16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

RATE-DISTORTION OPTIMIZED FRAME LEVEL RATE CONTROL FOR H.264/AVC

Xiang Li[12], *Norbert Oertel*[2], *Andreas Hutter*[2], *and André Kaup*[1]

[1]Chair of Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Erlangen, Germany
[2]Networks & Multimedia Communications, Siemens Corporate Technology, Munich, Germany

ABSTRACT

In this paper, a rate-distortion optimized frame level rate control algorithm is presented for H.264/AVC. To improve the performance in both distortion and rate, two techniques are developed. First, an adaptive frame layer rate-distortion optimization technique is included into the rate control module so that the average distortion is decreased. Second, the mapping from the bit budget to the quantization step is well considered at frame level by an accurate rate model. Consequently, a more precise control over the rate is achieved. Simulations show that the proposed rate control outperforms both the coding with fixed quantization step size (Fixed-QP) and the recommended rate control JVT-G012 in the H.264/AVC reference software. Compared with Fixed-QP, a gain up to 1.04 dB is observed and the maximal rate mismatch for seven sequences is only 0.61%. For JVT-G012, an improvement up to 0.42 dB is obtained while the average rate mismatch is reduced by 70%.

1. INTRODUCTION

Rate Control is always an important module in the practical video codec. Its main task is to regulate the coded bitstream to meet a target rate, which is normally achieved by adjusting the so called quantization step during the coding.

There are two essential problems in the rate control. The first one is how to assign bit budget for each coding unit whose granularity may be from a group of frames to a single macroblock. Generally, such allocations are according to the target rate and buffer status, such as TM5 [1]. Recently, to obtain a more efficient bit budget assignment, the complexity of the frame is also adaptively considered [2] so that the performance is further improved. Besides this, another key problem in rate control is how to determine the quantization step for a coding unit according to the bit budget before the real coding. Normally, to achieve such a mapping properly, a connection between the bit budget and quantization step is necessary. Therefore, a lot of rate models were proposed in the literature. Among them, the famous quadratic model [3] and recent linear model [2] are widely accepted.

Traditionally, most rate control algorithms focus on the two forementioned problems in order to approach the target rate accurately. However, to improve the overall performance, distortion should be considered as well. In fact, in modern video coding, such as H.264/AVC [4], there are a lot of candidate coding modes for each macroblock. Consequently, how to select a best mode, especially the selection under a rate constraint, will greatly affect the coding efficiency and final rate. At present, such a problem is considered by the rate-distortion optimization (RDO) technology proposed in [5, 6], where the optimization is directly dependent on the quantization step.

In theory, the overall coding efficiency can be further increased if any of the three forementioned aspects can be improved. Following this idea, a rate-distortion (R-D) optimized frame level rate control is presented for H.264/AVC in this paper, where the last two aspects are well considered. First, an adaptive frame layer RDO technique is included into the rate control module so that the average coding distortion is decreased. Second, the mapping from the bit budget to the quantization step is further improved based on the accurate rate model developed in adaptive frame layer RDO. Consequently, a more precise control over the rate is achieved. Simulations show that the proposed algorithm outperforms both the coding with fixed quantization step size (Fixed-QP) and the recommended rate control JVT-G012 [7] in the H.264/AVC reference software [8]. Compared with Fixed-QP, up to 1.04 dB gain in PSNR is observed and the maximal rate mismatch for seven sequences is only 0.61%. For JVT-G012, a gain up to 0.42 dB is obtained while the average rate mismatch is reduced by 70%.

The rest of this paper is organized as follows. First, the proposed adaptive frame layer RDO and mapping algorithm are discussed in Section 2 and Section 3, respectively. Then, the whole algorithm is summarized in Section 4. Subsequently, the performance of the proposed rate control is verified by the recent H.264/AVC reference software JM13.1 [8] in Section 5. Finally, the conclusion and future work are presented in Section 6.

2. ADAPTIVE FRAME LAYER RDO

In this section, the RDO algorithms in the literature are first reviewed. Then the proposed adaptive methods is discussed in detail.

2.1 RDO in the Literature

In general, the target of RDO is to minimize the distortion D for a given rate R_c by appropriate selections of coding parameters, namely

$$\begin{gathered}\min\{D\}\\ \text{subject to } R \leq R_c\end{gathered} \tag{1}$$

To solve such a constrained problem, (1) is converted to (2) by the Lagrange multiplier method.

$$\begin{gathered}\min\{J\}\\ \text{where } J = D + \lambda \cdot R\end{gathered} \tag{2}$$

Where J is the Lagrangian cost function and λ is the so-called Lagrange multiplier. Consequently, how to determine λ becomes a key problem in RDO.

Supposing R and D to be differentiable everywhere, the minimum J is given by setting its derivative to zero, i.e.

$$\frac{\mathrm{d}J}{\mathrm{d}R} = \frac{\mathrm{d}D}{\mathrm{d}R} + \lambda = 0 \tag{3}$$

yielding

$$\lambda = -\frac{\mathrm{d}D}{\mathrm{d}R} \tag{4}$$

In fact, (4) indicates that λ corresponds to the negative slope of the R-D curve. Intuitively, the overall performance of such a method will be highly dependent on the accuracy of the rate and distortion models.

So far the most widely used RDO algorithm was proposed in [5, 6]. Assuming a sufficiently high rate environment, R and D models were derived as

$$\begin{aligned} R &= a\log_2\left(\frac{b}{D}\right) \\ D &= \frac{(2Q)^2}{12} = \frac{Q^2}{3} \end{aligned} \tag{5}$$

where a and b are both constants, Q here is half the quantization step. Plugging (5) into (4), λ can be derived as

$$\lambda = c \cdot Q^2 \tag{6}$$

where c is a coding constant. In summary, this method is simple and efficient so that it has been adopted in the H.264/AVC reference software [8]. However, it still has some drawbacks. First, λ is only related to Q and no property of the input signal is considered, which means that it can not adapt to different video sequences dynamically. More important, the "high rate" assumption is not realistic all the time, which will result in a poor performance for both rate and distortion models, especially for low rate applications.

2.2 Proposed Frame Layer RDO

To be more general than the "high rate" assumption, new R and D models based on a well accepted assumption, i.e. Laplace distribution of transformed residuals [9], were proposed in our previous work [10, 11]. Following a similar idea, in this paper a more accurate rate model is derived by considering the statistics on skipped blocks. Consequently, a better frame layer RDO is achieved.

Suppose that the transformed residuals obey a zero-mean Laplace distribution, i.e.

$$\begin{aligned} f_{Lap}(x) &= \frac{\Lambda}{2}e^{-\Lambda|x|} \\ \Lambda &= \frac{\sqrt{2}}{\sigma} \end{aligned} \tag{7}$$

where x represents the transformed residual, σ is their standard deviation which indicates the property of the input sequence, and Λ here is called Laplace parameter [1].

Considering the uniform reconstruction quantizer used in H.264/AVC [4], the entropy H of the quantized transformed residuals can be derived

$$\begin{aligned} P_0 &= \int_{-(Q-\gamma Q)}^{Q-\gamma Q} f_{Lap}(x)dx \\ P_n &= \int_{nQ-\gamma Q}^{(n+1)Q-\gamma Q} f_{Lap}(x)dx \\ H &= -P_0 \cdot \log_2 P_0 - 2\sum_{n=1}^{\infty} P_n \cdot \log_2 P_n \end{aligned} \tag{8}$$

Where Q is the quantization step size, γQ represents the rounding offset and γ is between (0,1), such as 1/6 for H.264/AVC inter frame coding [8], and P_n indicates the probability of transformed residuals inside n^{th} quantization interval.

Since the skipped MB or even skipped 8x8 block is not coded as normal residuals [4], the quantized zeros in these blocks should not be counted when calculating the entropy H. Thus P_0 and P_n need to be corrected. Defining P_s as the probability of skipped 8x8 luma blocks, P_0 and P_n are revised to $P_0^* = (P_0 - P_s)/(1-P_s)$ and $P_n^* = P_n/(1-P_s)$, respectively. Consequently, the refined H^* is

$$H^* = (1-P_s)\cdot\left[-\frac{P_0-P_s}{1-P_s}\cdot\log_2\frac{P_0-P_s}{1-P_s} - 2\sum_{n=1}^{\infty}\frac{P_n}{1-P_s}\cdot\log_2\frac{P_n}{1-P_s}\right] \tag{9}$$

where the multiplier $(1-P_s)$ outside the bracket is to keep H^* as the average entropy for all quantized transformed residuals since only those of non-skipped blocks are counted in the entropy calculation inside the bracket. Moreover, if r is defined as P_s/P_0, H^* can be further simplified as

$$\begin{aligned} H^* = {} & H + P_0 \cdot [r \cdot \log_2 P_0 - (1-r)\cdot\log_2(1-r)] \\ & + (1 - r\cdot P_0)\cdot\log_2(1 - r\cdot P_0) \end{aligned} \tag{10}$$

where r is between (0,1) since the probability of skipped blocks should be no more than that of the quantized zero coefficients according to its physical meaning.

Notice that (10) actually describes an independent entropy coding while in practice a dependent one, such as runlength coding, is formally employed, some compensation has to be applied on (10) to achieve the rate model. From experiments, an exponential relationship is observed between R/H and $\Lambda \cdot Q$. Therefore, R is derived as

$$R = s \cdot H^* \cdot e^{-\xi\Lambda Q} \tag{11}$$

where s and ξ are both constants. According to comprehensive simulations, s is selected as 1.982 and 1.133 for intra and inter frames respectively, while ξ being 0.3 for CABAC and 0.35 for CAVLC.

Putting (7) - (10) into (11), the closed form of R is derived as in (12) (at the bottom of next page).

Similarly, the distortion model D can be derived according to the quantization process as in (13) [11].

$$D = \frac{\Lambda Q \cdot e^{\gamma\Lambda Q}(2 + \Lambda Q - 2\gamma\Lambda Q) + 2 - 2e^{\Lambda Q}}{\Lambda^2(1 - e^{\Lambda Q})} \tag{13}$$

Basically, the adaptive λ can be determined by taking (12) and (13) into (4), as

$$\lambda = -\frac{\mathrm{d}D}{\mathrm{d}R} = -\frac{\partial D/\partial Q}{\partial R/\partial Q} \tag{14}$$

Although the closed form of (14) is too long to be presented here, it can be easily computed by mathematical softwares or numerical methods.

In summary, by considering σ, the standard deviation of transformed residuals, the properties of input sequence are adaptively counted in the frame layer RDO process so that the overall coding efficiency is improved. As such an algorithm only adjusts λ, it can be easily applied. Moreover, the rate model in (12) is more accurate than previous ones, such as quadratic and linear models. Therefore, the mapping process in the rate control can be further improved based on it.

3. MAPPING ALGORITHM

In this section, the mapping algorithms in the literature are first reviewed. Then the proposed method is discussed in detail.

---

[1] Because of the one-to-one mapping between σ and Λ, the latter will be used in the following expressions and discussions for simplicity.

3.1 Mapping Algorithms in the Literature

In the literature, many methods were proposed to describe the relationship between the rate R and quantization step Q for rate control, such as the linear model in MPEG-2 TM5 [1], the quadratic model and its extensions in MPEG-4 VM5, H.263 TMN8, H.264 JVT-G012 [7, 12–14].

Recently, it is realized that the predictability of the control parameters is sometimes even more important than the model accuracy [15], and the linear model becomes popular again [2].

In general, most of the models focus on the rate for residuals since they traditionally take most of the bits. Nevertheless, in modern video coding standard, like H.263 [16] and H.264/AVC [4], more and more bits are spent on side information, such as MB modes and motion vectors. Accordingly, rate models on side information are proposed as well [17]. Unfortunately, most of them are empirical models and not accurate enough. Therefore, in today's rate control, the mapping process still mainly relies on the rate model for residuals.

3.2 Proposed Mapping Algorithm

In theory, for a given target rate R, the desired Q can be determined according to (12). However in practice, there are several problems worth mentioning.

First, Λ and r in (12) have to be predicted since they are both not available before the real coding. In the current implementation, they are estimated by the averages of those in five previous same-type frames, i.e.

$$\begin{aligned}\hat{\Lambda}^i &= \frac{1}{5}\sum_{n=1}^{5}\Lambda^{i-n} \\ \hat{r}^i &= \frac{1}{5}\sum_{n=1}^{5}r^{i-n}\end{aligned} \tag{15}$$

where the superscript ^{i} indicates the value for the i^{th} same-type frame, $\hat{\Lambda}^i$ and $\hat{r}^i$ represent the predictions on Λ^i and r^i, respectively.

Then with $\hat{\Lambda}^i$ and $\hat{r}^i$, Q^i can be derived as

$$Q^i = \arg\min\{|R^i_{BR}\cdot P^{i-1}_L - A\cdot F^{i-1}_L\cdot R(\hat{\Lambda}^i,\hat{r}^i,Q^i)|\} \tag{16}$$

where R_{BR} represents the bit budget for frame residuals assigned by JVT-G012, P_L which is defined as the percent of luma bits in all residuals is used to obtain the bit budget for luma since only luma is considered in the rate model for the current implementation, A is the frame size, F_L is a correction factor which is defined as the ratio between the real frame bits collected after coding and the estimated bits by (12), and $R(\hat{\Lambda}^i,\hat{r}^i,Q^i)$ denotes the rate model (12).

Nevertheless in practice, it is very hard to achieve a closed form solution to (16). Fortunately, in H.264/AVC, Q is not continuous but discrete. It is depend on the so-called quantization parameter QP [4], i.e.

$$Q = 2^{\frac{QP-12}{6}} \tag{17}$$

where QP is limited as an integer within (0,...,51) inclusively [4]. Therefore, the best Q and related QP can be easily found by testing all possible QP.

$$QP^i = \arg\min\left\{\begin{array}{c}|R^i_{BR}\cdot P^{i-1}_L - A\cdot F^{i-1}_L\cdot R(\hat{\Lambda}^i,\hat{r}^i,2^{\frac{QP-12}{6}})| \\ QP = 0,\cdots,51\end{array}\right\} \tag{18}$$

Sometimes, (18) is not able to provide a best mapping. When the real bits for a frame is far from the budget, a compensation in the next frame is necessary to avoid negative influence on the following coding. Currently, such a kind of compensation is empirically performed as follows. Define α as the ratio between the frame bit budget R_{BF} assigned by JVT-G012 and the real frame bits R_F collected after coding,

$$\alpha = R_{BF}/R_F \tag{19}$$

the accuracy of the mapping process can be easily evaluated. Ideally, α equaling one indicates a perfect mapping. On the contrary, a value far from one represents a bad mapping. In such a case, QP is refined as

$$QP^i \leftarrow \left\{\begin{array}{ll} QP^i+1 & \alpha^{i-1} < 0.75 \\ QP^i-1 & \alpha^{i-1} > 1.25\end{array}\right. \tag{20}$$

Finally, to avoid a sharp quality variation from the previous frame, QP is further clipped in the same way as that in JVT-G012 [7].

4. R-D OPTIMIZED RATE CONTROL

Basically, the framework of the proposed rate control is similar with JVT-G012 [7]. First, the bit budgets R_{BF} and R_{BR} for a whole frame and the frame residuals are allocated, respectively. Second, R_{BR} is mapped to the quantization step size Q with the algorithm proposed in Section 3.2. Then λ for adaptive RDO is derived according to Section 2.2. After that, the whole frame is coded with Q and λ just derived. Finally, the model parameters, such as Λ and r, are updated. Note that λ determination is on the frame level. Therefore the so-called chicken-and-egg dilemma between the RDO and MB level rate control [7] is automatically avoided.

For convenience, the proposed rate control is summarized as pseudo code in Algorithm 1. As the algorithm is embedded into H.264/AVC JM13.1 environment [8], the frame initialization, clipping on QP, encoding and finishing process are the same with those in JM13.1 except that in the encoding process the original λ and QP are replaced by the new derived ones.

From Algorithm 1, it can be also noticed that the computational complexity of the proposed algorithm is similar with JVT-G012. In fact, the most time consuming part in the proposed algorithm is the determination of Q. Normally, the number of Q values to be tested in (18) should be small because of the practical variation limitation on QP for successive frames. Furthermore, such calculations only occur at frame level. Consequently, no much computation is introduced by the proposed rate control.

$$\begin{aligned}R = \frac{Se^{-\xi\Lambda Q}}{\ln 2}\{&(1-e^{-(1-\gamma)\Lambda Q})[r\ln(1-e^{-(1-\gamma)\Lambda Q})-(1-r)\ln(1-r)]-(1-e^{-(1-\gamma)\Lambda Q})\ln(1-e^{-(1-\gamma)\Lambda Q}) \\ &+[1-r(1-e^{-(1-\gamma)\Lambda Q})]\cdot\ln[1-r(1-e^{-(1-\gamma)\Lambda Q})]+e^{-(1-\gamma)\Lambda Q}(\ln 2-\ln(1-e^{-\Lambda Q})-\gamma\Lambda Q+\frac{\Lambda Q}{1-e^{-\Lambda Q}})\}\end{aligned} \tag{12}$$

Algorithm 1: Encode one frame with the proposed R-D optimized frame level rate control

```
Initialize $i^{th}$ frame
Allocate bit budgets $R^i_{BF}$ and $R^i_{BR}$ for $i^{th}$ frame (by JVT-G012 Frame [7])
begin $Q^i$ calculation
    $QP^i \leftarrow \arg\min \left\{ \begin{array}{c} |R^i_{BR}P_L^{i-1} - AF_L^{i-1}R(\hat{\Lambda}^i, \hat{r}^i, 2^{\frac{QP-12}{6}})| \\ QP = 0, \cdots, 51 \end{array} \right\}$
    if $\alpha^{i-1} < 0.75$ then
        $QP^i \leftarrow QP^i + 1$
    else if $\alpha^{i-1} > 1.25$ then
        $QP^i \leftarrow QP^i - 1$
    $QP^i \leftarrow \mathrm{CLIP}(QP^i)$ [7]
    $Q^i \leftarrow 2^{\frac{QP^i-12}{6}}$
end
Calculate adaptive $\lambda^i$ (by Section 2.2)
Encode $i^{th}$ frame with $QP^i$ and $\lambda^i$
begin Update Process
    Update $\Lambda^i$ and $r^i$
    Calculate $\alpha^i$, $P^i_L$ and $F^i_L$
end
Finish $i^{th}$ frame
```

5. SIMULATIONS AND DISCUSSIONS

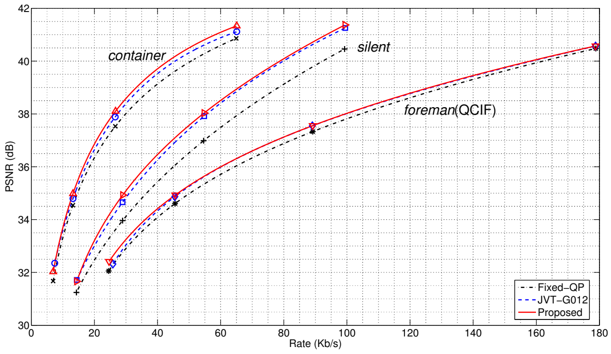
The proposed rate control is verified by the recent H.264/AVC reference software JM13.1 [8]. The test environment is derived from the simulation common conditions for coding efficiency experiments recommended by ITU-T VCEG [18]: seven QCIF/CIF sequences were coded with H.264/AVC Main Profile (CABAC) in IPPP mode (all P frames except the first I frame). To ease the comparison, each sequence was first coded by Fixed-QP with the recommended quantization parameters [18], i.e. QP=23, 28, 33 and 38. Then the real rates collected after Fixed-QP coding are taken as the target rates for the JVT-G012 and the proposed rate control. To avoid the affection by the initialization algorithm, the initial QP for both algorithms are set the same as those in Fixed-QP.

In practice, rate control affects not only the performance of one luminance component, but also those of two chrominance channels. Thus for a fair comparison, the coding efficiency of the all these three components should be jointly evaluated. As the test sequences are all in 4:2:0 format where the weighting among the three components is 4:1:1, a combined PSNR P as in (21) is used instead of three independent PSNRs in the following assessments and discussions for convenience.

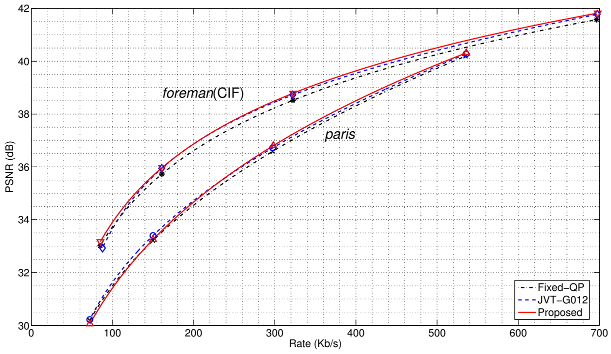
$$P = \frac{1}{6}(4P_Y + P_U + P_V) \qquad (21)$$

Here P_Y, P_U and P_V are PSNRs for one luminance and two chrominance channels, respectively.

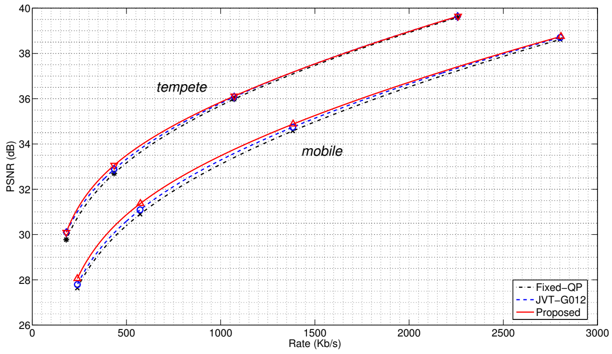
Table 1 shows the simulation results, where ΔP represents the gain over Fixed-QP, R_M shows the mismatch between the target and actual rate, G_{aver} denotes the average gains over Fixed-QP. Actually, such a kind of average gain indicates the difference between two R-D curves and is calculated with the tool recommended by VCEG [19].



(a) *container*, *foreman* (QCIF) and *silent*



(b) *paris* and *foreman* (CIF)



(c) *mobile* and *tempete*

Figure 1: R-D curves

Thanks for the adaptive frame layer RDO technique, the proposed rate control outperforms both Fixed-QP and JVT-G012 in distortion. A gain of up to 1.04 dB over Fixed-QP for *silent* and up to 0.42 dB over JVT-G012 for *foreman*(CIF) can be observed from the R-D curves shown in Fig. 1. On average, 0.37 dB and 0.10 dB gain over Fixed-QP and JVT-G012 are achieved, respectively.

More important, the proposed method shows a more accurate control over the rate because of the better mapping algorithm. On average, the mismatch between the target and actual rate is only 0.19%. Compared with JVT-G012, the average mismatch is reduced by 70%. In fact, the maximal mismatch by the proposed algorithm is only 0.61%. But for JVT-G012, the rate mismatches for container (QP=38) and foreman (QCIF,QP=38) are even up to 7.13% and 5.4% which are somewhat not acceptable in practice.

Table 1: Simulation Results

sequences	QP	Fixed-QP		JVT-G012			Proposed		
		P(dB)	R(Kb/s)	ΔP(dB)	R_M	G_{aver}(dB)	ΔP(dB)	R_M	G_{aver}(dB)
container (QCIF, 150 frms, 15 Hz)	23	40.86	64.94	0.25	0.25%	0.32	0.47	0.22%	0.49
	28	37.53	26.63	0.35	-0.14%		0.57	0.24%	
	33	34.54	13.20	0.25	0.23%		0.44	0.26%	
	38	31.67	6.94	0.68	**7.13%**		0.36	0.35%	
foreman (QCIF, 150 frms, 15 Hz)	23	40.48	178.84	0.09	-0.04%	0.19	0.10	-0.04%	0.25
	28	37.33	89.16	0.22	-0.12%		0.23	0.00%	
	33	34.61	45.55	0.26	-0.12%		0.32	0.15%	
	38	32.05	24.53	0.26	**5.40%**		0.36	0.49%	
silent (QCIF, 150 frms, 15 Hz)	23	40.46	99.26	0.80	0.22%	0.74	0.91	0.15%	0.90
	28	36.99	54.61	0.93	0.34%		1.05	0.27%	
	33	33.96	28.93	0.70	0.01%		0.96	0.15%	
	38	31.24	14.31	0.47	0.62%		0.42	0.61%	
paris (CIF, 150 frms, 15 Hz)	23	40.19	535.03	0.08	0.12%	0.13	0.13	0.16%	0.07
	28	36.59	297.24	0.13	0.06%		0.22	0.30%	
	33	33.21	149.86	0.19	0.11%		0.04	0.14%	
	38	30.19	72.10	0.03	0.00%		-0.13	0.27%	
foreman (CIF, 300 frms, 30 Hz)	23	41.57	696.54	0.20	0.11%	0.15	0.25	0.19%	0.24
	28	38.52	322.45	0.20	-0.09%		0.27	0.05%	
	33	35.73	160.96	0.22	-0.13%		0.25	0.21%	
	38	33.01	85.01	-0.09	3.24%		0.16	0.24%	
mobile (CIF, 300 frms, 30 Hz)	23	38.62	2801.38	0.09	0.03%	0.17	0.13	0.17%	0.37
	28	34.58	1382.63	0.17	0.00%		0.31	0.14%	
	33	30.90	572.63	0.19	-0.05%		0.47	0.23%	
	38	27.64	239.23	0.16	0.13%		0.42	0.06%	
tempete (CIF, 260 frms, 30 Hz)	23	39.59	2257.57	0.04	0.01%	0.15	0.05	0.11%	0.24
	28	35.98	1071.63	0.09	-0.02%		0.14	0.08%	
	33	32.70	434.57	0.18	0.00%		0.36	0.09%	
	38	29.78	180.60	0.31	0.19%		0.32	-0.03%	
average	-	-	-	0.27	0.63%	0.27	0.34	0.19%	0.37

6. CONCLUSION AND FUTURE WORK

In this paper, a R-D optimized frame level rate control is presented for H.264/AVC. To decrease the average distortion, an adaptive frame layer RDO method is included into the rate control. Moreover, a more accurate rate model is used to map the bit budget to the quantization step at frame level for the rate control. Simulations show that the proposed algorithm outperforms both the Fixed-QP and JVT-G012, i.e. up to 1.04 dB and 0.42 dB gains over Fixed-QP and JVT-G012, respectively. While at the same time, the average rate mismatch is reduced by 70% when compared with JVT-G012.

As mentioned in Section 1, there are two essential problems in the rate control, i.e. the bit budget allocation and the mapping algorithm between which only the latter is considered in this paper. Theoretically, a better budget allocation should result in another improvement.

REFERENCES

[1] "MPEG-2 Video Test Model 5," in *ISO/IEC JTC1/SC29/WG11 MPEG93/457*, 1993.

[2] Y. Liu, Z.G. Li, and Y.C. Soh, "A novel rate control scheme for low delay video communication of H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 1, pp. 68–78, 2007.

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

[4] JVT, *Advanced Video Coding (AVC) - 3rd Edition*, ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG-4 Part 10). 2004.

[5] G.J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74–90, 1998.

[6] T. Wiegand and B. Girod, "Lagrange multiplier selection in hybrid video coder control," in *IEEE Int. Conf. on Image Processing (ICIP)*, Thessaloniki, Greece, 2001, pp. 542–545, vol.3.

[7] Z.G. Li, F. Pan, K.P. Lim, G. Feng, X. Lin, and S. Rahardja, "Adaptive basic unit layer rate control for JVT (JVT-G012)," in *Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG*, Pattaya II ,Thailand, Mar. 2003.

[8] JVT, "H.264/AVC reference software (JM13.1)," http://iphome.hhi.de/suehring/tml/, Dec. 2007.

[9] E.Y. Lam and J.W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1661–1666, 2000.

[10] X. Li, N. Oertel, A. Hutter, and A. Kaup, "Advanced Lagrange multiplier selection for hybrid video coding," in *IEEE Int. Conf. on Multimedia & Expo (ICME)*, Beijing, China, Jul. 2007, pp. 364–367.

[11] X. Li, N. Oertel, A. Hutter, and A. Kaup, "Extended Lagrange multiplier selection for hybrid video coding using interframe correlation," in *Picture Coding Symposium (PCS)*, Lisbon, Portugal, Nov. 2007.

[12] Video Group, "MPEG-4 video verification model version 5.1," in *ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Asssociated Audio*, Maceio, Brazil, Nov. 1996.

[13] "Video Codec Test Model TMN8," in *ITU-T/SG15*, 1997.

[14] H. Wang and S. Kwong, "Rate-distortion optimization of rate control for H.264 with adaptive initial quantization parameter determination," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 140–144, 2008.

[15] J. Dong and N. Ling, "On model parameter estimation for H.264/AVC rate control," in *IEEE Int. Symposium on Circuits and Systems (ISCAS)*, New Oreleans, USA, May 2007, pp. 289–292.

[16] ITU-T, *Video Coding for Low Bitrate Communication*, ITU-T Rec. H.263 Version3 (H.263++). 2000.

[17] D.-K. Kwon, M.-Y. Shen, and C.J. Kuo, "Rate control for H.264 video with enhanced rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 5, pp. 517–529, 2007.

[18] T. Tan, G.J. Sullivan, and T. Wedi, "Recommended simulation common conditions for coding efficiency experiments (VCEG-AA10)," in *ITU-T SG16 Q.6 (VCEG)*, Nice, France, Oct. 2005.

[19] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves (VCEG-M33)," in *ITU-T SG16 Q.6 (VCEG)*, Austin, Texas, USA, Apr. 2001.