A New MPEG-2 Rate Control Scheme Using Scene Change Detection

Sanggyu Park, Youngsun Lee, and Hyunsik Chang

CONTENTS

ABSTRACT

We propose two new rate control schemes to improve MPEG-2 rate control in view of visual quality when scene changes happen. Two proposed schemes are characterized by real-time and non real-time improvement to reduce the impact of scene changes. We also propose a new target-bit prediction method using spatial activity of pictures and present a simple and efficient scene change detection scheme using signed difference of mean absolute difference (MAD). Computer simulation results show that the proposed real-time algorithm effectively alleviates visual quality degradation after scene changes. The proposed non real-time algorithm gives maximum 2 dB improvement in peak signal-to-noise ratio (PSNR) at a scene-changed picture, compared with MPEG-2 rate control scheme and it shows better quality than the real-time one.

I. INTRODUCTION

In order to transmit digital video signals over band-limited transmission channels or to store and retrieve digital video signals using digital storage equipment, it is necessary to employ compression techniques which reduce data band-width while minimizing the loss in picture quality [1]-[2]. The MPEG-2 standard [3], developed by ISO/IEC, is now being widely adopted in digital video applications such as video transmissions and broadcasting. It exploits the temporal correlation of a video sequence using motion-compensated prediction and the spatial correlation within a picture using discrete cosine transform. Temporal correlation can be removed using temporal prediction with motion compensation. MPEG-2 utilizes three kinds of temporal prediction and defines three types of pictures according to their prediction direction. They are intra-frame coded picture (I-picture), predictive coded picture (P-picture) and bi-directionally predictive coded picture (B-picture). After temporal prediction, remaining spatial correlation in prediction error signals can be further reduced by performing adaptive field/frame discrete cosine transform with adaptive quantization. Quantized coefficients are then variable length coded.

When a variable rate coded bit stream of video signals needs to be transmitted over a fixed rate channel, bit rate control is essential to regulate a bit stream. The most commonly used method for rate control is to adjust the quantization step size based on buffer fullness and the characteristics of a picture to be coded [4]-[6]. As the visual quality of reconstructed pictures heavily depends on the quantization step size, bit rate control is required both to adjust a bitstream and to stabilize the visual quality of reconstructed pictures.

MPEG-2 test model 5 (TM5) [7] presents a rate control scheme composed of target-bit allocation, rate control, and modulation. This scheme controls the bit rate in the hierarchy of coding layers, i.e., group of pictures (GOP), picture and macroblock. A GOP consists of at least one I-picture and P-pictures, and possibly some B-pictures. The rate control strategy in MPEG-2 TM5 works fairly well for normal video sequences. However it cannot handle scene changes efficiently. MPEG-2 TM5 rate control utilizes information obtained from previously coded pictures in estimating target-bits for currently coded pictures. However, if a scene change occurs, information obtained from previously coded pictures is no more useful and even can cause visual quality degradation in the pictures following the scene change. Scene changes may occur anywhere in a GOP. If a scene change happens to be a B- or a P-picture, the first P-picture after the scene change may need to be encoded with as many bits as an I-picture since a previous scene may not be a good reference for the first P-picture in a new scene. Nevertheless, MPEG-2 TM5 rate control treats the first P-picture of a new scene as an ordinary P-picture. Target-bits for a previous P-picture may not be enough for a scene-changed P-picture that is the first P-picture after a scene change. This can cause severe degradation in picture quality, which may become even worse in case of a transition from a simple scene to a complex scene since the scene-changed P-picture of a complex scene needs more coding bits than that of a simple scene. The poorly-coded P-picture will further affect the visual quality of the subsequent pictures within the same GOP.

Several methods to solve the scene change problem have been proposed and these can be classified into two groups. One is adapting the length of a GOP [8]-[10]. Scene-changed pictures are coded as I-pictures and the subsequent I-pictures are coded as P-pictures, which results in the variable length of a GOP. This approach is simple and efficient. However, this is based upon the assumption that there is at most one scene change within one GOP. Several scene changes in one GOP can cause buffer overflow due to the increasing number of bits. In the other approach, the size of a GOP remains fixed and appropriate target-bits are allocated to the scene-changed P-picture by reducing the number of bits to be assigned for the previous B-pictures that utilize the scene-changed P-picture as a reference picture [11]. This method has a drawback because scene-changed B-pictures suffer their own visual quality degradation due to lack of bits.

We propose two new approaches to the rate control that can reduce the impact of scene changes in view of visual quality. One is a real-time approach which is based on the adjustment of coding bits for two adjacent sub-groups of pictures (sub-GOPs). A sub-GOP is that we newly define for our presentation in this paper. One sub-GOP consists of at least one P-picture (or one I-picture) and possibly some B-pictures as shown in Fig. 1. It shows the series of the sub-GOPs in case that there are two B-pictures in one sub-GOP and there are four sub-GOPs in one GOP.

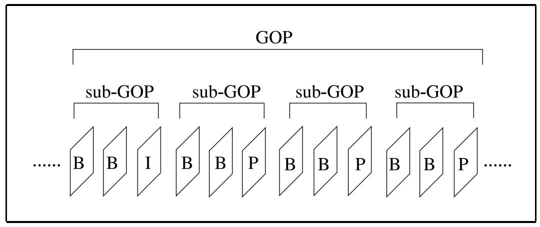

Fig. 1. The definition of a sub-GOP.

Our real-time approach looks ahead whether a scene change will happen in the subsequent sub-GOP before coding of the first picture in the sub-GOP (current sub-GOP) to be coded currently. Upon a scene change, it allocates appropriate target-bits among the pictures in such a way that more bits are assigned to the scene-changed picture than other pictures [12]. This is based upon that scene-changed pictures need relatively more bits than other pictures from the total visual quality point of view. The other is a non real-time approach that detects all the scene changes within a GOP and allocates target-bits for the scene-changed pictures using a proposed target-bit prediction scheme. In order to balance total amount of bits in a GOP, we adjust target-bits for other pictures after target-bit allocation for the scene-changed pictures, i.e. the target-bits for the remaining pictures are assigned using only the remaining bits in the corresponding GOP. Allocation of target-bits for the scene-changed pictures is performed GOP by GOP.

The key idea of our algorithms is to allocate sufficient coding bits for a scene-changed picture by reducing the coding bits to be assigned for other pictures with the balance of total amount of bits. The algorithms can effectively alleviate visual quality degradation in the subsequent pictures including a scene-changed picture and safely handle a buffer when scene changes happen. While the proposed real-time scheme and the proposed non real-time scheme are based upon the same idea, they have a big difference such that the former adjusts target-bits in two consecutive sub-GOPs but the latter adjusts target-bits in the whole GOP when scene changes occur. Our real-time scheme is useful in real-time applications such as video broadcasting even though it suffers the coding delay of one sub-GOP and our non real-time scheme is useful in non real-time applications such as video editing and recording.

Our approaches resort to accurate scene change detection prior to allocation of target-bits. Considering computational complexity and detection accuracy, we use a simple and accurate scene change detection scheme using signed difference of mean absolute difference (MAD), which is calculated in the original pixel domain between consecutive input pictures.

This paper is organized as follows. In section II, we outline MPEG-2 TM5 rate control and present the scene change detection scheme. In section III and IV, we present in detail our real-time and non real-time rate control methods. In section V, we present simulation results. Finally, we conclude the paper in section VI.

II. MPEG-2 RATE CONTROL AND SCENE CHANGE

1. MPEG-2 Rate Control

While the details of MPEG-2 TM5 rate control can be found in [7] and elsewhere, we now briefly review MPEG-2 TM5 rate control algorithm as it forms the basis for understanding our proposed algorithms. This algorithm hierarchically adjusts the bit rate in order of the coding hierarchy.

Target-bit Allocation : This step allocates target-bits for each GOP and each picture involved in the GOP. Before the first picture in a GOP is encoded, the number of remaining bits, R_j, is updated as follows :

$$R_j = R_{j-1} + (bit\_rate/\, picture\_rate) \times N \qquad (1)$$

where N is the number of pictures in the GOP, *bit_rate* is the available bit rate of a transmission channel, *picture_rate* is the frequency of pictures per second, and R_j is the available number of bits at the time of coding the j-th picture and becomes the total number of bits to be allocated for a GOP when the j-th picture is the first picture in the GOP.

After each picture is coded by using allocated target-bits, the respective complexity (X_i, X_p, X_b), which is the relative complexity measure of three kinds of previous pictures and used in allocating target-bits for the next picture, is calculated and R_j is updated as follows.

$$X_i = S_i Q_i \quad (2a)$$

$$X_p = S_p Q_p \quad (2b)$$

$$X_b = S_b Q_b \quad (2c)$$

$$R_j = R_{j-1} - S_{i,p,b} \quad (3)$$

where S_i, S_p, S_b are the number of bits generated after encoding each picture and Q_i, Q_p, Q_b are average quantization parameters computed by averaging actual quantization values used in coding of all the macroblocks for each type of picture.

Target-bits for each picture in the GOP are computed as follows :

$$T_i = \max\left\{ \frac{R}{1 + \frac{N_p X_p}{X_i K_p} + \frac{N_b X_b}{X_i K_b}},\ \frac{bit\_rate}{8 \times picture\_rate} \right\} \quad (4a)$$

$$T_p = \max\left\{ \frac{R}{N_p + \frac{N_b X_b K_p}{X_p K_b}},\ \frac{bit\_rate}{8 \times picture\_rate} \right\} \quad (4b)$$

$$T_b = \max\left\{ \frac{R}{N_b + \frac{N_p X_p K_b}{X_b K_p}},\ \frac{bit\_rate}{8 \times picture\_rate} \right\} \quad (4c)$$

where K_p and K_b are constants that depend on quantization parameters. N_p and N_b are the number of P-pictures and B-pictures remaining in the current GOP. The value of $bit\_rate/(8 \times picture\_rate)$ is the minimum number of target-bits to be defined in MPEG-2 TM5.

Rate Control : This step sets the reference value of quantization parameter for each macroblock in the picture by measuring virtual buffer fullness, which adjusts the amount of bits for each macroblock. Before encoding j-th macroblock, the virtual buffer fullness at the time of coding the j-th macroblock, is computed in the following :

$$d_j^x = d_0^x + B_{j-1} - (T_x / MB\_cnt) \times (j-1) \quad (5)$$

where x denotes the picture type (I-, P- or B-picture) and d_0^x is the initial values of virtual buffer fullness. B_j is the number of bits generated by encoding all macroblocks in the picture up to and including the j-th macroblock. MB_cnt is the number of macroblocks in the picture.

And the reference quantization parameter Q_j that affects the quantization step size after being modulated in the next step, is calculated for j-th macroblock as follows :

$$Q_j = (d_j \times 31 \times picture\_rate)/(2 \times bit\_rate) \quad (6)$$

The increase of d_j makes Q_j increase, which results in decrease of bits used for a macroblock, and *vice versa*.

Modulation : This step modulates the reference quantization parameter, Q_j using the spatial activity of a corresponding macroblock to derive the value of the quantization parameter, *mquant*, which is actually used in the quantization procedure.

There is a known problem with this scheme. The target-bits for an I-picture, P-pictures and B-pictures, T_i, T_p and T_b are set based upon only the information obtained from encoding of previous pictures, i.e. from X_i, X_p and X_b. This may not be proper for a picture in a new scene that differs significantly from the previous scene.

2. Scene Change

Scene change means the abrupt change of picture characteristics between consecutive pictures. If scene change occurs, information obtained from previously coded pictures is no more useful and new information is necessary for coding. In order to cope with the scene change problems efficiently, it is essential to check for the occurrence of a scene change in advance of coding a picture. If we detect the occurrence of scene change beforehand, we can apply several schemes as proposed in this paper to solve scene change problems.

Our rate control schemes are based upon an accurate scene change detection algorithm. It utilizes signed difference of MAD. The MAD of luminance components between consecutive input pictures is defined as :

$$\mathrm{MAD} = \frac{1}{MN}\left\{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} |\, f_n(i,\, j) - f_{n-1}(i,\, j)\,|\right\} \tag{7}$$

where $f_n\ (i,\ j)$ is the luminance value of the picture element at position $(i,\ j)$ in the n-th input frame, N is the number of pixels in a line and M is the number of lines in a picture. The MAD corresponds to the first-order derivative, df/dt, which has a large value where scene change occurs. The second-order derivative of f corresponds to signed difference of MAD in pictures, SDMAD, which can be described as follows :

$$\mathrm{SDMAD}(n) = \mathrm{MAD}(n) - \mathrm{MAD}(n-1) \tag{8}$$

Fig. 2 depicts MAD and SDMAD on consecutive pictures. An abrupt change in df/dt will result in a peak value in df^2/d^2t, which is easily detected by selecting a proper threshold value. According to our experience with the scene change detection scheme on real-world video sequences, the value of SDMAD shows large change upon scene change enough to allow using an insensitive threshold value. The scheme has an extremely low probability of false detection. By using SDMAD as the measure of scene change detection, accurate detection of scene change is possible.

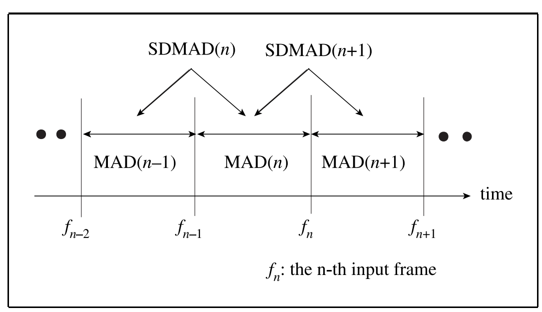

Fig. 2. The relation between MAD and SDMAD.

III. REAL-TIME IMPROVEMENT

Our proposed algorithm is the same as that in MPEG-2 TM5 if no scene change happens. However, when scene changes happen, the proposed scheme uses its own target-bit allocation method except in the case when scene-changed pictures are I-pictures. If a scene change happens, the first P-picture after the scene change may be given as many bits as an I-picture since many macroblocks in the scene-changed picture need to be coded in intra-mode. Since a poorly-coded scene-changed picture will further degrade the visual quality of the subsequent pictures within the same GOP, it is necessary to allocate sufficient target-bits for the scene-changed picture. Therefore, we allocate more target-bits to scene-changed pictures than normal P-pictures and less target-bits to pictures preceding the scene-changed pictures than usual.

Prior to bit allocation for a current coding picture, we check for a scene change on both the current sub-GOP and the next sub-GOP, where the current sub-GOP involves a picture to be coded currently and the next sub-GOP follows the current sub-GOP. If a scene change occurs in the next sub-GOP, less bits are assigned for the coding pictures in a current sub-GOP in order to assign relatively larger bits to a scene-changed picture. Fig. 3 is the pictorial description for the proposed real-time scheme. The bits saved from the sub-GOP before a scene change are allocated to the reference P-picture in the sub-GOP where a scene change happens. We assume that scene change does not occur consecutively on two consecutive sub-GOPs. The proposed algorithm can be applied regardless of the number of B-pictures between two reference pictures. In this paper we describe the proposed algorithm assuming that there are two B-pictures between two reference pictures.

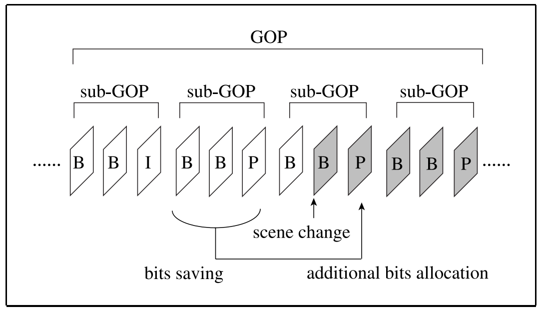

Fig. 3. The pictorial description of the proposed real-time scheme.

The proposed algorithm is described in detail as below :

When a scene change occurs in the next sub-GOP with no scene change in a current sub-GOP, we check the last picture type of the current sub-GOP. If the picture is an I-picture, target-bits for the coding pictures of a current sub-GOP are assigned as follows :

$$T_i\,^\wedge = T_i \times C1 \quad (9a)$$

$$RT_i = (1 - C1) \times T_i \quad (9b)$$

$$T_b\,^\wedge = T_b \times C2 \quad (10a)$$

$$RT_b = (1 - C2) \times T_b \quad (10b)$$

where $T_i\,^\wedge$ and $T_b\,^\wedge$ are the decreasing target-bits for an I-picture and a B-picture in the current sub-GOP. T_i and T_b are derived from (4a) and (4c). RT_i and RT_b are saved bits from an I-picture and a B-picture in a current sub-GOP, which will be used for a scene-changed picture in the next sub-GOP. $C1$ and $C2$ are weighting constants with $C1 = 0.7$ and $C2 = 0.8$ in this paper. These values are chosen through experiments with real-world test sequences

If the last picture of a current sub-GOP is a P-picture, the last picture type of the next sub-GOP is checked. If the picture is an I-picture, target-bits assigned for a current sub-GOP is the same as in MPEG-2 TM5. If the last picture of the next sub-GOP is a P-picture, target-bits for the coding pictures of a current sub-GOP are assigned as follows :

$$T_p{}^\wedge = T_p \times C1 \tag{11a}$$

$$RT_p = (1 - C1) \times T_p \tag{11b}$$

$$R_b{}^\wedge = T_b \times C2 \tag{12a}$$

$$RT_b = (1 - C2) \times T_b \tag{12b}$$

where $T_p{}^\wedge$ is the decreasing target-bits for a P-picture in the current sub-GOP. T_p is derived from (4b). RT_p is the bits saved from a P-picture in the current sub-GOP for the scene-changed picture in the next sub-GOP.

When no scene change occurs in the next sub-GOP with a scene change in the current sub-GOP, the picture type of the last picture of the current sub-GOP is checked. If the picture is an I-picture, target-bits assigned for the current sub-GOP is exactly the same as in MPEG-2 TM5. If the picture is a P-picture, the picture type of the last picture of the previous sub-GOP is checked, where the previous sub-GOP is followed by the current sub-GOP. If the last picture of the previous sub-GOP is a P-picture, target-bits for the coding pictures of a current sub-GOP are assigned as follows :

$$T_p{\sim} = T_p + RT_p + 2 \times RT_b \tag{13}$$

$$T_b{\sim} = T_b \tag{14}$$

where $T_p{\sim}$ and $T_b{\sim}$ are the increasing target-bits for a P-picture and a B-picture in the current sub-GOP. The value, 2 is derived from the number of B-pictures in the previous sub-GOP. If the picture is an I-picture, target-bits for the encoding pictures of the current sub-GOP are assigned as follows :

$$T_p{\sim} = T_p + RT_i + 2 \times RT_b \tag{15}$$

$$T_b{\sim} = T_b \tag{16}$$

We sacrifice the visual quality of pictures preceding scene-changed pictures which is less important than that of a scene-changed picture from the total visual quality point of view. Checking the occurrence of a scene change in the next sub-GOP causes a coding delay corresponding to one sub-GOP. Except for target-bit allocation, the rest of the procedure is exactly the same as that of MPEG-2 TM5.

IV. NON REAL-TIME IMPROVEMENT

In video editing and recording applications, edited sequences can cause several scene changes. In this case, scene changes can happen consecutively or be concentrated in a GOP. So several methods mentioned in section I and section III do not work well since they basically assume that sequences to be managed are normally stable with at most one scene change in one GOP.

Our proposed algorithm detects scene changes prior to coding the first picture in a GOP in order to assign sufficient coding bits for the scene-changed pictures. In case of the occurrence of scene changes, it calculates the proper target-bits for existing scene-changed pictures by using the proposed target-bit prediction method. Target-bits for the first I-picture after a scene change is also calculated using the same prediction method. Except for allocation of predicted target-bits for scene-changed pictures and I-pictures prior to coding the first picture in a GOP, our target-bit allocation scheme for the remaining pictures is the same as that of MPEG-2 TM5, i.e., the target-bit allocation procedure for the remaining pictures is progressed picture by picture. Additionally in case of transition from a complex scene to a simple scene, the same target-bits as a P-picture can be assigned to a B-picture that mainly uses only one-directional prediction from the a P-picture preceding the scene change. Otherwise, annoying visual quality degradation might occur at the B-picture preceding a scene-changed picture due to one-directional prediction and insufficient bits.

We present a method to predict target-bits for a scene-changed picture. In MPEG-2 TM5, target-bits for pictures are calculated from the complexity (X_i, X_p, X_b) in (2). However, if a scene change happens, the complexity (X_i, X_p, X_b) obtained from previous pictures is not useful anymore due to little temporal correlation between the previous and current scenes. So, if we obtain the more proper complexity instead of the complexity obtained from previous coded pictures, we can calculate more proper target-bits for a scene-changed picture than MPEG-2 TM5. We obtain the meaningful complexity by using our experimental results. From our experimental results using several sequences coded at a bit rate of 4 Mbit/s, we found that the complexity of I-pictures, X_i, is related to their activity at a bit rate of 4 Mbit/s as shown in Fig. 4. Based on our observations, we have an approximate expression in the following :

$$X_i = 2 \times 10^6 \times 10^{[(0.0005 \times act^2 + 12.7 \times act + 1840)/40000]} \tag{17}$$

where *act* is the activity of an I-picture given by

$$act = \sum_{k=1}^{\text{no\_of\_blocks}} \text{act\_mean}[k] / \text{ no\_of\_blocks}$$
$$\text{act\_mean}[k] = 1 + \text{variance of luminance pixel values in the } k\text{-th block} \tag{18}$$

So, by calculating *act* value of an I-picture, we can obtain X_i. Here, if we assume that the scene-changed picture can be treated as an I-picture, we can obtain the meaningful pseudo X_i for the scene-changed picture by using (17) with *act* value of the scene-changed picture. This assumption is quite acceptable in that many macroblocks in the scene-changed pictures will be encoded in intra-mode like as in I-pictures. Since the complexity of P-pictures or B-pictures has relatively small variations compared to that of I-pictures, the complexity calculated from previous coded pictures can be used as it is. Therefore, by using (4a) with pseudo X_i derived from (17) instead of X_i obtained from the previous coded pictures, we can predict the more proper target-bits to be assigned for a scene-changed picture.

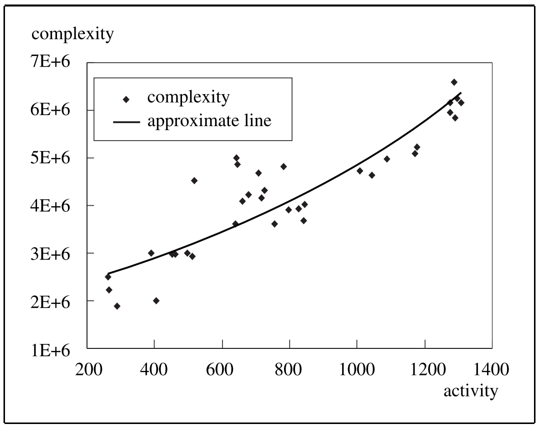

Fig. 4. The relation between activity and complexity of I-pictures (4 Mbit/s).

The target-bit allocation procedure of the non real-time scheme is described in a pseudo Pascal as below :

```
PROCEDURE TargetBitsAllocate_NonRealTime();
BEGIN
      IF(Scene changes exist in the previous GOP)
      THEN BEGIN
         Predict and allocate target-bits for an
         I-picture by using (4a) with Xi derived
         from (17);
      END;
      Check the GOP whether or not there are
      scene changes;
      IF(Scene change exists) THEN BEGIN
         Predict and allocate target-bits for the
         scene-changed pictures by using (4a)
         with pseudo Xi derived from (17);
      END;
      WHILE(NOT end of GOP) DO BEGIN
         Calculate remaining bits for other remain-
         ing pictures than an I-picture and scene-
         changed pictures;
         IF(Scene-changed picture or I-picture)
         THEN BEGIN
            Allocate the predicted target-bits;
         END ELSE THEN BEGIN
            Allocate target-bits by using
            remaining bits like as MPEG-2 TM5;
         END;
      END;
   END;
```

Since a scene-changed picture is the first picture of a new scene as the reference picture of subsequent pictures and largely affects the visual quality of subsequent pictures, we make a scene-changed picture use sufficient coding bits. Checking all scene changes in a GOP causes a coding delay of one GOP time, which makes the scheme practical only for non real-time applications. After the allocation of target-bits for each picture, the rest of rate control steps are performed like as those of MPEG-2 TM5.

V. SIMULATION RESULTS

Computer simulations are carried out on an edited test sequence of 69 frames containing five times of scene changes in order to test the performance of the proposed two schemes on the condition of scene changes. The

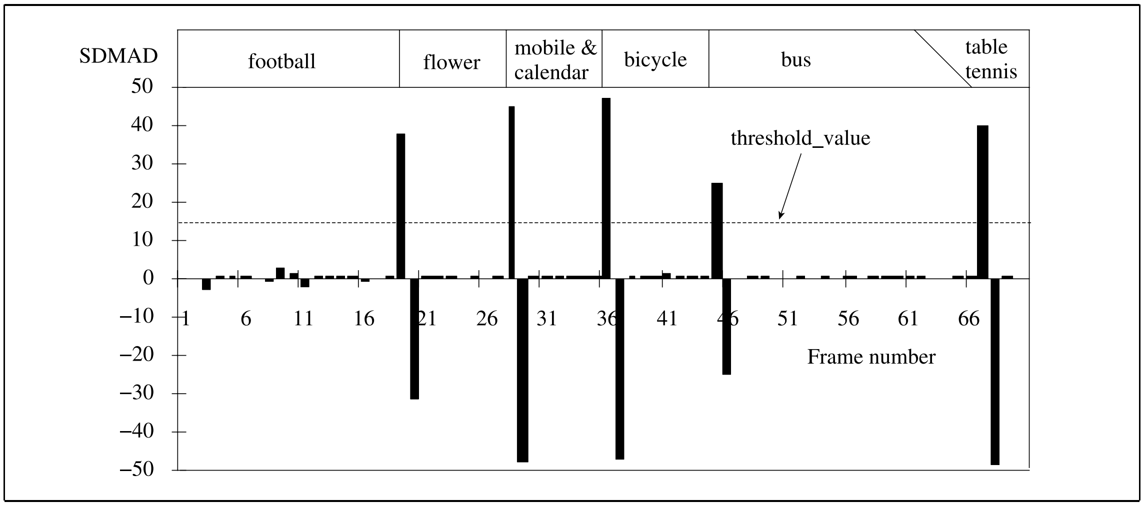

Fig. 5. The performance of scene change detection using the SDMAD.

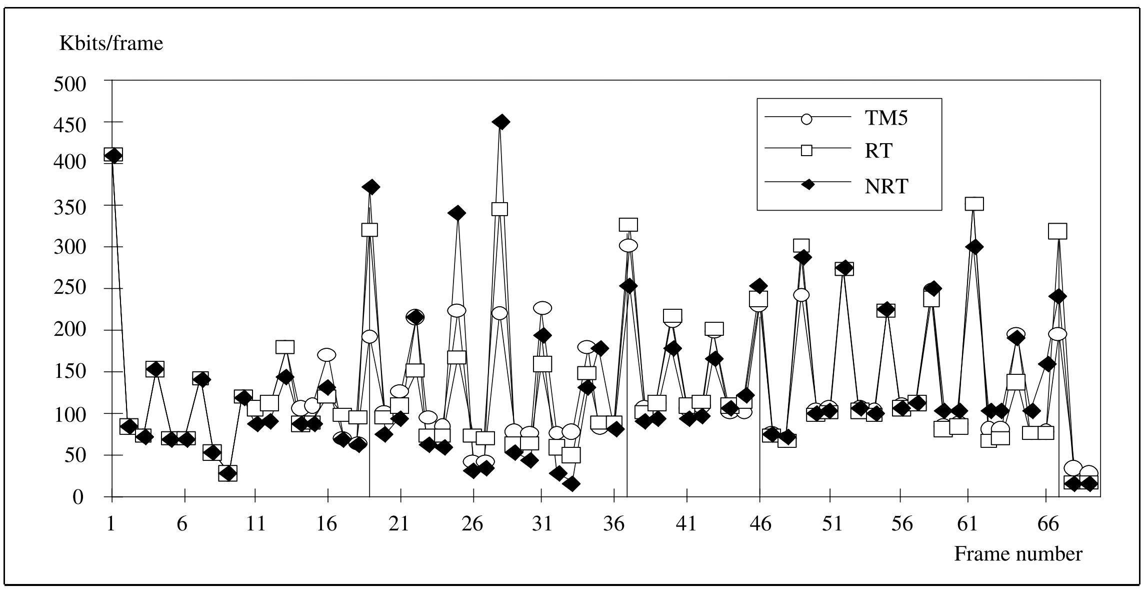

Fig. 6. The target-bit allocation of MPEG-2 TM5, the real-time scheme (RT), and the non real-time scheme (NRT).

edited sequence is sequentially composed of 18 frames from 'football' sequence, 9 frames from 'flower' sequence, 8 frames from 'mobile and calendar' sequence, 9 frames from 'bicycle' sequence, 22 frames from 'bus' sequence, and 3 frames from 'table tennis' sequence. The picture format of a sequence is 720 pixels by 480 lines per frame with chrominance com-

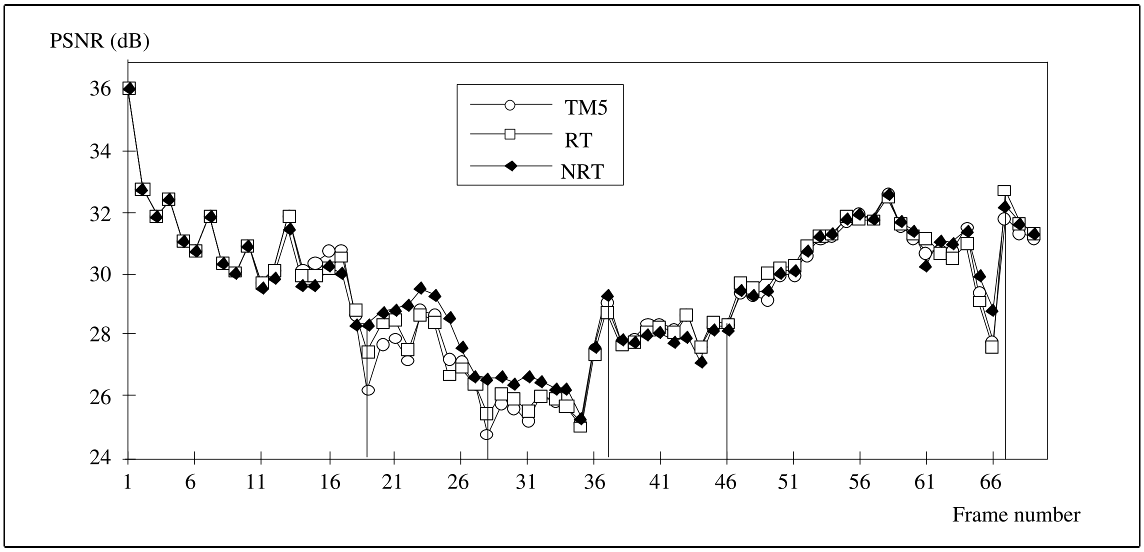

Fig. 7. The performance comparison of TM5, RT and NRT.

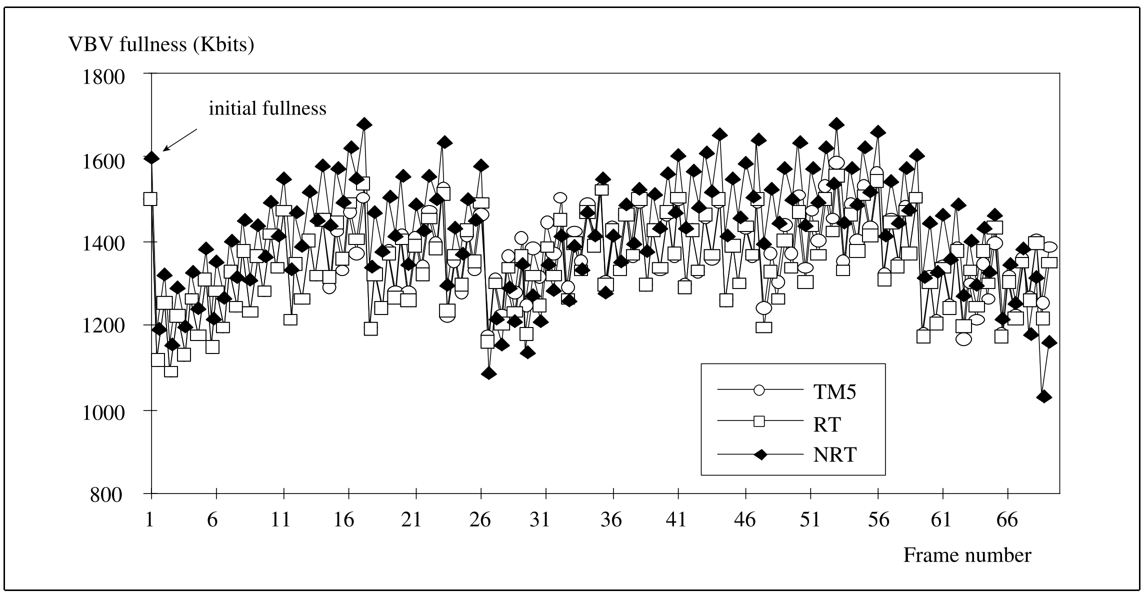

Fig. 8. The VBV fullness of TM5, RT and NRT.

ponents sub-sampled horizontally by a factor of 2, i.e., 4:2:2 picture format. After 4:2:2 to 4:2:0 picture format conversion, the test sequence is coded by using MPEG-2 video coding at a bit rate of 4 Mbit/s. The length of GOP is 12 as shown in Fig. 1 in this paper.

Fig. 5 shows the performance of the scene change detection algorithm. It detects scene change very accurately using the threshold value of 15. As the SDMAD has quite a small value when no scene change happens, the false detection can be easily avoided.

Fig. 6 shows the allocation of target-bits for MPEG-2 TM5 rate control, the proposed real-time scheme, and the proposed non real-time scheme. In the scene-changed pictures, proposed algorithms allocate more target-bits than MPEG-2 TM5 rate control algorithm, which results in alleviation of visual quality degradation in the scene-changed pictures and the subsequent pictures.

Fig. 7 shows the performance of each algorithm in PSNR of Y signals. While three schemes show similar performance in the transition from a complex scene to a simple scene, they provide quite different visual quality in the opposite transition. Our proposed schemes generally show better visual quality than MPEG-2 TM5 scheme through the whole edited sequence. Especially, the non real-time scheme maintains normal visual quality regardless of scene changes and provides 2 dB improvement in PSNR in the 19-th frame, compared with MPEG-2 TM5 rate control. Our visual comparison test for the subjective assessment shows that the real-time scheme effectively reduces visual quality degradation in the scene-changed picture and the subsequent pictures while it suffers a little visual quality degradation in the small number of pictures preceding the scene changes. The non real-time scheme maintains a normal visual quality even in the scene-changed pictures and shows better quality than the real-time one.

Fig. 8 shows video buffering verifier (VBV) fullness which is a model of physical decoder buffer fullness. It shows the VBV fullness for each scheme moves within the safe range in the whole edited sequence. The proposed schemes handle buffer fullness safely without overflow or underflow of a decoder buffer as well as MPEG-2 TM5 rate control even though target-bits are adjusted .

VI. CONCLUSIONS

We have presented two new rate control schemes to improve MPEG-2 TM5 rate control to handle scene changes better. The proposed schemes aim at real-time and non real-time improvement to reduce the impact of scene changes. We also have presented a new target-bit prediction method and a scene change detection scheme using the SDMAD. Simulation results show that the real-time algorithm effectively alleviates visual quality degradation after scene changes and the non real-time algorithm shows the best quality among the three schemes. As for buffer fullness, both of the proposed schemes handle the buffer as safely as MPEG-2 TM5 rate control scheme. We also demonstrated the scene change detection algorithm can detect scene changes very accurately.

ACKNOWLEDGMENT

The authors would like to thank Dr. Yong Han Kim and Dr. Jin Woong Kim for helpful discussions and the members of the Human Interface Department for their continuous encouragement.

REFERENCES

[1] J. O. Limb, C. B. Rubinstein, and J. E. Thompson, "Digital coding of color video signals-A review," *IEEE Trans. Commun.*, vol. COM-25, pp. 1349-1385, Nov. 1977.

[2] A. K. Jain, "Image data compression: A review," *Proc. IEEE*, pp. 1349-1389, Mar. 1981.

[3] ISO/IEC DIS 13818-2, MPEG-2 Video, 1994.

[4] W.H. Chen and W.K. Pratt, "Scene adaptive coder," *IEEE Trans. Commun.*, pp. 225-232, 1984.

[5] A. Puri and R. Aravind, "Motion-compensated video coding with adaptive perceptual quantization," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 1, no. 4, pp. 351-361, 1991.

[6] E. Viscito and C. Gonzales, "A video compression algorithm with adaptive bit allocation and quantization," *Proc. SPIE*, 1605, pp. 58-72, 1991.

[7] ISO/IEC JTC1/SC29/WG11, Test model 5, 1993.

[8] L. Wang, "Rate control for MPEG video coding," *Proc. SPIE*, 2501, pp. 53-61, 1995.

[9] C. Horne and A. Puri, "Video coding with adaptive quantization and rate control," *Proc. SPIE*, 1818, pp. 798-806, 1992.

[10] S. H. Jang and S. H. Park, "An adaptive rate control algorithm for DPCM/DCT hybrid video codec adopting bi-directional prediction," *Proc. SPIE*, 2094, pp. 1237-1248, 1993.

[11] S. G. Lee, J. H. Cho, S. G. Park, and S. W. Ra, "Scene adaptive bit rate control strategy," *Proc. SPIE*, 2501, pp. 1461-1469, 1995.

[12] J. Y. Nam, S. G. Park, Y. S. Lee, S. M. Lee, C. Ahn, and S. W. Ra, "An adaptive rate scheme with scene change detection," *Proc. Int'l Workshop on HDTV '94*, 1994.

Sanggyu Park was born in Taejon, Korea in 1960. He received B.S. and M.S. degrees in electronic engineering from Yonsei University in 1982 and 1985, respectively. He subsequently received Ph.D degree in electronic engineering from Chungnam National University in 1996. He joined ETRI in 1985 and has worked on the development of ISDN terminals, video and audio compression technologies, Digital TV and HDTV encoder systems. His current research interests include multi-media production and broadcasting.

Youngsun Lee was born in Pohang, Korea in 1966. She received B. S. and M. S. degrees in electronic engineering from Kyungpook National University in 1989 and 1991, respectively. Since 1991 she has been working at ETRI, on the development of Digital TV and HDTV encoder systems. Her research interests include image compression and its applications.

Hyunsik Chang was born in Kumi, Korea in 1968. He received B.S. and M.S. degrees in electronic engineering from Kyungpook National University in 1991 and 1993, respectively. He joined ETRI in 1993, where he currently works in video and audio communication. His research interests include video and audio compression.