JOINT ERROR-RESILIENT VIDEO SOURCE CODING AND FEC CODE RATE OPTIMIZATION FOR AN AWGN CHANNEL

Qing Song[⋆] *Arash Vosoughi*[†] *Pamela C. Cosman*[⋆] *Laurence B. Milstein*[⋆]

[⋆]UC San Diego, Dept. of Electrical and Comp. Engr., 9500 Gilman Dr, La Jolla, CA 92093-0407
[†]LG Electronics Mobile Research Lab, 2540 N. First Street Suite 300, San Jose, CA 95131

ABSTRACT

A joint source-channel rate-distortion (RD) optimization is proposed for video communication systems. The source coding and channel coding options are optimized by seeking the best trade-off between the estimated end-to-end distortion of a video packet and the sum of the number of source bits and forward error correction bits used to encode that packet. The proposed RD algorithm controls the total bit rate by using a Lagrange multiplier. Compared to conventional RD optimization schemes, which only optimize over the source coding modes of macroblocks, our proposed RD algorithm achieves superior performance over an AWGN channel.

Index Terms— Error resilience, rate distortion optimization, forward error correction

1. INTRODUCTION

Video packets can be corrupted when they are transmitted through a wireless channel. One way to preserve the quality is by using forward error correction (FEC) which uses channel coding to correct transmission errors. For a given type of FEC, more FEC bits provide stronger protection. Unequal protection of different video packets has been studied [1–6]. These approaches generally allocate more FEC bits to more important video packets.

Another approach to provide error control is error resilient source coding. One way is intra refresh: inserting intra-coded macroblocks or frames to limit error propagation. But intra coding usually costs more bits than inter coding. The work in [7] shows that inserting intra-coded macroblocks randomly increases the error resilience to packet losses, while a rate-distortion (RD) optimized technique provides more effective mode selection and further improves error resilience. The work in [8] describes an RD optimized technique which considers the distortion at the decoder. The distortion is computed recursively at pixel-level precision by considering quantization, error propagation and concealment in the model. Zhang et al. [9] proposed an end-to-end distortion estimation which works for H.264 at subpixel precision.

One problem of [7–9] is that the loss probability of all packets is assumed to be constant. Further, only the bits of the source coding are considered in the RD optimization (RDO). He et al. [10] optimized the FEC code rate and intra refreshing rate by minimizing the estimated distortion from an RD model they developed. Zhai et al. presented RD optimized error control which considers FEC in [11]. Though they packetize each macroblock into a packet, the complexity of solving the optimization problem is very high, so dynamic programming has to be applied. In this paper, we design a joint source-channel RD optimized algorithm for video communication. We assume the channel signal-to-noise ratio (SNR) is known by the encoder. The encoder selects the optimal quantization parameter (QP) and mode of each macroblock, as well as the optimal channel coding option (i.e., FEC code rate) of each packet sequentially. The overall bit rate, including both source bits and channel bits, is matched to a target bit rate by adjusting a Lagrange multiplier packet by packet.

The paper is organized as follows. In Section 2, an overview of our system is presented. In Section 3, we describe the problem formulation and explain the optimal solution. We discuss the numerical results in Section 4. Section 5 concludes the paper.

2. OVERVIEW OF THE PROPOSED JOINT-SOURCE-CHANNEL RDO

RD optimization plays a key role in all state-of-the-art video compression codecs such as H.264/AVC. Conventional RDO algorithms only consider source distortion and source coding, but do not handle channel impairments. In this paper, we show that we can improve RDO performance in an error-prone video transmission scenario, by considering the bits required by variable-rate FEC in the RDO process. We refer to our proposed RDO scheme as *joint source-channel RDO* (JSC-RDO). Fig. 1 shows a block diagram of the JSC-RDO scheme. At the transmitter, the raw video content is first compressed. The source bits are then protected against channel errors by adding FEC bits. The JSC-RDO controls the number of source bits by choosing from a set of available coding options such as the QP, coding mode, etc. It also determines the protection level by choosing an FEC code rate from a set of available code rates, each of which provides a different protection level. The resulting bit stream is then modulated and

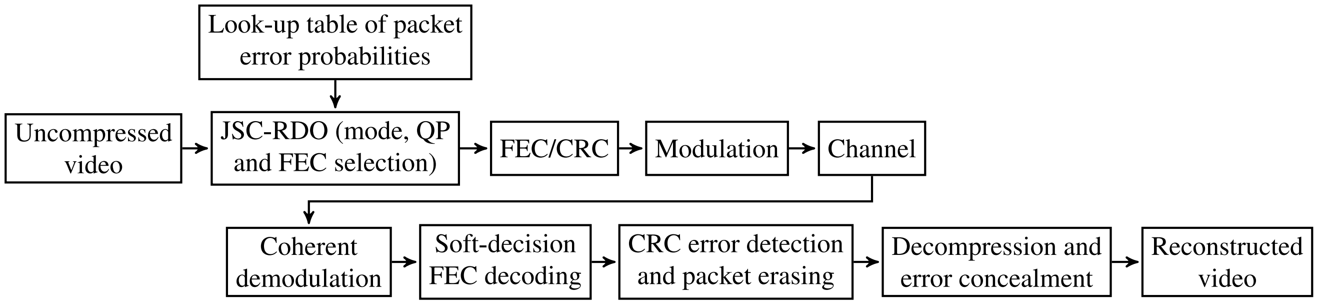

Fig. 1: System block diagram.

transmitted over the channel. At the receiver, channel decoding is done to detect and correct the erroneous bits. We assume an ideal CRC (cyclic redundancy check). If the CRC check fails, the whole packet is marked as undecodable. The intact slices are decoded by the video decoder, and the undecodable ones are concealed.

In this work, we consider video transmission over AWGN channel. We adopt binary phase shift keying (BPSK) modulation/demodulation for data transmission over the channel. We assume the energy per bit to noise power spectral density ratio (E_b/N_0) is known at the encoder. As we describe in the following sections, the proposed JSC-RDO scheme needs to have knowledge of the packet error probability to solve the optimization problem. The probabilities are computed by simulation and kept at the encoder in a look-up table.

3. PROBLEM FORMULATION

In the H.264/AVC standard, the QP and the mode are determined for each macroblock. Usually QP is first determined either by the user or by some rate control algorithm to meet the pre-defined bit rate. Then the mode is determined by RDO. For macroblock m, the conventional RDO [12–15] to select the optimal coding mode aims to minimize the Lagrangian cost $J_{MB}(m, o; q)$ of the macroblock: $J_{MB}(m, o; q) = D_{MB\_s}(m, o; q) + \lambda R_{MB\_s}(m, o; q)$, where D_{MB_s} is the source distortion, R_{MB_s} is the number of source bits needed to encode the macroblock with the mode o and given QP q, and λ is the Lagrange multiplier which controls the trade-off between distortion and rate.

An error-resilient RDO (ER-RDO) technique is proposed in [9]. The source distortion in the Lagrangian cost is replaced by the estimated end-to-end distortion $E[D_{MB}(m, o; q, p)]$, where p is the loss probability. The end-to-end distortion consists of source distortion D_{MB_s}, error propagated distortion D_{MB_ep} and error concealment distortion D_{MB_ec} of the macroblock: $E[D_{MB}(m, o; q, p)] = (1-p)\big(D_{MB\_s}(m, o; q)$ $+D_{MB\_ep}(REF, m_J)\big) + pD_{MB\_ec}(m)$, where REF is the reference frame for motion compensation, and m_J denotes the referenced blocks when mode o is selected. Each distortion is the sum of squared errors. The error propagated distortion is recursively calculated frame by frame. Frame copy, which means copying the co-located macroblocks in the previous frame, is used for error concealment, so the error concealment distortion depends only on the location of the macroblock.

Packets are protected by FEC in our system. Unlike [9], where only source bits are considered, we take into account both source and FEC bits. Moreover, the loss probability p is not constant as in [9], but depends on E_b/N_0, the FEC code rate and the size of the source packet. We will modify ER-RDO to include the channel impact as well as the source-channel bit allocation.

We assume a slice includes a row of macroblocks, so the sizes of slices vary in numbers of bits. Each slice is encapsulated into one packet, so all macroblocks of a slice are protected by the same FEC code rate. Let r be the FEC code rate of the slice, and p denote the loss probability of the slice when it is protected by r for a given E_b/N_0. Since the FEC code rate is determined at the slice level, we need to formulate the optimization problem at the slice level.

We denote the QP and mode of macroblock m by q_m and o_m, respectively. We define two vectors $\mathbf{q}$ and $\mathbf{o}$, which consist of the QPs and modes of all the macroblocks in the slice, i.e., $\mathbf{q} = [q_1, q_2, \cdots, q_M]$ and $\mathbf{o} = [o_1, o_2, \cdots, o_M]$, where M is the number of macroblocks in the slice. The estimated end-to-end distortion D_{SL} of the slice for a given p is obtained by summing up the distortion of all the macroblocks in this slice: $E[D_{SL}(\mathbf{q}, \mathbf{o}; p)] = \sum_{m=1}^{M} E[D_{MB}(m, q_m, o_m; p)]$. The total number of bits, R_{SL}, is the sum of the number of source bits and the number of FEC bits: $R_{SL}(\mathbf{q}, \mathbf{o}; r) = \sum_{m=1}^{M} \frac{R_{MB\_s}(m, q_m, o_m)}{r}$. The Lagrangian cost of the slice is given by

$$J_{SL}(\mathbf{q}, \mathbf{o}; r, p) = E[D_{SL}(\mathbf{q}, \mathbf{o}; p)] + \lambda R_{SL}(\mathbf{q}, \mathbf{o}; r) \quad (1)$$

We formulate our optimization problem as

$$\min_{\substack{\mathbf{q}\in\mathcal{Q}^M,\ \mathbf{o}\in\mathcal{O}^M \\ (r,p)\in\mathcal{R}\times\mathcal{P}}} J_{SL}(\mathbf{q}, \mathbf{o}; r, p), \quad (2)$$

where $\mathcal{Q}$ is the set of available QP, $\mathcal{O}$ is the set of available modes, $\mathcal{R}$ is the set of available FEC code rates, and $\mathcal{P}$ is the set of possible loss probabilities.

Since (r, p) is the same for the whole slice and is not chosen for each individual macroblock, we rewrite (2) as the minimization problem below:

$$\min_{(r,p)\in\mathcal{R}\times\mathcal{P}} \Big\{ \min_{\mathbf{q}\in\mathcal{Q}^M,\ \mathbf{o}\in\mathcal{O}^M} J_{SL}(\mathbf{q},\mathbf{o};r,p) \Big\}. \tag{3}$$

This shows that the optimization problem can be solved in two steps. For each $(r,p)\in\mathcal{R}\times\mathcal{P}$, the optimal QPs $\mathbf{q}^*$ and modes $\mathbf{o}^*$ are first obtained. Let $J^*_{SL}(r,p)$ be the slice Lagrangian cost when $\mathbf{q}^*$ and $\mathbf{o}^*$ are used to encode the slice, given (r,p). Then the optimal (r^*,p^*) is obtained by

$$(r^*,p^*) = \underset{(r,p)\in\mathcal{R}\times\mathcal{P}}{\arg\min}\ J^*_{SL}(r,p). \tag{4}$$

To find $\mathbf{q}^*$ and $\mathbf{o}^*$ for a given (r,p), we first define the Lagrangian cost of macroblock m as $\tilde{J}_{MB}(m,q_m,o_m;r,p) = E[D_{MB}(m,q_m,o_m;p)] + \lambda\frac{R_{MB\_s}(m,q_m,o_m)}{r}$. Then the Lagrangian cost of the slice can be written as the sum of the Lagrangian costs of all the macroblocks in the slice. To reduce the computational complexity to find the optimal QP and modes, the minimization of this sum is approximated by the problem below:

$$\sum_{m=1}^{M} \min_{\substack{q_m\in\mathcal{Q}\\ o_m\in\mathcal{O}}} \tilde{J}_{MB}(m,q_m,o_m;r,p). \tag{5}$$

This problem (5) can be solved by minimizing the Lagrangian cost of each macroblock for the given (r,p):

$$(q^*_m,o^*_m) = \underset{q_m\in\mathcal{Q},\, o_m\in\mathcal{O}}{\arg\min}\ \tilde{J}_{MB}(m,q_m,o_m;r,p),\ 1\le m\le M.$$

Note that it is possible $\mathbf{q}^*$ and $\mathbf{o}^*$ result in a source packet with an actual size $R^*_{SL\_s}$ bits such that, if that packet is protected by the code rate r, its loss probability would not be equal to p. Therefore, after we obtain $\mathbf{q}^*$ and $\mathbf{o}^*$, we find the actual loss probability $\hat{p}(p)$ which corresponds to the code rate r and the actual size $R^*_{SL\_s}$. The actual loss probability depends on the code rate, the assumed probability and video content, but for simplicity, we denote it as $\hat{p}(p)$. The slice Lagrangian cost J^*_{SL} for the assumed code rate r and the assumed loss probability p is computed by using the actual probability $\hat{p}(p)$:

$$J^*_{SL}(r,p) = E[D_{SL}(\mathbf{q}^*,\mathbf{o}^*,\hat{p}(p))] + \lambda\frac{R^*_{SL\_s}}{r}. \tag{6}$$

The optimal (r^*,p^*) is still obtained by (4). The corresponding actual loss probability, $\hat{p}(p^*)$, is used to update the error propagated distortion. The updating scheme of the error propagated distortion can be found in [9].

As we mentioned, the loss probability p depends on the size of the source packet, the FEC code rate, and E_b/N_0. We build a look-up table of packet loss probabilities by Monte-Carlo simulation. The loss probability entry $\mathrm{P}_{i,j}$ of the look-up table is estimated for a packet with source bits S_i which is protected by FEC code rate R_j under the given E_b/N_0. To find (r^*,p^*) of each slice, we perform an exhaustive search over all the loss probability entries in the look-up table under the given E_b/N_0. Given R_j and $\mathrm{P}_{i,j}$, we compute the best coding options for all the macroblocks of the current slice. We find the actual size is S_w, so we calculate the slice Lagrangian cost $J^*_{SL}(R_j,\mathrm{P}_{i,j})$ in (6) using $\hat{p}(\mathrm{P}_{i,j}) = \mathrm{P}_{w,j}$.

Though the RDO is formed as an unconstrained problem, it is still necessary to meet the overall target bit rate. The Lagrange multiplier controls the selection of the source coding options as well as the FEC code rate. So instead of computing the Lagrange multiplier as being dependent on QP, we determine the Lagrange multiplier by the method in [8, 16]. Suppose k slices have been encoded. $R_{sum}(k)$ denotes the total number of bits (including both source bits and FEC bits) that are used to encode these k slices, and R_{target} denotes the target number of bits to encode one slice. We update the Lagrange multiplier per slice via:

$$\lambda_{k+1} = \lambda_k\Big(1+\alpha\big(R_{sum}(k) - kR_{target}\big)\Big), \tag{7}$$

where α is given by $\alpha = \frac{1}{\beta R_{target}}$, and where β is a constant for the whole sequence [8].

4. NUMERICAL RESULTS

We implemented our proposed method by modifying the H.264/AVC codec JM reference software 15.1 [17]. For FEC, we use UMTS turbo codes [18]. The turbo encoder is composed of two recursive systematic convolutional encoders with constraint length 4, which are concatenated in parallel. The feedforward and feedback generators are 15 and 13, respectively, both in octal. The code rates we considered are $\mathcal{R_{FEC}} = \{\frac{8}{9},\frac{8}{10},\frac{8}{12},\frac{8}{14},\frac{8}{16},\frac{8}{18},\frac{8}{20},\frac{8}{22},\frac{8}{24}\}$, which are obtained by puncturing a mother code of rate $\frac{1}{3}$. Further, we include the options of "uncoded" and "discarding". "Uncoded" means the slice is not protected by FEC. "Discarding" means that the slice will not be transmitted and will be concealed at the decoder. If discarding a slice yields lower estimated distortion than sending it, the encoder would choose discarding. The saved bits would be allocated to the following slices. We provide the encoder with all the aforementioned code rate options, and name this scheme as unequal error protection (UEP). We compare UEP with several schemes:

1. EF-RDO+EEP: conventional RDO is used to select the mode of each macroblock. It is denoted error-free (EF) RDO since it does not consider channel impairments. Equal error protection (EEP) is adopted. All the packets are protected by the same FEC code rate r_E.
2. JSC-RDO+EEP: JSC-RDO is used to select coding options, but only one code rate r_E is provided to the encoder.
3. JSC-RDO+EEP+Discard: this scheme is similar to JSC-RDO+EEP. The difference is that the available FEC code rate options include one code rate r_E as well as discarding.

For each EEP scheme, we test each $r_E \in \mathcal{R_{FEC}}$ for 1,000 channel realizations to find the rate which yields the best PSNR. We test two CIF sequences, each including 100 frames at 30 fps. The GOP structure is IPPP with only one I-frame at the beginning of the video. We assume the slices

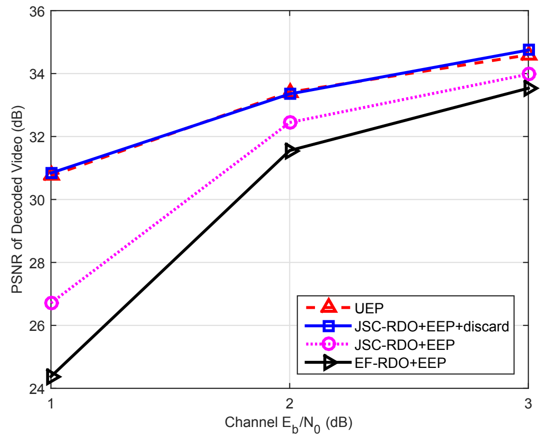

Fig. 2: PSNR vs. E_b/N_0: Hall at 256 kb/s.

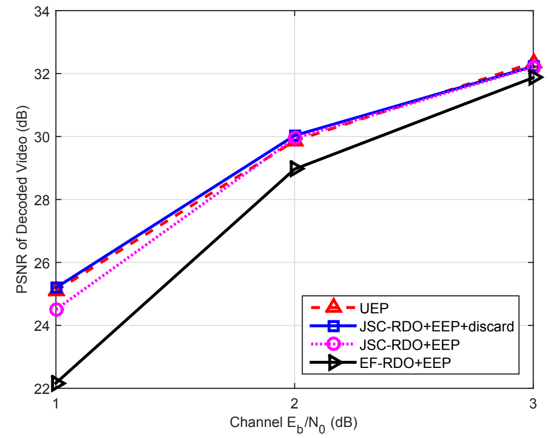

Fig. 3: PSNR vs. E_b/N_0: Foreman at 384 kb/s.

which contain the parameter sets and the first frame are received correctly.

Figs. 2-3 show the results for the sequences Hall and Foreman. For Hall, the performances of UEP and JSC-RDO+EEP+Discard are almost the same. They are both better than JSC-RDO+EEP, which is better than EF-RDO+EEP. When E_b/N_0 is 1 dB, UEP and JSC-RDO+EEP+ Discard outperform JSC-RDO+EEP by about 4 dB in PSNR, and outperform EF-RDO+EEP by about 6 dB. For Foreman, the performances of UEP, JSC-RDO+EEP+Discard and JSC-RDO+EEP are almost the same, which are better than EF-RDO+EEP. The PSNR gain by using JSC-RDO is about 3 dB when E_b/N_0 is 1 dB.

For JSC-RDO+EEP+Discard, the code rates giving the best performance for E_b/N_0 1 dB, 2 dB and 3 dB are $\frac{8}{24}$, $\frac{8}{16}$ and $\frac{8}{12}$, respectively, for both sequences. For UEP, the encoder selects almost the same code rates as JSC-RDO+EEP+Discard, even though 11 code rate options are provided. That explains why the performances of the two schemes are so similar. Note that the computational complexity of JSC-RDO+EEP+Discard is only about $\frac{1}{10}$ of UEP.

The results show that UEP and JSC-RDO+EEP+Discard perform better than JSC-RDO+EEP for Hall, but almost the same as JSC-RDO+EEP for Foreman. That is because Hall has a static background and the foreground objects move slowly. The slices which are mostly static are discarded by the encoder when discarding is allowed. These slices can be concealed very well at the decoder. The bits saved by these slices are allocated to the other slices which include more foreground information. For JSC-RDO+EEP, no discarding is allowed. Even if there is no motion information in these slices and all the macroblocks are coded in SKIP mode, the slice headers cost some bits, and these bits are protected by FEC. In Foreman, however, the camera is moving. Only a few slices are discarded by the encoder when discarding is available. So the performance is not improved by including the discarding option.

The results also show that all the schemes using JSC-RDO achieve better PSNR than using EF-RDO+EEP. EF-RDO does not consider any channel distortion. Any error in a frame can propagate to the following P frames. JSC-RDO, however, estimates the end-to-end distortion and inserts intra-coded macroblocks to limit the error propagation. The intra-coded macroblocks improve the robustness to channel errors.

Therefore, the advantage achieved by UEP and JSC-RDO+EEP+Discard is due to 1) JSC-RDO which provides error resilience by selecting intra mode, and 2) discarding slices which can be concealed well at the decoder and allocating the saved bits to the more important slices.

5. CONCLUSION

We proposed a RD optimization algorithm for video which jointly selects the source coding and channel coding options based on the channel SNR. While previous RDO algorithms extensively considered the error-free case, and later examined the error-prone case under the assumption of constant loss probability and no FEC, we allow variable loss probability and explicitly include the FEC in the optimization, including its effects both in the distortion computation and the bit rate cost. We showed that the schemes which use JSC-RDO significantly outperform the conventional RDO. Our results also show that the JSC-RDO+EEP+Discard scheme and UEP perform almost the same, while JSC-RDO+EEP+Discard is much more efficient than UEP. We found that the discarding option helps to improve the performance when there are areas with very little motion compared to the previous frame. This applies to video conferencing, where the camera is usually static. For real-time video communication systems, JSC-RDO+EEP+Discard is achievable without introducing too much computational complexity.

6. REFERENCES

[1] X. Yang, C. Zhu, Z. Li, X. Lin, G. Feng, S. Wu, and N. Ling, "Unequal loss protection for robust transmission of motion compensated video over the internet," *Signal Processing: Image Communication*, vol. 18, no. 3, p. 157167, Mar 2003.

[2] Y. C. Chang, S. W. Lee, and R. Komyia, "A fast forward error correction allocation algorithm for unequal error protection of video transmission over wireless channels," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 3, pp. 1066–1073, August 2008.

[3] T.-L. Lin and P. Cosman, "Efficient optimal RCPC code rate allocation with packet discarding for pre-encoded compressed video," *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 505–508, May 2010.

[4] Z. Luo, L. Song, S. Zheng, and N. Ling, "Raptor codes based unequal protection for compressed video according to packet priority," *IEEE Transactions on Multimedia*, vol. 15, no. 8, pp. 2208–2213, Dec 2013.

[5] J. Xiao, T. Tillo, C. Lin, and Y. Zhao, "Dynamic Sub-GOP forward error correction code for real-time video applications," *IEEE Transactions on Multimedia*, vol. 14, no. 4, pp. 1298–1308, Aug 2012.

[6] H. D. Pham and S. Vafi, "Optimized motion energy estimation for group of pictures in multi-level error protection of H.264/AVC video bitstreams," *IEEE Transactions on Multimedia*, vol. 16, no. 8, pp. 2118–2129, Dec 2014.

[7] G. Cote and F. Kossentini, "Optimal intra coding of blocks for robust video communication over the internet," *Signal Processing: Image Communication*, vol. 15, no. 1-2, pp. 25–34, Sep 1999.

[8] R. Zhang, S. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 966–976, June 2000.

[9] Y. Zhang, W. Gao, Y. Lu, Q. Huang, and D. Zhao, "Joint source-channel rate-distortion optimization for H.264 video coding over error-prone networks," *IEEE Transactions on Multimedia*, vol. 9, no. 3, pp. 445–454, April 2007.

[10] Z. He, J. Cai, and C. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 6, pp. 511–523, Jun 2002.

[11] F. Zhai, R. Beryy, T. Pappas, and A. Katsaggelos, "A rate-distortion optimized error control scheme for scalable video streaming over the internet," in *Proceedings of International Conference on Multimedia and Expo*, vol. 2, July 2003, pp. II–125–8 vol.2.

[12] G. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *Signal Processing Magazine, IEEE*, vol. 15, no. 6, pp. 74–90, Nov 1998.

[13] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 688–703, July 2003.

[14] K. Takagi, Y. Takishima, and Y. Nakajima, "A study on rate distortion optimization scheme for JVT coder," in *Proc. SPIE Visual Communications and Image Processing*, vol. 5150, July 2003, pp. 914–923.

[15] I. Choi, J. Lee, and B. Jeon, "Fast coding mode selection with rate-distortion optimization for MPEG-4 Part-10 AVC/H.264," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 12, pp. 1557–1561, Dec 2006.

[16] T. Wiegand, M. Lightstone, D. Mukherjee, T. Campbell, and S. Mitra, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 2, pp. 182–190, Apr 1996.

[17] *H.264/AVC Reference Software*. [Online]. Available: http://iphome.hhi.de/suehring/tml/

[18] European Telecommunications Standards Institute, "Universal mobile telecommunications system (UMTS): Multiplexing and channel coding (FDD)," *3GPP TS 125.212 version 3.4.0*, pp. 14–20, Sept. 23, 2000.