2012

Cross-layer Optimized Coding Mode Selection for Wireless Video Communications

Yun Ye
University of Nebraska-Lincoln, yye@huskers.unl.edu

Song Ci
University of Nebraska-Lincoln, sci2@unl.edu

Dalei Wu
University of Nebraska-Lincoln, dwu@huskers.unl.edu

2012 IEEE International Conference on Multimedia and Expo Workshops

Cross-layer Optimized Coding Mode Selection for Wireless Video Communications

Yun Ye, Song Ci
Department of Computer and Electronics Engineering,
University of Nebraska-Lincoln
Omaha, USA
yye@unlnotes.unl.edu, sci@engr.unl.edu

Dalei Wu
Mechatronics Research Lab
Massachusetts Institute of Technology
Cambridge, USA
daleiwu@mit.edu

Abstract— **This paper proposed a coding mode selection method for video transmission over wireless networks. Unlike previous mode selection methods disregarding channel distortion or assuming constant packet loss rate (PLR), this method includes a cross-layer controller to collect both source and channel information. The mode selection process is formulated as a delay constrained distortion minimization problem. The three components in the resulting Lagrange cost function, namely distortion, Lagrange multiplier and packet delay, are estimated with online channel information feedback. Suboptimal coding decision and physical layer modulation and coding scheme (MCS) are determined by the controller for each packet. In our experiment, three coding modes, intra, inter and down sampling, are tested under various channel conditions. Compared to conventional method, 3.6dB to 7.5dB average reduction in distortion is achieved under different channel condition, while down sampling further gains up to 2.2dB distortion reduction in low data rate transmission.**

Keywords-wireless communications; video codec; cross-layer control

I. INTRODUCTION

Video communications over wireless networks face various challenges including power and bandwidth constraint, random time-varying channel fading effect, network heterogeneity, and quality of service requirement [1]. Information from different network layers is necessary to be incorporated in data compression/coding procedure in order to achieve better system performance. Therefore, alongside with the development of source coding strategies [2-6], increasing dependence on cross-layer optimization is observed in emerging mobile multimedia applications [7-11].

In video compression, coding mode and corresponding quantization parameter (QP) selection is an important process for rate-distortion (RD) control. While QP selection under predefined GOP structure has been extensively studied [2, 10, 11], combining mode selection could enhance the coding results for its inherent adaptability [3]. Traditional mode selection methods mainly consider RD results obtained in application layer, i.e. given rate or buffer constraint, quantization distortion is minimized by optimal coding decision. A plethora of research work has been done in this area. In [3], the mode selection process is formulated as a Lagrange cost minimization problem solved by dynamic programming. Other information such as local edge or block boundary difference is also utilized to accelerate the coding decision process [4, 5]. The authors in [6] further explored the efficiency of multiresolutional coding by adaptively selecting among intra, inter and down sampling modes for each macro block (MB), and reported that better performance is acquired under low data rate, since smaller QPs could be chosen for down sampled MBs.

These source coding methods merely look into quantization induced distortion. In packetized video transmission over wireless networks, packet loss is also a major cause of distortion at receiver's side. In [12], the 'recursive optimal per-pixel estimate' (ROPE) method is presented for MB coding mode decision. Expected end-to-end distortion is estimated with certain PLR. This statistical model demonstrates a new way to adjust coding decision according to both source coding and channel distortion, whereas the impact of dynamic channel condition on RD results is not considered. Another joint source channel coding (JSCC) method introduced in [13] adopts random intra refreshing. Source coding distortion is modeled as a function of the intra MB refreshing rate, while channel distortion is calculated in a similar recursive fashion as is done in [12]. In this method, channel coding rate and forward error correction (FEC) are considered, yet some parameters need to be determined beforehand exclusively for each video, and the time varying channel condition is still ignored.

To cope with channel fading, the object based video coding method described in [8] calculates practical channel capacity in a Rayleigh fading channel. Distortion for intra and inter mode coding is estimated separately with resulting PLR. Discriminative coding decision is determined for shape and texture data under delay and transmission power constraints. The work in [11] takes into account physical layer MCS, and adaptively estimates PLR in a Rayleigh fading channel based on convolution coding and BPSK modulation. More flexible MCS configuration is applied in [9, 10] through cross-layer design with channel information feedback. However, these methods provide no specification for online coding mode selection. Exhaustive searching is time consuming; dynamic programming and random intra refreshing prearrange coding decision on consecutive packets/frames, and might not be suitable with online channel information feedback. Thus we propose to incorporate cross-layer design in coding decision. The mode selection process is formulated as a delay constrained distortion minimization problem. The suboptimal decision is carried out for each packet by a cross-layer controller with

978-0-7695-4729-9/12 $26.00 © 2012 IEEE
DOI 10.1109/ICMEW.2012.24

adaptive MCS configuration. In our experiment, intra or inter mode is selected for MBs in each packet under proper MCS, and down sampling is considered an alternative for inter coding in low data rate transmission.

The rest of the paper is organized as follows. Section II describes the delay constrained coding mode selection problem. Section III explains the system model of cross-layer design. Details of the implementation are included in Section IV. Section V presents experimental results. Conclusions are drawn in Section VI.

II. PROBLEM DESCRIPTION

Resource constrained video coding problem is usually formulated by a Lagrange cost function. In wireless video communications, the components in the cost function can be correlated to configuration in other layers in the network architecture. Therefore a cross-layer design is an ideal option to enhance overall system performance.

A. Background

In video coding, given resource constraint, coding decision is usually formulated as a distortion minimization problem. For example, in a block based encoder such as H.264/MPEG-4 AVC, coding mode (*mode*) and QP (*Q*) are selected for each MB under data rate *R* and frame delay limit *T*, according to resulted distortion *D* and data length *L*.

$$\underset{(Q,mode)}{arg\ min}\ \sum_{MBs} D(Q,mode) \quad s.t. \quad \sum_{MBs} \frac{L(Q,mode)}{R} \leq T \tag{1}$$

Classic solution to above formulation is the Lagrange cost function [2, 3]. The cost minimization process is implemented for each MB with an optimal Lagrange multiplier λ^*. λ^* can be obtained with bisection algorithm such that the condition in formula (1) is satisfied.

$$\underset{(Q,mode)}{arg\ min}\ D(Q,mode) + \lambda^* \cdot \frac{L(Q,mode)}{R} \tag{2}$$

In packetized video transmission over wireless networks, the end-to-end distortion consists of quantization distortion in source coding, and channel distortion caused by random packet loss. Several JSCC distortion estimation models have been proposed for optimal coding mode selection [8, 12-14]. However, these models either assume constant PLR and data rate, or adopt an aforehand strategy of making decision on several consecutive packets/frames. They fall short when finer adaptation is required [9, 10], where PLR is measured independently for each packet, and system configuration is adjustable accordingly.

Cross-layer control comes as a natural solution for this problem. Given channel state information (CSI), the controller is able to coordinate decision making in different layers in the network architecture. Since packet loss is related to various factors including source coding data length, path selection in network layer, physical layer MCS, as well as channel condition [8-11], optimal/suboptimal decision can be obtained through cross-layer cooperation.

When coding decision is provided by the cross-layer controller, the three terms in formula (2), namely distortion, Lagrange multiplier, and delay, need to be estimated independently for each packet based on CSI feedback. As to the distortion model, the ROPE algorithm [12] is desirable in making online mode selection for MBs contained in each packet. Specifically, for pixels in i-th MB in k-th packet, n-th frame, the end-to-end distortion between original pixel value f and the received reconstructed pixel value $\tilde{f}$ is estimated as follows.

$$\begin{aligned} &E(D) \\ &= \sum_{i \in MB} E\{[f(n,k,i) - \tilde{f}(n,k,i)]^2\} \\ &= \sum_{i \in MB} \left((f(n,k,i))^2 - 2f(n,k,i) \cdot E\{\tilde{f}(n,k,i)\} + E\{(\tilde{f}(n,k,i))^2\} \right) \end{aligned} \tag{3}$$

The cross-layer optimized solution based on this distortion model will be introduced in Section IV.

B. Problem Formulation

Inspired by previous work, we proposed to incorporate a cross-layer controller in solving the coding mode selection problem described in formula (1) and (2). The effect of physical layer MCS is focused in this work, since both PLR and data rate are affected by MCS [9, 10, 14]. Let (m, r) represents the modulation and FEC code pair as in HIPERLAN/2, IEEE 802.11a or IEEE 802.16 standards [15, 16], where m is the modulation order (bits per symbol constellation), and r is the FEC coding rate. Denoted by M the MCS affected decision, the coding mode selection problem can be formulated as follows.

$$\underset{(Q,mode,m,r)}{arg\ min} \sum_{k=1}^{K} \sum_{i=1}^{N} D_{n,k,i}^{M}(Q,mode) \quad s.t. \quad \sum_{k=1}^{K} \frac{\sum_{i=1}^{N} L_{n,k,i}(Q,mode)}{R_{n,k}^{M}} \leq T \tag{4}$$

where $L_{n,k,i}$ is the coded data length of i-th Mb in k-th packet, n-th frame, $R_{n,k}^{M}$ is the data rate limit for k-th packet in n-th frame, N is the total number of MBs in one packet, and K is the total packet in one frame. Corresponding coding decision for each packet can be expressed as following Lagrange cost minimization process. Coding decision and MCS configuration are determined together by the controller.

$$\underset{(Q,mode,m,r)}{arg\ min}\ \sum_{i=1}^{N} D_{n,k,i}^{M}(Q,mode) + \lambda_{n,k}^{M} \cdot \frac{\sum_{i=1}^{N} L_{n,k,i}(Q,mode)}{R_{n,k}^{M}} \tag{5}$$

Above formula demonstrates that all three components in the Lagrange cost function are affected by physical layer MCS. To maintain reasonable quality, the encoder may tend to select intra coding with small QP, and the MCS with smaller size constellation and higher FEC channel coding rate is preferred, while this will lead to unacceptable packet delay. On the other hand, higher modulation order and low rate FEC code speeds up transmission at the price of greater vulnerability to packet loss. Thereby coordination by the cross-layer controller can achieve system level optimization.

III. System Model

The system model of cross-layer optimized coding mode selection procedure is illustrated in Fig. 1. At the beginning of one frame, the controller collects frame information from application layer. For each packet in that frame, channel information feedback is gathered from physical and data link layer to update Lagrange multiplier and PLR. Then packet distortion and delay are calculated with refreshed PLR. Afterwards, the updated Lagrange multiplier, estimated packet delay and distortion are provided for decision making based on formula (5).

The focus of this model is online coding mode selection through adaptive MCS configuration by the controller. Dedications on other aspects in cross-layer design such as path selection and truncated ARQ can be referred to [9, 10]. Here we list some assumptions adopted in this work.

- The channel condition remains time invariant for one packet, but varies from packet to packet. A narrow-band block-fading channel with additive white Gaussian noise is simulated in the transmission process, where Rayleigh fading results in exponential distributed SNR:

$$\mathcal{F}(\gamma) = \frac{1}{\bar{\gamma}} e^{-\frac{\gamma}{\bar{\gamma}}} \tag{6}$$

  Here $\bar{\gamma}$ is the average received SNR, and is known by the transmitter.
- Perfect CSI is available to the receiver and is fed back to the transmitter without error and latency. This assumption could be approximately satisfied by using a fast feedback channel with powerful error control information as adopted in IEEE 802.16 [16].
- Receiver provides packet loss information to the transmitter. If the transmitter receives no loss feedback after $\triangle$ frames, a random packet loss is assumed for that packet based on its estimated PLR.
- A one-hop scenario is assumed in the transmission process, and other communication overhead is ignored.

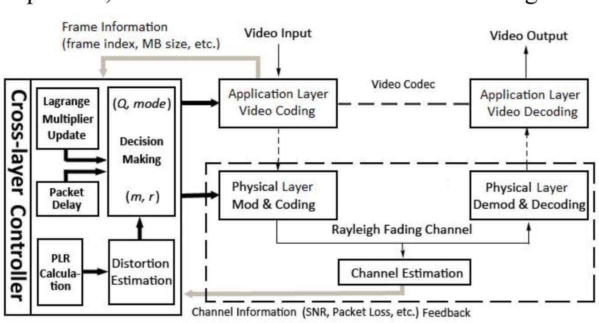


Figure 1. Cross-layer controller for coding mode selection.

IV. Implementation

This section provides detailed implementation to the coding mode selection problem previously stated. The three components of the Lagrange cost function (5) are estimated by the cross-layer controller separately, as depicted in Fig. 1.

A. Distortion

Distortion is estimated according to equation (3) for each packet with updated PLR.

1) PLR calculation

The PLR estimation method introduced in [7] models packet loss in video transmission over 802.11a WLAN networks, with a fixed packet length 8k bits.

$$p_{8kb}^{M} = \frac{1}{1+e^{\beta^{M}(\chi-\delta^{M})}} \tag{7}$$

Here χ is the signal to interference-plus-noise ratio (SINR) in dB. It is considered SNR when user interference noise is ignored. β^{M} and δ^{M} are constant parameters associated to each MCS denoted by *M*. Convolution coding is used as the FEC code. For variable packet size *L*, we adopt the approximate PLR calculation [10] as shown in equation (8). Four of the tested MCSs are listed in Table 1.

$$p^{M} = 1-(1-p_{8kb}^{M})^{L/8000} \tag{8}$$

TABLE I. Tested MCSs

Modulation, code rate	**β^{M} (dB-1)**	**δ^{M} (dB)**
MCS1 (64-QAM, 2/3)	0.625	18.2
MCS2 (16-QAM, 3/4)	0.352	15.1
MCS3 (QPSK, 1/2)	0.461	5.3
MCS4 (BPSK, 1/2)	0.640	2.3

2) Distortion estimation

In our experiment, in the presence of packet loss, error concealment at the receiver's decoder is performed in a way that lost MBs are replaced by co-located decoder reconstructed data in previous frame. Accordingly distortion is estimated for three coding modes with calculated PLR by equations (7) and (8). For pixels in an intra coded MB,

$$E\{\tilde{f}(n,k,i)\} = (1-p^{M})\cdot\hat{f}(n,k,i) + p^{M}\cdot E\{\tilde{f}(n-1,k,i)\} \tag{9}$$

$$E\{(\tilde{f}(n,k,i))^2\} = (1-p^{M})\cdot(\hat{f}(n,k,i))^2 + p^{M}\cdot E\{(\tilde{f}(n-1,k,i))^2\} \tag{10}$$

where $\hat{f}$ is the encoder reconstructed pixel value after quantization, and p^{M} is from equation (8). For pixels in an inter coded MB, with $\hat{e}$ the quantized residue data,

$$E\{\tilde{f}(n,k,i)\} = (1-p^{M})\cdot\left(\hat{e}(n,k,i) + E\{\tilde{f}(n-1,h,j)\}\right) + p^{M}\cdot E\{\tilde{f}(n-1,k,i)\} \tag{11}$$

$$E\{(\tilde{f}(n,k,i))^2\} = (1-p^{M})\cdot((\hat{e}(n,k,i))^2 + 2\hat{e}(n,k,i)\cdot E\{\tilde{f}(n-1,h,j)\} + E\{(\tilde{f}(n-1,h,j))^2\}) + p^{M}\cdot E\{(\tilde{f}(n-1,k,i))^2\} \tag{12}$$

$$\hat{e}(n,k,i) = \hat{f}(n,k,i) - \hat{f}(n-1,h,j) \quad (13)$$

where $\tilde{f}(n-1,h,j)$ is the expected pixel value in previous frame from which current pixel is predicted. When down sampling mode is selected, down sampled MBs are coded in intra mode,

$$E\{\tilde{f}(n,i)\} = (1-p^M)\cdot\check{f}(n,i) + p^M\cdot E\{\tilde{f}(n-1,i)\} \quad (14)$$

$$E\{(\tilde{f}(n,i))^2\} = (1-p^M)\cdot(\check{f}(n,i))^2 + p^M\cdot E\{(\tilde{f}(n-1,i))^2\} \quad (15)$$

Here $\check{f}$ denotes the up sampled value at the encoder after down sampling and quantization operations. The PDE based total variation up sampling method described in [17] is used to reconstruct coded data. If the transmitter has sufficient capability, more complex error concealment or quality enhancement methods, such as deblocking filtering and video super resolution [18] can also be applied with modified distortion estimation models.

Receiver loss feedback is available for the encoder to update the distortion estimation process. If the controller receives no loss feedback for a frame after transmission of △ frames, a random packet loss decision will be made in order to reduce the computation complexity.

B. Delay Constraint

The delay constraint T in formula (4) confines source coding data rate according to the channel capacity C, which is calculated with the channel bandwidth W and the dynamic channel SNR γ,

$$C = W\cdot\log_2(1+\gamma) \quad (16)$$

Under a specific MCS, the actual data rate limit is calculated with the modulation order m and the FEC code rate r.

$$R^M = C\cdot m\cdot r \quad (17)$$

Accordingly the controller calculates the packet delay with the coded packet data length and this data rate limit.

C. Lagrange Multiplier

The Lagrange multiplier λ is important to maintain acceptable quality and frame delay. Calculating an optimal λ for current frame is infeasible since channel feedback is unavailable for following packets. Instead we use a variant of the updating method adopted in [3, 12], based on previous estimated data rate by equations (16) and (17):

$$\lambda^M_{n,k+1} = \begin{cases} \lambda^M_{n,k}\cdot\left(1+\alpha\cdot\left(\sum_{j=1}^{k}\frac{K\cdot\sum_{i=1}^{N}L_{n,j,i}}{T\cdot R^M_{n,j}} - k\right)\right), \\ \qquad \text{if } \sum_{j=1}^{k}\frac{K\cdot\sum_{i=1}^{N}L_{n,j,i}}{T\cdot R^M_{n,j}} > k-\frac{1}{\alpha} \\ \beta\cdot\lambda^M_{n,k}\ , \qquad \text{else} \end{cases}$$
$$(n = 2,3,\ldots \quad k = 1,2,\ldots,K-1) \quad (18)$$

where α is a constant factor controlling update speed, and β is a positive number less than 1.0. MBs in the first packet of one frame are coded in intra mode according to slice delay constraint (T/K), and corresponding λ is set to a constant as an initial value. Suboptimal λs for following packets will be updated by formula (18).

D. Coding Mode Selection

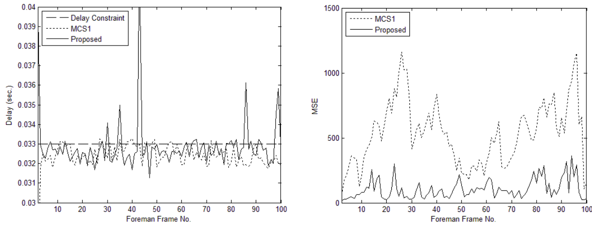
For each packet, the three components in formula (5) are estimated by the cross-layer controller with online channel information, as previously discussed. The MCS, coding mode and QP are determined by the controller such that the Lagrange cost is minimized. Several QPs are available for three different coding modes. The first frame is coded at full intra mode and is assumed to be correctly received with extra protection. MBs in following frames are coded in intra or inter/down sampling mode. Let S denote the MB size, N the total number of MBs in one slice/packet, △ the average recursion step for one slice, and U, V the number of optional values of QP and MCS mode. The computational complexity of this search is $O(S\cdot N\cdot\triangle\cdot U\cdot V)$ for one packet. To reduce the overhead, under each MCS, MBs are first coded with two available modes, each with the maximum and the minimum QPs. The MCS and coding mode corresponding to the lowest cost are chosen. Then QPs are searched under this configuration.

V. EXPERIMENTAL RESULTS

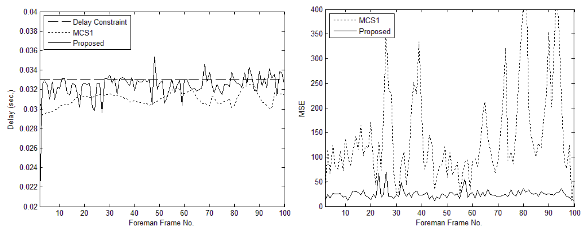
In this section, experiments are designed based on the H.264/AVC codec (JM12.2) [19]. Results with four QCIF sequence, *Foreman*, *Container*, *Carphone*, *Mother-daughter* [20], are selected for demonstration. Each video is encoded with MB size 16x16 in an IPPP GOP structure. The first frame is full intra coded and is assumed to be correctly received, while MBs are inter or intra coded (intra+inter) in following frames. When down sampling (DS) method is applied (intra+DS), a down sample rate of 2:1 is used, resulting in the block size 8x8. Available QPs for each coding mode range from 3 to 35. Encoded data of one slice are contained in one packet. Packet loss feedback maximum delay △ is 2 frames. The video frame transmission delay constraint T is fixed at 1/30 second. In channel simulation, signal bandwidth W is set to 100kHz ~24MHz, with average SNR $\bar{\gamma}$ varying from 10dB to 25dB.

A. Cross-layer optimized coding decision

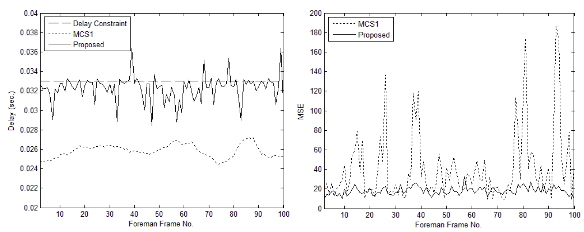
Fig. 2 shows the performance comparison on *Foreman* (intra+inter coded) between proposed method and the traditional mode selection method based on ROPE algorithm under fixed MCS. The bandwidth W = 100kHz, and SNR is set to 15dB, 20dB and 25dB. Note that higher bandwidth and SNR tend to generate greater channel capacity and lower PLR, sufficient for most MBs being coded in intra mode with small QP.



(a) $\bar{\gamma}$ = 15dB



(b) $\bar{\gamma}$ = 20dB



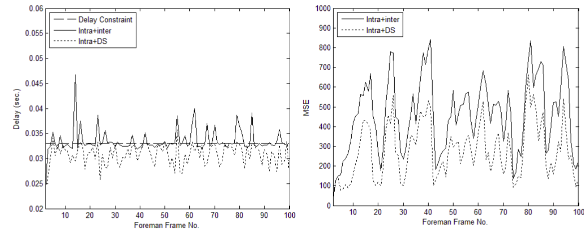
(c) $\bar{\gamma}$ = 25dB

Figure 2. Frame delay (left column) and distortion (right column) with W = 100kHz.

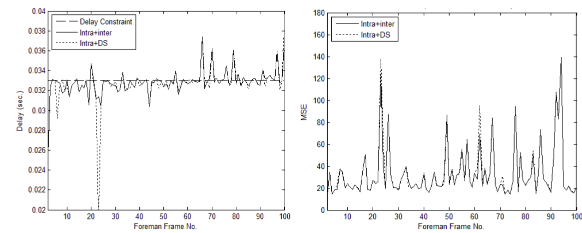
Under preset MCS, traditional mode selection is able to meet the delay requirement with Lagrange multiplier update. However, its distortion performance over dynamic lossy channel is inferior to the proposed method, especially when the channel condition is poor. With low signal bandwidth (W = 100kHz), an average 7.0dB, 7.5dB, 3.6dB distortion reduction in terms of MSE is observed by our method for corresponding $\bar{\gamma}$ = 15dB, 20dB, and 25dB, compared to the best result obtained with a fixed MCS (MCS1). Table 2 provides some details on average intra coding rate (the percentage of intra coded MBs), MSE and frame delay of *Foreman* coded with different methods.

TABLE II. PERFORMANCE COMPARISON (W = 100KHZ, $\bar{\gamma}$ = 20DB)

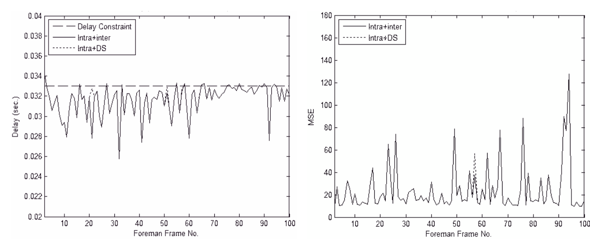
Coding scheme	Avg. intra coding rate	Avg. MSE	Avg. delay (sec.)
MCS1	1.0000	140.1749	0.0311
MCS2	1.0000	158.7511	0.0325
MCS3	0.6022	198.8942	0.0328
MCS4	0.2844	279.7757	0.0504
Proposed	0.8222	24.8175	0.0324



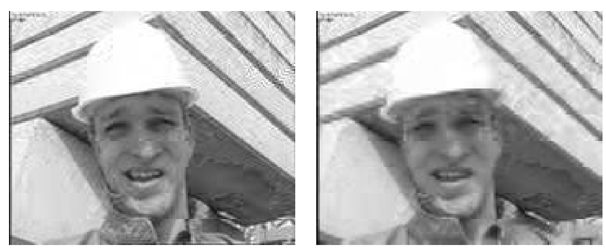
(a) W = 100kHz



(b) W = 500kHz



(c) W = 1MHz



(d) 44th frame coded at W=100kHz, $\bar{\gamma}$=10dB

Figure 3. (a) -(c) Frame delay (left column) and distortion (right column) with $\bar{\gamma}$ =10dB; (d) receiver reconstructed frame coded with intra+inter (left) and intra+DS (right).

B. *Down Sampling vs. Inter Coding*

Down sampling is tested to explore more coding options in low data rate transmission. Fig. 3 contains comparison results for intra+inter and intra+DS coding on *Foreman*, using $\bar{\gamma}$ = 10dB and different channel bandwidth.

Noticeable improvement is perceived in deteriorated channel condition with increased down sampling mode selection, gaining an average distortion reduction up to 2.2dB, while desirable delay performance is maintained, as illustrated in Fig.3 (a). Under better channel condition or on more stationary video (*Mother-daughter*), down sampling has similar or worse performance. This shows the potentials of down sampling in cross-layer optimized video coding for low rate transmission. Visual quality of the received data can be inspected in Fig. 3 (d). Table 3 provides more information on other tested video.

TABLE III. PERFORMANCE COMPARISON (W=100KHZ, $\bar{\Gamma}$=10DB)

Video	Coding mode	Avg. intra coding rate	Avg. MSE	Avg. delay (sec.)
Foreman	Intra+inter	0.3533	454.0071	0.0334
	Intra+DS	0.1667	274.9554	0.0308
Container	Intra+inter	0.3900	67.2637	0.0340
	Intra+DS	0.3456	60.4896	0.0321
Carphone	Intra+inter	0.3822	107.4395	0.0356
	Intra+DS	0.3844	95.7810	0.0310
Mother-daughter	Intra+inter	0.4300	28.1324	0.0345
	Intra+DS	0.5867	32.2609	0.0331

Some delay violation can be observed in Fig. 2 (a)-(c) and Fig. 3 (a)-(b). This is caused by the decision to transmit the packet under extremely poor network condition. An alternative is to simply drop the packet.

Video transmission over wireless networks is affected by the channel fading effect. Coding decision needs to take into account packet loss induced channel distortion. The proposed cross-layer model is capable of choosing suboptimal coding mode for each packet with adaptive MCS configuration. The mode selection process is formulated as a delay constrained distortion minimization problem. Each component in the Lagrange cost function is estimated by the cross-layer controller with online channel information feedback. Significant performance improvement is achieved due to better adaptation to the dynamic channel condition. Our work demonstrates the efficiency of cross-layer control in wireless video communications.

VI. CONCLUSIONS

Video transmission over wireless networks is affected by the channel fading effect. Coding decision needs to take into account packet loss induced channel distortion. The proposed cross-layer model is capable of choosing suboptimal coding mode for each packet with adaptive MCS configuration. The mode selection process is formulated as a delay constrained distortion minimization problem. Each component in the Lagrange cost function is estimated by the cross-layer controller with online channel information feedback. Significant performance improvement is achieved due to better adaptation to the dynamic channel condition. Our work demonstrates the efficiency of cross-layer control in wireless video communications.

ACKNOWLEDGMENT

The research is partially supported by NSF under Grant No. CCF-0830493 and Grant No. CNS-1145596.

REFERENCES

[1] Z. Han, H. Wang, D. O. Wu, J. Huang, and M. Van Der Schaar, "Cross-Layer Optimized Wireless Multimedia Communications," *Advances in Multimedia*, vol.2007, pp.1-2, 2007.

[2] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal Trellis-based Buffered Compression and Fast Approximations," *IEEE Transactions on Image Processing*, vol.3, no.1, pp.26-40, Jan. 1994.

[3] T. Wiegand, M. Lightstone, D. Mukherjee, T.G. Campbell, and S.K. Mitra, "Rate-distortion Optimized Mode Selection for Very Low Bit Rate Video Coding and the Emerging H.263 Standard," *CSVT*, vol.6, no.2, pp.182-190, Apr. 1996.

[4] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu, and S. Wu, "Fast Mode Decision Algorithm for Intra-prediction in H.264/AVC Video Coding," *CSVT*, vol.15, no.7, pp. 813- 822, July 2005.

[5] I. Choi, J. Lee, and B. Jeon, "Fast Coding Mode Selection With Rate-Distortion Optimization for MPEG-4 Part-10 AVC/H.264," *CSVT*, vol.16, no.12, pp.1557-1561, Dec. 2006.

[6] E.M. S. Uslubas and A.K. Katsaggelos, "A Resolution Adaptive Video Compression System," *Intelligent Multimedia Communication: Techniques and Applications*, vol. 280/2010: Springer Verlag, pp.167-194, 2010.

[7] D. Krishnaswamy and M. van der Schaar, "Adaptive Modulated Scalable Video Transmission over Wireless Networks with a Game Theoretic Approach," *6th IEEE Workshop on Multimedia Signal Processing*, pp.107-110, 2004.

[8] H. Wang, F. Zhai, Y. Eisenberg, and A.K. Katsaggelos, "Cost-distortion Optimized Unequal Error Protection for Object-Based Video Communications," *CSVT*, vol.15, no. 12, pp. 1505 - 1516, Dec. 2005.

[9] Y. Andreopoulos, N. Mastronarde, and M. van der Schaar, "Cross-Layer Optimized Video Streaming over Wireless Multi-hop Mesh Networks," *JSAC*, vol.24, no.11, pp.2104-2115, Nov. 2006.

[10] D. Wu, S. Ci and H. Wang, "Cross-layer Optimization for Video Summary Transmission over Wireless Networks," *JSAC*, vol.25, no.4, pp.841-850, May 2007.

[11] Z. Chen and D. Wu, "Rate-Distortion Optimized Cross-layer Rate Control in Wireless Video Communication," *CSVT*, 2011, accepted.

[12] R. Zhang, S.L. Regunathan, K. Rose, "Video Coding with Optimal Inter/Intra-Mode Switching for Packet Loss Resilience," *JSAC*, vol.18, pp. 966-976, Jun. 2000.

[13] Z. He, J. Cai and C.W. Chen, "Joint Source Channel Rate-Distortion Analysis for Adaptive Mode Selection and Rate Control in Wireless Video Coding," *CSVT*, vol. 12, pp. 511 -523, Jun 2002.

[14] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, " Analysis of Video Transmission over Lossy Channels," *JSAC,* vol. 18, pp. 1012 -1032, Jun. 2000.

[15] A. Doufexi, S. Armour, M. Butler, A. Nix, D. Bull, J. McGeehan, and P. Karlsson, "A Comparison of the HIPERLAN/2 and IEEE 802.11a Wireless LAN Standards," *IEEE Comm. Magazine*, vol.40, pp.172-180, May 2002.

[16] Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems, *IEEE Standard 802.16*, 2002.

[17] H. A. Aly and E. Dubois, "Image Up-Sampling Using Total-Variation Regularization with a New Observation Model," *IEEE Transactions on Image Processing*, vol. 14, pp. 1647 -1659, Oct. 2005.

[18] M. Shen, P. Xue, and C. Wang, "Down-Sampling Based Video Coding Using Super-Resolution Technique," *CSVT*, vol.21, no.6, pp.755-765, June 2011.

[19] H.264/AVC Reference Software, available online: http://iphome.hhi.de/suehring/tml/download/

[20] YUV Video Sequences, available online: http://trace.eas.asu.edu/yuv/