Adaptive Local Illumination Change Compensation Method for H.264/AVC-Based Multiview Video Coding

Jae-Ho Hur, Sukhee Cho, and Yung-Lyul Lee, *Senior Member, IEEE*

***Abstract*—In multiview video, illumination changes can occur due to imperfect camera calibration and variations of the camera position and direction. These characteristics can cause performance degradation in multiview video coding (MVC) that uses inter-view prediction by referring to the pictures obtained from the neighboring views. In order to overcome this problem, an adaptive local illumination change compensation method is proposed. In the proposed method, we compensate for the illumination changes in the macroblock (MB) unit between the current picture and the reference picture under the assumption that the DC component in the MB is influenced by the local illumination changes. By using the proposed method, the compression ratio in the multiview video coding was increased, and a 0.1$\sim$0.6 dB peak signal-to-noise ratio (PSNR) improvement was obtained compared with the case where the proposed method was not used.**

***Index Terms*—Block matching algorithm (BMA), H.264/AVC, illumination change compensation, inter-view prediction, mean-removed sum of absolute differences (MRSAD), multiview video coding (MVC).**

I. Introduction

MULTIVIEW video coding (MVC), which is being standardized in the joint video team (JVT) of the ITU-T video coding experts group (VCEG) and ISO/IEC moving picture experts group (MPEG), is expected to become a new video coding standard for the realization of future video applications such as 3D-TV and free viewpoint video [1]. The MVC group in the JVT has chosen the H.264/AVC [2]-based MVC method that was proposed by [3] as the MVC reference model, since this method showed better coding efficiency than H.264/AVC simulcast coding and the other methods that were submitted in response to the call for proposals made by the MPEG [4].

However, when illumination changes occur between pictures in the view-temporal direction, the displacement vector estimation (DVE), which is the motion vector estimation or disparity vector estimation in MVC, cannot be accurately performed, so that the error of the displacement vector (DV) and the amounts of the residual signals may be increased and, consequently, the coding efficiency may be decreased. The reason why this phenomenon is taken so seriously is that the performance of MVC is affected not only by local illumination changes due to the variations in the viewing position and direction, but also by the global signal changes due to imperfect camera calibration.

In order to compensate for these illumination changes, several methods [5]–[11], [15] have been proposed. Fundamentally, the algorithms in [5]–[7] use two factors, *viz.* the scale and offset, to compensate for the discrepancy induced by the illumination changes. In order to calculate these factors, the least square method is used in [5] and [6], and the variance and mean values are used in [7]. Also, the 2-D Lloyd max algorithm is used to efficiently represent the scale and offset factors in [7]. However, these algorithms require high computational complexity while achieving relatively low coding efficiency, making it difficult to use them in MVC.

In this paper, we propose a macroblock (MB)-based adaptive illumination change compensation (MBAIC) method for the luminance components in MVC. The illumination change is compensated for in the MB unit between the current picture and the reference picture under the assumption that the DC component in the MB is influenced by the local illumination change. The proposed method is applied to displacement vector estimation/compensation by using a modified block matching measure, that is, the so-called mean-removed sum of absolute differences (MRSAD).

The rest of the paper is organized as follows. In Section II, the H.264/AVC-based MVC reference model is introduced briefly. In Section III, the detailed algorithm of the proposed method is described and the integration of the proposed method into the MVC reference model is described. In Section IV, the performance of the proposed method is compared to that of the MVC reference model not using the proposed method by analyzing the various experimental results.

II. H.264/AVC-Based MVC

As already mentioned in the previous section, the H.264/AVC-based MVC method that is chosen as a reference model for the standardization of MVC has shown significant coding efficiency. Fig. 1 depicts an example of the H.264/AVC-based MVC structure, in which there are eight parallel views. As shown in Fig. 1, this structure utilizes the hierarchical B pictures, which not only improves the coding efficiency, but also provides temporal scalability. This structure can be divided into three kinds of picture sets, i.e., the picture

Manuscript received December 16, 2006. This work was supported in part by Seoul R&DB program (11098). This paper was recommended by Guest Editor Y. He.

J.-H. Hur and Y.-L. Lee are with the Computer Engineering Department, Sejong University, DMS Laboratory, Seoul 143-747, Korea (e-mail: jhhur@dms.sejong.ac.kr; yllee@sejong.ac.kr).

S. Cho is with the Electronics and Telecommunication Research Institute, Broadcasting Systems Research Group, Taegeon 305-707, Korea (e-mail: shee@etri.re.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2007.903774

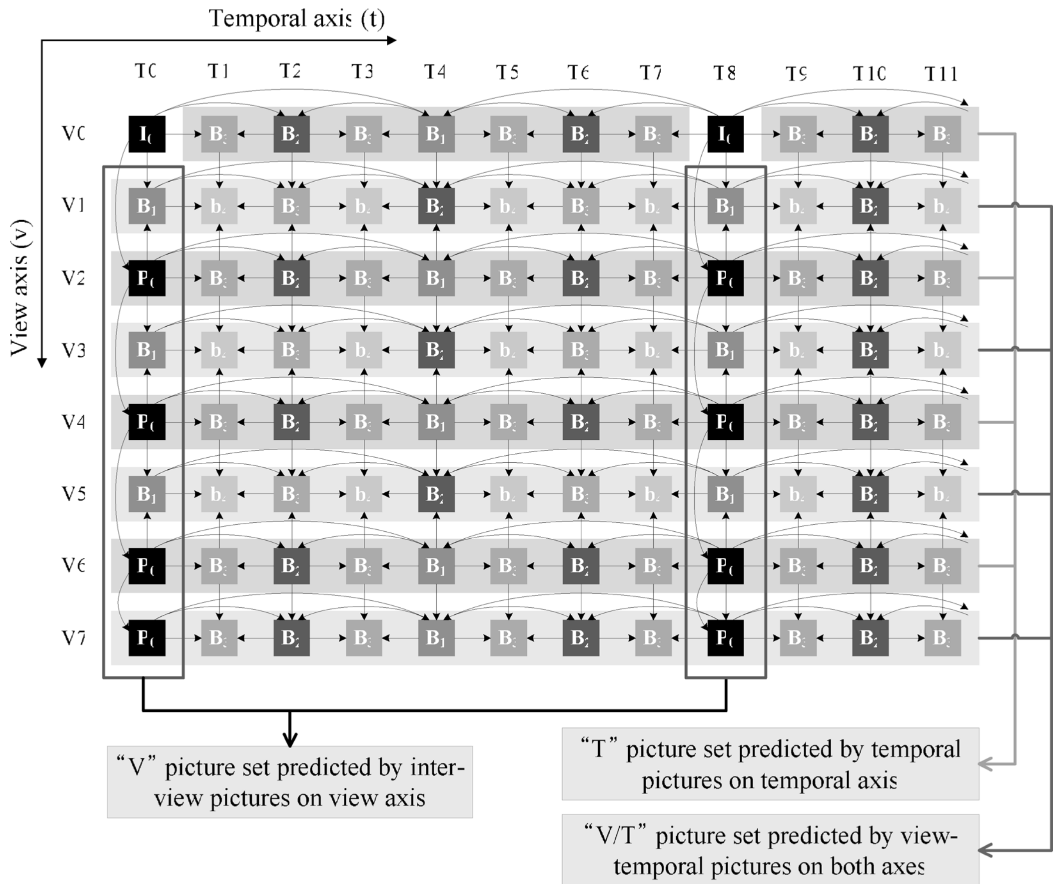


Fig. 1. H.264/AVC-based MVC structure on 1-D parallel arrangement with eight cameras.

set predicted by the inter-view pictures on the view axis, the picture set predicted by the temporal pictures on the temporal axis, and the picture set predicted by the view-temporal (spatio-temporal) pictures on the view and temporal axes. In this structure, the $\mathrm{I_k}$ and $\mathrm{P_k}$ pictures are only used at random access points. The $\mathrm{B_k}$ pictures are used between the I and P pictures that are random access points and are used in all other positions except at these random access points, where the subscript k means the temporal decomposition level. The $\mathrm{B_1}$ pictures on the temporal axis T0 are predicted spatially, the $\mathrm{B_k}$ pictures on the view axes V0, V2, V4, and V6 are predicted temporally, and the $\mathrm{B_k}$ pictures on the view axes V1, V3, V5, and V7 are predicted temporally and spatially. Our proposed method is integrated into the MVC structure shown in Fig. 1.

III. MB-Based Adaptive Illumination Change Compensation (MBAIC) Method

Basically, the proposed illumination change compensation method is based on a hybrid video coding structure and is performed adaptively in the MB unit. As shown in Fig. 2(a), the proposed encoder performs adaptive DVE using MB-based illumination change compensation and the prediction of the difference value of an illumination change (DVIC), which is an offset factor used for the MB-based illumination change compensation. By using adaptive DVE, the illumination change can be compensated for locally. Also, DVIC is efficiently coded through the proposed prediction method of DVIC. As shown in Fig. 2(b), the proposed decoder performs adaptive displacement and illumination change compensation.

A. Adaptive DVE Using the MB-Based Illumination Change Compensation

In the H.264/AVC-based MVC reference model, the sum of absolute differences (SAD) is used as a measure for the DVE. The current frame is denoted by $C(i,j)$ with spatial coordinates (i,j), where i and j are defined by the index of the column and row, respectively, and the reference frame is denoted by $R(i,j)$. The SAD calculation of the $S \times T$ blocks, *viz.* 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, and 4×4, is performed as follows:

$$\mathrm{SAD}^{(m,n)}(x,y) = \sum_{i=m}^{m+S-1} \sum_{j=n}^{n+T-1} |C(i,j) - R(i+x, j+y)| \tag{1}$$

where (x,y) represents a candidate DV and (m,n) represents the start position of each MB in the current frame.

When the illumination change occurs between different views in the multiview video, the correlation between these different views may be decreased. In this case, the conventional SAD measure cannot be used appropriately for inter-view prediction. In order to overcome this problem, an alternative measure is needed instead of SAD.

The proposed illumination change compensation method assumes that the discrepancy caused by the illumination change is equal to the discrepancy of the DC component between the

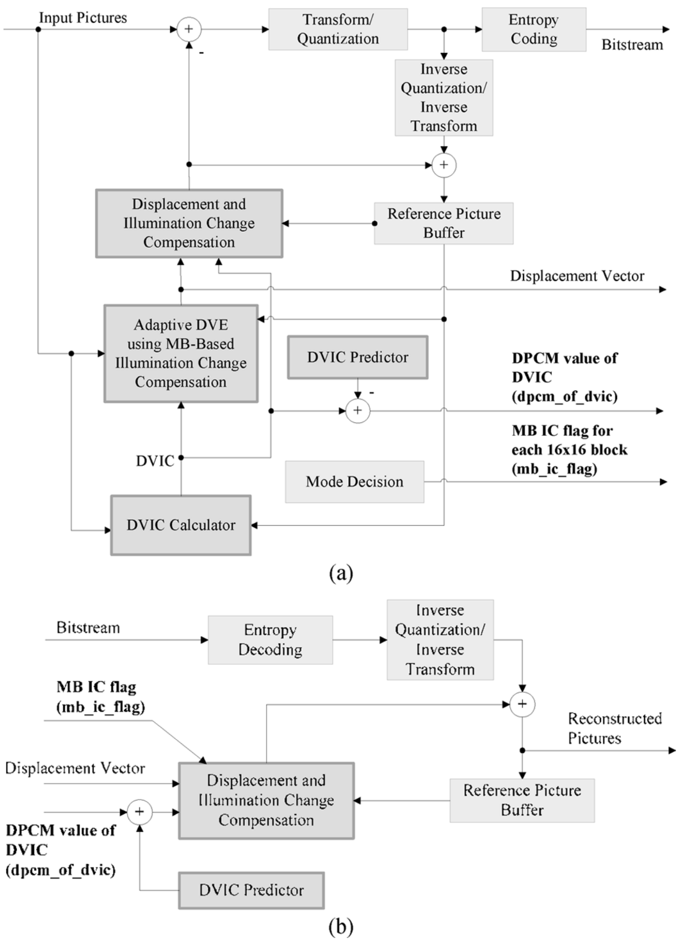


Fig. 2. Block diagram of the proposed MB-based adaptive illumination change compensation method. (a) Encoded. (b) Decoder.

current block and the reference block. In order to mitigate the discrepancy during DVE, the mean values of the current block and the reference block are subtracted from each other. By using this concept, the MRSAD is used as a measure for the DVE as follows:

$$M_C^{(m,n)} = \frac{1}{S \times T} \sum_{i=m}^{m+S-1} \sum_{j=n}^{n+T-1} C(i,j)$$
$$M_R^{(m,n)}(x,y) = \frac{1}{S \times T} \sum_{i=m+x}^{m+x+S-1} \sum_{j=n+y}^{n+y+T-1} R(i,j) \quad (2)$$

$$\begin{aligned}\mathrm{MRSAD}^{(m,n)}(x,y) &= \sum_{i=m}^{m+S-1} \sum_{j=n}^{n+T-1} \left| \left\{ C(i,j) - M_C^{(m,n)} \right\} \right. \\ &\quad \left. - \left\{ R(i+x,j+y) - M_R^{(m,n)}(x,y) \right\} \right| \\ &= \sum_{i=m}^{m+S-1} \sum_{j=n}^{n+T-1} \left| \{C(i,j) - R(i+x,j+y)\} \right. \\ &\quad \left. - \left\{ M_C^{(m,n)} - M_R^{(m,n)}(x,y) \right\} \right| \quad (3)\end{aligned}$$

where $M_C^{(m,n)}$ and $M_R^{(m,n)}(x,y)$ are the average value for all pixels in the current block and the reference block, respectively.

In the implementation of the proposed method, S and T are both set to 16, because the use of a variable block-size MRSAD does not give rise to any improvement of the coding efficiency which would offset the resulting increase of the computational complexity. In (2) and (3), $M_C^{(m,n)}$ is calculated only one time, while $M_R^{(m,n)}(x,y)$ is calculated for every candidate reference block pointed to by (x,y). The DV $(\tilde{x},\tilde{y})$ minimizing the MRSAD is computed, and the DVIC is calculated in the MB unit as follows:

$$\mathrm{DVIC}^{(m,n)} = \mathrm{Round}\left[M_C^{(m,n)} - M_R^{(m,n)}(\tilde{x},\tilde{y}) \right]. \quad (4)$$

The DVIC means an offset factor which is used for the compensation of the illumination change in the MB unit. By using the DVIC, the displacement and illumination change-compensated residual signal in the (m,n)th MB of the current frame is calculated as follows:

$$\begin{aligned}E^{(m,n)}(k,l) &= \left\{ C(m+k,n+l) - M_C^{(m,n)} \right\} \\ &\quad - \left\{ R(m+k+\tilde{x},n+l+\tilde{y}) - M_R^{(m,n)}(\tilde{x},\tilde{y}) \right\} \\ &= \{C(m+k,n+l) - R(m+k+\tilde{x},n+l+\tilde{y})\} \\ &\quad - \left\{ M_C^{(m,n)} - M_R^{(m,n)}(\tilde{x},\tilde{y}) \right\} \\ &= \{C(m+k,n+l) - R(m+k+\tilde{x},n+l+\tilde{y})\} \\ &\quad - \mathrm{DVIC}^{(m,n)} \quad (5)\end{aligned}$$

where $E^{(m,n)}(k,l)$ represents the displacement and illumination change-compensated residual signal to be transformed and quantized in the (m,n)th MB of the current frame, and $0 \le k \le 15$, $0 \le l \le 15$.

For the adaptive local illumination compensation, the SAD-based conventional coding procedure or the proposed MRSAD-based coding procedure can be used for coding the current MB, for which a 1-bit flag (mb_ic_flag) per MB is needed to signal to the decoder whether the proposed coding procedure is used or not. On the decoder side, if mb_ic_flag is 0, the illumination change compensation is not performed for the current MB, i.e., the conventional decoding procedure is performed. Otherwise, the displacement and illumination change compensation is performed as follows:

$$\begin{aligned}\widehat{C}^{(m,n)}(k,l) &= \left\{ \widehat{E}^{(m,n)}(k,l) + R(m+k+\tilde{x},n+l+\tilde{y}) \right\} \\ &\quad + \left\{ M_C^{(m,n)} - M_R^{(m,n)}(\tilde{x},\tilde{y}) \right\} \\ &= \left\{ \widehat{E}^{(m,n)}(k,l) + R(m+k+\tilde{x},n+l+\tilde{y}) \right\} \\ &\quad + \mathrm{DVIC}^{(m,n)} \quad (6)\end{aligned}$$

where $\widehat{E}^{(m,n)}(k,l)$ represents the displacement and illumination-compensated residual signal which is reconstructed by inverse quantization and inverse transform in the (m,n)th MB of the current frame, $\widehat{C}^{(m,n)}$ represents the reconstructed (m,n)th MB of the current frame, and $0 \le k \le 15$, $0 \le l \le 15$.

To enable the illumination change compensation in the decoder, the DVIC should be transmitted to the decoder in an appropriate way. For encoding the DVIC, the prediction error of

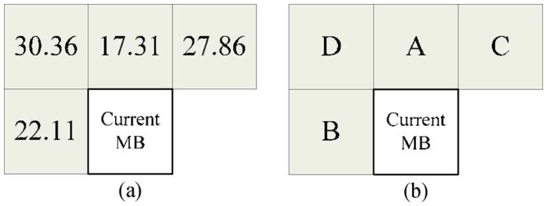


Fig. 3. (a) Average value of the MSE between the DVICs of the current MB and each neighbor MB for various test sequences. (b) Causal neighbor MBs used for the prediction of the current DVIC.

the DVIC is coded by means of the differential pulse code modulation (DPCM) of the DVIC and its prediction value, which is calculated by using the DVICs of the neighboring blocks of the current block. In the decoder, the DVIC is decoded through inverse DPCM with the prediction value which is calculated in the same way as that in the encoder. The prediction process of DVIC is explained in the next subsection.

B. Prediction Process of DVIC

The DVIC corresponds to a local illumination change affecting the DC component in each MB. Since the area over which the local illumination-change occurs is usually larger than the area of one MB, the magnitude of the local illumination change of the current MB may have a high correlation with that of the neighboring MBs. Therefore, in order to reduce the number of bits required to encode the DVIC, the DVIC is coded by DPCM with the predictor ($\mathrm{pred}_{\mathrm{DVIC}}$) obtained from the DVIC of the causal neighbor MBs, as shown in Fig. 3.

The prediction process was developed by analyzing the average value of the mean squared error (MSE) between the DVICs of the current MB and each neighboring MB for various test sequences. Fig. 3(a) shows the MSE average value of the DVICs between the current MB and neighboring MBs. Therefore, the neighbor block search order for the DVIC prediction for each neighboring block in Fig. 3(b) becomes A, B, C, and D, according to the smaller MSE average value of the neighboring blocks in Fig. 3(a). If there are no neighboring MBs whose reference index is equal to that of the current MB even if the mb_ic_flags of the neighboring MBs are equal to 1, median filtering is performed to calculate the value of $\mathrm{pred}_{\mathrm{DVIC}}$ by using three neighboring MBs (MBs A, B and C shown in Fig. 3(b)).

The prediction process of DVIC is as follows:

- Step 1) If the upper MB, A, of the current MB in Fig. 3(b) has a DVIC and its reference index is the same as that of the current MB, $\mathrm{pred}_{\mathrm{DVIC}}$ is set to the DVIC of MB A and the process is finalized.
- Step 2) If the left MB, B, of the current MB has a DVIC and its reference index is the same as that of the current MB, $\mathrm{pred}_{\mathrm{DVIC}}$ is set to the DVIC of MB B and the process is finalized.
- Step 3) If the upper-right MB, C, of the current MB has a DVIC and its reference index is the same as that of the current MB, $\mathrm{pred}_{\mathrm{DVIC}}$ is set to the DVIC of MB C and the process is finalized.
- Step 4) If the upper-left MB, D, of the current MB has a DVIC and its reference index is the same as that of the current MB, $\mathrm{pred}_{\mathrm{DVIC}}$ is set to the DVIC of MB D and the process is finalized.
- Step 5) If MB A, MB B and MB C have their own DVICs, then median filtering is performed with the three neighboring MBs. $\mathrm{pred}_{\mathrm{DVIC}}$ is set to the result of the median filtering and the process is finalized.
- Step 6) $\mathrm{pred}_{\mathrm{DVIC}}$ is set to 0.

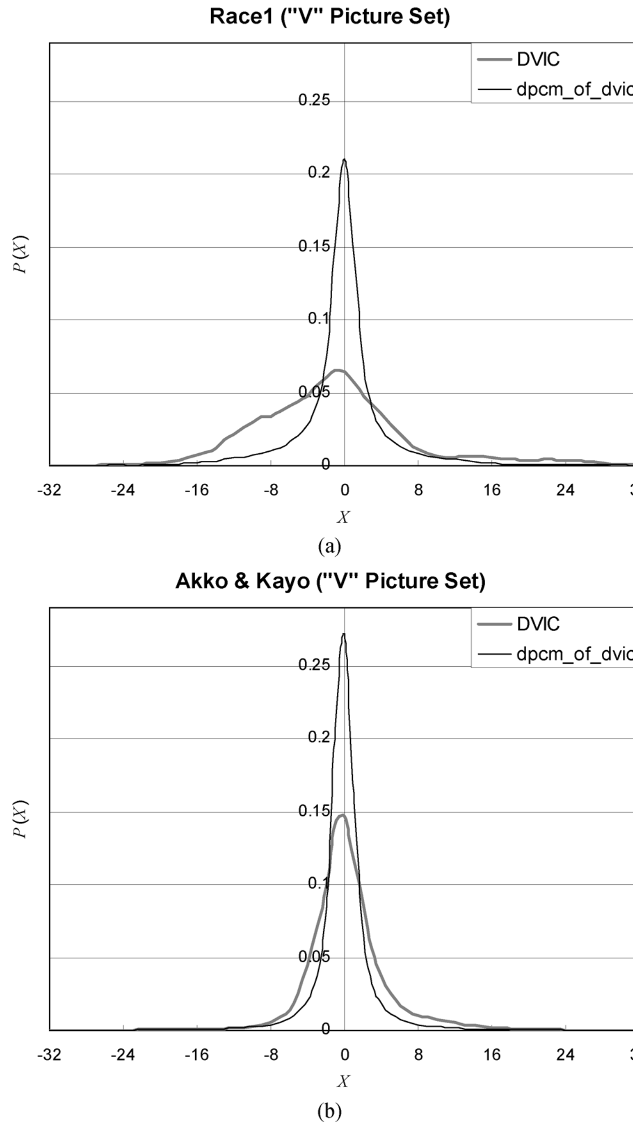


Fig. 4. Probability distribution of the DVIC and the prediction error of DVIC (dpcm_of_dvic) in the "V" picture set. (a) Race1. (b) Akko & Kayo.

After obtaining the prediction value of DVIC through the above steps, $\mathrm{pred}_{\mathrm{DVIC}}$ is subtracted from the DVIC and its prediction error (dpcm_of_dvic = DVIC-$\mathrm{pred}_{\mathrm{DVIC}}$) is coded by the entropy coding method.

Probability distribution of the DVIC and the prediction error of DVIC (dpcm_of_dvic) in the "V" picture set for Race1 and Akko & Kayo sequences are plotted in Fig. 4. As shown in Fig. 4, the distribution of dpcm_of_dvic is concentrated at zero, as if it were a zero-mean Laplacian distribution, while the distribution of DVIC is irregular. Therefore, the prediction error of DVIC is coded instead of the DVIC in this paper.

C. Integration of the MVC Reference Model and the Proposed MBAIC

To evaluate its performance, the proposed MBAIC method is integrated into the MVC reference model described in Section II. The proposed method is only applied to each 16×16 block for a luminance component (Y), even though the variable block-size adaptive illumination change compensation, *viz.* 16×8, 8×16, 8×8, 8×4, 4×8, and 4×4, in (2) and (3) could be applied. This is because additional information, such as the illumination compensation indication bit per block (not MB) and dpcm_of_dvic bits per block (not MB), would have to be transmitted to the decoder, with the result that no coding gain would be achieved, due to the increase in the number of these information-signaling bits. Also, there is very little coding gain when the proposed method is applied to the chrominance components (U and V).

The proposed method is applied to the three MB modes existing in H.264/AVC, *viz.* Inter 16×16 mode (in P or B slice), Skip mode (in P slice), and Direct 16×16 mode (in B slice). In the case of Skip mode, the DV is derived by the existing derivation process of the motion vector predictor specified in [2] and the DVIC is calculated using the average value of the DVICs of the neighboring MBs. This idea was proposed by Thomson/USC in [11], and was submitted to the MVC standardization of JVT as a joint proposal with our algorithm. In the case of Direct 16×16 mode, the DV that is derived by the existing derivation process of Direct mode [12] is used and the DVIC is calculated by (4), in which the mean value of the reference block is calculated with the derived DV. As a reference, the spatial Direct mode is used for our experiments.

In order to compensate for the local illumination changes adaptively, mb_ic_flag is needed for each Inter 16×16 and Direct 16×16 block mode. On the encoder side, the rate-constrained coder control [13] is used to decide whether the local illumination change compensation is needed for each MB or not. Although the current MB is determined as "Direct Skip", which is a skipped MB mode in the B slice, mb_ic_flag and dpcm_of_dvic need to be transmitted to the decoder for the current MB. mb_ic_flag and dpcm_of_dvic are encoded using the context-based adaptive binary arithmetic coding (CABAC) method [14].

IV. Experimental Results

To demonstrate the effectiveness of the proposed MBAIC method, several experimental results are provided. First, the performance of DVE is demonstrated by comparing DVE using SAD with DVE using MRSAD. Next, the rate-distortion (RD) performance is shown for each picture set depicted in Fig. 1 and for various MVC test sequences, in which the reference model (RM) with the proposed MBAIC method is compared to the RM with variable block-size adaptive illumination change compensation, the RM without any illumination change compensation, and the RM with the weighted prediction (WP) [15] adopted in H.264/AVC. Finally, a comparison of the encoding time complexity is shown.

For our experiments, we used the JSVM 6.5 reference software[1] that supports MVC mode and performed the experiments on various MVC test sequences with different image properties and camera arrangements of the YUV 4:2:0 format, as shown in Table I. Basically, the proposed method is performed for temporal prediction as well as inter-view prediction, and the coding tools specified in the Main and FRExt profiles, such as the deblocking filter, CABAC, and adaptive block transform (4×4 or 8×8 size) are used for the encoding simulation. For more information regarding the test conditions, refer to [3] and [16].

[1]From CVS Repository of JSVM.

TABLE I
Test Sequences With Different Image Properties and Different Camera Arrangements for Multiview Video Coding

Test Squences	Image Property	Camera Arrangement
Ballroom, Exit	640x480, 25fps (rectified)	8 cameras with 20cm spacing; 1D/parallel
Uli	1024x768, 25fps (non-rectified)	8 cameras with 20cm spacing; 1D/parallel convergent
Race1	640x480, 30fps (non-rectified)	8 cameras with 20cm spacing; 1D/parallel
Flamenco2	640x480, 30fps (non-rectified)	5 cameras with 20cm spacing; 2D/parallel (Cross)
Breakdancers	1024x768, 15fps (non-rectified)	8 cameras with 20cm spacing; 1D/arc
Rena	640x480, 30fps (rectified)	100 cameras with 5cm spacing; 1D/parallel
Akko & Kayo	640x480, 30fps (non-rectified)	100 cameras with 5cm horizontal and 20 cm vertical spacing; 2D array (3 vertical × 5 horizontal views)

A. Comparison of the Performances of DVE Using SAD and MRSAD

To compare the DVE using SAD with the DVE using MRSAD, the disparity vectors are extracted in the 16×16 block unit by using the T0/V2 picture of Fig. 1 as the reference frame and the T0/V3 picture of Fig. 1 as the current frame in the "Race1" sequence, as shown in Fig. 5(a) and (b), respectively. Fig. 5(c) and (d) depict two kinds of disparity vector field. Fig. 5(c) and (d) depict the disparity vector fields extracted by the DVEs using SAD and MRSAD, respectively. The disparity vector field depicted in Fig. 5(c) is very irregular and rough in some local areas. As we predicted, it can be seen that the local illumination change deteriorates the encoding performance by disturbing the DVE using SAD. On the other hand, the disparity vector field depicted in Fig. 5(d) is relatively smoother than that in Fig. 5(c) and looks more like a real disparity vector. Through these results, we can intuitively conclude that the DVE using MRSAD can improve the disparity vector coding by using a more smooth disparity vector field, and enhance the overall encoding performance by compensating for the local illumination changes when they occur.

B. Comparison of R-D Performance According to Picture Set

As already mentioned in Section II, the pictures coded by the H.264/AVC-based MVC reference model can be classified into three picture sets, as shown in Fig. 1, *viz.* the "V", "T", and "V/T" picture sets. The three R-D curves per sequence are shown in Fig. 6 for the RM and the RM with the proposed MBAIC (RM+MBAIC). In the "V" picture set, the highest coding gain

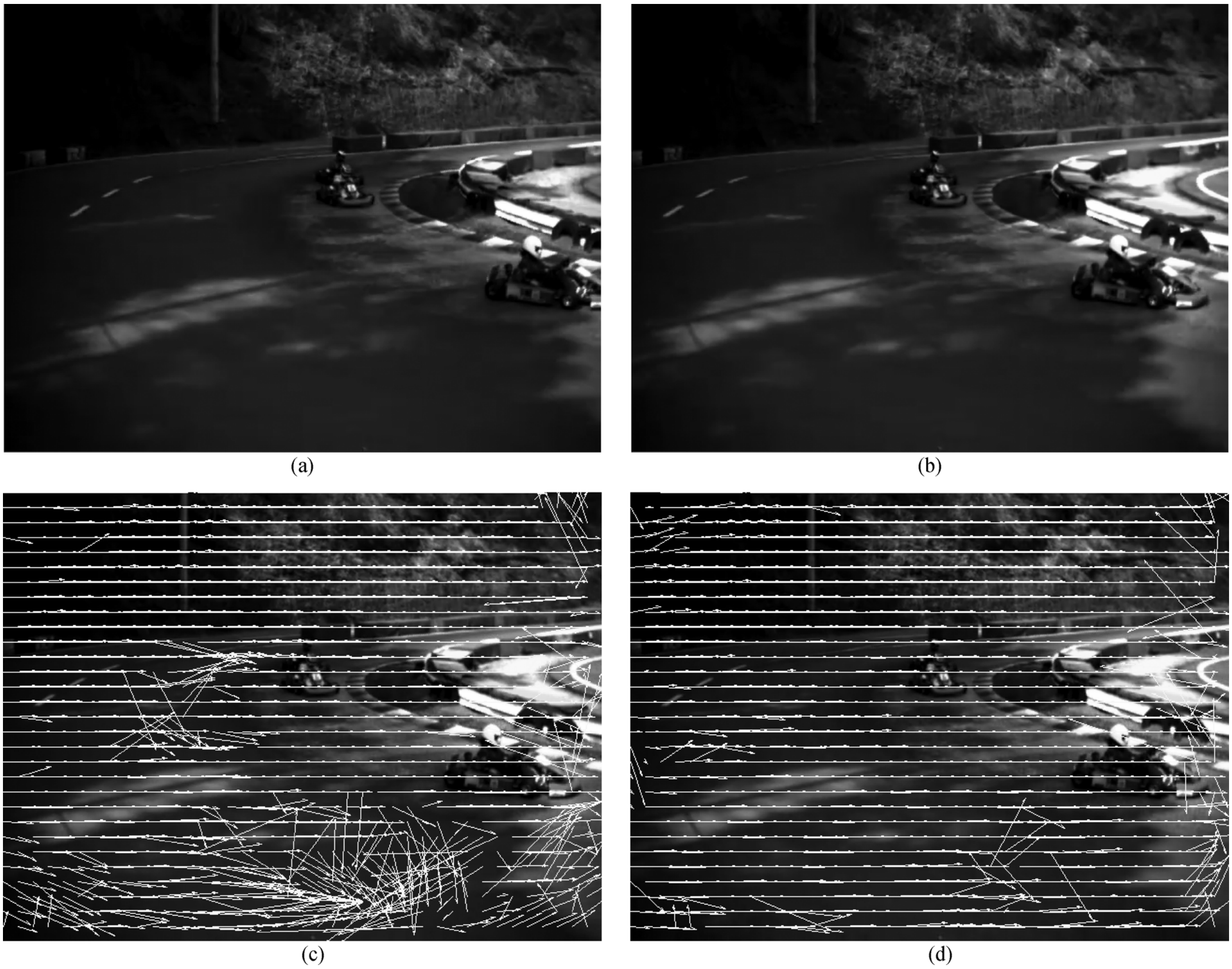


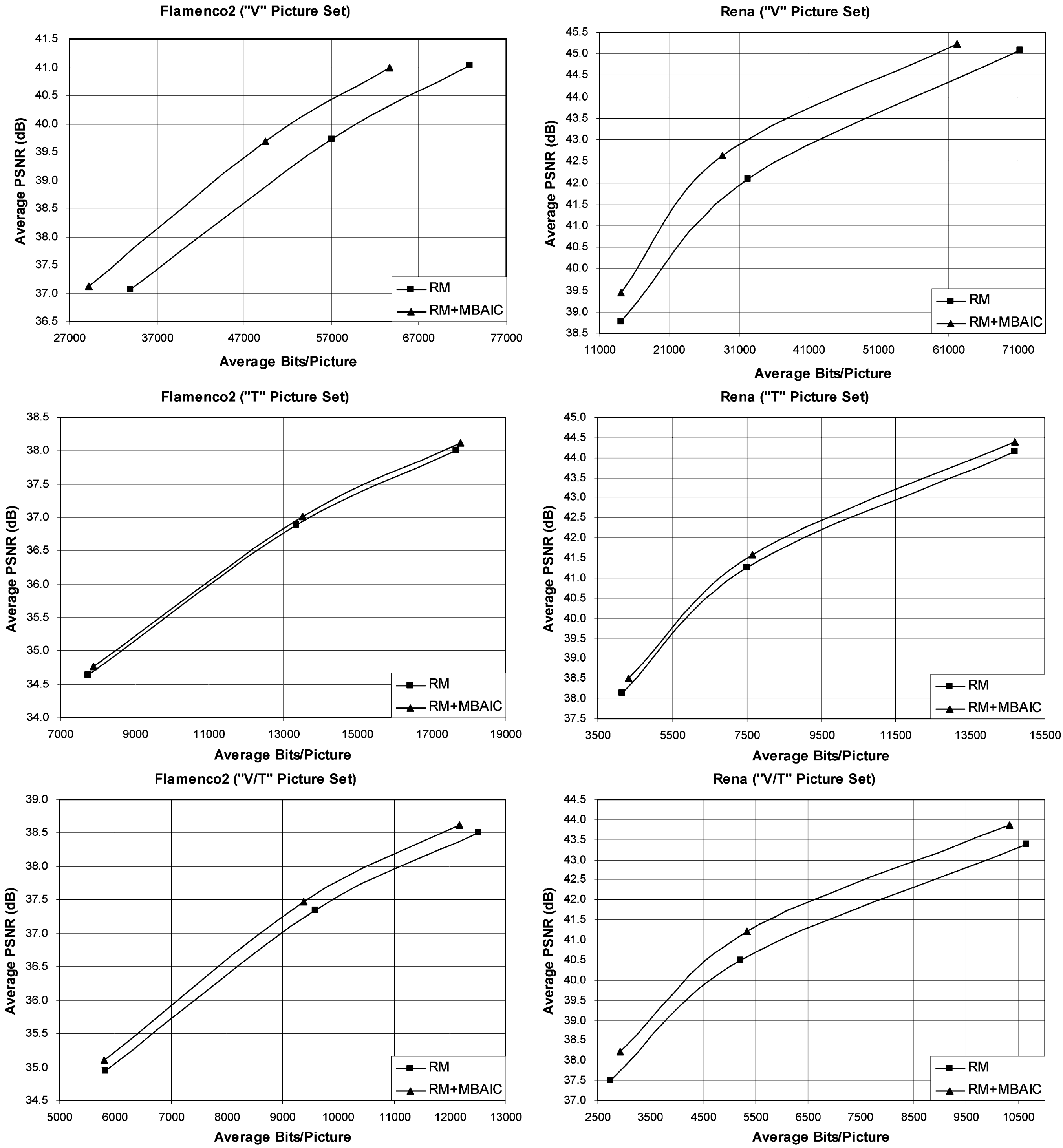
Fig. 5. Comparison of the DVE using SAD and MRSAD. (a) T0/V2 picture of Fig. 1 as the reference frame in the "Race1" sequence. (b) T0/V3 picture of Fig. 1 as the current frame in the "Race1" sequence. (c) Disparity vector field extracted by the DVE using SAD. (d) Disparity vector field extracted by the DVE using MRSAD.

obtained was about $0.2 \sim 1$ dB, because the discrepancy caused by the illumination change occurs mainly between different views. On the other hand, in the "T" and "V/T" picture sets, little or no gain was achieved with the proposed method, depending on the test sequence. The reason for this is that the test sequences used for our experiments have almost no discrepancy caused by illumination changes in the temporal direction, so that there is little gain in the "T" picture set. Also, the gains that could be obtained by the proposed method were reduced in the "V/T" picture set, because the temporal correlation is higher than the inter-view correlation in most regions of the picture (e.g., the background and nonmoving objects). However, the proposed method would work well if the illumination change occurred in the temporal direction.

C. *Comparison of* R-D *Performance for Various Test Sequences*

To demonstrate the effectiveness of the proposed MBAIC method over the whole sequence, the R-D curves are shown in Fig. 7. The RM with the proposed method is compared to the RM with the variable block-size adaptive illumination change compensation method (RM+VBAIC), the RM without any illumination change compensation method, and the RM with the WP (RM+WP), in which S and T are set to 8 or 16 in (2) and (3) for RM+VBAIC. The WP is a slice-based illumination change compensation tool used to alleviate the decrease in the coding performance induced by the fade-in/fade-out effect and the global illumination change in single-view video. It compensates for the illumination change between the current picture and the reference picture by using two factors, *viz.* the scale and offset. Three kinds of coding methods were investigated in the experiment using the WP. The first method is to use only scale factor (RM+WP(Scale)), the second method is to use only offset factor (RM+WP(Offset)), and the final method is to use multipass encoding, RM+WP(RDO). The best way is chosen among RM+WP(Scale), RM+WP(Offset), and RM without WP, by using the rate-distortion optimization technique in slice unit.

As shown in Fig. 7, the RM with the proposed MBAIC method shows the best R-D performance in comparison with the other methods. Gains were obtained of approximately $0.1 \sim 0.6$ dB compared to the RM and $0.1 \sim 0.4$ dB compared to the RM with the WP based on the multipass encoding. Also, the performance of the RM with the VBAIC shows quite similar results to the RM with the MBAIC. However, since the RM with the VBAIC requires more computational complexity than the RM with the MBAIC, the latter is proposed in this



(a) (b)

Fig. 6. R-D performance comparison of the RM and the RM with the proposed MBAIC method on three picture sets (“V”, “T”, and “V/T”) for two sequences. (a) Flamenco2. (b) Rena.

paper. Even though the WP based on the multipass encoding shows relatively good performance in Fig. 7, it needs high computational complexity because the encoding has to be performed three times per slice to calculate the R-D costs. On the other hand, the WP using either the scale or offset factor shows poor performance. It can be concluded that the global approach for illumination change compensation cannot adaptively compensate for the local illumination changes that occur in multiview video.

D. Comparison of Encoding Time Complexity

To show the complexity of the proposed method, the average encoding times for the RM, the RM with the proposed MBAIC method, and the RM with the WP based on the multipass

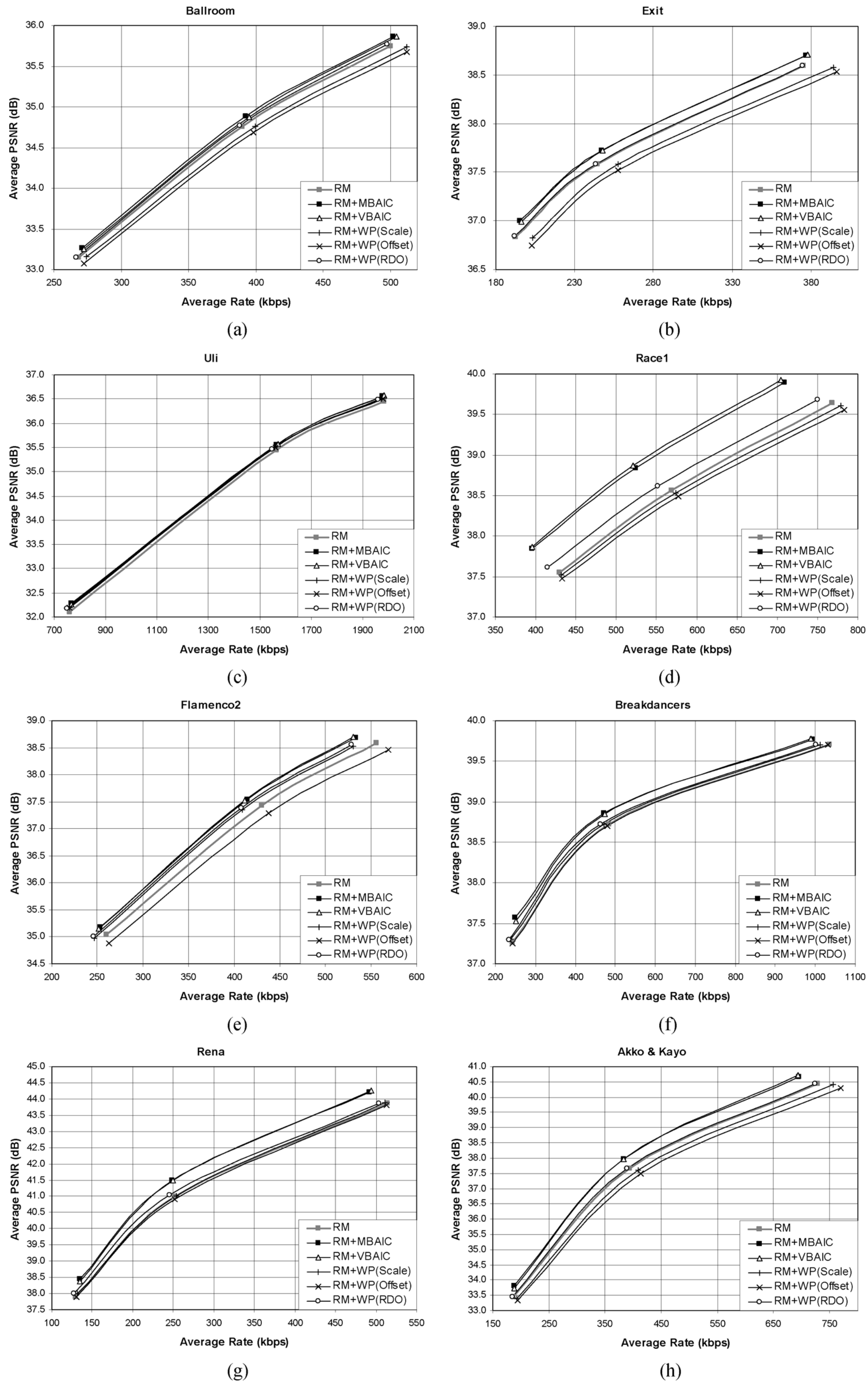


Fig. 7. R-D performance comparison of the RM, the RM with the proposed MBAIC method, the RM with VBAIC, and the RM with the WP for various test sequences. (a) Ballroom. (b) Exit. (c) Uli. (d) Race1. (e) Flamenco2. (f) Breakdancers. (g) Rena. (h) Akko & Kayo.

encoding were evaluated for all of the test sequences shown in Table I. The experiments were performed on a Pentium 4 2.66-GHz computer that has 1 GB of memory.

As shown in Fig. 8, the encoding time of the RM with the proposed method is about 1.2 times longer than that of the RM. On the other hand, the encoding time of the RM with the WP based on the multipass encoding is about 3.5 times longer than that of the RM. In the case where the proposed method is used, only two additional R-D cost calculations (Inter 16×16 and Direct 16×16) are needed in the MB unit, while in the case

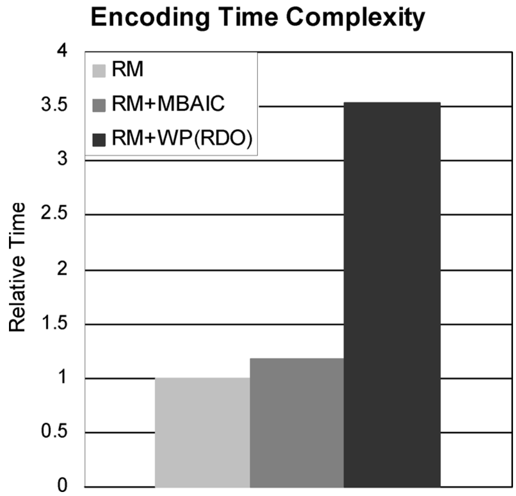


Fig. 8. Comparison of average encoding time complexity of the RM, the RM with the proposed MBAIC method, and the RM with the WP based on the multipass encoding for all the test sequences.

where the WP based on the multipass encoding is used, three times R-D cost calculations per slice are needed, so that the computational complexity is significantly increased.

According to these results, the proposed method works efficiently with relatively low computational complexity.

V. Conclusion

In this work, we have proposed an MB-based adaptive illumination change compensation method to alleviate the performance decrease caused by the illumination changes in MVC. To compensate for the local illumination changes, the proposed method is performed adaptively in the MB unit and uses MRSAD as a measure of the DVE as well as the existing SAD measure. Through the prediction process of DVIC, the side information used for transmitting the DVIC to the decoder is reduced to a greater extent.

It can be concluded from the experimental results that the discrepancy caused by the local illumination changes among views can deteriorate the high coding efficiency for MVC and this problem can be mitigated efficiently by using the proposed method with only a small increase of the encoding time complexity. The proposed method was adopted with Thomson/USC's method in the Joint Multiview Video Model (JMVM) 2.0 [17] in the JVT. Thomson/USC's method uses a slice-based illumination compensation flag and P Skip mode. A more detailed explanation of their method can be found in [10] and [11].

Acknowledgment

The authors would like to thank Dr. Y. Su and J.-H. Kim for their collaboration and verification of the experimental results concerning the H.264/AVC-based Multiview Video Coding.

References

[1] A. Smolic and D. McCutchen, "3DAV exploration of video-based rendering technology in MPEG," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 3, pp. 348–356, Mar. 2004.

[2] G. Sullivan, T. Wiegand, and A. Luthra, Draft of Version 4 of H.264/AVC (ITU-T Recommendations H.264 and ISO/IEC 14496–10 (MPEG-4 Part 10) Advanced Video Coding), ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16, Doc. JVT-N050d1, 2005.

[3] K. Mueller, P. Merkle, A. Smolic, and T. Wiegand, "Multiview Coding Using AVC," ISO/IEC JTC1/SC29/WG11, Bangkok, Thailand, Doc. M12945, 2006.

[4] "Subjective Test Results for the CfP on Multi-View Video Coding," ISO/IEC JTC1/SC29/WG11, Bangkok, Thailand, Doc. N7779, 2006.

[5] K. Kamikura, H. Watanabe, H. Jozawa, H. Kotera, and S. Ichinose, "Global brightness-variation compensation for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 12, pp. 988–1000, Dec. 1998.

[6] S. H. Kim and R.-H. Park, "Fast local motion-compensation algorithm for video sequence with brightness variations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 4, pp. 289–299, Apr. 2003.

[7] K. Sohn, Y. Kim, J. Seo, J. Yoon, G. Kim, C. Park, and J. Lee, "Results on CE2 Using IBDE for Multi-View Video Coding," ISO/IEC JTC1/SC29/WG11, Klagenfurt, Austria, Doc. M13543, 2006.

[8] Y.-L. Lee, W.-C. Sung, Y.-K. Lee, D.-Y. Kim, and J.-H. Hur, "Multi-View Video Coding Using Illumination Change-Adaptive Motion Estimation/Motion Compensation and 2D Direct Mode," ISO/IEC JTC1/SC29/WG11, Hong Kong, China, Doc. M11588, 2005.

[9] Y.-L. Lee, J.-H. Hur, Y.-K. Lee, S. Cho, H. Kwon, N. Hur, and J. Kim, "Results of CE2 on Multi-View Video Coding," ISO/IEC JTC1/SC29/WG11, Klagenfurt, Austria, Doc. M13498, 2006.

[10] J.-H. Kim, P.-L. Lai, A. Ortega, Y. Su, P. Yin, and C. Gomila, "Results of CE2 on Multi-View Video Coding," ISO/IEC JTC1/SC29/WG11, Klagenfurt, Austria, Doc. M13720, 2006.

[11] Y.-L. Lee, J.-H. Hur, Y.-K. Lee, K.-H. Han, S. Cho, N. Hur, J. Kim, J.-H. Kim, P.-L. Lai, A. Ortega, Y. Su, P. Yin, and C. Gomila, "CE11: Illumination Compensation," ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16, Doc. JVT-U052, 2006.

[12] A. M. Tourapis, F. Wu, and S. Li, "Direct mode coding for bipredictive slices in the H.264 standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 119–126, Jan. 2005.

[13] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 688–703, Jul. 2003.

[14] D. Marpe, T. Wiegand, and H. Schwarz, "Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 620–636, Jul. 2003.

[15] J. M. Boyce, "Weighted prediction in the H.264/MPEG AVC video coding standard," in *Proc. ISCAS*, May 2004, vol. 3, pp. III-789–92.

[16] Video and Test, "Updated Call for Proposals on Multi-View Video Coding," ISO/IEC JTC1/SC29/WG11, N7567, 2005.

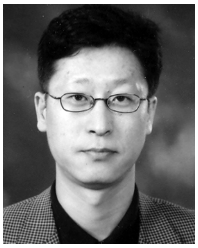
[17] A. Vetro, Y. Su, H. Kimata, and A. Smolic, "Joint Multiview Video Model (JMVM) 2.0," ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16, Doc. JVT-U207, 2006.

Jae-Ho Hur received the B.S. and M.S. degrees in computer engineering from Sejong University, Seoul, Korea, in 2005 and 2007, respectively.

He has been working for Pixtree Technologies, Seoul, Korea, since February 2007. His current research interests include video compression, image processing, 3-D video coding, and multimedia systems.

Sukhee Cho received the B.S. and M.S. degrees in computer science from Pukyong National University, Busan, Korea, in 1993 and 1995, respectively. She received the Ph.D. degree in electronics and information engineering from Yokohama National University, Yokohama, Japan, in 1999.

She is currently with the Radio and Broadcasting Research Division, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. Her research interests include stereoscopic video coding, multiviewpoint video coding (MVC), and 3DTV broadcasting systems.



Yung-Lyul Lee (SM'07) received the B.S. and M.S. degrees in electronic engineering from Sogang University, Seoul, Korea, in 1985 and 1987, respectively, and the Ph.D. degree in electrical and electronic engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1999.

He has been an Associate Professor with the School of Computer Engineering, Sejong University, Seoul, Korea, since 2001. He was a Principal Researcher at Samsung Electronics Co., Ltd., from 1987 to 2001. His current research interests include video compression, image processing, 3-D video coding, scalable video coding, and multimedia systems.

Dr. Lee received a Minister prize from the Ministry of Commerce, Industry and Energy in Korea in November 2006 and a Korea Science Technology Superiority paper prize in October 2006.