

LAGRANGIAN MULTIPLIER OPTIMIZATION USING CORRELATIONS IN RESIDUES

Zhenyu Liu^{†*} Dongsheng Wang[†] Junwei Zhou^{*} Takeshi Ikenaga^{†*}

[†]TNList, Tsinghua University, Beijing 100084, China (liuzhenyu73,wds@tsinghua.edu.cn)

^{*}Oracle Corporation, Redwood Shores, CA 94065 USA

^{†*}Graduate School of IPS, Waseda University, Kitakyushu 808-0135, Japan

ABSTRACT

Rate distortion optimization (RDO) algorithm plays the vital role in the up to date hybrid video codec H.264/AVC. The RDO algorithm of H.264/AVC reference software is built up by assuming that the transformed residues are memoryless variables. However, our experiments reveal that, for some sequences, the strong temporal correlations exist in the prediction residues. This paper extends the Lagrangian optimization techniques by modeling the transformed residues as the first-order Markov source and calibrating the distortion model with the piecewise approximation function. The proposed algorithms adjust the Lagrangian multiplier dynamically to improve the overall coding quality. Comprehensive experiments testify that, as compared with the JM reference software, our optimizations can achieve up to 1.875dB coding gain. Moreover, our algorithms possess more robust coding performance and introduce less computational overhead than the Laplace distribution based methods. The inherent short process latency makes it possible to cooperate our algorithms with rate control operation. Last but not least, the proposed approach is also useful for the emerging standard, HEVC.

Index Terms— Lagrangian Multiplier, Rate Distortion Optimization, Hybrid Video Coding, H.264/AVC, HEVC

1. INTRODUCTION

The archetypal block-based hybrid video coder, such as the state-of-art video coding standard H.264/AVC [1], subdivides the current frame into macroblocks (MBs) and derives the prediction signals from previously decoded pixels in the same or other frames, which are denoted as Intra or Inter motion compensation, respectively. To improve the accuracy of motion-compensation, the prediction signals of one MB in H.264/AVC can be either replicated from previously decoded frames, or generated by using one choice of the seven Inter prediction modes or the seventeen Intra prediction modes. In turn, judiciously selecting the best prediction mode is the critical problem for the operation control of the encoder. For the variable block size motion estimation, the prediction errors are ameliorated with the increase of the partition number, but in this process, the rate cost for the side information, including the motion vector differences and the reference frame indices etc., is deteriorated. Obviously, the distortion cost and the rate overhead must be considered jointly in the optimal mode selection algorithm.

As many other international standards, H.264/AVC merely stipulates the stream syntax for the decoder, and leaves the encoder control optimizations open. In the current video codec, the mode selection problem is normally denoted as rate-distortion optimization (RDO) [2], which aims at minimizing the distortion under the given

rate constraint. The classical solution converts the constrained problem to the corresponding unconstrained counterpart by introducing the Lagrangian multiplier. The RDO algorithm recommended by H.264/AVC reference encoding software minimizes the Lagrangian cost function, which is written as

$$\min\{J = D(I, Q) + \lambda_{\text{mode}} \cdot R(I, Q)\}, \quad (1)$$

where I represents the coding mode, Q is the quantization interval. The Lagrangian multiplier λ_{mode} is employed to evaluate the rate cost $R(I, Q)$ with respect to the current distortion $D(I, Q)$. According to the analysis of [3], the rate-distortion curve is convex hull. The minimum of the Lagrangian cost function (1) is achieved when its derivative is equal to zero, i.e.,

$$\lambda_{\text{mode}} = -\frac{dD}{dR}. \quad (2)$$

In comparison with the heuristic Lagrangian multiplier estimation methods[4], the analytical counterparts [2, 5] have the advantages in sense of the computational complexity and the estimation accuracy. Assuming that the transformed residual coefficients are memoryless and are uniformly distributed within each quantization interval, Sullivan deduced that the Lagrangian multiplier should be linear with Q^2 [2]. The experimental results complied well with the theoretical analysis, especially with small and moderate quantization interval values. However, it was also found that the experimental Lagrangian multiplier amplitudes of some sequences, such as *news*, were larger than the analytical value. The underlying reasons for this deviation were not investigated.

In literature [5], Li formulated the adaptive Lagrangian multiplier by using the Laplace distributed transform coefficients assumption, and applying the unilateral differential entropy. Whereas, the unilateral differential entropy based rate model cannot remain invariant under the coordinate transformation[3]. Furthermore, to maintain the accuracy of the proposed RD-models, a set of complex escape methods are devised in [5] to detect the mismatched cases. Even so, for the high resolution video sequences, the proposed algorithms in [5] still deteriorated the coding quality slightly. Moreover, it is hard to efficiently embed the above algorithms into rate control operation.

In this paper, we develop the dynamic Lagrangian multiplier algorithms by exploiting the inherent correlations in the prediction residues and using the piecewise approximated distortion model. The extensive experiments demonstrate our proposals can obtain up to 1.875dB quality gain in BDPSNR, while maintaining the stable performance and the low the computational complexity. The low processing latency makes our algorithms improving the coding gain as embedded in rate control operation.

The rest of this paper is organized as follows. In Section 2, we provide the first-order Markov chain based rate model and the piecewise approximated distortion model, and then, describe the overall

^{*}This work is funded by TNList cross-discipline foundation and Nature Science Foundation of China (60833004).

Table 1. $E(|r|)$ over QCIF, CIF and HD720p

QCIF	$E(r)$	CIF	$E(r)$	HD720p	$E(r)$
<i>Container</i>	0.90386	<i>Container</i>	0.84011	<i>City</i>	0.35556
<i>News</i>	0.88103	<i>News</i>	0.83592	<i>Crew</i>	0.34822
<i>Missam</i>	0.64677	<i>Waterfall</i>	0.75274	<i>Mobcal</i>	0.32867
<i>Foreman</i>	0.50221	<i>Foreman</i>	0.35548	<i>Harbour</i>	0.31972
<i>Football</i>	0.17383	<i>Mobile</i>	0.32185	<i>Parkrun</i>	0.08503

adaptive Lagrangian multiplier algorithms in detail. Experiments are presented in Section 3 to justify our proposals. Finally, conclusions are drawn in Section 4.

2. ADAPTIVE LAGRANGIAN MULTIPLIER

The current rate-distortion models (RD-models) adopted in H.263 and H.264 are developed on two conditions: First, the transformed residues are memoryless variables; Second, the source are uniformly distributed within each quantization interval (Q). However, it is observed that, in some test sequences, the strong dependencies exist among the residues in the successive frames, and the uniform distribution assumption can not stand with the increase of Q . In this section, we first investigate the correlations of the prediction residues, and construct the corresponding first-order Markov chain based rate model. Thereafter, the piecewise approximated distortion model is developed. Finally, the overall Lagrangian multiplier formulations from the proposed RD-models are derived.

2.1. First-Order Markov Chain Based Rate Model

For each 4×4 -block C , with the corresponding prediction block P indicated by the motion vector, the correlation criterion of one 4×4 -block prediction residues is defined as

$$r = \frac{16 \sum c_{ij} p_{ij} - \sum c_{ij} \sum p_{ij}}{\sqrt{16 \sum c_{ij}^2 - (\sum c_{ij})^2} \sqrt{16 \sum p_{ij}^2 - (\sum p_{ij})^2}}, \quad (3)$$

where $i, j \in [0, 3]$, c_{ij} and p_{ij} represents the prediction residues of C and P , respectively. Several representative QCIF, CIF and HD720p sequences are selected to testify the temporal-domain correlations of residues. The averaging $|r|$ values over 100 P-frames of 15 sequences, denoted as $E(|r|)$, are illustrated in Table 1. It is revealed that the following properties are possessed by prediction residues:

1. Naturally, the prediction residues of dormant MBs, such as *Container* and *News*, always have the strong dependency in temporal-domain.
2. The correlations of the MBs with the detailed textures or the high motions are weak.
3. The correlations of residues diminish with the augment of the frame size. For the HD720p sequences, the maximum value of $E(|r|)$ is even less than 0.36. This is because that compacting the CCD sensor size inherently increases the noise-to-signal ration of the captured pictures.

Let $\Phi_t(u, v)$ designate the $N \times N$ block to be encoded, and $\hat{\Phi}_t(u, v)$ is the corresponding decoded signals. $\hat{\Phi}_t(u, v)$ is composed of $\Phi_t(u, v)$ and the additional quantization noise D_t . Let $\Delta_{t+1}(u, v) = \Phi_{t+1}(u, v) - \Phi_t(u, v)$ and $\hat{\Delta}_{t+1}(u, v) = \Phi_{t+1}(u, v) - \hat{\Phi}_t(u, v)$ are zero-mean random variables. The

variances of $\hat{\Delta}_{t+1}(u, v)$ and $\Delta_{t+1}(u, v)$, which are labeled as $\delta_{\hat{\Delta}_{t+1}}^2(u, v)$ and $\delta_{\Delta_{t+1}}^2(u, v)$, can be formulated as

$$\delta_{\hat{\Delta}_{t+1}}^2(u, v) = \delta_{\Delta_{t+1}}^2(u, v) + D_t. \quad (4)$$

When $\hat{\Delta}_t(u, v)$ and $\hat{\Delta}_{t+1}(u, v)$ are viewed as the first-order Markov source with the correlation r , with the small quantization errors, T. Berger [3] deduced the explicit rate-distortion expression

$$R_t = \frac{1}{2} \ln \frac{(1-r^2)\delta_{\hat{\Delta}_t}^2(u, v)}{D_t}, \quad (5)$$

in which, $\delta_{\hat{\Delta}_t}^2(u, v)$ denominates the variance of $\hat{\Delta}_t(u, v)$. Equation (5) yields the Lagrangian multiplier as

$$\lambda_{\text{mode}} = 2D_t, \quad (6)$$

which is identical to the memoryless one deduced in literature [2].

On the other hand, for the larger values of quantization errors, the precise rate model employs the Clausen's function, which is expressed with the tabulation form [6]. To derive the closed form of rate model with respect to the current quantization interval Q_t , we simplify the investigation by following the motion estimation procedure adopted by the hybrid video coder. As pointed out by (4), the perturbation of Q_t affects the rates of not only the current t^{th} frame, but also the following $(t+1)^{\text{th}}$ frame, which uses the t^{th} decoded pixels as the prediction signals. The impact of the perturbation of Q_t to the rates is written as

$$\frac{dR}{dQ_t} = \frac{dR_t}{dQ_t} + \frac{dR_{t+1}}{dQ_t}, \quad (7)$$

in which, R_t and R_{t+1} represent the rates of the current and the next frames, respectively. In practice, increasing Q_t compresses the rates of the t^{th} frame R_t , whereas dilates the rates of the next frame R_{t+1} , coming from the increased $\delta_{\Delta_{t+1}}^2(u, v)$, as shown by (4). Supposing that

$$R_t = \frac{1}{2} \ln \left(a \cdot \frac{\delta_{\hat{\Delta}_t}^2(u, v)}{D_t} \right) \quad (8)$$

$$R_{t+1} = \frac{1}{2} \ln \left(a \cdot \frac{\delta_{\hat{\Delta}_{t+1}}^2(u, v)}{D_{t+1}} \right),$$

where, $a = 1$ for Gaussian distributed residues, and $a = \frac{e}{\pi}$ for Laplace distributed residues, substituting (4) and (8) into (7) yields

$$\lambda_{\text{mode}} = -\frac{dD_t}{dQ_t} \frac{dQ_t}{dR} = \frac{2D_t \delta_{\hat{\Delta}_{t+1}}^2(u, v)}{\delta_{\hat{\Delta}_{t+1}}^2(u, v) - D_t}. \quad (9)$$

It should be noted that D_{t+1} is determined by Q_{t+1} , and hence, dD_{t+1}/dQ_t is equal to zero in the deduction of (9).

2.2. Piecewise Approximated Distortion Model

Gaussian distribution and Laplace distribution are both generally adopted models for analyzing the transformed residues. In this section, we first investigate the distortion model based on both Laplace and Gaussian distributions via the numerical method, and then derive the piecewise approximated distortion model.

The whole distortion is achieved by summing up the error components in each quantization interval, which is formulated as

$$D_C = a\delta^2 \left[\int_0^{\frac{5}{6}\gamma} y^2 p(y) dy + \sum_{i=1}^N \int_{i\gamma - \frac{1}{6}\gamma}^{i\gamma + \frac{5}{6}\gamma} (y - i\gamma)^2 p(y) dy \right], \quad (10)$$

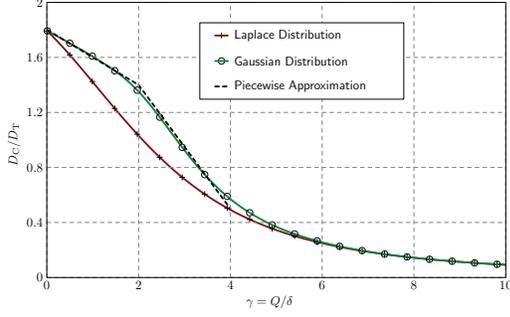


Fig. 1. Piecewise Approximated Analysis of Distortion Model D_C in which, δ represents the variance of transform coefficients, $\gamma = Q/\delta$, $p(y) = \exp(-0.5y^2)$ and $a = \sqrt{2/\pi}$ for Gaussian distribution, and $p(y) = \exp(-\sqrt{2}y)$ and $a = \sqrt{2}$ for Laplace distribution. For the original distortion model adopted by H.264/AVC is

$$D_T = 0.1088 \cdot Q^2 \approx 0.68 \cdot 2^{\frac{QP-12}{3}}, \quad (11)$$

the ratio between D_C and D_T is depicted by Fig.1.

In our experiments, the distortion model developed on the Gaussian distribution outperforms the Laplace based one. Therefore, we apply the following piecewise approximation to explicitly express the distortion model.

$$\tilde{D} = \begin{cases} (1.8 - 0.2\gamma) \cdot 0.68 \cdot 2^{\frac{QP-12}{3}} & \gamma \leq 2 \\ (2.3 - 0.45\gamma) \cdot 0.68 \cdot 2^{\frac{QP-12}{3}} & 2 < \gamma \leq 4 \end{cases} \quad (12)$$

In our distortion model, we neglect the cases when $\gamma > 4$. This is mainly because that, when the coefficient value is much less than Q , it is prone to be eliminated as zero during the quantization, which will play the trivial role in the Lagrangian optimization.

2.3. Adaptive Lagrangian Multiplier Algorithm

The proposed Lagrangian multiplier updating is conducted in an MB-by-MB manner for P- and B-frames. Namely, after the coding of one MB, with the obtained DCT coefficients and motion vectors of current MB, the Lagrangian multiplier is adjusted accordingly, and the updated Lagrangian multiplier will be applied in the following MB coding process. From (6), (9), and (12), the dynamic Lagrangian multiplier algorithm is described as Fig. 2.

$\delta_{\Delta t}^2(u, v)$, in which $u, v \in \{0, 1, 2, 3\}$, denotes the averaging variances of 4×4 -DCT coefficients indexed with (u, v) , which are derived from the 256×4 -DCT coefficients of current MB. The averaging motion vector amplitude of 16×4 -blocks in current MB is computed as

$$\begin{aligned} \overline{mvLi} &= \frac{1}{16} \sum_{j=0}^3 \sum_{k=0}^3 (|mvxLi(j, k)| + |mvyLi(j, k)|) \\ \overline{mv}_{\max} &= \max(\overline{mvL0}, \overline{mvL1}) \end{aligned} \quad (13)$$

in which, the subscript $Li(i \in \{0, 1\})$ denominates the reference picture list, i.e., list_0 or list_1. If \overline{mv}_{\max} is greater than one-pixel, the current macroblock is identified as the fast motion block, of which the DCT coefficients are assumed to possess the weak correlations in temporal-domain, and hence, the change of Lagrangian multiplier is equal to $s_0 \tilde{D}_t$; On the other hand, for the dormant MB, we apply the

```

n = 0; Δλ = 0;
Loop u and v
  If (δΔt2(u, v) > 0) {
    γ = Q / δΔt2(u, v); Derive D̃t from (12);
    If ((m̄vmax > 1-pel and γ ≤ 4) or (m̄vmax ≤ 1-pel and γ < 0.5)) {
      Δλ = Δλ + s0 D̃t; n ++;
    }
    Else If (m̄vmax ≤ 1-pel and γ ≤ 4) {
      Δλ = Δλ + s0 s1 D̃t * (δΔt2(u, v) / (δΔt2(u, v) - D̃t)); n ++;
    }
  }
End Loop
If (n > 0) λmode = (1 - w)λmode + wΔλ/n;

```

Fig. 2. Pseudo Coding of Updating Lagrangian Multiplier for the t^{th} MB (w is the reciprocal of MB number in one frame)

approach as shown in Section 2.1 to refine the Lagrangian multiplier value.

It should be noted the inherent hindrance in (9). That is, when encoding the t^{th} frame, it is infeasible to obtain $\delta_{\Delta t+1}^2(u, v)$, which belong to the following frame. Supposing the strong dependencies between the successive frames, we use $\delta_{\Delta t}^2(u, v)$ to approximate $\delta_{\Delta t+1}^2(u, v)$ in Fig.2. s_0 and s_1 are the empirical parameters. s_0 is inherited from the JM reference software, which is defined as

$$s_0 = \begin{cases} 1.0 & \text{P-frame} \\ \text{clip}\left(2.0, 4.0, \frac{QP-12}{6.0}\right) & \text{B-frame} \end{cases}, \quad (14)$$

where the function $\text{clip}(a, b, x)$ confines the value of x in the range of $[a, b]$. s_1 is determined by the slice type, the quantization parameter (QP), the successive B-frame number (β), \overline{mv}_{\max} , and γ , which is expressed as follows. With the increase of successive B-frame number β , the correlations of P-frame residues are weakened, and the magnitude of s_1 shrinks accordingly.

$$s_1 = \begin{cases} \text{B-Frame} \rightarrow \begin{cases} 1.0 & \gamma < 2 \text{ or } QP > 28 \\ 0.9 & \text{otherwise} \end{cases} \\ \text{P-Frame} \rightarrow \begin{cases} QP > 28 \rightarrow \begin{cases} \max(0.7, 1.0 - 0.05\beta) & \gamma < 2 \\ \max(0.7, 1.0 - 0.10\beta) & \gamma \geq 2 \end{cases} \\ QP \leq 28 \rightarrow \begin{cases} \overline{mv}_{\max} = 0 \rightarrow \begin{cases} \max(0.7, 1.0 - 0.1\beta) & \gamma < 2 \\ \max(0.7, 0.9 - 0.1\beta) & \gamma \geq 2 \end{cases} \\ \overline{mv}_{\max} \neq 0 \rightarrow \begin{cases} \max(0.5, 0.75 - 0.1\beta) & \gamma < 2 \\ \max(0.4, 0.7 - 0.15\beta) & \gamma \geq 2 \end{cases} \end{cases} \end{cases} \end{cases}, \quad (15)$$

3. EXPERIMENTAL RESULT

The following experiments integrated the proposed algorithms into the H.264/AVC reference software JM17.0, and the original method was used as the benchmark. The simulation conditions were defined according to the recommendations of literature [7]. Specifically, The test sequences were in YUV 4:2:0 format; We adopted a single slice per picture; The search ranges were defined as ± 16 for QCIF sequences, ± 32 for CIF sequences, ± 48 for SD576i sequences, and

Table 2. Coding Quality Analysis of QCIF and CIF

Seq. (QCIF)	BDBR (%)	BDPSNR (dB)	Seq. (CIF)	BDBR (%)	BDPSNR (dB)
<i>Container</i>	-17.539	+0.86667	<i>Container</i>	-14.424	+0.48022
<i>News</i>	-5.7871	+0.27214	<i>News</i>	-4.7685	+0.21186
<i>Hall</i>	-5.5695	+0.27547	<i>Akiyo</i>	-13.736	+0.58285
<i>Saleman</i>	-11.107	+0.43747	<i>Waterfall</i>	-7.3968	+0.24781
<i>Foreman</i>	-1.7447	+0.07044	<i>Foreman</i>	-0.9349	+0.03721
<i>Carphone</i>	-3.3232	+0.13423	<i>Tempete</i>	-0.7798	+0.02435
Average	-7.5118	+0.33427	Average	-7.0067	+0.26405

Table 3. Coding Quality Analysis of SD576i and HD720p

Seq. (SD576i)	BDBR (%)	BDPSNR (dB)	Seq. (720p)	BDBR (%)	BDPSNR (dB)
<i>Stockholm</i>	-2.1979	+0.08472	<i>City</i>	-2.5416	+0.08571
<i>Mobcal</i>	-1.7498	+0.06291	<i>Mobcal</i>	-1.7892	+0.04751
<i>Shields</i>	-1.0721	+0.03892	<i>Shields</i>	-0.0092	+0.00080
<i>Football</i>	-0.0188	+0.00119	<i>Harbour</i>	-2.7132	+0.08192
Average	-1.2597	+0.04694	Average	-1.7633	+0.05399

± 64 for HD720p sequences, respectively; The reference frame number was set as 5; The high complexity rate-distortion optimization (RDO_ON) and the GOP with one leading Intra prediction frame (I-frame) followed by the Inter prediction frames (P- or B-frames) were applied.

To evaluate the coding quality of the luminance and the chrominance components jointly, in the coding quality comparisons, we adopted the combined PSNR metric (P_{YUV}) [5] as follows

$$P_{YUV} = (4P_Y + P_U + P_V)/6. \quad (16)$$

The fixed QP experiments are carried out with $QP = \{28, 32, 36, 40\}$. The quantitative coding efficiency analysis was conducted on the basis of the average P_{YUV} gain (BDPSNR) and the average rate reduction (BDBR) [8], respectively. The (+) sign in BDPSNR and (-) sign in BDBR indicate the coding gain.

Twelve representative QCIF and CIF video sequences with 30-fps frame rate and various motion and texture features were tested to verify the proposed algorithms under baseline profile, in which one leading I-frame followed by 99 P-frames and CAVLC entropy coding were applied. The associated results in terms of BDPSNR and BDBR were summarized in Table 2. The averaged coding gains are 0.334dB and 0.264dB for QCIF and CIF sequences, respectively. BDPSNR represents the average PSNR gain. In fact, the peak PSNR gain is greater than the BDPSNR value. For example, the peak PSNR gain (1.06133dB) was obtained in *Container_QCIF*. It is observed that the coding gain was limited for the sequences with weakly correlated residues, such as *Foreman* and *Tempete*. This phenomenon was well consistent with the theoretical investigation. The averaging coding gain for low resolution sequences is 0.299dB, approaching the performance (0.34dB) of the work in literature [5].

Four SD576i sequences with 25-fps frame rate and four HD720p sequences with 50-fps frame rate, were used to verify the performance of our algorithms on the high resolution sequences. The adopted main profile settings include CABAC and GOP=IBBP (I+59P+118B). The coding quality results are shown as Table 3. As mentioned in Section 2.1, the high resolution sequences generally possess low correlations in residues, and this feature constrains the performance of our methods.

The prominent advantage of our proposals is that our algorithms can provide coding gain when embedded in rate-control. The cod-

Table 4. Coding Results with Rate Control On

Seq. (QCIF)	BDBR (%)	BDPSNR (dB)	Seq. (CIF)	BDBR (%)	BDPSNR (dB)
<i>Container</i>	-29.994	+1.87481	<i>Container</i>	-14.715	+0.48490
<i>News</i>	-5.8356	+0.28421	<i>News</i>	-3.8683	+0.17865
<i>Hall</i>	-4.0241	+0.20254	<i>Akiyo</i>	-2.6878	+0.09243
<i>Saleman</i>	-8.6066	+0.35446	<i>Waterfall</i>	-4.7754	+0.17292
<i>Foreman</i>	-1.4405	+0.05584	<i>Foreman</i>	-0.7499	+0.02775
<i>Carphone</i>	-3.0881	+0.13169	<i>Tempete</i>	-0.0083	+0.00053
Average	-8.8314	+0.4839	Average	-4.4675	+0.15953

ing results of representative sequences with rate control enabled are illustrated in Table 4.

As compared with the Laplace distribution based Lagrangian multiplier [5], the proposed Lagrangian multiplier optimization achieves the similar coding quality for the low resolution video sequences, while possessing the following advantages

- The computational complexity of our algorithms is lower than the counterpart. When employing fast full search and five reference frames, the additional coding time of our tests is less than 0.03%. In contrast, this merit is 0.05% in [5].
- Our algorithms have robust coding performance. Specifically, on average, 0.047dB and 0.054dB coding quality gains were obtained for our SD576i and HD720p tests, respectively. In contrast, the schemes proposed in [5] introduced an averaged -0.02dB coding quality loss.
- The prominence lies on that our Lagrangian multiplier schemes cooperate well with rate control operation, which is widely adopted in practical applications.

4. CONCLUSIONS

In this paper, we construct the adaptive Lagrangian multiplier by developing the first-order Markov chain based rate model and the piecewise approximated distortion model. The proposed Lagrangian multiplier algorithms universally contributes to the coding quality improvement. For low resolution video coding tests, up to 1.875dB gain in BDPSNR is achieved. As compared with the original Laplace distribution based Lagrangian multiplier schemes, our proposals have the advantages in sense of low computational intensity, robust coding performance, and being friendly to rate control operation.

5. REFERENCES

- [1] Thomas Wiegand et al., "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, July 2003.
- [2] Gary J. Sullivan and Thomas Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74–90, Nov. 1998.
- [3] Toby Berger, *Rate Distortion Theory*, Prentice Hall, 1971.
- [4] Thomas Wiegand et al., "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 2, pp. 182–190, Apr. 1996.
- [5] Xiang Li et al., "Laplace distribution based Lagrangian rate distortion optimization for hybrid video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 2, pp. 193–205, Feb. 2009.
- [6] Leonard Lewin, *Dilogarithms and associated functions*, Macdonal, 1958.
- [7] Gary J. Sullivan, "Recommended simulation common conditions for H.26L coding efficiency experiments on low-resolution progressive-scan source material," Santa Barbara, USA, Sept. 2001, VCEG-N81.
- [8] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," Austin, Texas, USA, Apr. 2001, VCEG-M33.