

Cross Layer Adaptation with QoS Guarantees for Wireless Scalable Video Streaming

Hsuan-Li Lin, Tung-Yu Wu, and Ching-Yao Huang, *Member, IEEE*

Abstract—In this paper, a cross-layer adaptation scheme is proposed for quality of service (QoS) provision in the scalable video streaming of high definition (HD) content. The cross-layer parameters, which contain the video rate, payload length of a packet, the mode of modulation and coding scheme (MCS), can be dynamically adapted to minimize distortion of a video streaming under the given delay bound. Based on the channel quality and rate-distortion parameters, the proposed scheme formulates the problem of parameter selection into an optimization problem. Simulation results show that our approach guarantees video quality under QoS constraints.

Index Terms—Cross-layer design, payload length adaptation, wireless video transmission, geometry programming.

I. INTRODUCTION

WITH recent development of wireless technologies, wireless high-definition (HD) video services become possible for home entertainment systems. However, there exist challenges to deliver HD video over fluctuated wireless channels. To support HD video, WiMedia ultra wideband (UWB) [1] which supports high data rates is considered. In addition, the scalable video coding (SVC) extension of H.264/AVC is applied to the HD video streaming for the high compression efficiency and video rate adaptation based on channel conditions and terminal types [2].

In this paper, the quality of service (QoS) of video streaming is defined as the video quality under a delay bound of each group of pictures (GOP) reception. Due to the limit in the receiver video buffer, the delay bound is relatively tight for each HD GOP reception. To meet the delay bound, the throughput improvement is essential, [3] and [4] provided a joint consideration of the payload length adaptation and the selection of modulation and coding scheme (MCS) to maximize the throughput. However the volatility of transmission rates would violate its delay bound and lead to fluctuations of video quality, which is undesirable from users' point of view. The volatility of transmission rates can be reduced by restricting the payload length and transmission modes to satisfy the packet error rate (PER) constraint [5,6]. However, the effective throughput also would be reduced by inefficient transmission modes. Hence, the retransmission policy is considered to improve the QoS in [7]. The authors proposed a scheme to satisfy the frame error rate, and the required transmission time of a video frame is minimized by properly adjusting the payload length and MCS.

Manuscript received April 6, 2012. The associate editor coordinating the review of this letter and approving it for publication was J. van de Beek.

The authors are with the Department of Electronics Engineering & the Institute of Electronics, National Chiao-Tung University, Taiwan (e-mail: x2322.ee95g@nctu.edu.tw).

Digital Object Identifier 10.1109/LCOMM.2012.070512.120760

In this paper, we propose a cross-layer optimization scheme which employs the concept of retransmission. Instead of maximizing effective throughput, the proposed scheme guarantees the HD video quality and smooth play-out of scalable video streaming by minimizing the video distortion under a delay bound. While SVC techniques enable the video rate adaptation, the proposed scheme can suggest a suitable video rate for the SVC extractor. It also provides the information of proper payload length and MCS for Medium Access Control (MAC) layer to improve the video quality. This cross layer design can ensure the quality of video transmission.

II. SYSTEM MODEL

Consider a GOP of a video stream needed to deliver from a transmitter to a user within a delay bound. The quality of a compressed video stream can be measured by mean squared error (MSE), which represents the differences between reconstructed pixel values and the original pixel values. The parametric rate-distortion (RD) model in [8] is adopted in our proposed optimization framework. For a given video content, the MSE of a video streaming in a GOP period can be estimated as:

$$D = D_0 + \frac{\theta_0}{V - V_0} \quad (1)$$

where D is the video distortion represented by MSE, and the peak signal-to-noise-ratio (PSNR) is given by $20 \cdot \log_{10} \frac{255}{\sqrt{D}}$. V is the video rate in a GOP period, D_0 and V_0 are the distortion offset and video rate offset respectively. While the video rate V (bits/sec) increases, the MSE distortion decreases non-linearly. The curve fitting parameters D_0 , θ_0 and V_0 in (1) can be obtained from the video extractor by the least-square method based on a collection of K empirical pairs of video rate V_k and distortion D_k [9]. These parameters are all positive constants in a GOP period and they are updated periodically to track time-varying video contents.

Suppose the data size of a GOP is $G = nV/f$, where n is the number of frames in a GOP and f denotes the frames per second (fps). Before transmitting to the receiver, the information data of size G is fragmented into N_f packets of L bytes. The transmission time of a L -byte fragmented packet at MCS-mode m is given by

$$T^m = \frac{8L}{R^m} + T_O^m \quad (2)$$

where R^m is transmission rate corresponding to MCS-mode m in bits per second, and T_O^m is the transmission time of overhead. This overhead consists of transmission time of layer headers and acknowledgement, either positive (ACK) or negative (NACK) acknowledgement. Immediate ACK (Imm-ACK) scheme [1] is considered in this paper.

In wireless video communication, there are two reasons for GOP reception failures: one is the packet errors due to fluctuation of the wireless channel, and the other one is the packet loss due to the transmission time of a GOP that exceeds the delay bound. If there is no constraint on delay bound, all error packets can be retransmitted, and the GOP will be successfully received eventually. However, the latency should be minimized for real-time services, and the delay bound should be considered. Hence, the outage rate P_e is defined as the probability that the GOP is not completely delivered to the receiver within the delay bound. For HD streaming, the delay bound could be tightly bounded by the limited size of buffer at receiver. We assume the wireless link is a memoryless packet erasure channel [4], so errors appear independently among packets. Considering all possible transmission paths with retransmission policy, the outage rate is:

$$P_e = \sum_{i=0}^{N_f-1} \binom{N_r}{i} P^i (1-P)^{N_r-i} \quad (3)$$

where P is the packet success rate (PSR), and N_r is the number of packets including retransmitted which would be accommodated within the delay bound. The difference between N_r and N_f is the upper limit of retransmission. The packet success rate for convolution code can be formulated [10] as

$$P = (1 - P_u^m)^{8L} \quad (4)$$

where P_u^m is the union bound of the first-event error probability corresponding to MCS-mode m and it is calculated based on the signal-to-noise-ratio (SNR). For other channel codes, the proposed scheme, e.g. block codes or turbo codes, can be also applied when the proper PSR estimation is available. From (2) and (3), the transmission time of a GOP should follow a delay bound, which is the transmission period allocated by a radio resource scheduler:

$$N_r \cdot T^m \leq \text{delaybound} \quad (5)$$

Any packet of the GOP beyond the delay bound is discarded in the transmission.

III. PROPOSED OPTIMIZATION SCHEME

A. Video Distortion Minimization with Target Outage Rate and Delay Bound

In this section, we investigate the scenario that a radio resource scheduler allocates a transmission period to a single user. By considering this scenario, the parameter adaptation of payload length L , video rate V , and MCS-mode m can be systematically solved to minimize the video distortion D with the target outage rate and delay bound. The target outage rate P_{out} could be properly selected for various video applications. Base on (1)-(5), the jointly optimization problem with fixed MCS-mode m is formulated as

$$\begin{aligned} \min_{V,L} \quad & D_0 + \frac{\theta_0}{V - V_0} \\ \text{s.t.} \quad & \sum_{i=0}^{N_f-1} \binom{N_r}{i} P^i (1-P)^{N_r-i} \leq P_{out} \\ & N_r \cdot T^m \leq \text{delaybound} \\ & V \geq V_0 \\ & L \geq 0. \end{aligned} \quad (6)$$

where $N_f = \frac{n \cdot V}{8f \cdot L}$, and P_{out} is the target outage rate. The minimum distortion can be achieved by having dynamic parameter setting to meet channel conditions. In order to solve the optimization problem, we transform the outage rate constraint and packet transmission time T^m in (6) into explicit forms in terms of design variables V , L and m . We can express N_r as a function in terms of N_f , P , and P_{out} to prevent the complexity introduced by integer programming. The left side of outage rate constraint is a cumulative distribution function (CDF) of a binomial distribution. The constraint could be approximated by a CDF of a standard normal distribution [11]:

$$P_{out} \geq \sum_{i=0}^x \binom{n}{i} P^i (1-P)^{n-i}, \text{ for } x \leq n \approx \begin{cases} \Phi(\sqrt{4x+3}\{1-P\} - \sqrt{4n-4x-1}P) \\ , \text{ for } 0.05 \leq P \leq 0.93; \\ \Phi(\sqrt{2x+1}\{1-P\} - 2\sqrt{n-x}P) \\ , \text{ for } P > 0.93. \end{cases} \quad (7)$$

where $x = N_f - 1$, $n = N_r$, and Φ is the CDF of the standard normal distribution. This approximation is valid to express N_r when $N_f > 10$. However, in HD video services, the data size of a GOP is often fragmented into many packets ($N_f \gg 10$). Given delay bound and fixing L , m and N_r , the high video rate in (6) with more packets leads the decreasing video distortion and the increasing outage rate. In other words, it achieves the minimum video distortion as the outage rate approaching to the upper bound of outage rate P_{out} . Therefore, the equality in (7) holds, and we can let $s = \Phi^{-1}(P_{out})$, where Φ^{-1} denotes the normal inverse CDF of normal. After a small amount of manipulation, (7) can be rewritten as:

$$N_r \approx N_f - \alpha + \frac{\left(s - \sqrt{(\beta N_f - 1)(1-P)}\right)^2}{4P} \quad (8)$$

where $\alpha = 3/4$, and $\beta = 4$ for $0.05 \leq P \leq 0.93$; $\alpha = 1$, and $\beta = 2$ for $P > 0.93$ from (7). By combining (2), (4), (6) and (8), the original problem in (6) becomes

$$\begin{aligned} \min_{V,L} \quad & D_0 + \frac{\theta_0}{V - V_0} \\ \text{s.t.} \quad & \left[N_f - \alpha + \frac{\left(s - \sqrt{(\beta N_f - 1)(1 - (1-p)^{8L})}\right)^2}{4(1-p)^{8L}} \right] \\ & \cdot \left(\frac{8L}{R^m} + T_O^m \right) \leq \text{delaybound} \\ & V \geq V_0 \\ & L \geq 0. \end{aligned} \quad (9)$$

where p represents the error probability P_u^m in (4). Note that s is given by the target outage rate.

Further, we formulate the constraints and objective function in (9) to *posynomials*, which is a summation of products of non-negative variables with positive coefficients and real exponents. Because the error probability p is small enough ($p < 10^{-5}$) in our application, $(1-p)^{8L}$ is approximated by $1 - 8Lp$. The α is neglected as $N_f \gg \alpha$, and the $\beta N_f - 1$ is approximated as βN_f . Therefore, we rewrite the delay bound

constraint in (9) as

$$\left[N_f + \frac{(s - \sqrt{8\beta N_f L p})^2}{4(1 - 8Lp)} \right] \cdot \left(\frac{8L}{R^m} + T_O^m \right) \leq \text{delaybound} \quad (10)$$

Note that s is negative, because the operating range of P_{out} is below 0.5. Therefore, $(s - \sqrt{8\beta N_f L p})^2$ is a posynomial. By introducing a new variable k , we can replace (10) by two posynomial constrains as shown in Appendix I:

$$\left[N_f + \left(s - \sqrt{8\beta N_f L p} \right) k \right] \cdot \left(\frac{8L}{R^m} + T_O^m \right) \leq \text{delaybound} \quad (11)$$

$$\frac{1}{4k} + 8Lp \leq 1 \quad (12)$$

Similarly, the objective function $D_0 + \frac{\theta_0}{V - V_0}$ can be replaced by

$$D_0 + \theta_0 q \leq D \quad (13)$$

$$\frac{1}{q} + V_0 \leq V \quad (14)$$

Combining (11)-(14), the optimization problem (9) is transformed into

$$\begin{aligned} & \min_{V, L, k, q} D_0 + \theta_0 q \\ & \text{s.t.} \left[N_f + \left(s - \sqrt{8\beta N_f L p} \right) k \right] \cdot \left(\frac{8L}{R^m} + T_O^m \right) \leq \text{delaybound} \\ & \frac{1}{4k} + 8Lp \leq 1 \\ & \frac{1}{q} + V_0 \leq V \\ & L \geq 0. \end{aligned} \quad (15)$$

where $N_f = \frac{nV}{8fL}$. Since all the variables and parameters are positive, the object function and all constraints in (15) are posynomials. As a result, if the MCS-mode m is fixed, the problem can be recognized as a geometric programming (GP) problem which can be converted to a convex optimization problem by a logarithmic change of variables and a logarithmic transformation of the objective and constraint functions. It can be solved efficiently with a global solution [12].

In summary, first we can solve the GP problem in (15) to acquire the optimal video rate and payload length for each MCS-mode. Secondly, since the number of MCS-modes is finite for implementation, we then find the best MCS-modes with the optimal video rate and payload length to minimize video distortion. These optimal decisions are designed with the delay bound, SNR, and target outage rate to ensure QoS of HD video streaming.

B. Single-user Transmission Scenario with Target Outage Rate and Constrained Video Distortion

In this subsection, we investigate the minimized transmission duration, which is allocated by a radio resource scheduler, to satisfy the target outage rate and video quality. By considering this scenario, the adaptation of the payload length, video rate, and MCS-mode can also be systematically solved. The transmission duration can be transformed from the delay bound constraint in (15). If the target video distortion is

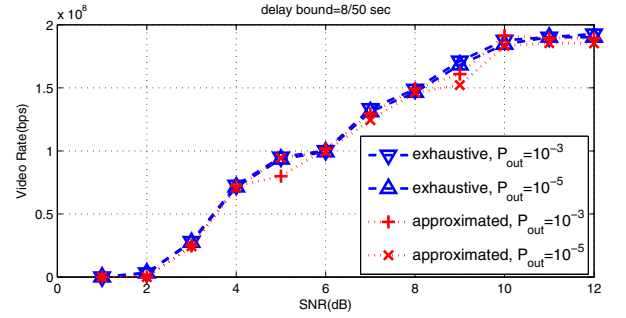


Fig. 1. Approximation of video rate adaptation with delay bound=8/50 seconds at $P_{out} = 10^{-3}$ and 10^{-5} respectively.

D_{target} , the optimization problem to minimize the transmission period with fixed MCS-mode m is given by:

$$\begin{aligned} & \min_L \left[\frac{nV_{target}}{8Lf} + \left(s - \sqrt{8\beta \frac{nV_{target}}{f} L p} \right) k \right] \cdot \left(\frac{8L}{R^m} + T_O^m \right) \\ & \text{s.t.} \frac{1}{4k} + 8Lp \leq 1 \\ & L \geq 0. \end{aligned} \quad (16)$$

where $V_{target} = \frac{\theta_0}{D_{target} - D_0} + V_0$. The objective function is a function of payload length, which trades the transmission time for outage rate. Because both the objective and constraints are posynomials, this adaptation is also a GP problem, and it can be solved globally with fixed MCS-mode m .

IV. RESULT AND DISCUSSION

In our simulation, the HD video sample (1280 by 720 pixels) is encoded using JSVM 9.19.9 reference code at a frame rate of 50 fps and GOP length of 8 frames [13]. We allocate the delay bound of a GOP as 8/50 seconds and 2/50 seconds respectively to investigate the video rate allocation and PSNR performance. The Maximum Likelihood (ML) estimate of SNR is used in this simulation as [14]. If the number of pilot samples N is large enough, which can be achieved by keeping tracking wireless channel indoor, the impact of channel SNR estimation error over additive white Gaussian noise (AWGN) channel could be negligible.

Fig. 1 shows the video rates of exact solution from (6) and the approximation from proposed scheme, respectively. It is clear the approximated video rate is close to the exact solution. We also observed the approximation errors increase slightly at high SNR values. It is because the approximation is accurate when the number of packets is large. However, the payload length trends to increase at high SNR values, so the number of packets in a GOP is fewer.

Fig. 2 shows the comparison of PSNR between the proposed scheme and one of common methodologies, the restricted PER method which set a PER constraint and decide the proper length of payload for maximizing throughput [5,6]. Consequently, for the same MCS mode, the PSNR of restricted PER method will saturate due to the constrained payload length. The proposed scheme makes trade-off between the volatility of throughput and video quality. In addition, while the restricted PER method hardly support the video streaming

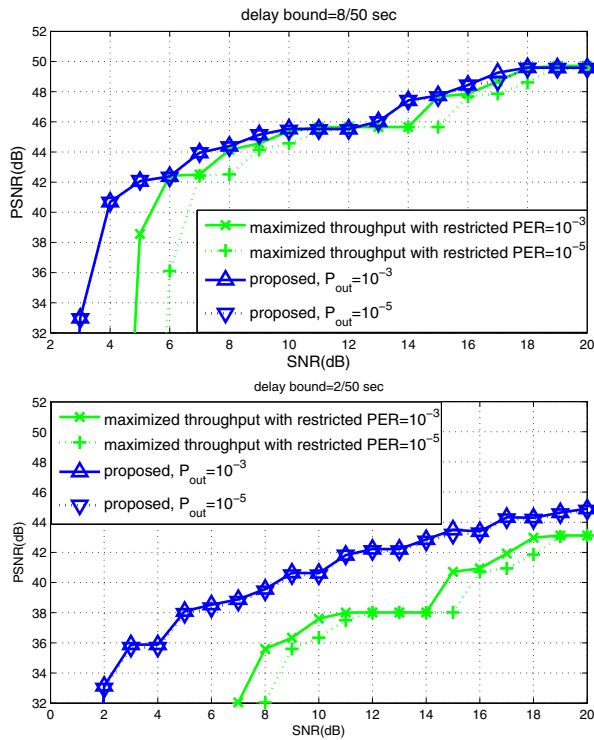


Fig. 2. PSNR under (a) delay bound=8/50 seconds and (b) delay bound=2/50 seconds at $P_{out} = 10^{-3}$ and 10^{-5} respectively.

at low SNR values (below 5 dB), the proposed scheme extends the operating range at low SNR values, which can enhance the radio coverage in the indoor environment.

V. CONCLUSION

We propose a cross-layer optimization scheme for HD scalable video streaming over wireless environment by jointly considering video content rate-distortion characteristics, channel conditions, transmission duration, and QoS constraints. We have shown that the proposed scheme can be approximated to a geometry programming problem, where the globally optimal solutions are guaranteed. The proposed scheme minimizes the video distortion by properly selecting the payload length, video rate, and MCS for applications with various target outage rate and delay bound. In addition, we analyze the minimized duration of transmission period under the target outage rate and target video distortion. Furthermore, the proposed scheme can enhance the radio coverage at low SNR values.

APPENDIX A

Claim 1 Let $r(x)$ be a monomial, $p(x)$, $q(x)$, $f(x)$ be generalized posynomials, and suppose $q(x) < r(x)$. The inequality

$$\frac{p(x)}{r(x)-q(x)} + f(x) \leq k \quad (17)$$

can be replaced with the following constraints.

$$p(x) \cdot t + f(x) \leq k \quad (18)$$

$$q(x) + \frac{1}{t} \leq r(x) \quad (19)$$

for $k, t \in \mathbb{R}^+$, where \mathbb{R}^+ denotes the set of nonnegative real numbers.

Proof : The feasible set of (17) is

$$A = \{x \mid \frac{p(x)}{r(x)-q(x)} \leq k - f(x), \forall x \in \mathbb{R}^+\}$$

And the feasible sets of (18) and (19) are

$$B = \{x \mid t \leq \frac{k - f(x)}{p(x)}, \forall x \in \mathbb{R}^+\}$$

$$C = \{x \mid t \geq \frac{1}{r(x)-q(x)}, \forall x \in \mathbb{R}^+\}$$

First, we show $A \subseteq B \cap C$. Let

$$D = \{x \mid t = \frac{1}{r(x)-q(x)}, \forall x \in \mathbb{R}^+\}$$

and $E = B \cap D$, hence

$$E = \{x \mid \frac{1}{r(x)-q(x)} \leq \frac{k - f(x)}{p(x)}, \forall x \in \mathbb{R}^+\} = A$$

Since $D \subseteq C$, $E \subseteq B \cap C$, hence $A \subseteq B \cap C$. Second, we show $A \supseteq B \cap C$. Let

$$F = B \cap C = \{x \mid \frac{k - f(x)}{p(x)} \leq t \leq \frac{1}{r(x)-q(x)}, \forall x \in \mathbb{R}^+\},$$

and if $x \in F$, then $\frac{k - f(x)}{p(x)} \leq \frac{1}{r(x)-q(x)}$ holds for all $x \in \mathbb{R}^+$, so $x \in A$. Hence $A \supseteq B \cap C$. As a result, $A = B \cap C$, and the proof is then completed. ■

REFERENCES

- [1] Standard ECMA-368: High Rate Ultra Wideband PHY and MAC Standard, ECMA International, Dec. 2005.
- [2] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.
- [3] T. Yoo and A. Goldsmith, "Throughput optimization using adaptive techniques," Wireless Systems Lab, Stanford University, CA, USA, Tech. Rep., 2004.
- [4] D. Qiao, S. Choi, and K. G. Shin, "Goodput analysis and link adaptation for IEEE 802.11a wireless LAN," *IEEE Trans. Mobile Comput.*, vol. 1, no. 4, pp. 580–589, June 2002.
- [5] S. Choudhury and J. D. Gibson, "Throughput optimization for wireless LANs in the presence of packet error rate constraints," *IEEE Commun. Lett.*, vol. 12, no. 1, Jan. 2008.
- [6] H.-H. Juan, H.-C. Huang, C. Huang, and T. Chiang, "Cross-layer mobile WiMAX MAC designs for the H.264/AVC scalable video coding," *ACM Wireless Networks*, vol. 16, no. 1, pp. 113–123, Jan. 2010.
- [7] T. Y. Wu, T. T. Chuang, and C. Y. Huang, "Optimal transmission of high definition video transmission in WiMedia systems," *ACM Wireless Networks*, vol. 17, no. 2, pp. 291–303, 2011.
- [8] K. Stuhlmüller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 1012–1032, June 2000.
- [9] X. Zhu, T. Schierl, T. Wiegand, and B. Girod, "Distributed media-aware rate allocation for video multicast over wireless networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 9, pp. 1181–1192, Sep. 2011.
- [10] M. B. Pursley and D. J. Taipale, "Error probabilities for spread spectrum packet radio with convolutional codes and Viterbi decoding," *IEEE Trans. Commun.*, vol. COM-35, pp. 1–12, Jan. 1987.
- [11] W. Molenaar, "Approximations to the poisson, binomial and hypergeometric distribution functions," Mathematical Centre Tracts, no. 31 Mathematisch Centrum, Amsterdam, 1970.
- [12] S. P. Boyd, S. J. Kim, L. Vandenbergh, and A. Hassibi, *A Tutorial on Geometric Programming*, Information Systems Laboratory, Dept. of Elect. Eng., Stanford Univ., 2004.
- [13] JVT, "H.264 SVC reference software (JSVM 9.16.9) and manual," CVS sever at garcon.ient.rwth-aachen.de, Jan. 2010.
- [14] M. Mohammad and R. M. Buehrer, "On the impact of SNR estimation error on adaptive modulation," *IEEE Commun. Lett.*, vol. 9, no. 6, pp. 490–492, June 2005.