BITRATE ALLOCATION FOR MULTIPLE VIDEO STREAMS AT COMPETITIVE EQUILIBRIA

Mayank Tiwari[1], *Theodore Groves*[2], *and Pamela Cosman*[1]

[1]Department of Electrical and Computer Engineering, [2] Department of Economics,
University of California, San Diego, La Jolla, CA 92093-0407
Technical Area: I-7 (Image and Video Coding)

ABSTRACT

Methods for multiplexing video streams often rely on identifying the relative complexity of the video sequences to improve the combined overall quality. In such methods, the quality of high motion videos is improved at the expense of reduction in the quality of low motion videos. In our approach, we use the Edgeworth box solution for competitive equilibrium to simultaneously improve the quality of all the video streams. The proposed method not only uses information about the differing complexity of the video streams at every time-slot but also the differing complexity of one stream over time. All the video sequences do at least as well as individual encoding.

1. INTRODUCTION

Applications where multiple compressed video streams are transmitted simultaneously through a shared channel include direct broadcast satellite, cable TV, video-on-demand service, and video surveillance. In existing methods for transmitting multiple video streams [1–3], improving the overall quality is the goal. However, not all video streams benefit from multiplexing processes. Generally, the quality of high complexity videos improves at the expense of reduced quality of low complexity videos.

In this paper, we constrain the video multiplexing process so that no video stream will suffer quality degradation. All videos will do at least as well as what they achieve by not participating in the multiplexing process. The method selects an expected efficient, Pareto optimal (PO), allocation of bitrate for multiple videos. By computing the expected competitive allocation in the Edgeworth box, a common tool in economics for equilibrium analysis, we find a point where all users perform better or at least as well as what they can achieve independently.

The rest of the paper is organized as follows: Section 2 describes the Edgeworth box for solving for competitive equilibria. Section 3 describes the application of the Edgeworth box solution in bitrate allocation for multiple video streams. Results and conclusions are given in Section 4.

2. EDGEWORTH BOX FOR COMPETITIVE EQUILIBRIUM

The Edgeworth box (EWB) [4] is a graphical tool for exhibiting PO allocation and illustrating a competitive (Walrasian) equilibrium in a pure exchange economy [5], in which no production is possible and the commodities that are ultimately consumed are those that individual users possess as initial endowments. The users trade these endowments among themselves in a market for mutual advantage.

Let there be two users ($i = 1, 2$) and two goods ($j = 1, 2$). User i's consumption vector is $x_i = (x_i^1, x_i^2)$, i.e., user i's consumption of good j is $x_i^j \geq 0$. Each user i is initially endowed with an amount $c_i^j \geq 0$ of good j. The total endowment of good j in the economy is denoted by $\bar{c}^j = c_1^j + c_2^j$, assumed strictly positive. An allocation $x \in \mathbf{R}_+^4$ is an assignment of a non-negative consumption vector to each user: $x = (x_1, x_2) = ((x_1^1, x_1^2), (x_2^1, x_2^2))$. We say that an allocation is nonwasteful and feasible if $x_1^j + x_2^j = \bar{c}^j$ (the total consumption of each good is equal to the economy's aggregate endowment of it).

In the EWB, user 1's quantities are measured with the southwest corner as the origin (O_1) as shown in Figure 1. User 2's quantities are measured using the northeast corner as the origin (O_2). For both users, the horizontal dimension measures quantities of good 1 and the vertical dimension measures quantities of good 2. The width and height of the box are $\bar{c}^1$ and $\bar{c}^2$, the economy's total endowment of goods 1 and 2. Any point in the box represents a division of the total endowment between user 1 and 2. Given $\bar{c}_i = (c_i^1, c_i^2)$, user i can calculate its utility $U_i(\bar{c}_i)$. A curve can be drawn for different x_i while keeping the utility at $U_i(\bar{c}_i)$. Such curves, known as *indifference curves*, can be drawn for different initial endowments for each user as shown in Figure 1. We assume these curves are convex. For any user, the utility increases when we move away from its origin. If we draw indifference

This work was supported by the National Science Foundation, the Center for Wireless Communications at UCSD, and the UC Discovery Grant program of the State of California.

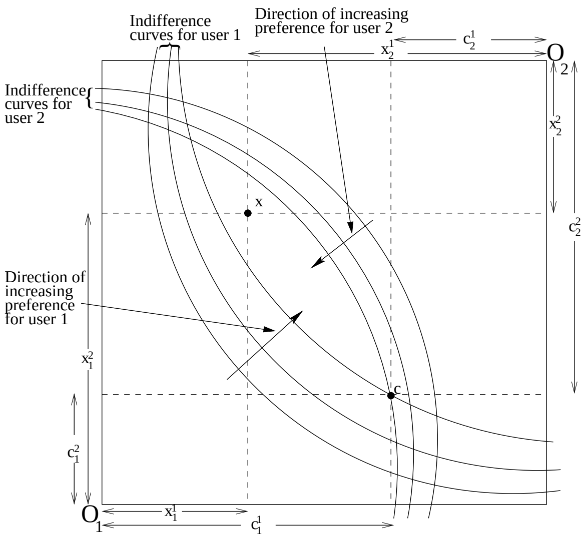


Fig. 1. An Edgeworth box

curves for both users in the box, the points where the indifference curves for both users are tangential to each other are PO allocations [5]. The set of all PO allocations is known as the *Pareto set*. The part of the Pareto set where both users do at least as well as at their initial endowments is called the *contract curve* (Figure 2). Any bargaining between the users should result in some point on the contract curve; these are the only points at which both users do at least as well as at their initial endowments and for which there is no alternative trade that can make both users better off [5].

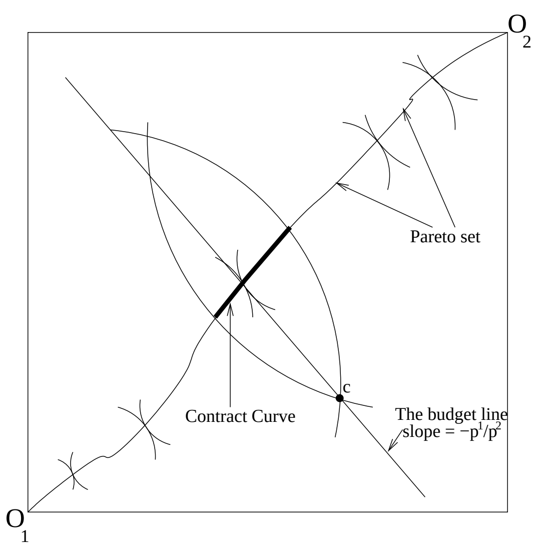


Fig. 2. The Pareto set and the contract curve in the EWB

Suppose users can buy or sell these goods in the market for prices p^1 and p^2. For any price system $p = (p^1, p^2)$ and initial endowments, the budget set for user i is:

$$B_i(p) = \{x_i \in \mathbf{R}_+^2 : p.x_i \le p.c_i\} \tag{1}$$

A competitive equilibrium for an EWB economy is a price vector p^* and an allocation $x^* = (x_1^*, x_2^*)$ in the EWB such that for i = 1,2,

$$U_i(x_i^*) \ge U_i(x_i') \qquad \forall\ x_i' \in B_i(p^*) \tag{2}$$

At an equilibrium, each user i's demanded bundle at price vector p^* is x_i^* and one user's net demand for a good is exactly matched by the other's net supply. The intersection of the budget line and contract curve, where the budget line is also tangential to the indifference curve for both the users on the contract curve, will result in the competitive equilibrium. At this equilibrium point, both users are better off compared to their initial endowment. This is shown in Figure 2. More details about the EWB and competitive equilibrium can be found in [5].

3. VIDEO MULTIPLEXING USING THE EW BOX

We extend the concept of the EWB from two users to N video users. The two goods are the bits available in two time-slots (TS). We consider one TS as one GOP; however, one can choose TS at any level. The complexity of the EWB increases with box dimension and there are many TS in each video stream. So, we reduce the problem to two TS for each user. We will use the terms GOP and TS interchangeably. We generate the rate distortion (RD) curve for each TS by calculating the mean squared error (MSE) at different bitrates. Note that the complexity of generating the RD curve can be further reduced by using the method described in [6]. Suppose that 1000 bits are available in TS 1 and in TS 2. Suppose also that user 1 and user 2 each has an initial endowment of 500 bits in each TS. If the RD curves for the two users are such that giving user 1 600 bits in TS 1 and 400 in TS 2 (with vice versa for user 2) produces a more favorable total MSE than the equal initial endowment, then the EWB approach would favor this allocation over the initial one. While this is the basic idea behind our approach, often adjacent TS have similar RD curves. Therefore, little benefit can be gained by trading bits between adjacent TS for two users. One would like to trade between the current encoding TS and some other TS widely separated in time. Since we may not know the specific RD curve for some distant GOP, instead of this, we will consider trades between the current encoding TS and some expected or approximate RD curve for the future. The RD curve for user i in TS j is fitted by

$$D_i^j(R_i^j) = a_i^j + \frac{b_i^j}{R_i^j} \tag{3}$$

where R_i^j is the number of bits and D_i^j is the MSE distortion for TS j in video stream i. a_i^j, b_i^j are the coefficients for generating this curve-fitting model and we use the least squares approach to find these coefficients. Note that the above function is convex. Other curve-fitting models are available in the literature [7]. Let the utility for user i be

$$U_i(x_i^1, x_i^2) = -(a_i^1 + \frac{b_i^1}{x_i^1} + a_i^2 + \frac{b_i^2}{x_i^2}) \quad (4)$$

that is, the negative sum of the MSE in both TS. The utility is a convex function. Let the initial endowment for user i be $\bar{c}_i$. Then the indifference curve through the initial endowment for user i can be derived as

$$-(a_i^1 + \frac{b_i^1}{x_i^1} + a_i^2 + \frac{b_i^2}{x_i^2}) = -(a_i^1 + \frac{b_i^1}{c_i^1} + a_i^2 + \frac{b_i^2}{c_i^2}) \quad (5)$$

for different combinations of x_i. A competitive equilibrium is found by solving

$$\max_{x_i^1, x_i^2} U_i(x_i^1, x_i^2) \; s.t. \; p^1 x_i^1 + p^2 x_i^2 = p^1 c_i^1 + p^2 c_i^2, \; \forall i = 1 \, to \, N \quad (6)$$

and

$$\sum_{i=1}^{N} x_i^j = \sum_{i=1}^{N} c_i^j \qquad \forall j = 1, 2 \quad (7)$$

The Lagrangian expression for user i is

$$L_i = U_i(x_i^1, x_i^2) + \lambda_i (p^1 c_i^1 + p^2 c_i^2 - p^1 x_i^1 - p^2 x_i^2) \quad (8)$$

By differentiating L_i with respect to x_i^1, x_i^2, and λ_i, equating the results to 0, we get

$$\sum_{i=1}^{N} \sqrt{\frac{b_i^j}{p^j}} \frac{p^1 c_i^1 + p^2 c_i^2}{\sqrt{p^1 b_i^1} + \sqrt{p^2 b_i^2}} = \sum_{i=1}^{N} c_i^j \qquad \forall j = 1, 2 \quad (9)$$

To determine the competitive equilibrium, we need to find the slope $-p^1/p^2$. Therefore, we assume $p^1 = 1$ and solve Eq. 9 numerically for p^2. With p^2, we find x_i^j which is the solution for competitive equilibrium.

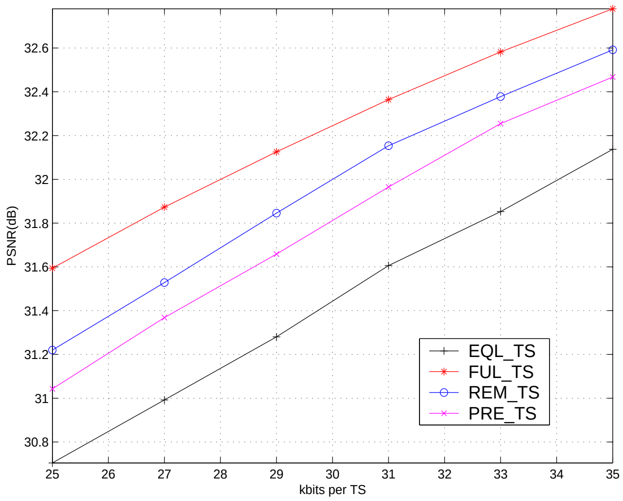
First, we consider the constant bit allocation for each TS (**EQL_TS**). Here each video in every TS receives an equal number of bits to encode its video. This rate control method is applied in many video standards such as H.264/AVC [8]. We now compare the competitive equilibrium bit allocation for various video streams. The first TS is always considered to be the current TS that we are encoding. If we assume that we have some information about the future, like the average RD curves for future TS, then we can use such information for trading bits for the current TS with the average of remaining TS (**REM_TS**). This is an *ex ante* approximation model where we assume some information about the future. If we have no future information then we predict the future by looking at the previous TS (*ex post*) with the assumption that the average RD curve of previous TS will be similar to that of the future. We trade bits for the current TS with the average of previous TS (**PRE_TS**). The performance of PRE_TS depends on how exactly the past TS represents the future TS. Both REM_TS and PRE_TS are solved for competitive equilibrium. For comparison, if each user has full information about its RD curves in all TS, then it can divide the bits among all the TS based on their relative complexity. All video streams use this criteria for bit allocation for their TS independently. Since the total number of bits is constant for each TS, we normalize the number of bits produced in each video stream by the total available bits for a TS (**FUL_TS**). In this paper, these four bit allocation methods are compared for video multiplexing.

4. RESULTS

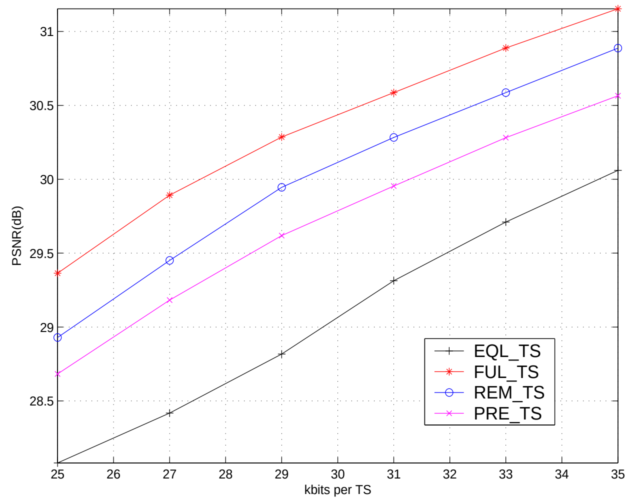
The simulation was performed using the baseline profile of H.264/AVC [9] reference software JM 11.0 [10]. The 30-second test videos containing varying types of scenes and motion were taken from a 72 minute travel documentary at a resolution of 176×120 pixels and 30 frames per second. The GOP size is 15 frames (I-P-P-P). The frames inside a TS are encoded using H.264 rate control [8]. The coding parameters such as resolution or GOP size can be changed for any appropriate application and the results are expected to be similar.

Figure 3 shows the result of multiplexing four video streams. The four curves in each plot represent the multiplexing methods described previously. Each plot shows PSNR versus bitrate (ranging from 25-35 kbits per TS (50-70 kbps)). We calculate the MSE of each frame and average across all frames of a video then convert to PSNR. The performance of EQL_TS is worst in all videos, and this is the method used in most video standards for GOP level rate control. For archived video we know RD curves for all TS and we see that FUL_TS performs the best. The PSNR gain over EQL_TS varies from 0.25-0.43 dB for g12 to 1.1-1.5 dB for g9. However, this method cannot be used for real-time video multiplexing. If a user knows the average RD curve for future TS, then this is sufficient to improve the video quality as shown by REM_TS. This method finds the competitive equilibrium point for the current TS when compared to its average of remaining TS. This method improves the quality of each video stream from 0.18-0.34 dB for g12 to 0.82-1.13 dB for g9 over the EQL_TS method. Finally, we assume that we have no prior knowledge about the video and we predict the future RD curves by looking at the previous TS. Again we compute the competitive equilibrium for the current TS when compared to the average of the previous TS (PRE_TS curve in the figure). This method improves the PSNR from 0.11-0.23 dB for g12 to 0.50-0.80 dB for g9. Similarly, Figure 4 shows the results for two videos (the two most extreme cases) when six video streams are multiplexed together using the methods described above. The PSNR of all the six videos improves from these multiplexing methods for a wide range of bitrates. g9 video again performs the best while the performance for REM_TS, PRE_TS, and FUL_TS are the same in g5 video.

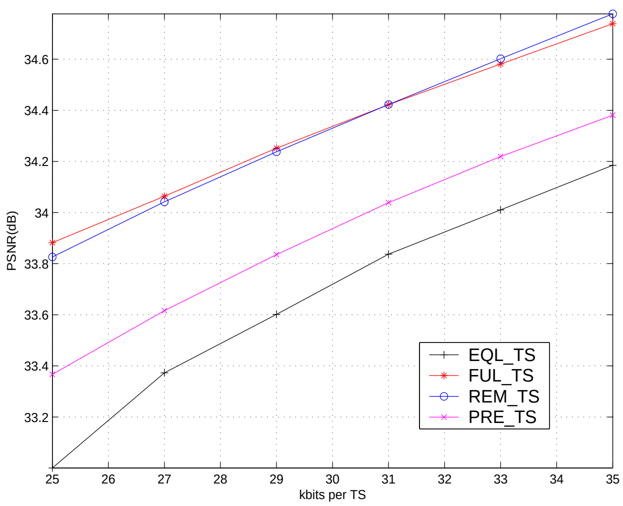
We note that the largest PSNR gain is achieved by finding the competitive equilibrium when there is a lot of fluctuation



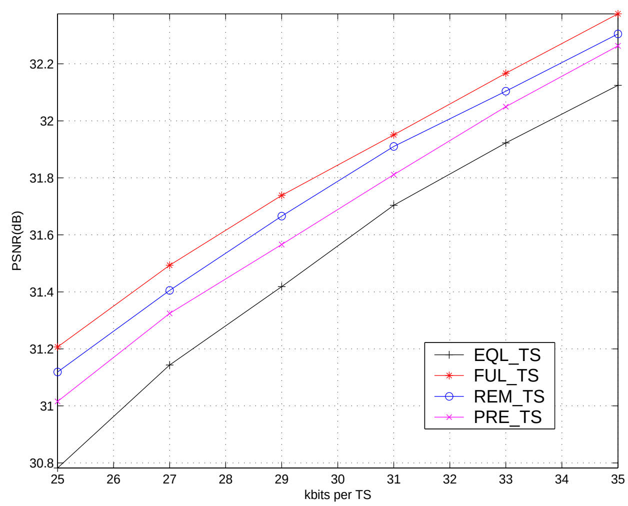
(a) g1 video stream



(b) g9 video stream



(c) g10 video stream

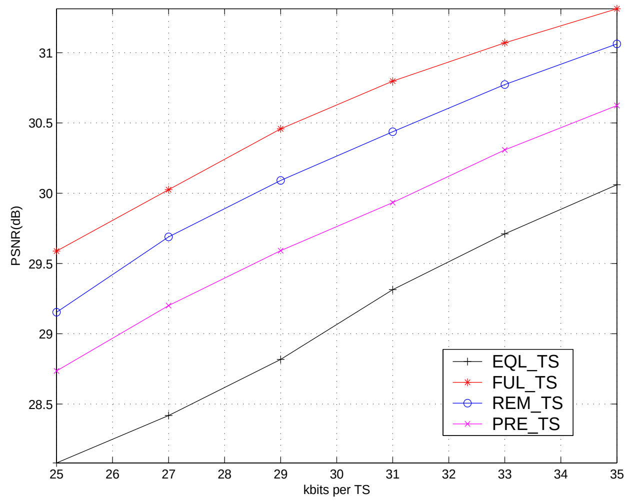


(d) g12 video stream

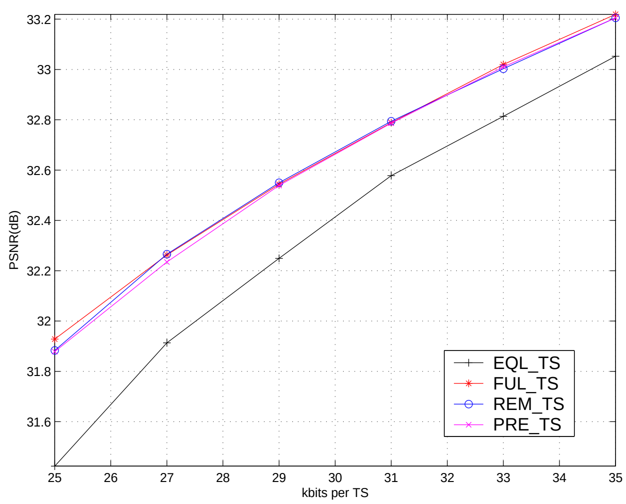
Fig. 3. PSNR variation with bitrate for four multiplexed video streams

in the video motion, for example g9. Conversely, the PSNR gain is low if the motion fluctuation in a video stream is low, for example g12. Most of the video streams have a lot of motion fluctuation and scene change, so multiplexing them by computing the competitive equilibrium will improve the quality. The performance of PRE_TS depends on how accurate is the representation of future TS from past TS. As can be seen from Figure 3, all the video streams gain from the multiplexing process. The PSNR gain varies from one video to another, depending on the content. The multiplexing method using the competitive equilibrium borrows bits from a low motion TS of a video and gives these bits to another video in the same TS with the promise of taking it back when the need arises. So, the multiplexing method exchanges bits between video streams as well as across the TS. This leads to another observation that the quality fluctuation for each video stream is reduced because the high motion TS gets more bits than the low motion TS instead of getting the same number of bits for all TS. Figure 5 shows the PSNR fluctuation for g9 for all the multiplexing methods. The EQL_TS method has the highest fluctuation and FUL_TS method has the lowest. In the end, all the videos receive equal numbers of bits in the multiplexing method unlike previous methods for video multiplexing where some videos get more bits than the other videos. By changing the encoding technique inside a GOP (e.g., using multiple reference frame prediction or using hierarchical B-frames), along with these multiplexing methods, the overall video quality can be expected to further improve.

In conclusion, we discussed four methods for multiplexing video streams. We proposed two novel methods of multiplexing video streams using the EWB solution for finding competitive equilibrium. The results show PSNR improve-



(a) g9 video stream



(b) g5 video stream

Fig. 4. PSNR variation with bitrate for six multiplexed video streams

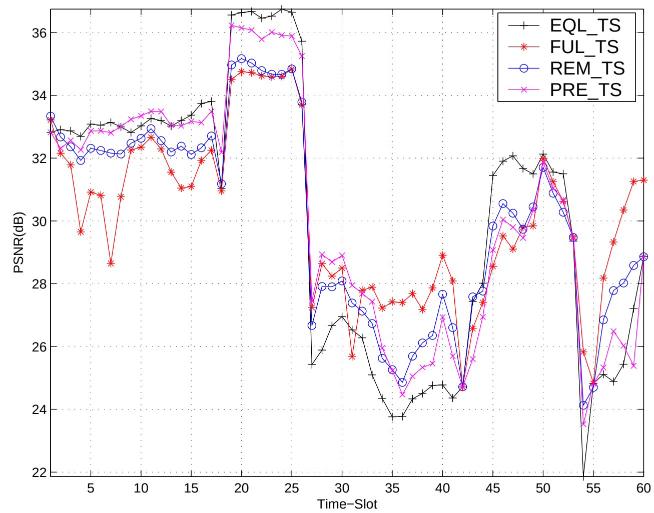


Fig. 5. Variation of PSNR with TS for g9 video at 50 kbps

ment for all video streams unlike previous methods [1–3] where the quality of some videos is improved while deteriorating the quality of other videos. The PSNR gain is greater for videos with higher motion fluctuation. In future work, we would like to examine better estimation of future frames from the previous frames.

5. REFERENCES

[1] L. Wang and A. Vincent, “Bit allocation and constraints for joint coding of multiple video programs,” *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 9, no. 6, pp. 949–959, Sept. 1999.

[2] G. Su and M. Wu, “Efficient bandwidth resource allocation for low-delay multiuser video streaming,” *IEEE Trans. on Cir. and Sys. for Video Tech.*, vol. 15, no. 9, pp. 1124–1137, Sept. 2005.

[3] M. Tiwari, T. Groves, and P. Cosman, “Multiplexing video streams using dual-frame video coding,” *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Mar. 2008.

[4] F. Edgeworth, “Mathematical psychics: An essay on the application of mathematics to the moral sciences,” *London: C.Kegan Paul and Co.*, 1881.

[5] A. Mas-Colell, M. Whinston, and J. Green, “Microeconomic theory,” *Oxford University Press Inc.*, 1995.

[6] L. Lin and A. Ortega, “Bit-rare control using piecewise approximated rate-distortion characteristics,” *IEEE Trans. on Cir. and Sys. for Video Tech.*, vol. 8, no. 4, pp. 446–459, Aug. 1998.

[7] K. Stuhlmüller, N. Färber, M. Link, and B. Girod, “Analysis of video transmission over lossy channels,” *IEEE Journal on Selected Areas in Comm.*, vol. 18, no. 6, pp. 1012–1032, June 2000.

[8] K.-P. Lim, G. Sullivan, and T. Wiegand, “Text description of joint model reference encoding methods and decoding concealment methods,” *Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-K049*, Mar. 2004.

[9] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” *IEEE Transactions on Circuits and Systems for Video Tech.*, vol. 13, no. 7, pp. 560–576, July 2003.

[10] “H.264/AVC ref. software,” http://iphome.hhi.de/suehring/tml.