Advanced Real-time Rate Control in H.264

Chi-Wah Wong, Oscar C. Au, Raymond Chi-Wing Wong+*
Hong Kong Univ. of Science and Technology, Hong Kong +the Chinese University of Hong Kong, Hong Kong
Email: {dickywcw, eeau}@ust.hk Email: {cwwong@cse.cuhk.edu.hk}*

Abstract

Most existing rate control schemes in the literature use one rate model and calculate quantization parameters of the macro-blocks (MB), regardless of MB types. In advanced video coding standards such as H.264, MBs belong to more advanced MB types, such as skipped and non-skipped MBs. In non-skipped MBs, the encoder determines whether each of 8x8 luminance sub-blocks and 4x4 chrominance sub-block of a MB is to be encoded, giving the different number of sub-blocks at each MB encoding times. As a result, a traditional single rate model is insufficient to represent each MB accurately. In this work, it is found that different MB types have different rate behavior. Under different conditions of MB types, we establish novel different rate models and distortion models. Our rate control scheme is proposed based on these models. The experimental results suggest that our scheme can achieve PSNR gain over JM10.2 and TMN8.

Index Terms- *Rate Control, H.264, rate and distortion models*

1. Introduction

Standard video systems, such as H.261/3/4 and MPEG, exploit the spatial, temporal and statistical redundancies in the source video. Since the level of redundancy changes from frame to frame, the number of bits per frame is variable, even if the same quantization parameters are used for all frames. Therefore, a buffer is required to smooth out the variable video output rate and provide a constant video output rate. Rate control is used to prevent the buffer from over-flowing (resulting in frame skipping) or/and under-flowing (resulting in low channel utilization) in order to achieve good video quality. For real-time video communications such as video conferencing, it is more challenging as the rate control is required to satisfy the low-delay constraints, especially in low bit rate channels.

In the literature, most existing rate control schemes ([1], [2], [3]) calculate quantization parameters of macro-blocks (MB) based on one rate model, regardless of MB types. In advanced video coding standards such as H.264 [2], MBs belongs to more advanced MB types, such as skipped and non-skipped MBs, which is normally ignored in traditional rate control schemes. As a result, one rate model is less accurate under the same assumption of each MB. In this work, we present a novel model-based rate control scheme for encoders in real-time video communications. This work focuses on doing rate control for inter-coded frames (i.e. P-frame), which is used mostly in low-delay video communication. Using Lagrange optimization, we minimize the distortion subject to the target bit constraint and obtain formulas that indicate how to choose the quantization parameters.

This paper is organized as follows. In the following section, we describe the MB characteristic in H.264. In section 3, we propose novel rate and distortion models. Based on these models, we find the optimal quantization step size that minimizes distortion subject to the target bit constraint. In section 4, our proposed rate control scheme is described. Then the experiments are conducted to evaluate the performance. Finally, the conclusion is made.

2. MB Characteristic in H.264

In H.264, frames are divided into *N* macro-blocks of 16x16 luminance samples each, with two corresponding 8x8 chrominance samples. In QCIF picture format, there are 99 macro-blocks for each frame. A number of consecutive macro-blocks in raster-scan order can be grouped into slices, representing independent coding units to be decoded without referencing other slices of the same frame.

Given that the whole frame is adopted as a unit slice, the frame header is encoded and *N* MBs are processed one by one. The resulting MB syntax is MB header followed by MB residue data. In P-frame, the MB header basically consists of run-length, MB mode, motion vector data, coded block pattern (*CBP*) and change of quantization parameter. When the MB header starts to be encoded, the run-length indicates the number of skipped macro-blocks that are made by copying the co-located picture information from the last decoded frame. Table 1 shows the relative percentage of the number of skipped MBs (MB_s) and non-skipped MBs (MB_N) in H.264. The experimental conditions are described as follows. Software version is JM 10.2, the picture format is QCIF, the encoded frame rate is 10fps, the structure of GOP is IPPP, maximum search range is 16, the number of reference frame is 1 and the entropy coding method is UVLC.

Video Sequence	*QP*	MB_s (%)	MB_N (%)
Akiyo	15	43.4	56.6
	35	85.3	14.7
	45	95.9	4.1
Foreman	15	0.1	99.9
	35	30.8	69.2
	45	61.0	39.0
Stefan	15	0.2	99.8
	35	17.8	82.2
	45	47.2	52.8

Table 1: Relative percentage of the number of skipped MBs and non-skipped MBs in H.264

It is observed that for any video sequences, the percentage of skipped MBs increases with *QP* as skipped MBs can save more bits with reasonable video quality. It is also noticed that fast-motion video sequence such as "Stefan" requires more non-skipped MBs compared with other sequences at any given *QP* because the use of dominant skipped MBs cannot give reasonable video quality in fast-motion sequences.

In MB header, *CBP* determines the number of Y/UV sub-blocks and their encoded bits. Four bits of 6-bit *CBP* (called *CBPY* [2]) indicates whether each of 4 8x8 luminance (Y) sub-blocks contains non-zero coefficients. In binary representation, the values "0" and "1" represent that the corresponding 8x8 sub-block has no coefficient and non-zero coefficients respectively. In chrominance (UV) sub-blocks, there are three possible CBP (called *nc*) ((1) no chrominance coefficients at all, (2) Only DC coefficients, (3) DC and AC coefficients). Table 2 shows the percentage of zero Y ($MB_{N,\bar{Y}}$), non-zero Y ($MB_{N,Y}$), zero UV ($MB_{N,\overline{UV}}$) and non-zero UV ($MB_{N,UV}$) MBs in the non-skipped mode.

It is observed that the percentage of $MB_{N,\bar{Y}}$ and $MB_{N,\overline{UV}}$ increases with *QP* for any video sequences. In these MBs, the Y/UV sub-blocks are skipped for quantization and encoding. Only MB header is required for processing. It is also noticed that the percentages of $MB_{N,Y}$ and $MB_{N,UV}$ are higher in

fast-motion "Stefan" sequence since the use of dominant $MB_{N,\bar{Y}}$ and $MB_{N,\overline{UV}}$ does not give reasonable video quality. From the above results, it is implied that each MB has different characteristic, including skipped and non-skipped MBs. In non-skipped MBs, the number of Y and UV sub-blocks can change based on *CBP* parameters. Therefore, a standalone rate model cannot be sufficient when a wide variety of MB characteristics is used in advanced video standards such as H.264.

Video Sequence	*QP*	Non-skipped MB (%)			
		$MB_{N,\bar{Y}}$	$MB_{N,Y}$	$MB_{N,\overline{UV}}$	$MB_{N,UV}$
Akiyo	15	29.1	70.9	27.5	72.5
	35	13.5	86.5	87.6	12.4
	45	47.7	52.3	89.1	10.9
Foreman	15	0.9	99.1	6.8	93.2
	35	25.5	74.5	79.3	20.7
	45	56.7	40.3	81.4	18.6
Stefan	15	1.2	98.8	4.7	95.3
	35	12.9	87.1	38.5	61.5
	45	35.5	64.5	70.9	29.1

Table 2: Percentage of zero Y, non-zero Y, zero UV and non-zero UV MBs in the non-skipped mode

3. Optimization

3.1. Rate Model for the residue frame

As discussed earlier, the number of encoded bits in a MB depends on either skipped or non-skipped MB. In a non-skipped MB, the coded block pattern (*CBP*) also determines bit contribution on non-zero luminance and chrominance coefficients of the MB. Depending on MB data characteristic, the quadratic rate model is adopted ([1], [3], [5]) and the new rate model of the *i*-th MB is shown in Table 3. Some existing rate control schemes [6] use the linear rate model instead. In the following, the quadratic rate model is addressed.

MB type:	Property (determined by CBP):	Rate Model, R_i (without overhead)
Non-skipped (MB_N)	Non-zero Y ($MB_{N,Y}$)	$A_{Y,i}K_Y\sigma_{Y,i}^2/Q_i^2$
	Zero Y ($MB_{N,\bar{Y}}$)	0
	Non-zero UV ($MB_{N,UV}$)	$A_{UV,i}K_{UV}\sigma_{UV,i}^2/Q_i^2$
	Zero UV ($MB_{N,\overline{UV}}$)	0
Skipped (MB_S)	Zero Y and Zero UV	0

Table 3: Rate model for the *i*-th MB

Where $A_{Y,i}$ is the number of pixels for Y coefficients of non-zero 8x8 blocks in a MB, $A_{UV,i}$ is the number of pixels for UV coefficients of non-zero DC and AC blocks in a MB, K_Y is the model parameter of Y coefficient in rate model, K_{UV} is the model parameter of UV coefficient in rate model, $\sigma_{Y,i}$ is standard deviation of the residue for Y coefficient of non-zero 8x8 blocks in the *i*-th MB, $\sigma_{UV,i}$ is standard deviation of the residue for UV coefficient of non-zero DC and AC blocks in the *i*-th MB, and Q_i is quantization step size of the *i*-th MB.

$A_{Y,i}$ is set to 256, 192, 128 or 64 for 4, 3, 2, or 1 8x8 non-zero Y blocks in a MB respectively. Similarly, $A_{UV,i}$ is set to 8 or 128 for non-zero DC blocks or non-zero UV blocks respectively in accordance with available *CBP* settings [2]. K_Y and K_{UV} are updated after encoding $MB_{N,Y}$ and $MB_{N,UV}$ respectively, as described later. It is noticed that the rate model (without overhead) gives non-zero value for $MB_{N,Y}$ and $MB_{N,UV}$. Otherwise, the model is set to zero due to no bit contribution on $MB_{N,\bar{Y}}$, $MB_{N,\overline{UV}}$ and MB_S.

The rate model (including overhead) within a frame is given by:

$$R = \sum_{i\in MB_{N,Y}} A_{Y,i}K_Y\frac{\sigma_{Y,i}^2}{Q_i^2} + \sum_{i\in MB_{N,UV}} A_{UV,i}K_{UV}\frac{\sigma_{UV,i}^2}{Q_i^2} + C \quad (1)$$

$$C = \sum_{i\in MB_{N,Y}\cap MB_{N,UV}} C_{Y,UV,i} + \sum_{i\in MB_{N,Y}\cap MB_{N,\overline{UV}}} C_{Y,\overline{UV},i} + \sum_{i\in MB_{N,\bar{Y}}\cap MB_{N,UV}} C_{\bar{Y},UV,i} + \sum_{i\in MB_{N,\bar{Y}}\cap MB_{N,\overline{UV}}} C_{\bar{Y},\overline{UV},i}$$

where $C_{Y,UV,i}$, $C_{Y,\overline{UV},i}$, $C_{\bar{Y},UV,i}$ and $C_{\bar{Y},\overline{UV},i}$ are the overhead of the *i*-th MB, belonging to the sets $MB_{N,Y}\cap MB_{N,UV}$, $MB_{N,Y}\cap MB_{N,\overline{UV}}$, $MB_{N,\bar{Y}}\cap MB_{N,UV}$ and $MB_{N,\bar{Y}}\cap MB_{N,\overline{UV}}$ respectively. They are adopted in their own because they belong to different *CBP* values and give different overhead bits, making more accurate overhead prediction. For skipped MBs MB_S, the MB overhead is set to zero as these MBs do not introduce overhead bits.

3.2. Distortion Model for the residue frame

The following simplified distortion model, defined in the similar way as [3], of the *i*-th MB is given as follows:

$$D_i = \alpha_i^2 Q_i^2/12 \quad (2)$$

It is known that the distortion is $Q^2/12$ for uniform quantization of uniform distribution [3]. In DCT coefficients likely following the Laplacian distribution [3], the distortion is not only dependent on Q, but also α_i, which increases with $\sigma_{Y,i}$ and $\sigma_{UV,i}$. It is observed that the distortion increases with Q_i and α_i. Although our model does not estimate the distortion actually, this simplified expression gives important advantages. As seen later, the closed form of optimal quantization factor is found and low quantization overhead (ΔQP) is made by introducing the term α_i. Also, no distortion model parameter is required to be updated. Some previous works (e.g. [6]) propose the distortion model derived from distortion definition with some approximations. Such model requires computation-demanding operations to update the model parameters due to the high complexity function of such kind of model.

The distortion model within a frame is given by:

$$D = \frac{1}{N_{Y\cup UV}}\sum_{i\in MB_{N,Y}\cup MB_{N,UV}} \alpha_i^2\frac{Q_i^2}{12} + \frac{1}{N_{\bar{Y}\cap\overline{UV}}}\sum_{i\in MB_{N,\bar{Y}}\cap MB_{N,\overline{UV}}} D_{\bar{Y}\cap\overline{UV},i} + \frac{1}{N_S}\sum_{i\in MB_S} D_{S,i} \quad (3)$$

where $N_{Y\cup UV}$, $N_{\bar{Y}\cap\overline{UV}}$ and N_S are the number of MBs belonging to either $MB_{N,Y}$ or $MB_{N,UV}$ in a frame, the number of MBs belonging to both $MB_{N,\bar{Y}}$ and $MB_{N,\overline{UV}}$ in a frame and the number of skipped MBs in a frame respectively. $D_{\bar{Y}\cap\overline{UV},i}$ and $D_{S,i}$ are distortion of the *i*-th MB belonging to both $MB_{N,\bar{Y}}$ and $MB_{N,\overline{UV}}$, and distortion of the *i*-th skipped MB in a frame respectively. It is noticed that change in Q_i affects the distortion of non-zero MBs (including any non-zero Y and UV blocks). As no Q_i is required for coding in the skipped MBs MB_S and non-skipped MBs (belonging to both $MB_{N,\bar{Y}}$ and $MB_{N,\overline{UV}}$), their distortion depends on the reconstructed MBs which copied from their co-located MBs and MV-offset MBs

in the previous frame respectively.

3.3. Quantization Step Size Optimization

In this section, we derive formulas for optimal quantization step size to minimize MSE distortion subject to the target bit constraint based on our models. In our problem, the important thing in real-time rate control is how to choose the quantization step size. The quantization step sizes are chosen based on the optimization formula. The original problem is

$$\{Q_i^* | i = 1,2,...,N\} = \underset{\{Q_i | i=1,2,...,N\}, R=B}{\arg\min} D \tag{4}$$

where Q_i^* is the optimal quantization step size of the *i*-th MB, *B* is the target number of bits for the frame and *N* is the number of MBs in a frame.

Using Lagrange optimization, the problem becomes

$$\{Q_i^*, \lambda^* | i = 1,2,...,N\} = \underset{\{Q_i, \lambda | i=1,2,...,N\}}{\arg\min} \left\{ \begin{array}{l} \frac{1}{N_{Y\cup UV}} \sum_{i \in MB_{N,Y} \cup MB_{N,UV}} \alpha_i^2 \frac{Q_i^2}{12} \\ + \frac{1}{N_{\overline{Y} \cap \overline{UV}}} \sum_{i \in MB_{N,\overline{Y} \cap MB_{N,\overline{UV}}}} D_{\overline{Y} \cap \overline{UV},i} + \frac{1}{N_S} \sum_{i \in MB_S} D_{S,i} \\ + \lambda \left[\sum_{i \in MB_{N,Y}} A_{Y,i} K_Y \frac{\sigma_{Y,i}^2}{Q_i^2} + \sum_{i \in MB_{N,UV}} A_{UV,i} K_{UV} \frac{\sigma_{UV,i}^2}{Q_i^2} + C - B \right] \end{array} \right\} \tag{5}$$

The expression for the optimization quantization step size Q_i^*, which is the key for our proposed rate control, is obtained as follows:

$$Q_i^* = \sqrt{\sigma_i S / [\alpha_i (B - C)]} \tag{6}$$

wher $S = S_Y + S_{UV}$, $S_Y = K_Y \sum_{j \in N_Y} \frac{A_{Y,j} \sigma_{Y,j}^2 \alpha_j}{\sigma_j}$, $S_{UV} = K_{UV} \sum_{j \in N_{UV}} \frac{A_{UV,j} \sigma_{UV,j}^2 \alpha_j}{\sigma_j}$

and $\sigma_i = \sqrt{A_{Y,i} K_Y \sigma_{Y,i}^2 + A_{UV,i} K_{UV} \sigma_{UV,i}^2}$

Similar to [3], α_i is defined as:

$$\alpha_i = \begin{cases} \sigma_i + [2B(1-\sigma_i)]/AN & B/AN < 0.5 \\ 1 & , otherwise \end{cases} \tag{7}$$

where *A* is the total number of pixels (including Y and UV coefficients) in a MB (i.e. 384) and *N* is the number of MBs in a frame (i.e. 99 in QCIF format). At lower bit rates, α_i will linearly approach the corresponding σ_i and progressively reduce the dynamic range of Q_i^* in Eq. (6), resulting in small overhead bits on ΔQP to be encoded in MB header.

4. Advanced Rate Control (ARC)

In this section, we will introduce our proposed rate control algorithm. We assume that the first frame is intra-coded (I-frame) with a fixed quantization parameter and all subsequent frames are encoded as P-frames. This means that they are predicted from the corresponding previous decoded frames using motion compensation and the residue is obtained. First, we do the rate control in the frame layer, which is similar to TMN8 [3]. After that, we do the rate control in macro-block level.

4.1. Frame-Layer Rate Control

The encoder buffer size *W* is updated before the current frame is encoded with the following formula:

$$W = \max(W_{prev} + B' - R/F, 0) \tag{8}$$

where W_{prev} is the previous number of bits in the buffer (initially set to zero), *B'* is the actual number of bits used for the encoded previous frame, *R* is the channel bit rate (bit per sec), and *F* is the frame rate (frame per sec).

After updating the buffer size, if *W* is larger than or equal to the predefined threshold *M (=R/F)*, the encoder skips encoding the frames until *W* is smaller than *M*. This means that buffer overflow will not occur at the cost of frame skipping.

The target number of bits *B* for the current frame is estimated as:

$$B = (R/F) - \Delta \tag{9}$$

where $\Delta = \begin{cases} W/F & W > 0.1M \\ W - 0.1M & , otherwise \end{cases}$

The buffer size *W* keeps the low target buffer level (i.e. *0.1M*) for real-time rate control with low communication delay.

4.2. Macro-block Layer Control

In H.264, quantization parameters are used in both rate control scheme and rate-distortion optimization (RDO), which results in the chicken and egg dilemma when rate control is used . RDO gives residue information (e.g. MAD) from given *QP* whereas rate control gives desired *QP* from given residue information. In [5], this problem is solved by a prediction method based on co-located residue information of the previous frame due to high correlation. In other words, $\sigma_{Y,i}$, $\sigma_{UV,i}$, σ_i, $A_{Y,i}$ and $A_{UV,i}$ and α_i of MBs, belonging either $MB_{N,Y}$ or $MB_{N,UV}$, in the current frame can be predicted and obtained with reference to the co-located MB information of the previous frame. Then coding mode of each MB is determined according to the most updated *QP*. The algorithm is shown as follows:

For each P-frame {

Compute all of $\sigma_{Y,i}$, $\sigma_{UV,i}$, σ_i, $A_{Y,i}$, $A_{UV,i}$ and α_i in a frame

Compute B_1=B, ($S_{Y,1}$ and $S_{UV,1}$) in Eq. (6), and C_1 in Eq. (1)

Set $S'_{Y,1}$, $S'_{UV,1}$, $\overline{S}'_{Y,1}$, and $\overline{S}'_{UV,1}$ be 0

For each MB (*i=1 to N*) {

(i) Compute $Q_i^* = \sqrt{\frac{\sigma_i (S_{Y,i} + S_{UV,i})}{\alpha_i (B_i - C_i)}}$ based on Eq. (6)

(ii) Use the corresponding QP^*_i

(iii) $QP_i = \max\{\min\{QP^*_i, QP_{i-1} + \Delta QP_{max}\}, QP_{i-1} + \Delta QP_{min}\}$

(iv) Quantize and encode the *i*-th MB

(v) Compute the following:

(a) $S'_{Y,i+1} = S'_{Y,i} + (A_i K_Y \sigma_{Y,i}^2 \alpha_i)/\sigma_i$ *if* $i \in MB_{N,Y}$

(b) $S'_{UV,i+1} = S'_{UV,i} + (A_{UV} K_{UV} \sigma_{UV,i}^2 \alpha_i)/\sigma_i$ *if* $i \in MB_{N,UV}$

(c) $\overline{S}'_{Y,i+1} = \overline{S}'_{Y,i} + (\overline{A}_{Y,i} K_Y \overline{\sigma}_{Y,i}^2 \overline{\alpha}_i)/\overline{\sigma}_i$ *if* $i \in \overline{MB}_{N,Y}$

(d) $\overline{S}'_{UV,i+1} = \overline{S}'_{UV,i} + (\overline{A}_{UV,i} K_{UV} \overline{\sigma}_{UV,i}^2 \overline{\alpha}_i)/\overline{\sigma}_i$ *if* $i \in \overline{MB}_{N,UV}$

(e) $S_{Y,i+1} = S_{Y,i} - S'_{Y,i+1}$ and $S_{UV,i+1} = S_{UV,i} - S'_{UV,i+1}$

(vi) C_{i+1}=C_i – the *i*-th MB overhead bits of the previous frame

(vii) $B_{i+1} = B_i$ – *(actual current bits, R')*

(viii) Update the rate model parameters K_Y and K_{UV} based on Table 3

}

}

where $\sigma_{Y,i}$, $\sigma_{UV,i}$, σ_i $A_{Y,i}$, $A_{UV,i}$, α_i, $MB_{N,Y}$ and $MB_{N,UV}$ are obtained in the co-located *i*-th MB in the previous frame whereas $\overline{\sigma}_{Y,i}$, $\overline{\sigma}_{UV,i}$, $\overline{\sigma}_i$, $\overline{A}_{Y,i}$, $\overline{A}_{UV,i}$, $\overline{\alpha}_i$, $\overline{MB}_{N,Y}$ and $\overline{MB}_{N,UV}$ are obtained after the *i*-th MB is encoded in the current frame. In the beginning of the first P-frame, $\sigma_{Y,i}$, $\sigma_{UV,i}$, $A_{Y,i}$, $A_{UV,i}$, K_Y and K_{UV} are initialized to 1, 1, 256, 128, 0.5 and 0.5 respectively because there is no co-located MBs. σ_i and α_i are then calculated accordingly. When the *i*-th MB starts to be encoded in the current frame, the optimal Q_i^* is calculated

(using Eq. (6)) and its corresponding QP^*_i is obtained. In case of B_i-$C_i < 0$, the QP^*_i are used from the previous one. Under the QP range constraint between $QP_{i-1} + \Delta QP_{max}$ and $QP + \Delta QP_{min}$, resulting QP is used to quantize and encode the i-th MB where QP_{i-1} is the QP of the last encoded MB. QP_{i-1} for the first MB of the current frame is chosen to be the average QP of the previous frame. If $\left(\overline{S}'_{Y,i} + \overline{S}'_{UV,i}\right) / \left(S'_{Y,i} + S'_{UV,i}\right) > 1.2$ for all the MBs except the first MB, then the residue energy is much larger in the current frame than the previous one and scene change may happen. It is not recommended that smaller QP is chosen. In this case, $\Delta QP_{min}=0$. Otherwise, $\Delta QP_{min}=-3$ to have more flexible QP choices. On the other hand, if $\left(\overline{S}'_{Y,i} + \overline{S}'_{UV,i}\right) / \left(S'_{Y,i} + S'_{UV,i}\right) < 0.8$ for all the MBs except the first MB, then the residue energy of the current frame is much smaller and buffer under-flow may happen. It is not recommended that larger QP is chosen. In this case, $\Delta QP_{max}=0$. Otherwise, $\Delta QP_{max}=3$. After encoding the i-th MB, the parameters (including $S'_{Y,i+1}$, $S'_{UV,i+1}$, $\overline{S}'_{Y,i+1}$, and $\overline{S}'_{UV,i+1}$) are updated accordingly as described above. The rate model parameters K_Y ($= \left(R'_Y Q_i^2\right) / \left(\overline{A}_{Y,i} \overline{\sigma}^2_{Y,i}\right)$) and K_{UV} ($K_{UV} = \left(R'_{UV} Q_i^2\right) / \left(\overline{A}_{UV,i} \overline{\sigma}^2_{UV,i}\right)$) are updated in a similar way to [3] where R'_Y and R'_{UV} are actual bits of the i-th MB for Y and UV sub-blocks in the current frame respectively.

5. Experimental Results

We implemented the rate control scheme in a JVT JM 10.2 version [4]. In the following experiments, we compare the proposed rate control algorithm with the rate control used in JM 10.2 and the rate control algorithm developed in TMN8 [3]. In the test, the first frame was intra-coded (I-frame) with $QP=31$ and several frames were skipped after the first frame to decrease the number of bits in the buffer below $M=R/F$. Then the remaining frames were all inter-coded (P frames).This means that the number of skipped frames is the same in JM10.2, TMN8 and our proposed schemes (for fair comparison). Afterwards, they use their own MB-layer schemes. The proposed algorithm and TMN8 use the same residue information (e.g. σ_i), as described in Section 4.2 whereas JM10.2 uses their original MB-layer scheme. The proposed algorithms, JM10.2 and TMN8 were simulated on some QCIF test sequences with a frame rate of 10fps and various target bit rates. Here are the test conditions. MV resolution is 1/4 pel. Hadamard is "OFF". RD optimization is "OFF". Search range is "$\pm$ 16". Restrict search range is "0". Reference frames is "1" and symbol mode is "UVLC".

Table 4 shows the actual encoded bit rates achieved by JM10.2, TMN8 and the proposed rate control. It is verified that these rate control methods can achieve the target bit rates. The error between target bit rate and actual bit rate is below 0.2%. Table 5 shows the comparison of PSNR of the reconstructed pictures for JM10.2, TMN8 and the proposed rate control. A gain in PSNR by the proposed rate control over JM10.2 and TMN8 is observed, ranging from +0.13dB to +0.86dB.This is probably because our novel rate models are more accurate and suitable for different MB types. This suggests that our rate models provide good quality with different MB types. It is also noticed that JM10.2 has PSNR gain over TMN8 because the previous residue frame is used only in TMN8, causing less accurate rate model for this simulation. Figure 1 shows comparison of PSNR against frame number in "Fmn128". It is observed that instantaneous PSNR is higher in our proposed algorithm at most of time due to more accurate rate models with different MB types.

6. Conclusion

In advanced video coding standard such as H.264, MBs belongs to more advanced MB types, which were not considered in traditional rate control schemes. As a result, a traditional rate model is less accurate under the same assumption of each MBs. In this work, it is found that different MB types cause different rate behavior. Under different conditions of MB types, we establish novel rate models and distortion models. Our rate control scheme is proposed based on these models. The experimental results suggest that our scheme can achieve PSNR gain over JM10.2 and TMN8.

7. References

[1] X. Yang, W. Lin, Z. Lu, X. Lin, S. Rahardja, E. Ong and S. Yao, "Rate Control for Videophone Using Local Perceptual Cues", *IEEE Trans. Circu. Syst. Video Tech.*, vol. 15, pp.496-507, 2005.

[2] T. Wiegand, "Working Draft Number 2, Revision 8(WD-2 rev 8)", JVT-B118r8, *ISO/IEC MPEG & ITU-T-T VCEG*, Geneva, Switzerland, 29 Jan.-29 Feb.,2002

[3] J. Ribas-Corbera and S. Lei, "Rate Control in DCT Video Coding for Low-Delay Communications", *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 9, pp.172-p.185,1999

[4] "JVT JM10.2", http://iphome.hhi.de/suehring/tml/, 2006

[5] Z. G. Li, W. Gao, F. Pan, S. W. Ma, G. N. Feng, K. P. Lim, X. Lin, S. Rahardja, H. Lu and Yan Lu, "Adaptive Rate Control with HRD Consideration", *ISO-IEC/JTC1/SC29/WG11/ JVT-H014*, Geneva, Switzerland, 20-26 May, 2003.

[6] Z. He and S. K. Mitra, "Zero-domain Bit Allocation and Rate Control for real time video coding", *in Proc. IEEE Int. Conf. Image Processing(ICIP)*, vol. 3, pp. 546-549, 2001

Test Name	Video Se-quence	Target bit (kbps)	Encoded bits (kbps)		
			JM10.2	TMN8	Proposed
Aki24	"Akiyo"	24	24.05	24.03	24.01
Fmn48	"Foreman"	48	48.07	48.05	48.04
Fmn128	"Foreman"	128	128.14	128.12	128.13
ctg256	"Coast-guard"	256	255.63	255.70	254.64
Sil24	"Silent"	24	24.04	24.03	24.02
Stf256	"Stefan"	256	256.26	256.24	256.21

Table 4: Comparison of bit rate achieved by JM10.2, TMN8 and the proposed rate control

Test Name	PSNR (dB)			PSNR Gain (dB) over	
	JM10.2	TMN8	Proposed	JM10.2	TMN8
Aki24	38.84	38.76	39.01	+0.17	+0.25
Fmn48	32.01	31.93	32.24	+0.23	+0.31
Fmn128	36.63	36.58	37.44	+0.81	+0.86
ctg256	37.17	37.10	37.46	+0.29	+0.36
Sil24	31.91	31.84	32.04	+0.13	+0.20
Stf256	33.52	33.44	33.85	+0.33	+0.41

Table 5: Comparison of average PSNR for JM10.2, TMN8 and the proposed rate control

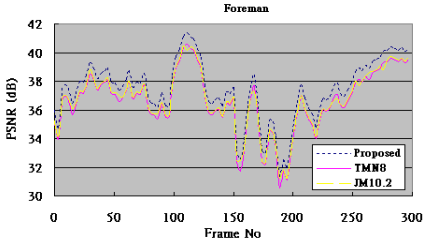

Figure 1: Comparison of PSNR against frame number in "Fmn128"