Adaptive Rate-Distortion Optimization for H.264

Kwan-Jung Oh and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST),
1 Oryong-dong Buk-gu, Gwangju, 500-712, Korea
{kjoh81,hoyo}@gist.ac.kr

Abstract. In video coding standards, such as MPEG-4 and H.263, one important question is how to determine motion vectors for motion compensation in the INTER mode. Usually the sum of absolute differences (SAD) or the sum of squared differences (SSD) is employed as a matching criterion. Although these criteria are related to the distortion, they do not consider the bit rate appropriately. If we want to consider both the rate and the distortion, a Lagrangian technique targeting for rate-distortion optimization (RDO) is a good alternative. Even if H.264 used the RDO scheme to decide the best macroblock mode among several candidates, H.264 employs only one RDO model for all macroblocks. Since the characteristics of each macroblock is different, each macroblock should have its own RDO model. In this paper, we propose an adaptive rate-distortion optimization algorithm for H.264. We regulate the Lagrangian multiplier according to the picture type and characteristics of each macroblock.

Keywords: Adaptive RDO, Lagrangian Multiplier, H.264.

1 Introduction

Motion-compensated transform video coding, also called as hybrid video coding, provides a good combination of data compression tools. In various video coding standards, motion vectors are determined by the sum of absolute differences (SAD) or the sum of squared differences (SSD), related to the distortion of the motion prediction. However, these criteria do not consider the bit rate. Thus, different matching criteria have been proposed to consider both the distortion and the bit rate [1].

A common way of formulating such a criterion is the Lagrangian optimization, which is adopted in H.264 for selecting the best macroblock mode. Usually, the Lagrangian multiplier λ is only defined as a function of the quantization parameter (QP). However, the optimal choice of λ should depend on the characteristics of each macroblock as well as QP [2].

The current rate-distortion optimization (RDO) model used in H.264 is applied to each macroblock, but it does not provide the optimization of the whole sequence. However, this problem can be improved by using an adaptive RDO model. There have been previous efforts that incorporate perceptual characteristics into video coding. However, they have focused on the perceptual distortion which is related to the human visual system (HVS). Despite of slightly improved performance with respect to HVS, they cannot provide good performance in terms of the rate and distortion [2].

Y.-S. Ho and H.J. Kim (Eds.): PCM 2005, Part II, LNCS 3768, pp. 617 – 628, 2005.

In this paper, we propose an adaptive rate-distortion optimization algorithm for H.264. After we find a Lagrangian multiplier for each picture through several experiments, we regulate the Lagrangian multiplier for each macroblock considering characteristics of each macroblock. We take the distortion variance as the characteristics of each macroblock.

The distortion variance is obtained after the motion estimation process. During motion estimation, we calculate distortions for several different modes. We get the distortion variance from the distortions of 16×16, 16×8, and 8×16 modes. The distortion variance reflects the characteristics of each macroblock. If the distortion variance is small, this macroblock belongs to a flat area. While we give more weights to the rate part of the RDO model when the macroblock belongs to a flat area, we give more weights to the distortion part of the RDO model when the macroblock belongs to a complex area. In this manner, we can optimize the rate and the distortion for the whole sequence more efficiently.

This paper is organized as follows. After the Lagrangian optimization in hybrid video coding is explained in Section 2, we propose an adaptive RDO algorithm in Section 3. In Section4, experimental results show effectiveness of the proposed algorithm, and we conclude this paper in Section 5.

2 Lagrangian Optimization

2.1 Optimization Using Lagrangian Techniques

Consider K source samples that are collected from the set $S = (S_1,\cdots,S_k)$. Each source sample S_k can be quantized using several possible coding options that are indicated by an index out of the set $O_k = (O_{k1},\cdots,O_k)$. Let $I_k \in O_k$ be the selected index for a code S_k.

The coding options assigned to the elements in S are given by the components in the set $I = (I_1,\cdots,I_k)$. The problem of finding the combination of coding options that minimizes the distortion subject to a given rate constraint R_C can be formulated as

$$\begin{aligned} &\min_I D(S,I) \\ &subject\ to\ \ R(S,I) \le R_C \end{aligned} \tag{1}$$

where $D(S, I)$ and $R(S, I)$ represent the total distortion and bit rate, respectively. These parameters result from the quantization of S with a particular combination of coding options I. In practice, rather than solving the constrained problem in Eq. (1), an unconstrained formulation is employed, that is

$$\begin{aligned} &I^* = \arg\min_I J(S,I \mid \lambda) \\ &with\ \ J(S,I \mid \lambda) = D(S,I) + \lambda \cdot R(S,I) \end{aligned} \tag{2}$$

and $\lambda \geq 0$ being Lagrange parameter. I^* in Eq. (2) is optimal in the sense that if a rate constraint R_C corresponds to λ. The total distortion $D(S, I^*)$ is minimum for all combinations of coding options with bit rate less or equal to R_C [3].

Assume that the additive distortion and rate measures only depend on the choice of the parameter corresponding to each sample. Then, a simplified Lagrangian cost function can be computed by using

$$J(S_k, I \mid \lambda) = J(S_k, I_k \mid \lambda) \tag{3}$$

In this case, the optimization problem reduces to

$$\min_I \sum_{k=1}^{K} J(S_k, I \mid \lambda) = \sum_{k=1}^{K} \min_{I_k} J(S_k, I_k \mid \lambda) \tag{4}$$

and can be easily solved by independently selecting the coding option for each $S_k \in S$. For this particular scenario, the problem formulation is equivalent to the bit-allocation problem for an arbitrary set of quantizers, proposed by Shoham and Gersho [4].

2.2 Lagrangian Optimization in Hybrid Video Coding

The Lagrangian technique can be used for the motion estimation. The motion estimation is so heavy process that we do not employ the Lagrangian technique for the motion estimation. However, the efficiency of the macroblock mode decision can be improved by Lagrangian technique. In previous video coding standards, the macroblock mode is determined by using the previously coded macroblock [5]. However, the coding mode for each macroblock should be determined using the Lagrangian cost function. Assume that Lagrangian parameter λ_{MODE} and the quantizer value Q were given. The Lagrangian mode decision for a macroblock S_k proceeds by minimizing

$$J_{MODE}(S_k, I_k \mid Q, \lambda_{MODE}) = D_{REC}(S_k, I_k \mid Q) + \lambda_{MODE} R_{REC}(S_k, I_k \mid Q) \tag{5}$$

where the macroblock mode I_k varies as the sets of possible macroblock modes for the various standards.

MPEG-2: SKIP, 16×16, INTRA
H.263/MPEG-4: SKIP, 16×16, 8×8, INTRA
H.264: SKIP, 16×16, 16×8, 8×16, P8×8, I16×16, I4×4

H.264 additionally provides the following sets of sub-macroblock types for P8×8:

8×8, 8×4, 4×8, 4×4

The distortion $D_{REC}(S_k, I_k|Q)$ and the rate $R_{REC}(S_k, I_k|Q)$ for the various modes are computed as follows: For INTRA modes, the corresponding 4×4 blocks (H.264) or 8×8 blocks (MPEG-2, H.263/MPEG-4) of the macroblock S_k are processed by transformation and subsequent quantization. The distortion $D_{REC}(S_k, \text{INTRA}|Q)$ is measured by calculating SSD between the reconstructed (s') and the original (s) macroblock pixels

$$SSD = \sum_{(x,y)\in A} \left| s[x,y,t] - s'[x,y,t] \right|^2 \tag{6}$$

where A is the subject macroblock. The rate $R_{REC}(S_k, \text{INTRA}|Q)$ is the rate that results after entropy coding.

For SKIP mode, the distortion $D_{REC}(S_k$, INTRA|$Q)$ and the rate $R_{REC}(S_k$, INTRA|$Q)$ do not depend on the current quantizer value. The distortion is determined by SSD between the current picture and the value of the inferred INTER prediction.

The computation of the Lagrangian costs for INTER modes is much more demanding than for INTRA and SKIP modes. This is because of the block motion estimation step. Given the Lagrangian parameter λ_{MOTION} and the decoded reference picture *s'*, the rate-constrained motion estimation for a block S_i is performed by minimizing the Lagrangian cost function

$$m_i = \arg\min_{m \in M} \{ D_{DFD}(S_i, m) + \lambda_{MOTION} R_{MOTION}(S_i, m) \} \tag{7}$$

where *m* is the set of possible coding modes. Eq. (7) has the distortion term given by

$$D_{DFD}(S_i, m) = \sum_{(x,y) \in A_i} \left| s[x, y, t] - s'[x - m_x, y - m_y, t - m_t] \right|^P \tag{8}$$

with p=1 for SAD and p=2 for SSD. $R_{MOTION}(S_i, m)$ is the number of bits used to transmit all the components of the motion vector (m_x, m_y), and m_t when multiple reference frames are used. The search range *M* is ±32 integer pixel positions horizontally and vertically and either one or more previously decoded pictures are referenced. Depending on SSD or SAD, the Lagrangian parameter λ_{MOTION} has to be adjusted.

The Lagrangian parameter λ_{MODE} for H.263/MPEG-4 is obtained by the following equation:

$$\lambda_{MODE} = 0.85 \cdot Q^2{}_{H.263} \tag{9}$$

The corresponding λ_{MOTION} for SAD or SSD is as follows, respectively:
For SAD

$$\lambda_{MOTION} = \sqrt{\lambda_{MODE}} \tag{10}$$

In case of SSD,

$$\lambda_{MOTION} = \lambda_{MODE} \tag{11}$$

By conducting the same experiment that leads to the relationship in Eq. (8) again for H.264, λ_{MODE} is obtained as follows:

$$\lambda_{MODE} = 0.85 \cdot 2^{(Q_{H.264} - 12)/3} \tag{12}$$

The corresponding λ_{MOTION} for H.264 is calculated by Eq. (10) and Eq. (11).

Following equation is the cost function which is used in H.264

$$\begin{aligned} &J_{MODE}(s, r, MODE \mid \lambda_{MODE}) = \\ &SSD(s, r, MODE) + \lambda_{MODE} \cdot R(s, r, MODE) \end{aligned} \tag{13}$$

where *s* and *r* represent the current block and the reference block, respectively. *MODE* represents the various macroblock modes [6].

3 Adaptive Rate-Distortion Optimization

Since H.264 encoder employs many sophisticated schemes in the coding procedure, the H.264 video coding standard achieves much higher coding efficiency than the previous video coding standards such as H.263 and MPEG-4. One important scheme is variable block size motion estimation and mode decision. Generally, the motion estimation is performed on the macroblock level, thus each macroblock needs one motion vector which can lead to a minimum block matching error.

However, when the macroblock contains multiple objects and every object moves in different directions or when the macroblock lies on the boundary of a moving object, only one motion vector will not be enough to represent real motions. It will result in serious prediction error. In order to improve the prediction accuracy, H.264 uses seven different modes which are SKIP, 16×16, 16×8, 8×16, P8×8, I16×16, and I4×4. Using these various macroblock modes, the efficiency of the motion estimation and the motion compensation of H.264 is improved. Figure 1 shows these modes.

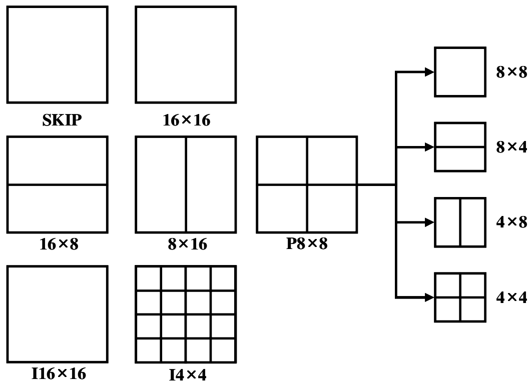

Fig. 1. Various Macroblock Modes

A problem of the mode selection is how to select the best macroblock mode among several modes. The Lagrangian cost function in Eq. (13) provides the solution for this problem. During the encoding process, all macroblock modes are examined and the resulting rate and the distortions are calculated. The mode that has the minimum Lagrangian cost is selected as the best mode for the macroblock [7].

However, as we can see from the Lagrangian function in Eq. (13), there is no parameter which reflects characteristics of a given macroblock. H.264 uses only one RDO model for all macroblocks for whole sequence. Although the current RDO model provides the best result for each macroblock, it does not lead to the optimization of whole sequence. Since the each picture type and characteristics of each macroblock are different, RDO model need to be changed to the adaptive RDO model considering picture type and characteristics of macroblock. After we find the proper Lagrangian multiplier for each picture, we expand it into macroblock level.

3.1 Adaptive Lagrangian Multiplier for Each Picture

In hybrid video coding, the structure of GOP (group of picture) influences the whole coding efficiency. GOP is consists of the one I picture and several other kinds of pictures. IPPP⋯ and IBBP⋯ are good examples of the GOP structures. Picture types also influence the coding efficiency. Among them, I picture is most important since it used as a reference picture for P picture. So, it is not too much to say that I picture runs the coding efficiency of the given GOP. In order to evaluate the influencing power of I picture, we employ the first 90 frames from the FOREMAN sequence in QCIF format 176×144 and QP is 28. GOP structure is IPPP… and intra period is 30. Search range is ±32. Table 1 shows the simulation results. We yield the results for whole sequence by changing the Lagrangian multiplier from 0.1 to 0.8 for I picture.

Table 1. Influencing Power of I Picture for FOREMAN

Lagrangian Multiplier	PSNR (dB) of I Picture	Bit Rate (bits) of I Picture	PSNR (dB)	Bit Rate (kbits/s)
Original (0.85)	36.763	24,112	35.837	116.69
0.1	37.893	31,632	36.067	122.69
0.2	37.579	28,387	36.011	119.60
0.3	37.302	26,232	35.954	117.58
0.4	37.139	25,440	35.968	118.13
0.5	37.017	24,880	35.903	116.87
0.6	36.907	24,501	35.900	116.91
0.7	36.861	24,397	35.889	116.71
0.8	36.722	24,083	35.852	116.17

As we can see, the smaller Lagrangian multiplier leads a better quality but needs larger bits. We have done extensive experiments to obtain the proper Lagrangian multiplier for I and P picture. We yield the results by changing the Lagrangian multiplier from 0.2 to 0.4 for I picture and from 0.9 to 1.1 for P picture.

Table 2. Lagrangian Multiplier and Its Coding Efficiency for FOREMAN

Lagrangian Multiplier	PSNR (dB)	Bit Rate (kbits/s)	ΔPSNR (dB)	ΔBit Rate (kbits/s)
Original (0.85, 0.85)	35.837	116.69	0	0
0.2, 0.9	35.941	118.16	0.104	1.47
0.2, 1.0	35.804	114.63	-0.033	-2.06
0.2, 1.1	35.750	112.99	-0.087	-3.70
0.3, 0.9	35.878	116.02	0.041	-0.67
0.3, 1.0	35.762	113.11	-0.075	-3.58
0.3, 1.1	35.654	110.46	-0.183	-6.23
0.4, 0.9	35.867	115.57	-0.030	-1.12
0.4, 1.0	35.737	112.17	-0.100	-4.52
0.4, 1.1	35.640	110.60	-0.197	-6.09

As we can see, if we well select the Lagrangian multiplier for each picture, we can get a better coding efficiency. A better coding efficiency means that we can get better quality in spite of using less bits. Through the extensive experiments, we set the Lagrangian multiplier which shows the best coding efficiency for several test sequences.

To derive the general Lagrangian multiplier for picture, we investigate the best Lagrangian multiplier for eight QCIF sequences. Table 3 shows the proposed Lagrangian multiplier and comparison of coding efficiency between H.264 and proposed algorithm. We use the first 90 frames from the eight test video sequences in QCIF format and QP = 28. Other test conditions are same with the test conditions of previous experiments.

Table 3. Proposed Lagrangian Multiplier and Its Coding Efficiency

Test Sequences	Lagrangian Multiplier (I , P Picture)	H.264 PSNR (dB)	Proposed PSNR (dB)	H.264 Bit Rate (kbits/s)	Proposed Bit Rate (kbits/s)
AKIYO	0.3, 1.0	38.725	38.873	38.85	38.64
CARPHONE	0.3, 0.9	37.486	37.538	100.88	100.76
CONTAINER	0.3, 1.1	36.461	36.545	49.37	48.31
FOREMAN	0.4, 0.9	37.837	35.867	116.69	115.57
MOBILE	0.2, 0.9	33.607	33.633	421.89	418.03
MOTHER & DAUGHTER	0.2, 1.0	36.599	36.626	92.54	91.78
NEWS	0.3, 1.0	37.063	37.293	83.74	83.57
SALESMAN	0.3, 1.1	35.825	36.045	74.57	74.37
Average	0.2875,0.975				

From Table 3 we can know most Lagrangian multipliers are similar and most results show better coding efficiency than H.264 coding efficiency. The PSNR values are increased and bit rates are decreased. From the results of previous experiments, we can get the general Lagrangian multiplier for I picture and P picture in case of QP=28. The values are 0.2875 and 0.9875. However these values are adaptively changed by QP. If QP is increased Lagrangian multiplier for I picture should be increased and Lagrangian multiplier for P picture should be decreased. In this manner, we experiment for several QP and we can obtain following equation from these results.

$$\begin{aligned} &\textit{Lagrangian multiplier for I Picture} \\ &\lambda_{MODE} = \frac{QP}{\alpha} \cdot 2^{(Q_{H.264}-12)/3} \\ &\textit{Lagrangian multiplier for P Picture} \\ &\lambda_{MODE} = \left(1.2 - \frac{QP}{\beta}\right) \cdot 2^{(Q_{H.264}-12)/3} \end{aligned} \tag{14}$$

where, α is 97 and β is 132. Since I picture use more bits than original I picture, we give more weights to the rate part of the RDO model for P pictures.

3.2 Adaptive Lagrangian Multiplier for Each Macroblock

In the previous section, we proposed the adaptive Lagrangian multiplier for picture. Even though we proposed adaptive RDO model for each picture, the macroblocks in same picture still use a same RDO model. Now, we derive the adaptive RDO model for each macroblock considering the characteristics of each macroblock.

We change the previous RDO model as follows:

$$\begin{aligned} & J_{MODE}(s, r, MODE \mid \lambda_{MODE}) = \\ & SSD(s, r, MODE) + \chi \cdot \lambda_{MODE} \cdot R(s, r, MODE) \end{aligned} \tag{15}$$

where χ is the new parameter, which reflects the characteristics of the macroblock. To derive the χ, firstly we introduce a parameter which is called log-scaled standard deviation (LSD). *LSD* is calculated for 16×16, 16×8, and 8×16 modes.

$$LSD = \log\left(\sqrt{\frac{1}{3}\sum_{MODE=0}^{2}(M_{distortion} - D_{MODE})^2}\right) \tag{16}$$

where $M_{distortion}$ is the mean value of the distortion corresponding to each mode, and D_{MODE} is the distortion of a given mode. By using LSD, χ is calculated by

$$\begin{aligned} & if\,(LSD_{cur} < LSD_{mean}) \\ & \qquad \chi = 1 - \frac{|LSD_{mean} - LSD_{cur}|}{20\delta} \\ & else \\ & \qquad \chi = 1 + \frac{|LSD_{mean} - LSD_{cur}|}{20\delta} \end{aligned} \tag{17}$$

where LSD_{cur} represents the *LSD* value for current macroblock and LSD_{mean} is the mean value of the LSD until the previous macroblock. Through this procedure, we obtain χ which is around 1. Since the rate is more sensitive with respect to RDO than the distortion, we use different denominator. δ reflects the characteristics of the given sequence. δ is derived by

$$\delta = \frac{I_{bitrate}}{V \times H} + \sqrt{\frac{255^2}{10^{\frac{PSNR}{10}}}} \tag{18}$$

where $I_{bitrate}$ is the number of coded bits for first I-frame. V and H represent vertical size and horizontal size of the given image, respectively. So the first term of the right-hand side of Eq. (18) is the average value of the bit per pixel for I-frame and the second term is the square-rooted MSE. The reason why δ is defined as Eq. (18) is because each sequence has different characteristics. If motional characteristics of a sequence is monotonous, δ is small or otherwise, δ shall be large.

In this section, we propose the adaptive rate-distortion optimization algorithm for each macroblock. The proposed RDO model depends on χ as well as on the previous parameters such as QP, distortion, and bit rate. In this way, we can save the bits in the flat area and can assign more bits to the complex area. In the latter case, although the proposed algorithm needs more bits, it can be compensated in next frames through the motion compensation. Smaller amount of distortion leads to better motion estimation and better motion estimation leads to the bit saving. Figure 2 shows the flow diagram of the proposed algorithm.

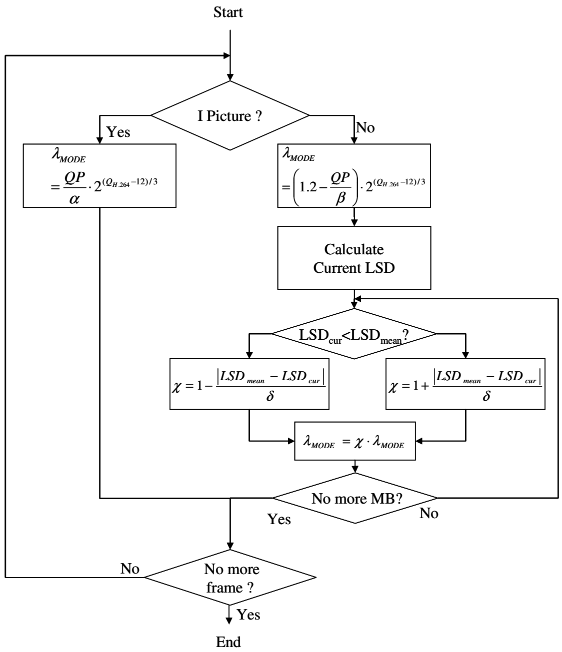

Fig. 2. Flow Diagram of the Proposed Algorithm

4 Experimental Results and Analysis

In order to evaluate the performance of the proposed algorithm, we use the first 120 frames from the five test video sequences (AKIYO, FOREMAN, MOBILE, NEWS, and SALESMAN) in QCIF format 176×144. JM 9.5 is used to conduct the experiments [8]. The Hadamard transform, CABAC, and reconstruction filter are enabled. In the motion estimation, five reference frames are enabled with the maximum search range ±32 and the motion vector resolution is 1/4 pixel. The frame rate is 30 fps and the frame coding structure is IPPP⋯P. Intra period is 30. The experiments are performed for four quantization parameters QP=28, 32, 36, and 40. We perform the two experiments. At first, we experiment about the efficiency of the adaptive RDO model for each picture. Then we combine this scheme with the adaptive RDO model for each macroblock.

The PSNR value and bit rate comparison between the H.264 and the proposed algorithm for the adaptive RDO model for each picture are tabulated in Table 5. As we can see, most results of the proposed algorithm show better performance compared with H.264 standard.

Table 5. Comparison for PSNR Values and Bit Rates

Test Sequences	Quantization Parameter	PSNR (dB)		Bit Rate (kbits/s)	
		H.264	Proposed	H.264	Proposed
AKIYO	QP=28	38.690	38.872	38.02	38.09
	QP=32	35.781	36.049	23.45	24.03
	QP=36	33.209	33.423	15.35	15.70
	QP=40	30.634	30.993	10.54	10.88
FOREMAN	QP=28	35.929	35.879	121.99	119.45
	QP=32	33.386	33.363	69.02	68.23
	QP=36	31.013	31.035	42.38	41.83
	QP=40	28.651	28.694	27.54	27.23
MOBILE	QP=28	33.580	33.493	434.02	422.47
	QP=32	30.179	30.199	221.52	217.10
	QP=36	27.231	27.321	119.39	118.24
	QP=40	24.124	24.407	72.43	71.56
NEWS	QP=28	37.076	37.284	88.01	88.27
	QP=32	33.971	34.222	54.62	54.90
	QP=36	31.108	31.305	33.69	33.99
	QP=40	28.358	28.606	20.45	20.66
SALESMAN	QP=28	35.839	36.177	75.70	77.13
	QP=32	32.821	33.110	43.91	45.24
	QP=36	30.192	30.462	25.39	26.09
	QP=40	27.742	28.019	14.30	14.62

Figure 3 shows the rate distortion curves for the MOBILE and NEWS. Left is the worst case and right is the best case among the results. As we can see, the rate distortion curves of the proposed algorithm located upper than the rate distortion curves of the H.264. This means that the proposed algorithm shows the better performance with respect to PSNR value and the bit rate.

The PSNR value and bit rate comparison between the H.264 and the proposed algorithm for CIF format 352×288 are tabulated in Table 6. We use the first 120 frames

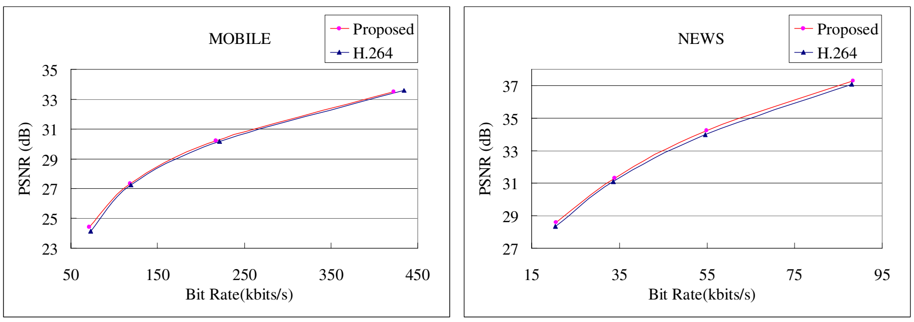

Fig. 3. Rate Distortion Curves for MOBILE and NEWS

Table 6. Comparison for PSNR Values and Bit Rates

Test Sequences	Quantization Parameter	PSNR (dB)		Bit Rate (kbits/s)	
		H.264	Proposed	H.264	Proposed
AKIYO	QP=28	40.109	40.225	107.61	106.81
	QP=32	37.555	37.764	62.32	62.69
	QP=36	35.130	35.424	38.69	39.28
	QP=40	32.573	32.877	24.61	25.23
FOREMAN	QP=28	37.076	37.040	362.77	355.89
	QP=32	34.689	34.683	210.78	207.03
	QP=36	32.462	32.480	129.04	127.85
	QP=40	30.277	30.346	83.07	82.91
MOBILE	QP=28	34.381	34.321	1718.54	1701.66
	QP=32	31.064	31.023	884.26	873.67
	QP=36	28.159	28.193	504.13	491.55
	QP=40	24.955	25.045	260.48	255.29
NEWS	QP=28	38.469	38.566	237.96	237.46
	QP=32	35.660	35.833	145.80	146.40
	QP=36	32.990	33.205	89.92	90.20
	QP=40	30.274	30.543	54.47	55.12
SALESMAN	QP=28	36.160	36.340	239.16	242.53
	QP=32	33.660	33.877	129.61	132.11
	QP=36	31.265	31.536	74.16	76.50
	QP=40	28.912	29.159	42.28	43.32

from the same five test video sequences. As we can see, the proposed algorithm shows better performance compared to H.264 standard.

Figure 4 shows the rate distortion curves for the FOREMAN and AKIYO. Left is the worst case and right is the best case among the results. As we can see, the rate distortion curves of the proposed algorithm located upper than the rate distortion curves of the H.264. Also the proposed algorithm shows the better performance in case of CIF.

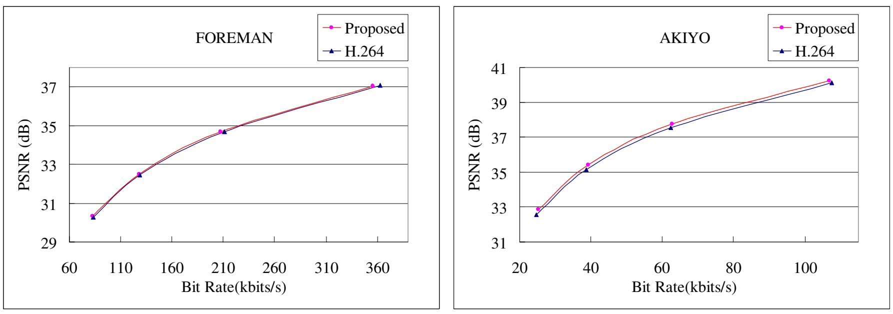

Fig. 4. Rate Distortion Curves for FOREMAN and AKIYO

5 Conclusions

In this paper, we have proposed an adaptive rate-distortion optimization algorithm for H.264. The proposed algorithm added the weighting factor to the previous RDO model. This weighting factor is changed according to the picture type and the standard deviation of the distortion for the current macroblock mode. In the picture level, we give more weights to the distortion part for I picture and we give more weight to the rate part of the RDO model for P picture. In the macroblock level, we give more weights to the rate part of the RDO model when the macroblock belongs to a flat area, and we give more weights to the distortion part of the RDO model when the macroblock belongs to a complex area. Experimental results show that the proposed algorithm achieves the better result compared to H.264. We obtain higher PSNR values though we use less or similar bit rate.

Acknowledgements

This work was supported in part by Gwangju Institute of Science and Technology (GIST), in part by the Ministry of Information and Communication (MIC) through the Realistic Broadcasting Research Center (RBRC), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

References

1. Pekka, S., Janne, H., Olli, S.: Selection of the Lagrange Multiplier for Block-Based Motion Estimation Criteria. Proceeding of International Conference on Acoustics, Speech, and Signal Processing, Vol. 3 (2004) 325-328
2. Tsai, C.J., Tang, C.W., Chen, C.H., Yu, Y.H.: Adaptive Rate-Distortion Optimization using Perceptual Hints. Proceeding of International Congress on Mathematical Education (2004) 667-670
3. Sullivan, G.J., Baker, R.L.: Rate-Distortion Optimized Motion Compensation for Compression Using Fixed or Variable Size Blocks. Global Telecommunications Conference, Vol. 1 (1991) 85-90
4. Shoham, Y., Gersho, A.: Efficient Bit Allocation for an Arbitrary Set of Quantizers. IEEE Transaction on Speech and Signal Processing, Vol. 36, No. 9 (1988) 1445-1453
5. Wiegand, T., Lightstone, M., Mukherjee, D., Campbell, T.G., Mitra, S.K.: Rate-Distortion Optimized Mode Selection for Very Low Bit Rate Video Coding and the Emerging H.263 Standard. IEEE Transaction on Circuit and System for Video Technology, Vol. 6, No. 2 (1996) 182-190
6. Wiegand, T., Girod, B.: Lagrangian Multiplier Selection in Hybrid Video Coder Control. Proceeding of International Conference on Image Processing (2001) 542-545
7. ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC: Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification. JVT Document JVT G050 (2003)
8. JVT Reference Software Version 9.5: http://iphome.hhi.de/suehring/tml/download/jm_old