1998 International Computer Symposium
Workshop on Image Processing and Character Recognition
December 17-19, 1998, N.C.K.U., Tainan, Taiwan, R.O.C.

A New Rate Control Scheme for H.263 Images

Chi-Chang Kuo * *and Jin-Jang Leou* #

* Digital Video-Audio Signal Processing Tech. Department
Digital Video Tech. Division, Computer & Communication Research Labs.
ITRI, Hsinchu 310, Republic of China

Department of Computer Science and Information Engineering
National Chung Cheng University, Chiayi, Taiwan 621, Republic of China

ABSTRACT

In some image/video applications, the variable bit rate (VBR) image bitstream will be transmitted over a constant bit rate (CBR) transmission channel, in which a channel buffer is employed. In this study, a new rate control scheme for H.263 images is proposed. By using the proposed rate control procedure for the frame layer, the target bits for each image frame is obtained. Then the proposed just-noticeable-distortion (JND) preprocessing is applied on all the INTRA mode MBs. A new bit estimation model using the average weighted square sum of the zig-zag ordered DCT (discrete cosine transform) coefficients is proposed. The optimal quantization parameter (QP) of each MB (macroblock) within the MB layer is determined by using the Lagrange multiplier. Based on the simulation results obtained in this study, the proposed approach can meet the target bit rate more accurately, keep a lower buffer fullness, and have a better temporal resolution than TMN9, whereas the PSNR of the proposed approach is as good as that of TMN9.

1. INTRODUCTION

In some image/video applications [1], a variable bit rate (VBR) image bitstream will be transmitted over a constant bit rate (CBR) transmission channel, in which a channel buffer is usually employed to smooth out the bit rate variation of the VBR image bitstream. It is greatly desirable to design a feasible rate control scheme to convert the VBR bitstream into the CBR transmission channel and prevent the channel buffer from overflow or underflow. In this study, a new rate control scheme for compressed VBR H.263 images [3] transmitted over CBR channels is proposed.

Rate control is used to achieve a target bit rate with a consistent visual perceptual quality, which can be generally separated into two steps: (1) to allocate a target bit for each MB (macroblock), GOB (group of blocks), frame, or GOP (group of pictures), (2) to apply a rate control scheme to meet the target bit. For target bit allocation, it is important to distribute an adequate bit budget for each allocation unit with an identical perceptual quality. A simple allocation method is proposed in the H.263 Test Model Near-term version 5 (TMN5) [4]. For each frame in TMN5, the target bit is set to the total constrained bit rate divided by the picture frame rate. MPEG Test Mode 5 (TM5) [5] allocates bits to I-, P-, and B-pictures according to some global complexity measures of previous encoded pictures, such as product of the coded bits and the average quantization parameter. Additionally, several researches developed the approaches for rate control [6] – [10].

The bit rate control strategy by varying the quantization scale can refer the perceptual tolerance of the quantization noise in the human visual system. It is known that visual sensitivity of human observers decreases at and adjacent to large luminance changes [11]. Weber's law shows that the perceived brightness of an object stays relatively constant, independent of illumination with very light or very dark background [11]. It follows that human vision has limited spatial and temporal resolution. Hence, the quantization artifacts will be less noticeable in more activity regions of an image sequence that has a high number of luminance edges or contains a significant amount of motion [12]. The quantization scale can be employed to control the bit rate according to the content activities of an image sequence and the planned buffer fullness [5], [13]. One type of rate control approaches uses an explicit bit model to predict the number of compressed bits when a certain quantization scale is used [14]-[15]. Ribas-Corbera and Neuhoff [15] derived a bit estimation model by exploiting the differential entropy concept. This model has been embedded into the rate control method of H.263 Test Model Near-term version 9 (TMN9) [16]. Many rate control algorithms can not handle scene changes efficiently. Lee and Dickinson [17] used the dynamic GOP structure for bit rate control by taking advantage of the temporal masking in human vision. Lin and Wu [18] proposed a content-based bit rate control scheme.

To maintain high image quality at low bit rates, an efficient coding algorithm should eliminate not only the spatial and temporal redundancies, but also the perceptual redundancy from video signals. The perceptual redundancy is due to the nonlinear responses of the human visual system (HVS) [11]. Chou and Li [19] presented a perceptually tuned subband image coder based on the measure of just-noticeable-distortion (JND) profile. The JND profile provides each signal being coded with a visibility threshold of distortion, below which

Correspondence author: Dr. Jin-Jang Leou, Tel: 886-5-2720411 Ext. 5310, 6014, Fax: 886-5-2720859, E-mail: jjleou@cs.ccu.edu.tw

reconstruction errors are rendered imperceptible. Incorporating the threshold sensitivities due to background luminance and texture masking effect, the JND profile is estimated from analyzing local properties of image signals. The full-band JND profile is decomposed into component JND profiles of different frequency subbands. In this study, a new rate control scheme for H.263 images is proposed, in which human visual (perceptual) properties are employed.

In the remainder of this paper, the proposed rate control scheme is addressed in Section 2. Simulation results are included in Section 3, followed by concluding remarks.

2. PROPOSED RATE CONTROL SCHEME FOR H.263 IMAGES

The proposed rate control scheme embedded in the H.263 coder is shown in Fig. 1. For the frame layer, according to the channel buffer fullness and the target channel rate, the proposed approach first estimates the appropriate number of target bits allocated to each image frame, F_j.

Each input frame, F_j, is partitioned into nonoverlapping macroblocks (MBs) ($X_{i,j}$, the ith macroblock of the frame F_j). For each MB, the best mating macroblock within a specific tracking range is determined by the motion estimation and compensation procedure from the previously reconstructed frame, $\hat{F}_{j-1}$, which is stored in the one frame delay memory. In general, due to the temporal correlation between two successive frames without scene change, the current MB can be fairly well represented by its prediction (a motion-compensated MB in $\hat{F}_{j-1}$), i.e., the current MB is coded in the INTER mode. Otherwise, the current MB is coded in the INTRA mode. The decision of INTRA/INTER mode coding for each MB is determined by a mode selection procedure.

Within the MB layer, for the luminance block of each INTRA mode coded MB, the proposed JND preprocessing procedure is employed to remove its perceptual redundancy. Based on the statistical activities of the 64 DCT (discrete cosine transform) coefficients measured by a weighting function, the target bit, and the channel buffer fullness, the proposed rate control procedure for the MB layer is used to determine the quantization parameter (QP) for the current MB. The quantized DCT coefficients are variable-length coded and packaged with the header information. Finally, the resulting number of coded bits and the quantization parameter QP of the current MB are fed backward to update the corresponding parameters.

2.1 Proposed Rate Control Approach for Frame Layer

The objective of rate control for the frame layer is to allocate a bit budget for the current encoding frame according to the channel buffer fullness and the target channel rate. The proposed rate control procedure for the frame layer is simply given by:

$$B = \frac{(R - W)}{F}, \tag{1}$$

where B is the target bits of the current frame, R is the channel bit rate in bit/sec, F is the target frame rate in frame/sec, and W is the number of bits occupied in the channel buffer.

2.2 Proposed JND Preprocessing for INTRA Coded Macroblocks

The perceptual redundancy in INTRA coded MBs as well as the number of coded bits required can be reduced by applying the perceptual coding technique [11], [19]. Before encoding each INTRA mode MB, the JND profile for the four 8 × 8 luminance blocks in the MB are given by [26]:

$$\begin{aligned} &\mathrm{JND}_{j,k}(x,y) \\ &= \max\{f_1(bg(x,y), mg(x,y)), f_2(bg(x,y))\}, \\ &\text{for } 0 \le x < 8, \text{ and } 0 \le y < 8, \end{aligned} \tag{2}$$

where $\mathrm{JND}_{j,k}(x,y)$ is the JND profile of pixel (x,y) within the jth block of the kth MB in the current frame. To obtain the JND-embedded INTRA MBs, denoted by JNDE MBs, the luminance value of the ith pixel of the kth MB is modified by:

$$P_{L,k}^{\mathrm{JNDE}}(i) = \begin{cases} \max\{A, \overline{P}_{L,k}\}, \text{if } P_{L,k}(i) > \overline{P}_{L,k}, \\ \min\{A, \overline{P}_{L,k}\}, \text{otherwise}, \end{cases} \tag{3}$$

$$A = P_{L,k}(i) + B \cdot \mathrm{JND}_k^W(i), \tag{4}$$

$$B = \begin{cases} +1, & \text{if } P_{L,k}(i) \le \overline{P}_{L,k}, \\ -1, & \text{otherwise}, \end{cases} \tag{5}$$

where $P_{L,k}(i)$ is the original luminance value of the ith pixel of the kth MB, $P_{L,k}^{\mathrm{JNDE}}(i)$ is the corresponding JNDE pixel luminance value, $\overline{P}_{L,k}(i)$ is the average luminance value of the corresponding block, and $\mathrm{JND}_k^W(i)$ is the weighted version of $\mathrm{JND}_{j,k}(x,y)$ in the jth block of the kth MB, i.e.,

$$\mathrm{JND}_k^W(i) = \begin{cases} \mathrm{JND}_{1,k}(x,y) \cdot WM(x,y), \\ \quad \text{if } i\text{th pixel is the } (x,y) \text{ position of subblock 1}, \\ \mathrm{JND}_{2,k}(x,y) \cdot WM(x,y), \\ \quad \text{if } i\text{th pixel is the } (x,y) \text{ position of subblock 2}, \\ \mathrm{JND}_{3,k}(x,y) \cdot WM(x,y), \\ \quad \text{if } i\text{th pixel is the } (x,y) \text{ position of subblock 3}, \\ \mathrm{JND}_{4,k}(x,y) \cdot WM(x,y), \\ \quad \text{if } i\text{th pixel is the } (x,y) \text{ position of subblock 4}, \end{cases} \tag{6}$$

where $WM(x,y)$ is the weighted mask for JND profile, as shown in Fig. 2(a). Eq. (3) is used to adjust the pixel values in INTRA MBs closer to the mean value of blocks. Adjusting the original luminance values with a JND rule, it guarantees that:

$$P_{L,k}^{\mathrm{JNDE}}(i) - P_{L,k}(i) \le \mathrm{JND}_k^W(i).$$

It is noted that if the difference between the original and reconstructed values is below JND, the error visibility is rendered imperceptibly by the HVS.

2.3 Proposed Rate Control Approach for Macroblock

Layer

The bit estimation model is used to predict the required number of compressed bits for an MB when a certain quantization step size is used. In this study, a modified version of the bit estimation model derived in [15] is employed, which is the low bit rate model for scalar quantization and entropy coding and given by:

$$B_k = K\frac{\omega_k^2}{Q_k^2} + H_k, \quad (7)$$

where $Q_k = 2QP_k$ is the quantization step size for the kth MB in H.263 syntax, K is a constant depending on the statistical properties of the kth MB, H_k is the bits required to encode the motion vector and the header information of the kth MB, and ω_k^2 is the average weighted sum square of the zig-zag ordered DCT coefficients (except the DC term if the kth MB is INTRA coded), i.e.,

$$\omega_k^2 = \frac{1}{a}\sum_{j=1}^{6}\sum_{i=b}^{8\cdot 8}\left(f_{j,k}(i)\right)^2 \cdot c(i), \quad (8)$$

$$a = 6\sum_{i=b}^{8\cdot 8} c(i),$$

$$b = \begin{cases} 1, & \text{if the } k\text{th MB is INTER coded,} \\ 2, & \text{if the } k\text{th MB is INTRA coded,} \end{cases}$$

$$c(i) = \frac{(i-1)}{63} + 1,$$

where $f_{j,k}(i)$ is the ith zig-zag ordered DCT coefficient of the jth block of the kth MB. Based on the properties of the uniform quantizer used in H.263, the distortion measure for the image frame containing N MBs is given by:

$$D = \sum_{k=1}^{N} D_k = \sum_{k=1}^{N}\frac{Q_k^2}{12}, \quad (9)$$

where D_k is the distortion of the kth MB in the image frame.

The constrained optimization problem is to optimally select $\{QP_1, QP_2, \ldots, QP_N\}$ for the N MBs within an image frame to minimize the distortion D, subject to the given frame bit budget constraint B, i.e.,

$$\min_{Q_1, Q_2, \ldots, Q_N} (D_1 + D_2 + \ldots + D_N), \quad \text{subject to} (B_1 + B_2 + \ldots + B_N) \le B. \quad (10)$$

Eq. (10) can be transformed into an equivalent unconstrained optimization problem by using the Lagrange multiplier, i.e., to minimize the Lagrangian cost:

$$\min_{Q_1, Q_2, \ldots, Q_N} J_\lambda(Q_1, Q_2, \ldots, Q_N)$$
$$= \min_{Q_1, Q_2, \ldots, Q_N}\left[\begin{matrix}(D_1 + D_2 + \ldots + D_N) \\ + \lambda(B_1 + B_2 + \ldots + B_N)\end{matrix}\right]$$
$$= \min_{Q_1, Q_2, \ldots, Q_N}\left[\sum_{k=1}^{N}\frac{Q_k^2}{12} + \lambda\left(\sum_{k=1}^{N}\left(K\frac{\omega_k^2}{Q_k^2} + H_k\right)\right)\right]$$
$$= \min_{Q_1, Q_2, \ldots, Q_N}\left[\sum_{k=1}^{N}\frac{Q_k^2}{12} + \lambda\left(\sum_{k=1}^{N}\left(K\frac{\omega_k^2}{Q_k^2}\right) + H\right)\right], \quad (11)$$

where $J_\lambda(Q_1, Q_2, \ldots, Q_N)$ is the Lagrangian cost function, and $H = H_1 + H_2 + \ldots + H_N$ is the total number of bits required to encode the header information and motion vectors for the all of the MBs in the image frame.

Taking the partial derivative of Eq. (11) with respect to Q_k and equating it to zero yields, upon simplification,

$$Q_k^2 = \sqrt{12\lambda K}\cdot\omega_k. \quad (12)$$

Substituting Eq. (12) into the bit budget constraint, we have:

$$B = \sum_{k=1}^{N}\left(K\frac{\omega_k^2}{Q_k^2}\right) + H = \sum_{k=1}^{N}\left(K\frac{\omega_k^2}{\sqrt{12\lambda K}\cdot\omega_k}\right) + H,$$

or equivalently:

$$\sqrt{12\lambda K} = \frac{K}{B - H}\sum_{k=1}^{N}\omega_k. \quad (13)$$

Combining Eqs. (12) and (13), the optimized quantization parameter is:

$$Q_k = \sqrt{\frac{K}{B - H}\omega_k\sum_{m=1}^{N}\omega_m}, \qquad k = 1, 2, \ldots, N. \quad (14)$$

Moreover, if the former $(k-1)$ MBs in an image frame have been already quantized and encoded, the optimized quantization parameter QP for the kth MB becomes:

$$Q_k = \sqrt{\frac{K}{\tilde{B}_k - \tilde{H}_k}\omega_k\sum_{m=k}^{N}\omega_m}, \quad (15)$$

where $\tilde{H}_k$ is the total number of bits required to encode the header information and motion vectors for the remaining $(N-k+1)$ MBs in the image frame, and $\tilde{B}_k$ is the number of bits available for them.

If the quantization step size is changed, five bits are required. To reduce the overhead of five bits in low bit rate coding, the similar technique in [19] is employed, i.e.,

$$Q_k = \sqrt{\frac{K}{\tilde{B}_k - \tilde{H}_k}\frac{\omega_k}{\alpha_k}\sum_{m=k}^{N}(\alpha_m \cdot \omega_m)}, \quad (16)$$

$$\alpha_k = \begin{cases} \max\left(\frac{2B}{16\cdot 16 + 2\cdot 8\cdot 8}(1-\omega_k) + \omega_k,\ 1\right), \\ \text{if } \frac{B}{16\cdot 16 + 2\cdot 8\cdot 8} < 0.5, \\ 1, \quad \text{otherwise.} \end{cases} \quad (17)$$

The proposed approach for the MB layer is to select the optimal QP for each MB in the current image frame, which is described as follows.

Step 1. Initialization

Let $i = 1$, $j = 0$, $\tilde{B}_1 = B$, $\tilde{H}_1 = H$, $N_1 = N$, $K = K_1 = K_{pre}$, $S_1 = \sum_{k=1}^{N}\alpha_k\omega_k$, and H be given by:

$$H = \sum_{k=1}^{N} H_k + 50 + N_{INTRA}\cdot 48$$
$$= \sum_{k=1}^{N}(B_{k,MV} + 1) + 50 + N_{INTRA}\cdot 48$$
$$= \sum_{k=1}^{N}(B_{k,MV}) + N + 50 + N_{INTRA}\cdot 48,$$

where $B_{k,MV}$ denotes the number of bits required to encode the motion vector of the kth MB, and N_{INTRA} denotes the number of INTRA coded MBs in the image frame. If the current frame is the first frame to be encoded, set $K_1 = 128$; otherwise, set $K_1 = K_{pre}$.

Step 2. Compute Optimized Q for ith MB

If $L = (\tilde{B}_i - \tilde{H}_i) \le 0$, set $Q_i = 62$; otherwise, compute:

$$Q_i = \sqrt{\frac{K}{L}\frac{\omega_i}{\alpha_i}} S_i.$$

Step 3. Find QP and Encode Macroblock

Let $QP = \text{round}(Q_i/2)$, rounding to the nearest integer. If $QP > 31$, compute:

$f^*_{j,i}(m) = f_{j,i}(m) \cdot SM(m),\ 1 \le m \le 64,$

where $f_{j,i}(m)$ denotes the mth DCT coefficient of the jth block of the ith MB that is INTRA mode coded, $f^*_{j,i}(m)$ denotes the corresponding modified value, and the $SM(m)$ matrix is shown in Fig. 2(b) [11]. After the truncation, set $QP = 31$. Set $DQUANT = QP - QP_{pre}$.

Within the H.263 standard, $DQUANT$ must be within {-2, -1, 0, 1, 2}. If $DQUANT > 2$, set $DQUANT = 2$. If $DQUANT < TH_{mode}$, the coding mode of the current MB is changed from INTER to INTRA, where

$$TH_{mode} = \begin{cases} -2, & \text{if } 1 \le QP_{pre} < 5, \\ -4, & \text{if } 5 \le QP_{pre} < 10, \\ -5, & \text{if } 10 \le QP_{pre} < 20, \\ -10, & \text{otherwise.} \end{cases}$$

Set $QP = QP_{pre} + DQUANT$. Set $QP_{pre} = QP$.

Step 4. Update Counters

Let B'_i denotes the actual number of bits used to encode the ith MB, and H'_i be the number of bits for the header of ith MB, i.e.,

$$H'_i = \begin{cases} 1 + 50 + B_{i,MV} + I, & \text{if the } i\text{th MB is the first MB in the image frame,} \\ 1 + B_{i,MV} + I, & \text{otherwise,} \end{cases}$$

$$I = \begin{cases} 48, & \text{if the } i\text{th MB is INTRA coded,} \\ 0, & \text{otherwise.} \end{cases}$$

Compute:

$\tilde{B}_{i+1} = \tilde{B}_i - B'_i$, $\tilde{H}_{i+1} = \tilde{H}_i - H'_i$, $S_{i+1} = S_i - \alpha_i\omega_i$, and $N_{i+1} = N_i - 1$.

Step 5. Update Model Parameters K

The model parameter measured for the ith MB is:

$$\hat{K} = \frac{B'_{LC,i}(2QP)^2}{\omega_i^2},$$

$$B'_{LC,i} = \begin{cases} B'_i - H'_i, & \text{if } B'_i > 1, \\ 3, & \text{otherwise,} \end{cases}$$

where $B'_{LC,i}$ is the number of bits required to encode the luminance and chrominance components of the ith MB. If $\hat{K} > 0$ and $\hat{K} \le TH_{\hat{K}}$, set $j = j + 1$ and compute $\tilde{K}_j = \tilde{K}_{j+1}(j-1)/j + \hat{K}/j$. $TH_{\hat{K}}$ is a modified threshold given by [18]: $TH_{\hat{K}} = 256 \cdot \pi \log_2 e$. Finally, K is given by:

$K = \tilde{K}_j(i/N) + K_1(N-i)/N$.

Step 6.

If $i = N$, set $K_{pre} = K$ and go to Step 1 to process the next image frame; otherwise, set $i = i + 1$, and go to Step 2.

3. SIMULATION RESULTS

The proposed rate control scheme for H.263 images has been implemented on a Pentium-II 300 PC using the ANSI-C programming language. Seven QCIF (176×144 pixels) test image sequences, "Claire," "Foreman," "Grandmother," "Miss America," "Mother and Daughter," "Salesman," and "Trevor" are used to evaluate the performance of the proposed approach.

The peak signal-to-noise ratio (PSNR) is employed in this study as the first objective performance measure for each component (Y, C_B, and C_R) of H.263 images. The PSNR of the ith image frame within an image sequence, denoted by PSNR_i, is defined as:

$$\text{PSNR}_i = (4 \times \text{PSNR}_{Y,i} + \text{PSNR}_{C_B,i} + \text{PSNR}_{C_R,i})/6, \quad (18)$$

where $\text{PSNR}_{Y,i}$, $\text{PSNR}_{C_B,i}$, and $\text{PSNR}_{C_R,i}$ are the PSNRs of the Y, C_B, and C_R components of the ith image frame, respectively. The average PSNR of an image sequence, denoted by PSNR_{SEQ}, is defined as:

$$\text{PSNR}_{\text{SEQ}} = \sum_{i=1}^{n} \text{PSNR}_i / n, \quad (19)$$

where n is the total number of image frames within the sequence. Moreover, the peak signal-to-perceptible-noise ratio (PSPNR) of the ith image frame, denoted by $\text{PSPNR}_{Y,i}$, is given by:

$$\text{PSPNR}_{Y,i} = 10 \cdot \log_{10} \frac{255^2}{\frac{1}{H \cdot W}\sum_{x=0}^{H-1}\sum_{y=0}^{W-1}\left(\left(|p(x,y) - \hat{p}(x,y)| - \text{JND}(x,y) \right)^2 \cdot \delta(x,y) \right)}, \quad (20)$$

$$\delta(x,y) = \begin{cases} 1, & \text{if } |\hat{p}(x,y) - p(x,y)| > \text{JND}(x,y), \\ 0, & \text{if } |\hat{p}(x,y) - p(x,y)| \le \text{JND}(x,y), \end{cases} \quad (21)$$

where H and W are the height and width of the ith image frame, $p(x,y)$, $\hat{p}(x,y)$ and $\text{JND}(x,y)$ denote the original, reconstructed, and JND profile of the image pixel (x, y) within the ith image frame, respectively. $\text{PSPNR}_{Y,i}$ is used to evaluate the visual quality of the luminance (Y) component for each image frame in term of the perceptual distortion energy. The average PSPNR of the Y component of an image sequence, denoted by $\text{PSPNR}_{\text{SEQ,Y}}$, can be similarly defined. To illustrate the effectiveness of the proposed JND preprocessing for INTRA coded MBs. Table 1 shows the coding results (with QP=15) for each first image frame of three among the seven test image sequences without and with the proposed JND preprocessing. All MBs in the image frame are INTRA mode coded.

The seven test image sequences are tested at three target channel rates (R), 56000 bps, 48000 bps, and 33600 bps with the frame rate (F) being 10 fps. To compare the performances between the proposed rate control scheme and TMN9, three criteria, namely: target bits accuracy, channel buffer fullness (delay), and average PSNR, are employed. The accuracy of achieving the target bits of the ith P-frame, denoted by $A_{\mathrm{F},i}$, is given by:

$$A_{\mathrm{F},i} = \frac{B'_{\mathrm{F},i} - B_{\mathrm{F},i}}{B_{\mathrm{F},i}},$$

where $B'_{\mathrm{F},i}$ is the actual number of bits used to encode the ith P-frame, and $B_{\mathrm{F},i}$ is the target bits of the ith P-frame. The average accuracy of achieving the target bits of an image sequence, denoted by $A_{\mathrm{F,SEQ}}$, is given by:

$$A_{\mathrm{F,SEQ}} = \frac{1}{n}\sum_{i=1}^{n} A_{\mathrm{F},i},$$

where n is the number of P-frames within the sequence. The accuracy of achieving the target bits for the channel of the ith P-frame, denoted by $A_{\mathrm{C},i}$, is given by:

$$A_{\mathrm{C},i} = \frac{B_{\mathrm{F},i} - M}{M},$$

where $M = R/F$ is the target channel bits per frame. The average accuracy of achieving the target bits for the channel of an image sequence, denoted by $A_{\mathrm{C,SEQ}}$, is given by:

$$A_{\mathrm{C,SEQ}} = \frac{1}{n}\sum_{i=1}^{n} A_{\mathrm{C},i}.$$

In terms of $A_{\mathrm{F,SEQ}}$, $A_{\mathrm{C,SEQ}}$, the average bit rate, and the number of skipped image frames, the performance comparison between TMN9 and the proposed approach for three among the seven test image sequences are listed in Table 2. In terms of the number of bits per frame, the performance comparison between TMN9 and the proposed approach for the "Claire" sequence at channel rate 33.6 kbps is shown in Fig. 3. In terms of the channel buffer fullness, the performance comparison between TMN9 and the proposed approach for the "Claire" sequence at channel rate 33.6 kbps is shown in Fig. 4. The horizontal straight line indicates the value of $M = R/F$ (the threshold used for frame skipping). In terms of $\mathrm{PSNR_{SEQ}}$, $\mathrm{PSNR_{SEQ,Y}}$, $\mathrm{PSNR_{SEQ,C_B}}$, and $\mathrm{PSNR_{SEQ,C_R}}$ (dB), the performance comparison between TMN9 and the proposed approach for three among the seven test image sequences are listed in Table 3.

4. CONCLUDING REMARKS

Based on the simulation results obtained in this study, several observations can be found. (1) The larger the channel rate is, the better the rate control result will be. (2) The less activities of the image contents of an image sequence are, the more accuracy of achieving the target bits will be. (3) Based on the simulation results shown in Table 1, the JNDE images need less bits (7%-15%) than the original images, whereas the JNDE images have higher PSPNR values. (4) Based on the simulation results shown in Table 2, the proposed rate control scheme achieves a bit rate closer to the target channel rate than TMN9 does, especially for the test sequences containing quick movements or scene changed images. (5) Based on the simulation results shown in Fig. 4, the proposed rate control scheme achieves a nearly constant of bits per frame with a well accuracy of target bits, whileas TMN9 oscillates in many test sequences. (6) Based on the simulation results shown in Fig. 5, the buffer fullness of the proposed approach is almost lower than that of TMN9. (7) Based on the simulation results shown in Table 3, the PSNR values of the proposed approach are comparable to that of TMN9.

In this study, a new rate control scheme for H.263 images is proposed. Based on the simulation results obtained in this study, the proposed approach can meet the target bit rate more accurately, keep a lower buffer fullness, and a better temporal resolution than TMN9, whereas the PSNR of the proposed approach is as good as that of TMN9. This shows the feasibility of the proposed approach.

REFERENCES

[1] J. L. Mitchell, W. B. Pennebaker, C. E. Fogg, and D. J. LeGall (Eds.), *MPEG Video Compression Standard.* New York: Chapman & Hall, 1997.

[2] W. Ding, "Joint encoder and channel rate control of VBR video over ATM networks," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 7, no. 2, pp. 266-278, 1997.

[3] ITU-T, "Video coding for low bitrate communication," Recommendation H.263, 1996.

[4] ITU-T SG 15 Expert's Group for Very Low Bitrate Visual Telephony, "Video codec test model for the near-term 5 (TMN5)," 1995.

[5] ISO/IEC JTC1/SC29/WG11, "Test model 5 (TM5)," Document N400, 1993.

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

[7] T. Chiang and Y. Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 7, no. 1, pp. 246-250, 1997.

[8] ISO/IEC JTC1/SC29/WG11, "H.263 anchors for MPEG-4 tests," Document MPEG95/N0322, 1995.

[9] C. F. Chang and J. S. Wang, "A stable buffer control strategy for MPEG coding," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 7, no. 6, pp. 920-924, 1997.

[10] C. T. Chen and A. Wong, "A self-governing rate buffer control strategy for pseudoconstant bit rate video coding," *IEEE Trans. on Image Processing,* vol. 2, no. 1, pp. 50-59, 1993.

[11] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proceedings of the IEEE,* vol. 81, no. 10, pp. 1385-1422, 1993.

[12] M. R. Pickering and J. F. Arnold, "A perceptually efficient VBR rate control algorithm," *IEEE*

Trans. on Image Processing, vol. 3, no. 5, pp. 527-532, 1994.

[13] K. W. Chun, K. W. Lim, H. D. Cho, and J. B. Ra, "An adaptive perceptual quantization algorithm for video coding," *IEEE Trans. on Consumer Electronics,* vol. 39, no. 3, pp. 555-558, 1993.

[14] Y. S. Saw, P. M. Grant, J. M. Hannah, and B. Mulgrew, "Nonlinear predictive rate control for constant bit rate MPEG video coders," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing,* Munich, Germany, 1997, vol. 4, pp. 2641-2644.

[15] J. Ribas-Corbera and D. L. Neuhoff, "On the optimal motion vector accuracy for block-based motion-compensated video coders," in *Proceedings of the SPIE Conference on Digital Video Compression: Algorithms and Technologies,* San Jose, USA, 1996, vol. 2668, pp. 302-314.

[16] ITU-T/SG16 Video Coding Experts Group, "Video codec test model, near-term, version 9 (TMN9)," Document Q15-C-15, 1997.

[17] J. Lee and B. W. Dickinson, "Temporally adaptive motion interpolation exploiting temporal masking in visual perception," *IEEE Trans. on Image Processing,* vol. 3, no. 5, pp. 513-526, 1994.

[18] C. H. Lin and J. L. Wu, "Content-based rate control scheme for very low bit-rate video coding," *IEEE Trans. on Consumer Electronics,* vol. 43, no. 2, pp. 123-133, 1997.

[19] C. H. Chou and Y. C. Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," *IEEE Trans. on Circuits and Systems for Video Tech.,* vol. 5, no. 6, pp. 467-476, 1995.

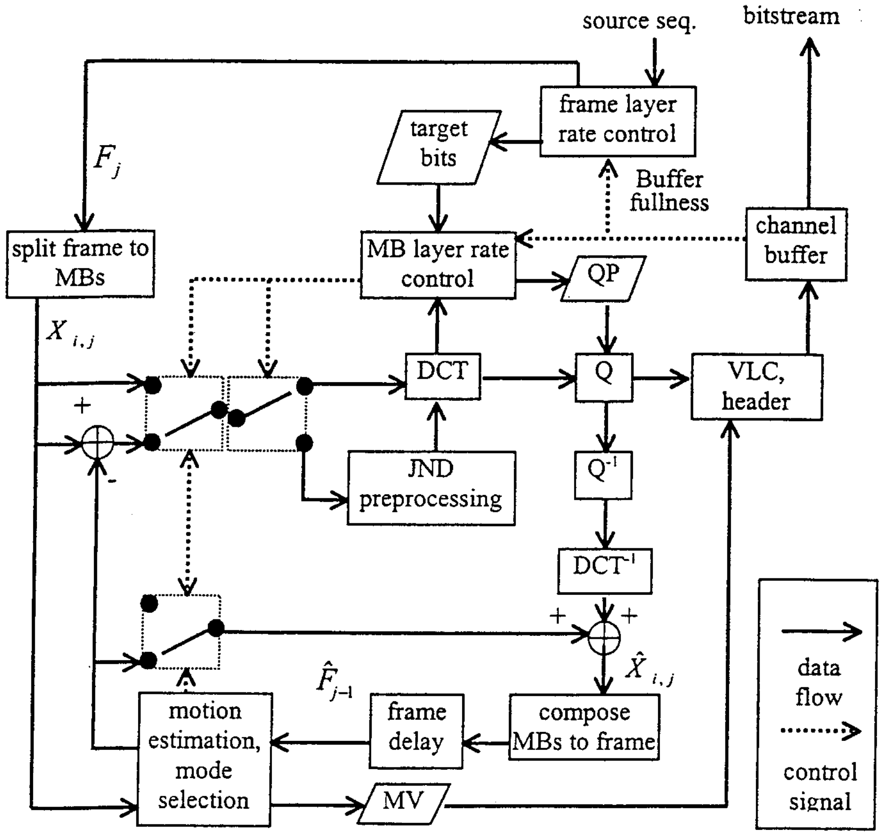

Fig. 1. The proposed rate control scheme embedded in H.263 coder.

0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8
0.8	1	1	1	1	1	1	0.4
0.4	1	1	1	1	1	1	0.8
0.8	1	1	1	1	1	1	0.4
0.4	1	1	1	1	1	1	0.8
0.8	1	1	1	1	1	1	0.4
0.4	1	1	1	1	1	1	0.8
0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4

(a)

1	1	0.8	0.8	0.8	0.5	0.5	0.5
1	1	0.8	0.8	0.8	0.5	0.5	0.5
0.8	0.8	0.8	0.8	0.5	0.5	0.5	0.5
0.8	0.8	0.8	0.5	0.5	0.5	0.5	0.5
0.8	0.8	0.5	0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

(b)

Fig. 2. (a) The weighted mask, $WM(x, y)$, for the JND profile, and (b) the $SM(m)$ matrix for modifying the DCT coefficients of the INTRA MB.

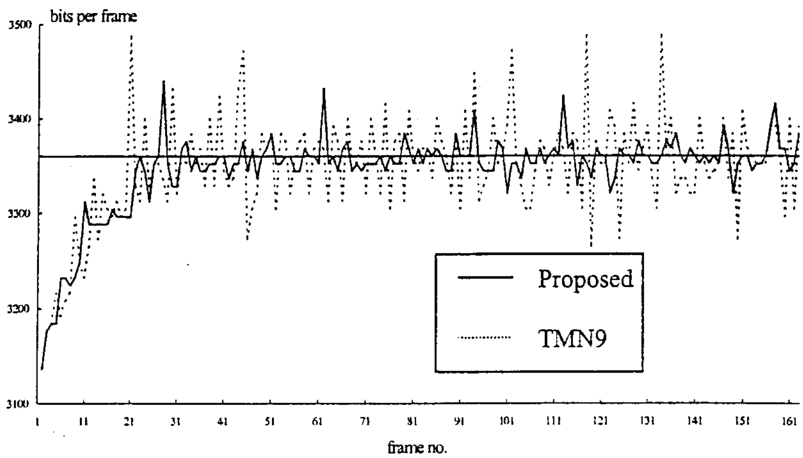

Fig. 3. The number of bits per frame for the "Claire" sequence by TMN9 and the proposed approach at channel bit rate 33.6 kbps. The horizontal straight line indicates the target channel bits per frame ($M = R / F$).

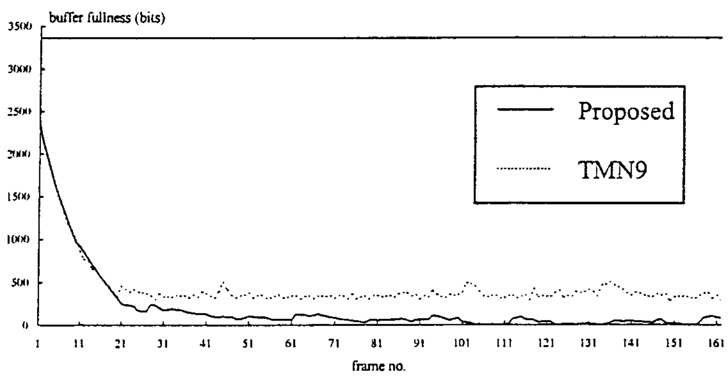

Fig. 4. The channel buffer fullness per frame for the "Claire" sequence by TMN9 and the proposed approach at channel bit rate 33.6 kbps. The horizontal straight line indicates the threshold used for frame skipping ($M = R / F$).

Table 1. The coding results (with QP=15) in terms of PSNR, PSPNR, bits for the first image frame of three among the seven test image sequences, and total bit reduction (%) of the proposed JND preprocessing.

Image sequence	Original images		JNDE images			bit reduction (%)
	$PSNR_{Y,0}$	Bits	$PSNR_{Y,0}$	$PSPNR_{Y,0}$	Bits	
Claire	34.73	12444	32.22	38.36	11480	7.75
Foreman	30.82	16176	29.26	33.28	14448	10.68
Grandmother	31.77	11080	30.10	35.80	9600	13.36

Table 2. The coding results in terms of the average bit rate, $A_{F,SEQ}$, $A_{C,SEQ}$, and the number of the skipped image frames for three among the seven test image sequences by TMN9 and the proposed approach.

Image sequence	Target channel bit rate (kbps)	TMN9				Proposed			
		Obtained average bit rate (kbps)	$A_{F,SEQ}$ (%)	$A_{C,SEQ}$ (%)	No. of skipped frames	Obtained average bit rate (kbps)	$A_{F,SEQ}$ (%)	$A_{C,SEQ}$ (%)	No. of skipped frames
Claire	33.60	33.62	0.68	1.23	0	33.60	0.39	0.71	0
	48.00	48.03	0.63	1.01	0	48.00	0.31	0.56	0
	56.00	56.04	0.54	0.80	0	56.01	0.30	0.32	0
Foreman	33.60	33.63	4.90	4.64	3	33.61	1.06	1.49	0
	48.00	48.04	1.36	1.36	0	48.01	0.74	0.62	0
	56.00	56.04	1.07	1.38	0	56.00	0.57	0.95	0
Grandmother	33.60	33.61	0.93	1.37	0	33.60	0.69	0.80	0
	48.00	48.01	0.90	1.30	0	48.00	0.45	0.56	0
	56.00	56.04	0.86	1.46	0	55.99	0.36	0.68	0

Table 3. The simulation results in terms of $PSNR_{SEQ}$, $PSNR_{SEQ,Y}$, $PSNR_{SEQ,C_B}$, and $PSNR_{SEQ,C_R}$ (dB) for three among the seven test image sequences by TMN9 and the proposed approach.

Image sequence	Target channel bit rate (kbps)	TMN9				Proposed			
		Average PSNR	Average PSNR of Y	Average PSNR of C_B	Average PSNR of C_R	Average PSNR	Average PSNR of Y	Average PSNR of C_B	Average PSNR of C_R
Claire	33.60	40.31	39.70	40.38	42.65	40.29	39.68	40.30	42.72
	48.00	41.94	41.37	41.98	44.18	41.75	41.20	41.68	44.04
	56.00	42.64	42.11	42.70	44.72	42.41	41.89	42.30	44.63
Foreman	33.60	31.25	29.10	35.33	35.78	31.03	28.85	35.15	35.62
	48.00	32.37	30.40	36.01	36.60	32.38	30.43	36.06	36.52
	56.00	32.92	31.01	36.04	37.03	32.90	31.01	36.38	37.02
Grandmother	33.60	37.99	36.80	40.18	40.57	37.98	36.83	40.11	40.45
	48.00	39.39	38.31	41.38	41.69	39.28	38.23	41.22	41.56
	56.00	40.07	39.06	41.92	42.28	39.86	38.86	41.67	42.04