produces a large MC error. Especially, in object boundaries, most of the detected MVs yield large MC errors, which is illustrated in Figure 2. In the second step of the proposed algorithms in Figure 1(a), we try to find more appropriate MVs under the assumption that there exist up to three nonoverlapping regions with different motions in a block. This classification step uses the DFD between the original image and the compensated image. The region classification is done pixelwise as in the sliced BMA (SBMA) [21]:

$$
\begin{cases}
R_1 = \{(x, y) \mid d(x, y, t) > \alpha\} \\
R_2 = \{(x, y) \mid d(x, y, t) < -\alpha\} \\
R_3 = \{(x, y) \mid |d(x, y, t)| \le \alpha\},
\end{cases} \tag{2}
$$

where $d(x, y, t) = f(x, y, t) - f(x - u^*, y - v^*, t - 1)$ denotes the DFD with the detected MV (u^*, v^*) and α signifies the positive threshold. Note that the SBMA classifies pixels based on the frame difference (FD), whereas the proposed algorithm (DFD) classifies pixels based on the DFD. Dominant or secondary local regions near object boundaries show the DFD larger than α. Region R_3 represents regions with small intensity changes when motion occurs, usually in the interior of an object. Especially, regions R_1 and R_2 correspond to covered and uncovered regions, respectively.

Figure 3 shows block classification of (2), in which 8×8 blocks and $\alpha = 10$ are used. The block with more than one region is represented by black (gray level 0), whereas the block consisting of a single region is represented by white (gray level 255). Note that the black blocks represent blocks in object boundaries with large local motions.

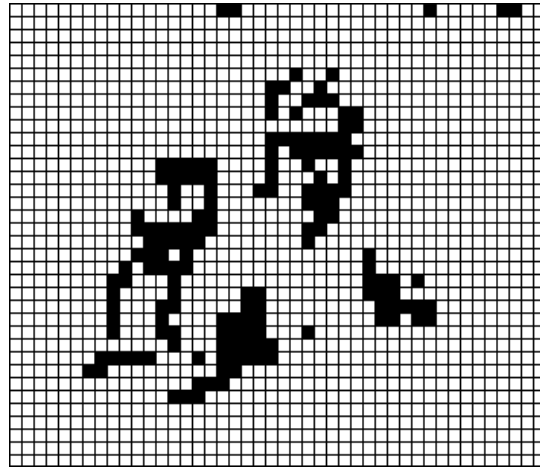Figure 2. Absolute MC error (Salesman sequence, 5th frame).

Figure 3. Block classification (Salesman sequence, 352×288, 5th frame, 8×8 blocks).

Figure 4(a) shows the pixels classified by (2) with $\alpha = 10$, in which selected regions R_1, R_2, and R_3 are represented by gray levels 0, 128, and 255, respectively. Most of the black blocks in Figure 3 are classified as R_1 and R_2 with up to three nonoverlapping regions, in which each

region has more accurate MVs. Figure 4(b) shows percentages of each region R_1, R_2, and R_3 as a function of the frame number of the Salesman sequence with $\alpha = 10$. Percentage values of R_1 and R_2 are indicated along the left vertical axis, whereas the percentage value of R_3 is indicated along the right vertical axis. We observe that most of the regions (94-99%) are included in region R_3, and the regions (1-6%) sensitive to motions are included in regions R_1 and R_2. Note that regions R_1 and R_2 have a similar percentage in most of the frames, which also can be confirmed in Figure 4(a) with the two regions (gray levels 0 and 128) adjacent to each other. The influence of the threshold value on the performance will be discussed in Section 2.2.

The MC errors are utilized to distinguish the characteristics of an object and motions, giving meaningful classification and thus good results. For region sequences R_1 and R_2, more reliable MVs of a block are obtained by separately applying the regionwise ME/MC, which is equivalent to the conventional BMA, to each region. Note that the refined ME method for each region uses the same block size and search range as the first ME. However, we need to give a different process for region sequence R_3. Since the region sequence R_3 nearly corresponds to regions with small intensity changes when motion occurs, another ME/MC hardly gives results different from those of the first step in Figure 1(a). That is, another ME process nearly does not give the refined MVs for this region. Thus, for this region we use the OBMC [16-17] without a new ME process to reduce the computational complexity and the blocking artifact. The OBMC of the other region sequences R_1 and R_2, corresponding to covered and uncovered regions, respectively, may degrade the details due to the interaction of the neighboring blocks. Thus, the OBMC applied to the specific region, i.e., R_3 region only, reduces the blocking artifact caused by the block-based ME/MC and simultaneously reduces the degradation caused by the OBMC [19-20].
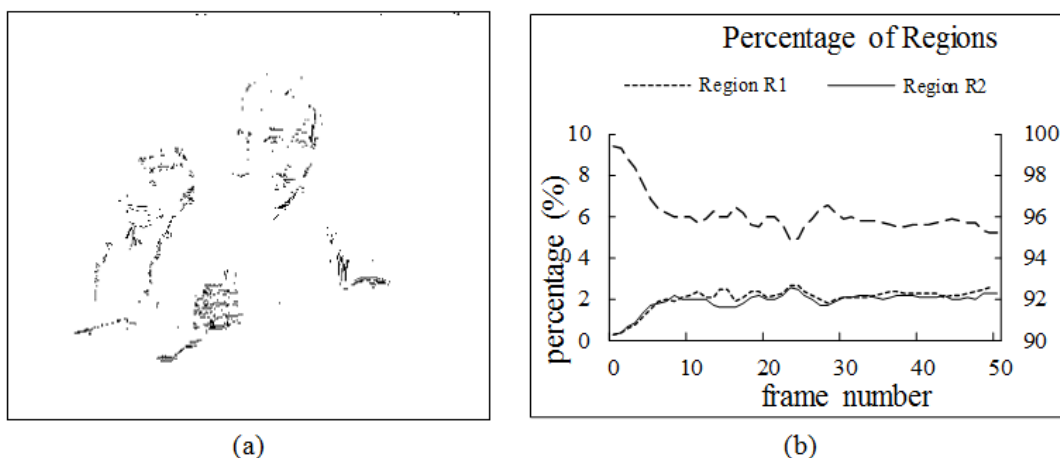

Figure 4. Region classification (Salesman sequence, 352×288): (a) Region image (5th frame), and (b) Region ratio (50 frames).

Finally, the MCCM shows the confidence between the above two MC processes in the first and third steps, as illustrated in Figure 1(a). The separate ME/MC processes for the each region give more accurate and reliable MVs and pixel values. Thus, the classification of the region in the second step affects the results of the region-based ME/MC in the third step. The optimal selection of threshold α is not easy, and the isolated pixel may be found for the specific threshold value. These isolated pixels may degrade the results of the regionwise ME/MC. To reduce the effect of the threshold decision and isolated pixels in the region, the MCCM is defined by

$$\text{MCCM}(x, y) = \begin{cases} f(x-\hat{u}, y-\hat{v}, t-1), & \left|\hat{d}(x, y, t)\right| \le \left|d(x, y, t)\right| \\ f(x-u^*, y-v^*, t-1), & \left|\hat{d}(x, y, t)\right| > \left|d(x, y, t)\right|, \end{cases} \tag{3}$$

where $\hat{d}(x, y, t) = f(x, y, t) - f(x-\hat{u}, y-\hat{v}, t-1)$ denotes the DFD with the refined MV $(\hat{u}, \hat{v})$ obtained by the region-based ME in the third step. In consequence, we obtain a more accurate and natural reconstructed image sequence.

For a small motion, applying the Taylor series expansion to the DFD gives

$$
\begin{aligned}
f(x, y, t) &- f(x-u, y-v, t-1) \\
&= f(x, y, t) - \left[f(x, y, t) - \frac{\partial f}{\partial x} u - \frac{\partial f}{\partial y} v - \frac{\partial f}{\partial t} \right] \\
&= \frac{\partial f}{\partial x} u + \frac{\partial f}{\partial y} v + \frac{\partial f}{\partial t} \\
&\cong f_x u + f_y v + f_t,
\end{aligned}
\tag{4}
$$

where f_x, f_y, and f_t denote the partial derivatives of f with respect to x, y, and t, respectively. We can generalize our algorithm under the brightness constancy constraint assumption (BCCT) using (4), yielding the proposed algorithm (BCCT).

In the proposed algorithm (BCCT), assuming that the pixel intensity is constant along the motion trajectory, we extend the assumption that intensity of each region in a block is preserved along the motion trajectory. Using the BCCT in (4), we first detect the approximate MV (u^*, v^*) of the block, as described in the first step of the proposed algorithm (DFD). With (u^*, v^*), the MC errors can be expressed in terms of the BCCT in (4) and each pixel in a block is classified by (2).

Figure 5 shows the peak signal to noise ratio (PSNR) of the reconstructed image (5th frame of the Salesman sequence) by four ME methods as a function of α. Note that the three algorithms (two proposed algorithms and SBMA) have the similar characteristics. Also, the proposed algorithms give a higher PSNR than the SBMA algorithm for most of the threshold values. It is possible and advantageous to make the threshold α adaptive to block features such as edge information or local variances
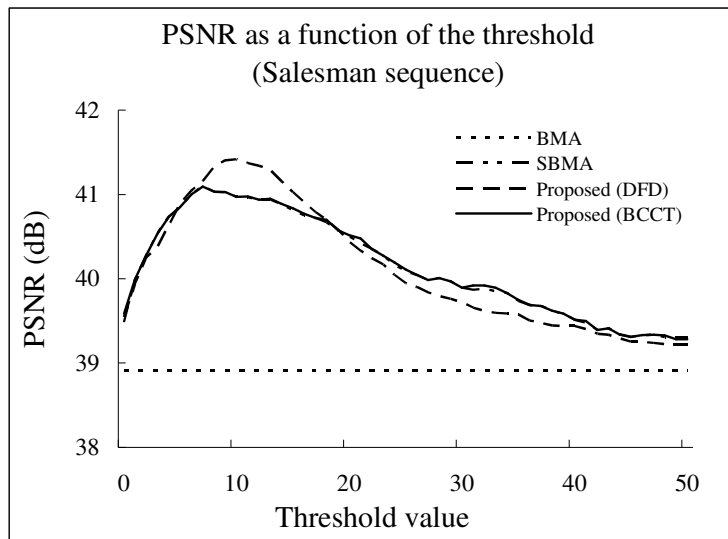
Figure 5. Threshold selection in the proposed algorithms (Salesman sequence, 352×288, 5th frame).

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, we show the effectiveness of the proposed algorithms by computer simulation with several test image sequences. Figure 6 shows test image sequences used in experiments to compare the performance of the proposed algorithms with that of the conventional methods including the SBMA [21]. Figure 6(a) shows the 2nd frame of the 352×240 Football sequence consisting of 50 frames, Figure 6(b) shows the 30th frame of the 352×288 Calendar sequence consisting of 50 frames, and Figure 6(c) shows the 12th frame of the 352×288 Salesman sequence consisting of 50 frames. Figures 6(a) and 6(b) have a lot of local motions, whereas Figure 6(c) contains less local motions. Figure 6(b) has more details than Figures 6(a) and 6(c).
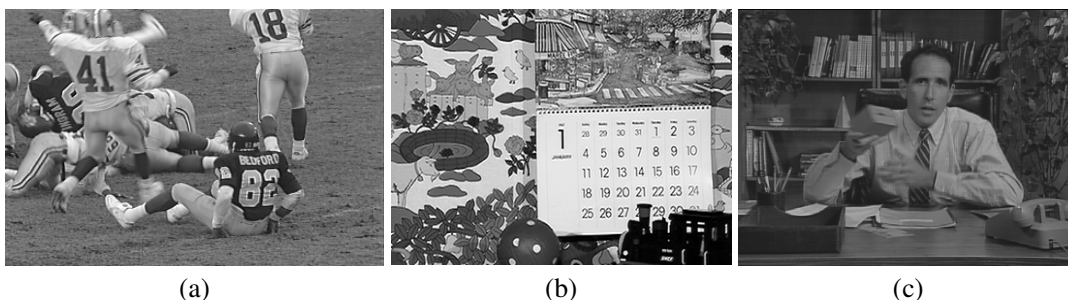

(a)   (b)   (c)

Figure 6. Image sequences used in experiments: (a) Football sequence (352×240, 2nd frame), (b) Calendar sequence (352×288, 30th frame), and (c) Salesman sequence (352×288, 12th frame).

For performance comparison of each algorithm, Figures 7, 8, and 9 show the absolute MC error, the reconstruction image, and the enlarged part of the reconstruction image of Figures 6(a), 6(b), and 6(c), respectively. Note that only the proposed algorithm (DFD) is compared, in which the proposed algorithm (BCCT) suitable for small motions gives worse results than the proposed algorithm (DFD). Figures 7(a), 7(b), and 7(c) show the absolute MC errors by the BMA, the

SBMA [21], and the proposed algorithm (DFD), respectively, in which 8×8 blocks and 31×31 search area are assumed. Figures 7(d), 7(e), and 7(f) illustrate the reconstruction images of Figure 6(a) by the BMA, the SBMA, and the proposed algorithm (DFD), respectively. Figures 7(g), 7(h), and 7(i) show the enlarged images of Figures 7(d), 7(e), and 7(f), respectively. Similarly, Figures 8 and 9 are illustrated to compare the performance of three algorithms for Figures 6(b) and 6(c), respectively.

In Figures 7(a), 7(b), and 7(c), the absolute MC error is illustrated and the darker region represents the larger magnitude. Most of large absolute MC errors in Figures 7(a), 7(b), and 7(c) are found at boundaries of objects, and absolute MC errors of the proposed algorithm (DFD) in Figure 7(c) are the least of three absolute MC error images in Figures 7(a), 7(b), and 7(c). Blocking artifacts near boundaries of objects are reduced in Figure 7(f), compared to those in Figures 7(d) and 7(e). For easy comparison, we enlarge a portion, which shows large blocking artifacts, of the reconstruction images. Figure 7(i) by the proposed algorithm (DFD) shows the least blocking artifacts among Figures 7(g), 7(h), and 7(i), and especially the numbers and name on the back and the line of clothes are clearer than those in Figures 7(g) and 7(h).

(a)   (b)   (c)
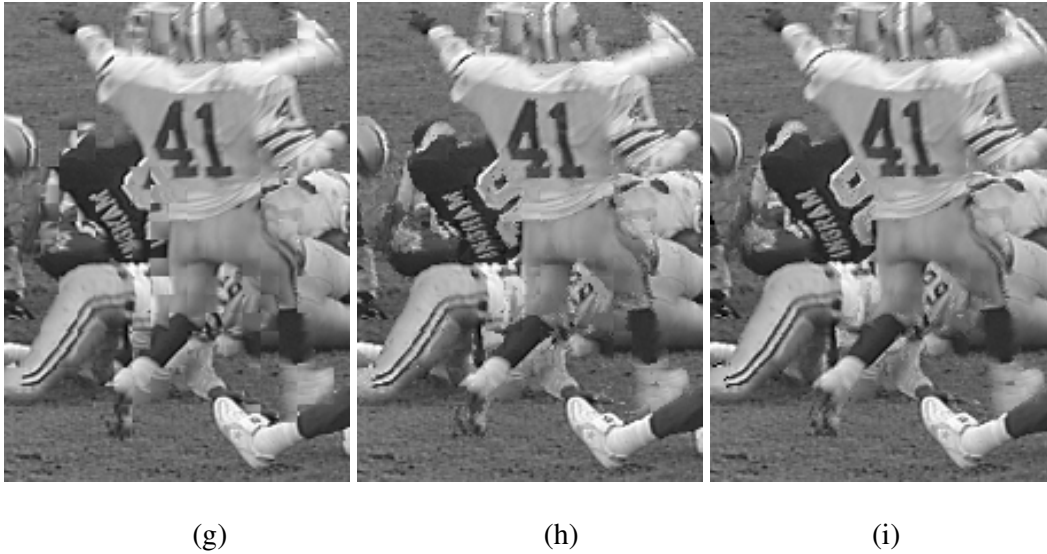

(d)   (e)   (f)

(g)   (h)   (i)
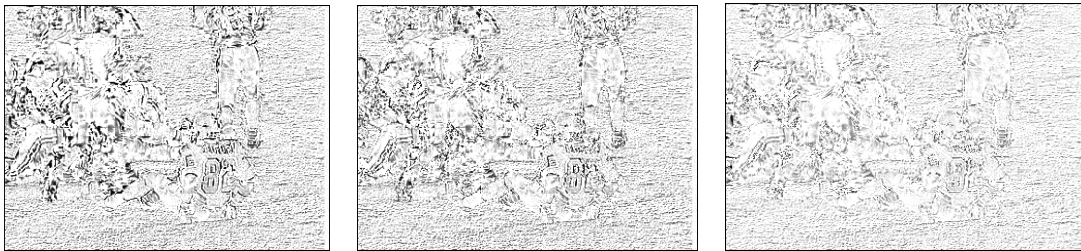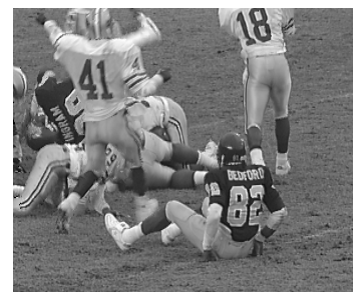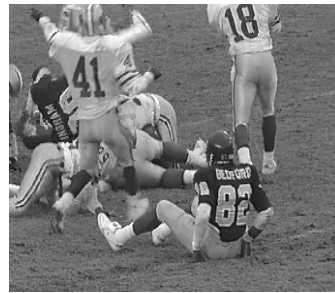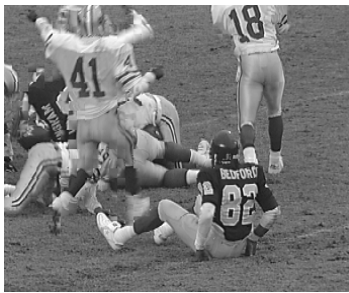
Figure 7. Football sequence (352×240, 2nd frame): (a), (b), and (c) Absolute MC errors by the BMA, the SBMA, and the proposed algorithm (DFD), respectively, (d), (e), and (f) Images reconstructed by the BMA, the SBMA, and the proposed algorithm (DFD), respectively, (g), (h), and (i) Enlarged regions of (d), (e), and (f), respectively.

As in Figures 7(a), 7(b), and 7(c), most of large absolute MC errors in Figures 8(a), 8(b), and 8(c) are found at boundaries of objects, and absolute MC errors in Figure 8(c) are smaller than those in Figures 8(a) and 8(b). In Figures 8(g), 8(h), and 8(i), which are the enlarged images of Figures 8(d), 8(e), and 8(f), respectively, blocking artifacts in Figure 8(i) are less than those in Figures 8(g) and 8(h). Since Figure 6(b) has a number of details, the overall reduction effects of absolute MC errors in Figures 8(g), 8(h), and 8(i) are less significant, compared with Figures 7(g), 7(h), and 7(i). However, numbers in the calendar in Figure 8(i) are certainly clearer than those in Figures 8(g) and 8(h).

(a)   (b)   (c)

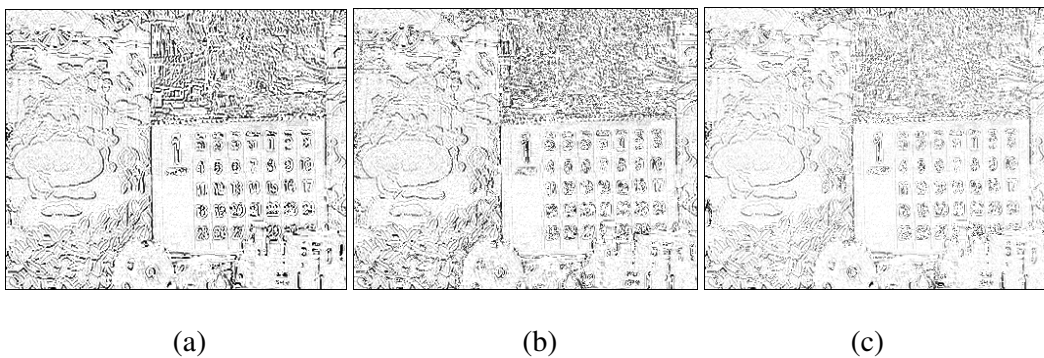(d)   (e)   (f)

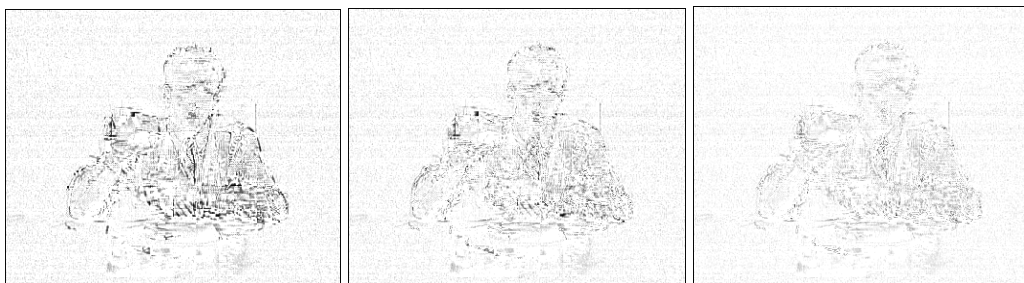(g)   (h)   (i)

Figure 8. Calendar sequence (352×288, 30th frame): (a), (b), and (c) Absolute MC errors by the BMA, the SBMA, and the proposed algorithm (DFD), respectively, (d), (e), and (f) Images reconstructed by the BMA, the SBMA, and the proposed algorithm (DFD), respectively, (g), (h), and (i) Enlarged regions of (d), (e), and (f), respectively.

In Figures 9(a), 9(b), and 9(c), we observe that the absolute MC errors in Figure 9(c) are less noticeable than those in Figures 9(a) and 9(b). Blocking artifacts in Figure 9(i) are smaller than those in Figures 9(g) and 9(h). Note that the hand of a man, a tape, and the object behind the man in Figure 9(i) are clear.

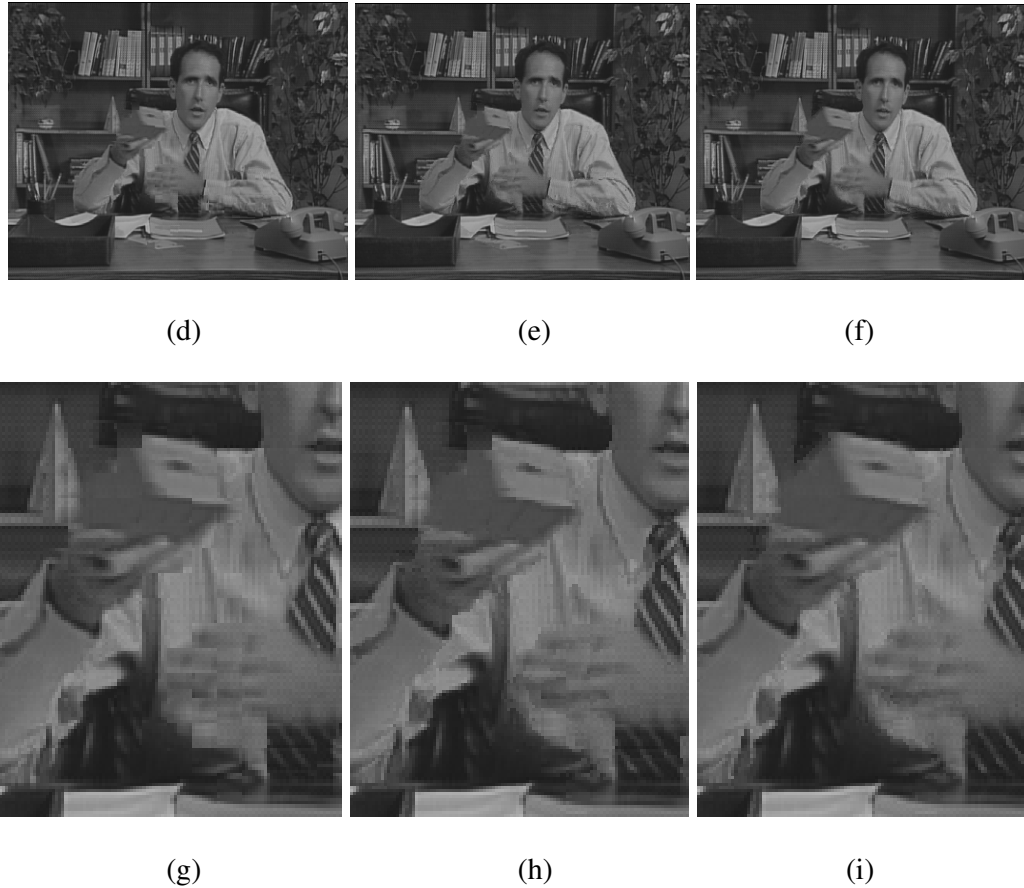(a)   (b)   (c)

(d)   (e)   (f)

(g)   (h)   (i)

Figure 9. Salesman sequence (352×288, 12th frame): (a), (b), and (c) Absolute MC errors by the BMA, the SBMA, and the proposed algorithm (DFD), respectively, (d), (e), and (f) Images reconstructed by the BMA, the SBMA, and the proposed algorithm (DFD), respectively, (g), (h), and (i) Enlarged regions of (d), (e), and (f), respectively.

Figure 10 shows the PSNR comparison of four ME algorithms (BMA, SBMA, and two proposed algorithms). Figures 10(a), 10(b), and 10(c) are the PSNR graphs for the Football, Calendar, and Salesman sequences, respectively. In all of Figures 10(a), 10(b), and 10(c), the proposed algorithms give higher PSNRs than the conventional algorithms. Especially, the PSNR of the proposed algorithm (DFD) shows the best results among the algorithms considered for comparison. Note that the PSNR of the proposed algorithm (DFD) is higher than that of the proposed algorithm (BCCT). The proposed algorithm (BCCT) is suitable for video sequences with small and simple motions because it is derived using the first-order Taylor series expansion
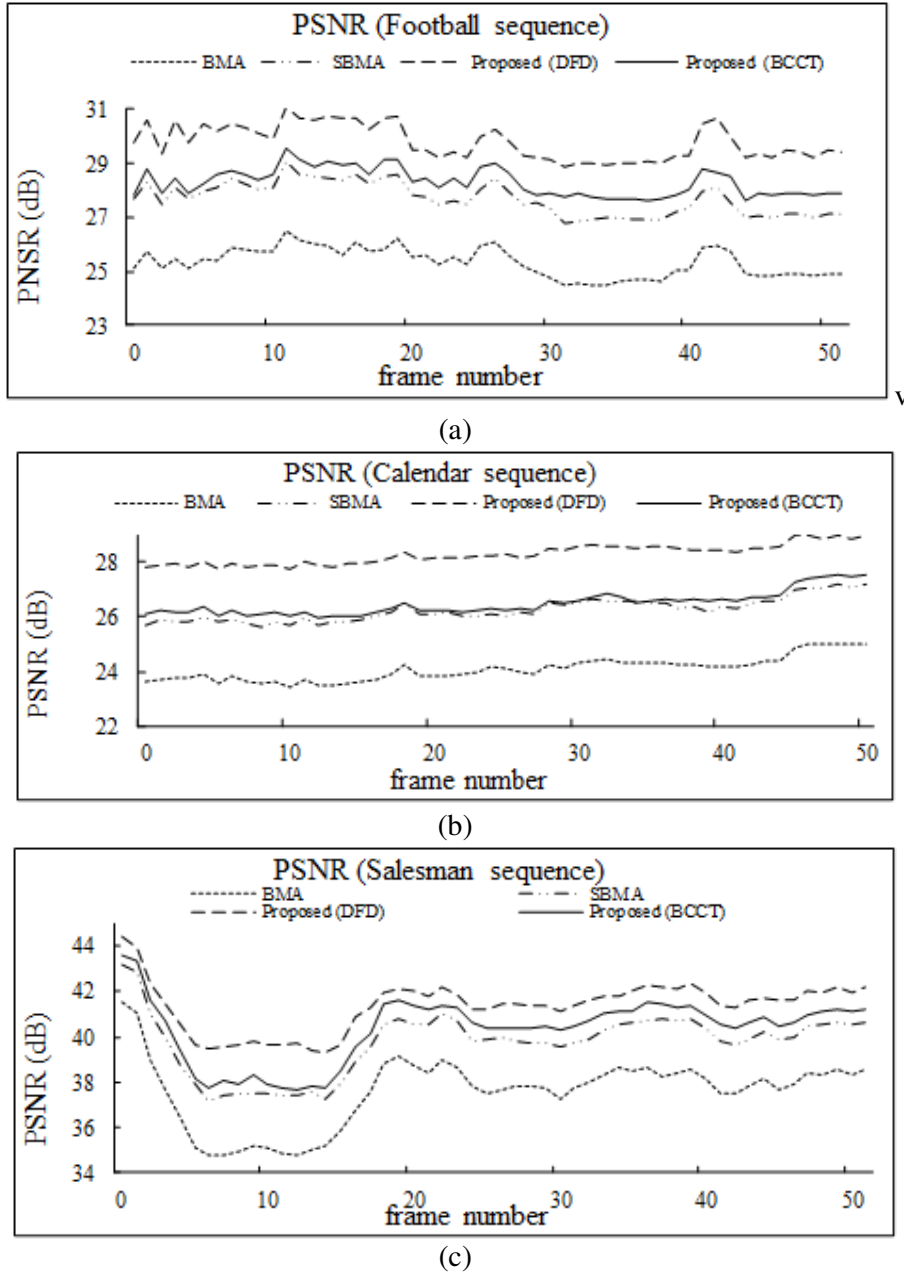
(a)

(b)

(c)

Figure 10. Performance comparison in terms of the PSNR: (a) Football sequence (352×240, 50 frames), (b) Calendar sequence (352×288, 50 frames), and (c) Salesman sequence (352×288, 50 frames).

Figure 11 illustrates the PSNR comparison of three regions R_1, R_2, and R_3, by the BMA and the proposed algorithm (DFD), for the Football sequence. The pixels belong to regions R_1 and R_2 are regions near the boundaries of the moving objects. Figures 11(a), 11(b), and 11(c) show the PSNR graphs of regions R_1, R_2, and R_3, respectively. The PSNR difference between the BMA and the proposed algorithm (DFD) is large in Figures 11(a) and 11(b). This fact results from the regionwise ME/MC, i.e., the proposed algorithm (DFD) separately considers the regions that have

13

the high and similar motions. And, the PSNR difference between the BMA and the proposed algorithm (DFD) is small in Figure 11(c), compared with Figures 11(a) and 11(b). Since region R_3 consists of pixels that have small motions, the MVs of these areas have similar values in both the BMA and the proposed algorithm (DFD). However, the OBMC used in this region improves the quality of the reconstruction of region R_3. The improved quality of each region gives better quality over the overall reconstruction image.
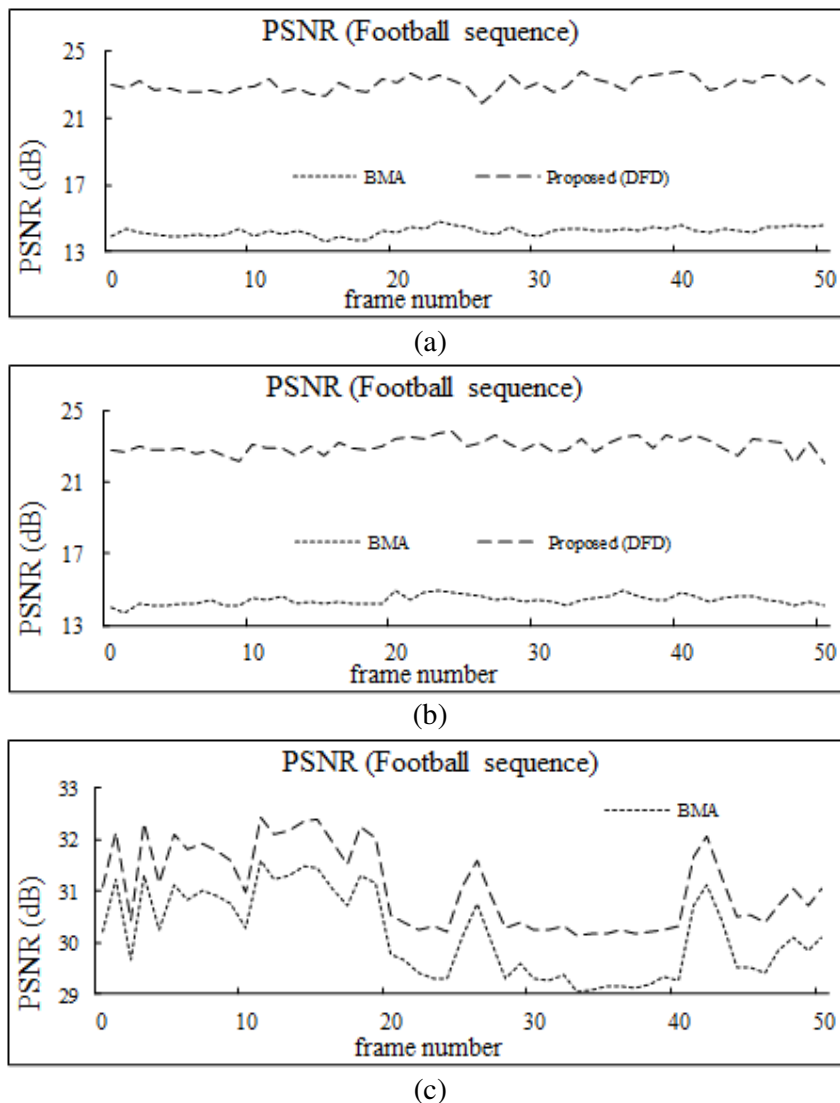

(a)

(b)

(c)

Figure 11. PSNR comparison of each region (Football sequence, 352×240, 50 frames): (a) Region R_1, (b) Region R_2, and (c) Region R_3.

Figure 12 shows the PSNR graph, in which the performance enhancement by the MCCM defined in (3) is illustrated for the Football sequence. Elimination of isolated pixels gives higher PSNRs. Note that MCCM reduces the effects of the isolated pixels, thus improving the accuracy of the ME process.
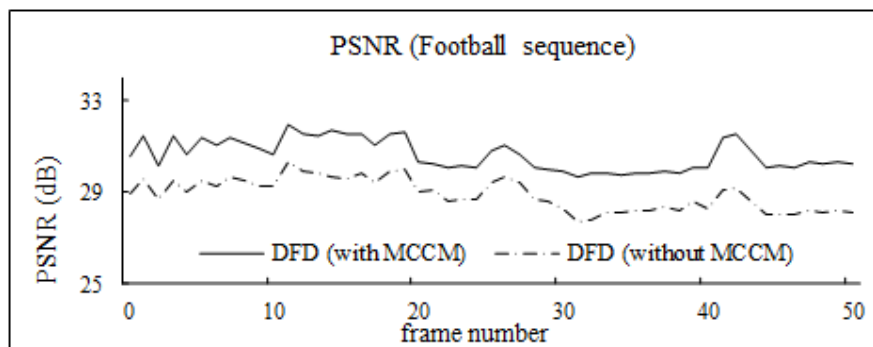
Figure 12. PSNR comparison with effect of the isolated pixels (Football sequence, 352×240, 50 frames).

Table 1 compares the computation time of the conventional BMA, the SBMA, and the proposed algorithm (DFD) for three experimental images in Figure 6. The relative computation time listed in Table 1 is defined by the ratio of the average computation time of each method per frame with respect to that of the BMA, on the PC with 3.01GHz Pentium IV (1GB RAM, Visual C++ compiler). Note that the proposed algorithm (DFD) requires extra time for the pixelwise classification step. Further research will focus on the reduction of the computational load of the proposed algorithm (DFD).

Table 1. Comparison of the relative computation time.

	BMA	SBMA	Proposed (DFD)	Remarks
Football	1.00	1.34	2.21	352×240, 50 frames
Calendar	1.00	1.36	2.27	352×288, 50 frames
Salesman	1.00	1.26	2.21	352×288, 50 frames

Table 2 compares the performance of the proposed algorithm (DFD), with different block size and search area, for three test images in Figure 6. Note that the performance is represented in terms of the average PSNR of the sequence of 50 frames each.

Table 2. PSNR comparison of the proposed algorithm (DFD) for different block size and search area (unit: dB).

Block size	Search area	Football	Calendar	Salesman
8×8	17×17	30.421	27.417	40.394
8×8	31×31	31.047	28.090	40.912
16×16	17×17	27.580	25.351	38.875
16×16	31×31	28.162	25.665	39.207
Remarks		352×240, 50 frames	352×288, 50 frames	352×288, 50 frames

4. CONCLUSIONS

In this paper, we propose two BMA-based ME algorithms based on pixelwise classification of the MC error: 1) DFD and 2) BCCT. We attempt to reduce the blocking artifacts especially near the boundaries of objects. The proposed algorithms classify each block into up to three nonoverlapping regions by the MC error of the reconstruction image. The larger the absolute MC error, the worse the quality of the resulting image. Thus, we classify the pixels with large absolute MC error as significant elements to improve the ME/MC performance, and classify those pixels. The classified regions have more accurate MVs, thus we can obtain the improved results. Also, for the pixels with small absolute MC error the OBMC is used to reduce the blocking artifact. The proposed algorithm (DFD) gives better results than the proposed algorithm (BCCT) for test sequences with large motions.

Simulation results with several test sequences show the improved performance of the proposed algorithms, especially in object boundaries. Especially, the regions having large MC errors are reconstructed with relatively high PSNRs. Also, using up to three nonoverlapping regions, the proposed algorithms can effectively segment objects and background. They can be effectively applied to accurate ME for video-based applications. Further research will be focused on the extension of the proposed algorithms to color image sequences.

ACKNOWLEDGEMENTS

REFERENCES

[1]   F. H. Jamil, A. Chekima, R. R. Porle, O. Ahmad, N. Parimon, BMA performance of video coding for motion estimation, in Proc. 2012 Third Int. Conf. Intelligent Systems Modeling and Simulation (ISMS), 2012, pp. 287-290.

[2]   H.M. Musmann, P. Pirsch, H.J. Gravoert, Advances in picture coding, Proc. IEEE 73 (1985) 523–548.

[3]   ISO/IEC 13818-2, Generic Coding of Moving Pictures and Associated Audio Information, 1995.

[4]   ISO/IEC 14496-2, Information Technology-Coding of Audio-Visual Objects Part 2: Visual, 1999.

[5]   ISO/IEC JTC1/SC29/WG11 MPEG93/N457, MPEG-2 Test Model Version 5, 1993.

[6]   ITU-T Recommendation H.263, Video Coding for Low Bitrate Communications, 1995.

[7]   S. Yang, Y.-Y. Jung, Y.H. Lee, R.-H. Park, Motion compensation assisted motion adaptive interlaced-to-progressive conversion, IEEE Trans. Circuits Syst. Video Technol. 13 (2004) 1138–1148.

[8]   K.A. Bugwadia, E.D. Petajan, N.N. Puri, Progressive-scan rate up-conversion of 24/30 source materials for HDTV, IEEE Trans. Consum. Electron. 42 (1996) 312–321.

[9]   S.-J. Kang, S. Yoo, and Y. H. Kim, Dual motion estimation for frame rate up-conversion, IEEE Trans. Circuits Syst. Video Technol. 20 (2010) 1909–1914.

[10] B.-W. Jeon, G.-I. Lee, S.-H. Lee, R.-H. Park, Coarse-to-fine frame interpolation for frame rate-conversion using pyramid structure, IEEE Trans. Consum. Electron. 49 (2003) 499–508.

[11] S. Fujiwara, A. Taguchi, Motion-compensated frame rate up-conversion based on block matching algorithm with multi-size blocks, in Proc. 2005 Int. Symp. Intelligent Signal Processing and Communication Systems, 2005, pp. 353-356.

[12] B. Tao, M.T. Orchard, A parametric solution for optimal overlapped block motion compensation, IEEE Trans. Image Proc. 10 (2001) 341–350.

[13] M.T. Orchard, G.J. Sullivan, Overlapped block motion compensation: An estimation-theoretic approach, IEEE Trans. Image Proc. 3 (1994) 693–699.

[14] Y.-W. Chen, W.-H. Peng, Parametric OBMC for pixel-adaptive temporal prediction on irregular motion sampling grids, IEEE Trans. Circuits Syst. Video Technol. 22 (2012) 113–127.

[15] S. Nogaki, M. Ohta, An overlapped block motion compensation for high quality motion picture coding. in: Proc. IEEE ISCS, 1992, pp. 184–187.

[16] J.K. Su, R.M. Mersereau, Motion estimation methods for overlapped block motion compensation, IEEE Trans. Image Processing 9 (2000) 1509-1521.

[17] J.K. Su, R.M. Mersereau, Non-iterative rate-constrained motion estimation for OBMC, in: Proc. IEEE ICIP, 1997, pp. 33–36.

[18] R. Rajagopalan, E. Feig, M.T. Orchard, Motion optimization of ordered blocks for overlapped block motion compensation, IEEE Trans. Circuits Syst. Video Technol. 8 (1998) 119-123.

[19] B.-D. Choi, J.-W. Han, C.-S. Kim, S.-J. Ko, Motion-compensated frame interpolation using bilateral motion estimation and adaptive overlapped block motion compensation, IEEE Trans. Circuits Syst. Video Technol. 17 (2007) 407–416.

[20] W. Woo, A. Ortega, Overlapped block disparity compensation with adaptive windows for stereo image coding, IEEE Trans. Circuits Syst. Video Technol. 10 (2000) 194–200.

[21] E.D. Sciascio, C. Guarangnella, Object oriented motion estimation by sliced-block matching algorithms, in: Proc. Int. Conf. Pattern Recognition, 2000, pp. 857–860.

Authors

Jun-Yong Kim received the B.S. degree from Sogang University in 2004. He is working toward the M.S. degree in electronic engineering from Sogang University. His current research interests are image processing and resolution enhancement.

Rae-Hong Park was born in Seoul, Korea, in 1954. He received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, Korea, in 1976 and 1979, respectively, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1981 and 1984, respectively. In 1984, he joined the faculty of the Department of Electronic Engineering, School of Engineering, Sogang University, Seoul, Korea, where he is currently a Professor. In 1990, he spent his sabbatical year as a Visiting Associate Professor with the Computer Vision Laboratory, Center for Automation Research, University of Maryland at College Park. In 2001 and 2004, he spent sabbatical semesters at Digital Media Research and Development Center, Samsung Electronics Co., Ltd. (DTV image/video enhancement). His current research interests are computer vision, pattern recognition, and video communication. He served as Editor for the Korea Institute of Telematics and Electronics (KITE) *Journal of Electronics Engineering* from 1995 to 1996. Dr. Park was the recipient of a 1990 Post-Doctoral Fellowship presented by the Korea Science and Engineering Foundation (KOSEF), the 1987 Academic Award presented by the KITE, and the 2000 Haedong Paper Award presented by the Institute of Electronics Engineers of Korea (IEEK), the 1997 First Sogang Academic Award, and the 1999 Professor Achievement Excellence Award presented by Sogang University.

Seungjoon Yang received the B.S. degree from Seoul National University, Seoul, Korea, in 1990, and the M.S. and Ph.D. degrees from the University of Wisconsin, Madison, in 1993 and 2000, respectively, all in electrical engineering. He was with the Digital Media Research and Development Center, Samsung Electronics Company, Ltd., Seoul, from September 2000 to August 2008. He is currently with the School of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology, Ulsan, Korea. His current research interests include image processing, estimation theory, and multi-rate systems.
Dr. Yang received the Samsung Award for the Best Technology Achievement of the Year in 2008 for his work on the premium digital television platform project.