

LA-RDO Based Error Resilient Coding Using Multi-Layer Lagrange Multiplier Selection in SVC

Muhammad Shoaib, Anni Cai

Multimedia Communication Centre

Beijing University of Posts and Telecommunications, P.R. China

moshoaib@gmail.com, annicai@bupt.edu.cn

Abstract— Loss-aware rate distortion optimized (LA-RDO) MB mode decision has been introduced in scalable video coding (SVC) to minimize the quality degradation impact of transmission errors. It aims to optimize rate-distortion performance considering transmission error and resulted error propagation. However, Lagrange multiplier (λ) used in LA-RDO based mode decision does not consider the correlations between layers. Consequently, λ is not optimal for multi-layer scenario in error resilient coding. In this paper, modified λ selection method in multi-layer scenario, for mode decision of LA-RDO based error resilient coding is proposed. Simulation results show that significant PSNR gain is reportedly achieved for various tested sequences compared to JSVM 9.8 with quality and spatial scalabilities for IPPP coding, as well as for hierarchical B picture coding.

Keywords— error resilient coding; error concealment; Lagrange multiplier; loss-aware rate distortion optimization; scalable video coding

I. INTRODUCTION

Error resilient coding, one of the error control method in joint scalable video model (JSVM), is an encoding technique at the encoder side that helps decoder in recovery or concealment from potential channel errors. The objective of error resilient coding is to design a scheme which can achieve the minimum end-to-end distortion under certain rate. In SVC, when encoding a macroblock (MB) in an enhancement layer picture, the traditional MB coding modes inherited from single-layer coding as well as new inter-layer prediction modes can be used. Similar as in single-layer coding, MB mode selection in SVC also affect the error resilient performance of the encoded bit-stream.

The Lagrange multiplier (λ) based rate-distortion optimization (RDO) technique has been extensively used in video coding to select the best mode for the given bit-rate with minimum distortion by appropriate selection of coding parameters. RDO multi-layer encoder control is presented in [1] to trade-off the base layer (BL) and enhancement layer (EL) coding efficiencies by joint optimization of BL and EL coding parameters. Extending from [1], multi-layer Lagrange multiplier selection for SVC is presented in [2] by considering correlation between layers.

In [3] RDO for single-layer coding in a packet loss environment is presented. In this method, given the target packet loss rate (PLR), the 4x4 block-base error prorogation map for a picture is calculated, and the map is taken into account to perform mode decision for the picture in the later. In [4], single-layer mode decision method of [3] is extended to multi-layer SVC scenario.

In [4] the Lagrange multiplier, used to calculate optimized mode in a packet loss environment, is directly applied in a multi-layer SVC environment, as in the conventional single-layer coding. However, such a direct application is not very efficient since the correlation between layers is not considered in loss-aware rate distortion optimization (LA-RDO) based λ selection. Exploiting strong correlation between two consecutive layers, considerable gain should be achieved using multi-layer Lagrange multiplier (ML- λ) selection for mode decision in an error prone environment. In order to achieve better performance in both quality and spatial scalabilities in a packet loss environment while keeping low computation cost, ML- λ parameter should also be considered, to calculate the cost of LA-RDO based MB mode decision.

The rest of this paper is organized as follows. First, analysis of the multi-layer Lagrange multiplier selection, LA-RDO based algorithm, and the proposed implementation are presented in details in Section II. Simulation results and related discussions are presented in Section III. Finally, Section IV concludes the paper.

II. THE MULTI-LAYER LAGRANGE MULTIPLIER SELECTION, LA-RDO BASED ALGORITHM AND THE PROPOSED IMPLEMENTATION

A. Multi-Layer Lagrange Multiplier Selection

The λ based rate-distortion optimization (RDO) techniques have been extensively used in hybrid video coding. Its target is to minimize the distortion D_S for the given rate R_S by appropriate selection of coding parameters [5].

$$C_{MODE}(p) = D_S(p) + \lambda_{MODE}(p) \cdot R_S(p) \quad (1)$$

where C_{MODE} denotes the Lagrangian cost function for mode selection, p is the coding parameter, and λ_{MODE} is the Lagrangian multiplier. Supposing D_S and R_S are differentiable

everywhere and are related with the quantization Q , the total differentials of rate and distortion, assuming a sufficiently high rate environment, are given as [2]:

$$\begin{aligned} dR_s &= \frac{\partial R_s}{\partial Q} dQ = \frac{-2a}{Q \ln 2} dQ \\ dD_s &= \frac{\partial D_s}{\partial Q} dQ = \frac{Q}{6} dQ \end{aligned} \quad (2)$$

The minimization of C_{MODE} is given by setting its derivative to zero in (1), and λ_{MODE} becomes:

$$\lambda_{MODE}(Q) = -\frac{dD_s(Q)}{dR_s(Q)} = c.Q^2 \quad (3)$$

where $c = \ln 2 / (12a)$ is a constant which is experimentally suggested to be 0.85 [6] though other proposed 0.68 [7].

In practice, this selection method is efficient for single-layer coding. An empirical multi-layer λ selection method proposed for quality SVC in [8]. Moreover, multi-pass RDO method for SVC presented in [1] in which two neighbouring layers were jointly optimized while keeping λ for each layer unchanged. In this paper we follow the method in [2, 9] with multi-layer λ in scalable video coding, in which all layers are jointly optimized in order to achieve the best performance. Without loss of generality, two layers scenario is considered here. The minimization of rate-distortion cost is extended as:

$$\min\{C\} = \min\{(1-w).C_0 + w.C_1\} \quad (4)$$

where C_0 and C_1 represent cost for BL and EL respectively, and $w \in [0, 1]$ represents layer weighting factor which controls the trade-off between BL and EL coding efficiencies. More concretely, C in (4) is written as:

$$\begin{aligned} \min\{C\} &= \min_{\{p_0, p_1 | p_0\}} (1-w).(D_0(p_0) + \lambda_0.R_0(p_0)) \\ &+ w.(D_1(p_1 | p_0) + \lambda_1.(R_1(p_1 | p_0) + R_0(p_0))) \end{aligned} \quad (5)$$

where p_0 and p_1 represent the product space of parameter selection, λ_0 and λ_1 represent Lagrange multipliers, D_0 and D_1 represent distortions with corresponding rates R_0 and R_1 for BL and EL respectively. It should be noted that, for the calculation of C_1 , R_0 is included along with R_1 in the second term in (5), since accumulated bit-rate instead of self bit-rate for EL layer is used to evaluate the coding efficiency. Moreover, for spatial scalability frame size of EL is also considered as (5) represents rate and distortion at pixel level, so a resolution factor is introduced in EL cost. Relating the product space p_0 and p_1 in rate-distortion with the quantization step Q , (5) is extended to:

$$C = (1-w).(D(Q_0) + \lambda_0.R(Q_0)) + w.(D(Q_1).\beta + \lambda_1.(R(Q_1).\beta + R(Q_0))) \quad (6)$$

where Q_0 and Q_1 are quantization steps for the BL and EL respectively, and β represents the resolution ratio between two layers. As well known, for the same content at different resolutions, bit-rate ratio may be approximated by the resolution ratio although the two ratios are not precisely the same. Applying this assumption in (6), to minimize the cost, setting derivative to zero, λ_1 can be solved as [2]:

$$\lambda_1 = \frac{Q_1((1-w)(\ln(2)Q_0^2 - 12a\lambda_0) + \ln(2)\beta w Q_1 Q_0)}{12aw(\beta Q_0 + Q_1)} \quad (7)$$

Since there is no neighbouring lower layer for layer zero, λ_0 is determined by single-layer λ selection method i.e. $\lambda_0 = c.Q_0^2$. With optimized BL, setting $w=1$, λ_1 can be reduced as:

$$\lambda_1 = \frac{\beta.Q_0}{\beta.Q_0 + Q_1} (c.Q_1^2) = \frac{\beta.Q_0}{\beta.Q_0 + Q_1} \lambda^*(Q_1) \quad (8)$$

where $\lambda^*(Q_1)$ is actually the Lagrange multiplier evaluated using single-layer λ selection method as in (3). In general, λ_1 in (8) extended to layer $i+1$:

$$\lambda_{i+1} = \frac{\beta.Q_i}{\beta.Q_i + Q_{i+1}} \lambda^*(Q_{i+1}) \quad (9)$$

where Q_i and Q_{i+1} are the quantization steps for layer i and layer $i+1$ respectively. According to the specification in [10], the quantization steps Q_i and Q_{i+1} are derived as:

$$\begin{aligned} Q_i &= 2^{(QP_i - 12)/6} \\ Q_{i+1} &= 2^{(QP_{i+1} - 12)/6} \end{aligned} \quad (10)$$

where QP_i and QP_{i+1} are the quantization parameters. Applying (10) into (9), Lagrange multiplier λ_{i+1} for layer $i+1$ leads to:

$$\lambda_{i+1} = \frac{\beta.2^{\Delta QP_{i,i+1}/6}}{\beta.2^{\Delta QP_{i,i+1}/6} + 1} \lambda^*(Q_{i+1}) \quad (11)$$

where $\Delta QP_{i,i+1} = QP_i - QP_{i+1}$ describes the difference between the quantization parameters for layers i and $i+1$.

B. Loss-aware Rate Distortion Optimized MB Mode Decision

In scalable coding with multiple layers, the MB mode decision for the base layer pictures is exactly the same as the H.264/AVC single-layer method. For a slice in an enhancement layer picture, if no inter-layer prediction is used, the single-layer method is adopted, with the used loss rate being the loss

rate of the current layer. Otherwise, if inter-layer prediction is used, new MB modes that use inter-layer texture, motion or residual prediction can be employed.

In LA-RDO based MB mode selection, the distortion of the reconstructed MB resulted from possible packet losses with the given coding rate is estimated for all candidate modes, and the mode with the minimal RDO cost is selected. However, in existing JSVM software [4], λ used for LA-RDO based coding is identical for all layers. Exploiting strong inter-layer prediction correlation between neighbouring layers, we propose a modified λ i.e. multi-layer Lagrange multiplier (ML- λ) for enhancement layers. To encode the bit-stream in an error prone environment, besides rate-distortion and target packet loss in the transmission channel, the proposed ML- λ should also be considered to select the best mode.

In the following, existing LA-RDO for multi-layer coding with same λ for each layer in [4] is extended to propose LA-RDO with ML- λ , using the correlation between neighbouring layers in error prone environment.

Let the current layer containing the current MB be l_c , the lower layer containing the collocated MB used for inter-layer prediction of the current MB be l_{c-1} , the further lower layer containing the MB used for inter-layer prediction of the collocated MB in l_{c-1} be l_{c-2} , ..., and the lowest layer containing an inter-layer dependent block for the current MB as l_0 , and let the packet loss rates at $l_c, l_{c-1}, \dots, l_0$ layers be $p_{l_c}, p_{l_{c-1}}, \dots, p_{l_0}$, respectively. For a current slice that may use inter-layer prediction, it is assumed that the current-layer MB would be decoded only if the current MB and all the dependent lower-layer blocks are received and can be decoded correctly, otherwise the slice is concealed. For a slice that does not use inter-layer prediction, the content MB would be decoded as long as it is received.

1) Distortion Calculation: The overall distortion of the m^{th} MB in the n^{th} picture in layer l_c with the candidate coding option o is represented by [4]:

$$D(n, m, o) = \left(\prod_{i=0}^c (1 - p_{l_i}) \right) (D_s(n, m, o) + D_{ep_ref}(n, m, o)) + \left(1 - \prod_{i=0}^c (1 - p_{l_i}) \right) D_{ec}(n, m) \quad (12)$$

where $D_s(n, m, o)$ and $D_{ep_ref}(n, m, o)$ denote the source coding distortion and the error propagation distortion respectively; and $D_{ec}(n, m)$ denotes the error concealment distortion in case the MB is lost. For the calculation of the error propagation distortion $D_{ep_ref}(n, m, o)$, a distortion map D_{ep} for each picture on a block-basis (e.g. 4x4 luma samples) is defined. Given the distortion map, $D_{ep_ref}(n, m, o)$ is calculated as:

$$D_{ep_ref}(n, m, o) = \sum_{k=1}^K D_{ep_ref}(n, m, k, o) = \sum_{k=1}^K \sum_{l=1}^4 w_l D_{ep}(n_l, m_l, k_l, o_l) \quad (13)$$

where K is the number of blocks in one MB, and $D_{ep_ref}(n, m, k, o)$ denotes the error propagation distortion of the k^{th} block in the current MB. $D_{ep_ref}(n, m, k, o)$ is calculated as the weighted average of the error propagation distortion $D_{ep}(n_l, m_l, k_l, o_l)$ of the blocks k_l in MB m_l of reference frame n_l that are referenced by the current block. The weight w_l of each reference block is proportional to the area that is being used as reference.

The distortion map of higher layer is derived using distortion map of lower layer. When the current layer l_c is of a higher spatial resolution, the distortion map of the lower layer l_{c-1} is first up-sampled. For example, if the resolution is changed by a factor of 2 for both the width and the height, then each value in the distortion map of lower layer is simply up-sampled to be a 2 by 2 block of identical values. The distortion map with the optimal coding mode o^* is defined as [4]:

$$D_{ep}(n, m, k) = \left(\prod_{i=0}^c (1 - p_{l_i}) \right) D_{ep_ref}(n, m, k, o^*) + \left(1 - \prod_{i=0}^c (1 - p_{l_i}) \right) (D_{ec_rec}(n, m, k, o^*) + D_{ec_ep}(n, m, k)) \quad (14)$$

where $D_{ep_ref}(n, m, k, o^*)$ is the error propagation distortion of k^{th} block in collocated MB of lower layer in case of texture prediction, or of the same layer in the case of motion and residual prediction. $D_{ec_rec}(n, m, k, o^*)$ is the distortion between the error-concealed block and the reconstructed block, and $D_{ec_ep}(n, m, k)$ is the distortion due to error concealment and the error propagation distortion in the reference picture that is used for error concealment.

The distortion map of various MB modes using inter-layer predictions in SVC are presented below [4]:

Texture Prediction: In this mode, distortion can be propagated from the lower layer. Then the distortion map of the k^{th} block in the current MB is calculated as in (14). Note that here $D_{ep_ref}(n, m, k, o^*)$ is the distortion map of the k^{th} block in the collocated MB in the lower layer l_{c-1} .

Motion Prediction: The motion prediction in JSVM use the motion vector field, reference indices and MB partitioning of the lower layer for the corresponding MB in the current layer. The inter-prediction process still uses the reference pictures in the same layer. For a block that uses inter-layer motion prediction and does not use bi-prediction, the distortion map of the k^{th} block is calculated the same as in (14).

For a block that uses inter-layer motion prediction and also uses bi-prediction, the distortion map of the k^{th} block in the current MB is:

$$\begin{aligned}
D_{ep}(n, m, k) = & w_{r0} \times \left(\prod_{i=0}^c (1 - p_{l,i}) \right) D_{ep_ref_r0}(n, m, k, o^*) \\
& + \left(1 - \prod_{i=0}^c (1 - p_{l,i}) \right) (D_{ec_rec}(n, m, k, o^*) + D_{ec_ep}(n, m, k)) \\
& + w_{r1} \times \left(\prod_{i=0}^c (1 - p_{l,i}) \right) D_{ep_ref_r1}(n, m, k, o^*) \\
& + \left(1 - \prod_{i=0}^c (1 - p_{l,i}) \right) (D_{ec_rec}(n, m, k, o^*) + D_{ec_ep}(n, m, k))
\end{aligned} \quad (15)$$

where $D_{ep_ref_r0}(n, m, k, o^*)$ and $D_{ep_ref_r1}(n, m, k, o^*)$ are the distortion maps of the k^{th} block calculated from the reference pictures in the same layer and w_{r0} and w_{r1} are the weights of the reference pictures used for bi-prediction, respectively.

Residual Prediction: If the low layer is received, and the residue of the low layer can be decoded correctly, there is no error propagation. Otherwise, the error concealment is performed. Therefore, (14) and (15) can also be used to derive the distortion map for an MB mode using inter-layer residual prediction.

No Inter-Layer Prediction: For an inter-coded block, (14) and (15) are used to generate the distortion map, while for an intra-coded block:

$$\begin{aligned}
D_{ep}(n, m, k) = & \left(1 - \prod_{i=0}^c (1 - p_{l,i}) \right) (D_{ec_rec}(n, m, k, o^*) \\
& + D_{ec_ep}(n, m, k))
\end{aligned} \quad (16)$$

2) Lagrange Multiplier Selection: Substituting distortion map of k^{th} block $D_{ep}(n, m, k)$ for different modes into (13), we get error propagation distortion of MB due to reference picture $D_{ep_ref}(n, m, o)$. The cost of coding mode is derived by substituting distortion from (12) into (1), we get:

$$\begin{aligned}
C = & \left(\prod_{i=0}^c (1 - p_{l,i}) \right) (D_s(n, m, o) + D_{ep_ref}(n, m, o)) \\
& + \left(1 - \prod_{i=0}^c (1 - p_{l,i}) \right) D_{ec}(n, m) + \lambda_{ed} R
\end{aligned} \quad (17)$$

where λ_{ed} is the Lagrange multiplier in the error distortion environment.

C. Proposed Implementation

The minimization of mode cost C is given by setting its derivatives to R be zero, then λ_{ed} in (17) becomes:

$$\lambda_{ed} = - \left(\prod_{i=0}^c (1 - p_{l,i}) \right) \frac{dD_s(n, m, o)}{dR} = \left(\prod_{i=0}^c (1 - p_{l,i}) \right) \lambda_{ef,c} \quad (18)$$

here $\lambda_{ef,c} = -dD_s(n, m, o)/dR$, is the multi-layer Lagrange multiplier for layer l_c in the error free environment. $\lambda_{ef,c}$ is related with quantization parameter and can be expressed as given in (11):

$$\lambda_{ef,c} = \frac{\beta \cdot 2^{\Delta QP_{c-1,c}/6}}{\beta \cdot 2^{\Delta QP_{c-1,c}/6} + 1} \lambda_{ef}^*(Q_c) \quad (19)$$

where $\Delta QP_{c-1,c} = QP_{c-1} - QP_c$ is the difference of quantization parameter between layer l_c and the corresponding lower layer l_{c-1} . $\lambda_{ef}^*(Q_c)$ is the Lagrange multiplier for layer l_c corresponding to quantization Q_c using existing single-layer coding method adopted in JSVM in the error free environment. Plugging $\lambda_{ef,c}$ from (19) into (18), λ_{ed} for l_c layer becomes:

$$\lambda_{ed} = \prod_{i=0}^c (1 - p_{l,i}) \frac{\beta \cdot 2^{\Delta QP_{c-1,c}/6}}{\beta \cdot 2^{\Delta QP_{c-1,c}/6} + 1} \lambda_{ef}^*(Q_c) \quad (20)$$

Finally, the cost of coding mode using proposed multi-layer Lagrange multiplier (ML- λ) in error prone environment is evaluated by plugging λ_{ed} from (20) into (17), we get:

$$\begin{aligned}
C = & \left(\prod_{i=0}^c (1 - p_{l,i}) \right) (D_s(n, m, o) + D_{ep_ref}(n, m, o)) \\
& + \left(1 - \prod_{i=0}^c (1 - p_{l,i}) \right) D_{ec}(n, m) \\
& + \left(\prod_{i=0}^c (1 - p_{l,i}) \right) \frac{\beta \cdot 2^{\Delta QP_{c-1,c}/6}}{\beta \cdot 2^{\Delta QP_{c-1,c}/6} + 1} \lambda_{ef}^*(Q_c) R
\end{aligned} \quad (21)$$

III. SIMULATION RESULTS AND RELATED DISCUSSIONS

To demonstrate the performance of the proposed method, 200 frames (YUV, 4:2:0, progressive, intra period -1) of City, Crew, Harbour and Soccer sequences as specified in [11] are used. After generating bit stream, packet-loss simulator is used to simulate packet loss during transmission. Error pattern is used in simulator as proposed in [12]. Relative gain and coding efficiency are measured by calculating average difference between two RD curves as assigned in Bjontegaard metric [13]. Other simulation conditions are as follows:

- JSVM version 9.8 [14]
- LA-RDO and closed loop are enabled.
- Low delay application (IPPP coding), and high delay application (hierarchical B picture with group of pictures 16) are tested separately.
- Two layers sliding window, base layer (BL) and enhancement layer (EL) are used. For quality scalability, resolution ratio $\beta=1$ (QCIF:QCIF) is used, while for spatial scalability resolution ratio $\beta=4$ (QCIF:CIF) is used for BL and EL respectively, as error concealment at the decoder supports maximum two layers only.
- BLSkip method is performed at the decoder for error concealment [15].

- Bit-stream is decoded after passing through simulator with 10% packet loss rate (PLR) for BL and 20% PLR for EL. The same PLRs are used for target LA-RDO encoding.

The comparison of “JSVM” software is made with the proposed “ML- λ ” method in error prone environment.

A. Results for Low Delay Application

In this sub-section, the performance of the proposed algorithm is evaluated for inter-prediction (IPPP coding). Both quality scalability and spatial scalability are evaluated separately.

1) **Quality Scalability:** The performance of the proposed algorithm is evaluated for quality scalability with IPPP coding, where quantization parameter $QP_0 = (24, 28, 32, 36)$ and $QP_1 = (20, 24, 28, 32)$ are used for BL and EL respectively. On average, PSNR gain of 2.10 dB (corresponding to 21.78% bit-rate saving) is achieved using the proposed algorithm.

2) **Spatial Scalability:** The performance of the proposed algorithm is evaluated for spatial scalability with IPPP coding, where $QP_0 = (30, 34, 36, 38)$ and $QP_1 = (32, 36, 38, 40)$ are used for BL and EL respectively. On average, PSNR gain of 0.77 dB (corresponding to 16.56% bit-rate saving) is achieved using the proposed algorithm.

Fig. 1 shows rate-distortion (RD) curves of city sequence (top) for quality scalability and harbour sequence (bottom) for spatial scalability with low delay application.

B. Results for High Delay Application

In this sub-section, the performance of the proposed algorithm is evaluated for hierarchical B (H-B) picture coding with group of picture (GOP) size of 16. Both quality scalability and spatial scalability are evaluated separately.

1) **Quality Scalability:** The performance of the proposed algorithm is evaluated for quality scalability with H-B picture coding, where $QP_0 = (28, 32, 36, 40)$ and $QP_1 = (22, 26, 30, 34)$ are used for BL and EL respectively. On average, PSNR gain of 0.48 dB (corresponding to 6.13% bit-rate saving) is achieved using the proposed algorithm.

2) **Spatial Scalability:** Finally, the performance of the proposed algorithm is evaluated for spatial scalability with H-B picture coding, where $QP_0 = (26, 30, 34, 36)$ and $QP_1 = (30, 34, 38, 40)$ are used for BL and EL respectively. On average, PSNR gain of 0.31 dB (corresponding to 5.25% bit-rate saving) is achieved using the proposed algorithm.

Fig. 2 shows RD curves of crew sequence (top) for quality scalability and soccer sequence (bottom) for spatial scalability with high delay application.

Table 1 and Table 2 present PSNR-Y gains (in dB) and bit-rates saving (in %) over JSVM software of all sequences for both quality and spatial scalabilities in IPPP as well as in hierarchical B picture environments.

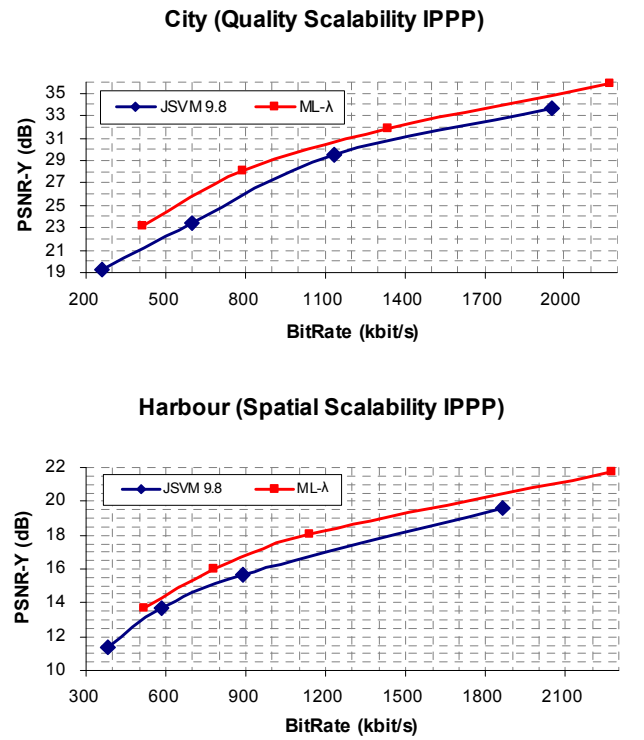


Figure 1. RD curves for quality scalability (top) and spatial scalability (bottom) in low delay application

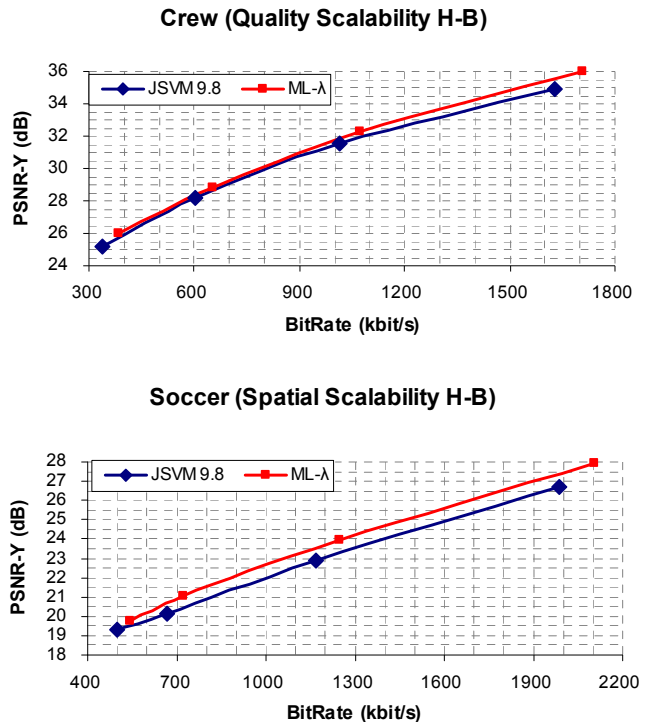


Figure 2. RD curves for quality scalability (top) and spatial scalability (bottom) in high delay application

Table 1. PSNR-Y Gains over JSVM 9.8

Sequences	IPPP Coding		Hierarchical-B Picture	
	Quality Scalability	Spatial Scalability	Quality Scalability	Spatial Scalability
City	1.8092	0.7574	0.3184	0.4476
Crew	1.7485	0.6428	0.3184	0.1282
Harbour	1.9725	1.0338	0.0155	0.0684
Soccer	2.8868	0.6514	1.2814	0.6137
Average	2.1043	0.7714	0.4834	0.3145

Table 2. Bit-Rate Savings (%) over JSVM 9.8

Sequences	IPPP Coding		Hierarchical-B Picture	
	Quality Scalability	Spatial Scalability	Quality Scalability	Spatial Scalability
City	20.16	14.90	4.58	7.85
Crew	21.08	16.52	4.64	2.26
Harbour	22.82	18.21	0.11	1.49
Soccer	23.09	16.28	15.21	9.42
Average	21.78	16.56	6.13	5.25

Comparing with low delay scenario (IPPP coding structure), the proposed algorithm shows marginal gain for high delay scenario (H-B picture). It is mainly because inter-coded MBs from the base layer are not fully decoded and reconstructed, so correlation information are not fully exploited. Instead, only the residual prediction of the base layer inter-coded MBs is decoded and may be used to predict enhancement layer residual. Since, the temporal prediction from the same layer is relatively efficient for B frames in H-B coding as compared to inter-layer prediction, so inter-layer prediction is less frequently used. In such a case, the proposed algorithm which focuses the correlation between layers can only provide a comparable performance to the JSVM.

In the second observation, the proposed algorithm gains less with spatial scalability as compared with quality scalability. It is mainly because distortion and loss of MB from lower layer having low resolutions will propagate to the collocated MBs of higher layers with corresponding higher resolutions; as a result distortion effects at higher layers will be multiplied by the resolution ratio (β). Consequently picture quality is relatively degraded, as error concealment will also be performed to the collocated MBs of higher layers; even though collocated MBs of higher layers have been received and can be decoded correctly.

IV. CONCLUSIONS

In this paper, multi-layer Lagrange multiplier selection method is proposed for mode decision of LA-RDO based error resilient coding in SVC. Exploiting strong correlations between two consecutive layers, considerable gain can be achieved by joint optimization of base layer and enhancement layer for mode decision in an erroneous channel condition. Comparing with JSVM software with LA-

RDO enabled encoding, the proposed algorithm achieves better PSNR-Y gain and bit-rate saving. Simulation results show that not only average PSNR-Y gains of 2.10 dB and 0.77 dB are obtained for quality and spatial scalabilities respectively in a low delay scenario, but also 0.48 dB and 0.31 dB are obtained for quality and spatial scalabilities respectively in a high delay scenario. Future work is required to limit the impact of error propagation distortion with consequent error concealment within layer in erroneous ML- λ selection. Moreover, efficient use of inter-layer in H-B picture, and joint optimization for more than two layers are interesting topics.

REFERENCES

- [1] H. Schwarz and T. Wiegand, "R-D optimized multi-layer encoder control for SVC," in *IEEE Int. Conf. on Image Process. (ICIP)*, San Antonio, USA, Sep. 2007, pp. II-281-284.
- [2] X. Li, P. Amon, A. Hutter, and A. Kaup, "Lagrange multiplier selection for rate-distortion optimization in SVC," *ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6 JVT-AD021.doc*, Geneva, CH, Jan. 2009.
- [3] Y. Zhang, W. Gao, H. Sun, Q. Huang, and Y. Lu, "Error resilience video coding in H.264 encoder with potential distortion tracking," *IEEE Int. Conf. on Image Process. (ICIP)*, vol. 1, Singapore, Oct. 2004, pp. 163-166.
- [4] Y. Guo, Y.-K. Wang, and H. Li, "Error resilient mode decision in scalable video coding," *ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6 JVT-R057*, Bangkok, Thailand, Jan. 2006.
- [5] T. Wiegand and B. Girod, "Parameter selection in Lagrangian hybrid video coder control," [Online]. Available: http://ip.hhi.de/imagecom_G1/assets/pdfs/clp.pdf
- [6] T. Wiegand and B. Girod, "Lagrange multiplier selection in hybrid video coder control," in *IEEE Int. Conf. on Image Process. (ICIP)*, Thessaloniki, Greece, 2001, pp. 542-545, vol.3.
- [7] K. Takagi, "Lagrange multiplier and RD-characteristics (JVT-C084)," in *JVT Meeting (Joint Video Team of ISO/IEC MPEG & ITU-T VCEG)*, Virginia, USA, May 2002.
- [8] H. Roodaki, H.R. Rabiee, and M. Ghanbari, "Performance enhancement of H.264 codec by layered coding," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Process (ICASSP)*, Las Vegas, USA, Apr. 2008, pp. 1145-1148.
- [9] H. Schwarz and T. Wiegand, "Further results for an RD-optimized multi-loop SVC encoder (JVT-W071)," in *JVT Meeting (Joint Video Team of ISO/IEC MPEG & ITU-T VCEG)*, San Jose, USA, Apr. 21-27 2007.
- [10] H.264, "Series: H audiovisual and multimedia systems, Infrastructure of audiovisual services-coding of moving videos," *ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG-4 Part 10)*, 2007 Corrigendum 1(01/2009).
- [11] M. Wien and H. Schwarz, "Testing conditions for SVC coding efficiency and JSVM performance evaluation (JVT-Q205)," in *JVT Meeting (Joint Video Team of ISO/IEC MPEG & ITU-T VCEG)*, Poznan, Poland, Jul. 23-29 2005.
- [12] S. Wenger, "Error patterns for Internet experiments," TU Berlin, *Doc. VCEG-Q15-I-16r1*, New Jersey, Oct. 1999.
- [13] S. Pateux and J. Jung, "An Excel add-in for computing Bjontegaard metric and its evolution additional performance analysis," Orange-France Telecom Research and Development, *Doc. VCEG-AE07*, Marrakech, Morocco, Jan. 2007.
- [14] JVT, "H.264/SVC reference software (JSVM 9.8) and manual," CVS sever at garcon.ient.rwth-aachen.de, Oct. 2007.
- [15] Y. Chen, K. Xie, F. Zhang, P. Pandit, and J. Boyce, "Frame loss error concealment for SVC," *Jour. of Zhejiang Univer.Science A.*, pp. 677-683, Hangzhou, China, Apr. 2006.