Efficient Optimal RCPC Code Rate Allocation With Packet Discarding for Pre-Encoded Compressed Video

Ting-Lan Lin, *Student Member, IEEE*, and Pamela C. Cosman, *Fellow, IEEE*

***Abstract*—In an error-prone communication channel, more important video packets should be assigned stronger channel codes. With various packet sizes and distortions for each packet, we use the subgradient method to search in the dual domain for the optimal RCPC channel code rate allocation for each packet, to minimize the end-to-end video quality degradation for an AWGN channel. We exploit the advantage of not sending or not coding packets of lower importance.**

***Index Terms*—Packet discarding, packet protection, RCPC codes, unequal error protection, video coding.**

I. Introduction

COMPRESSED video packets can have different impact on video quality when lost. For example, due to error propagation, lost reference frames cause more damage than lost nonreference frames. Therefore, error-handling measures should be tailored to different components of the video. For example, in [1], H.264 Flexible Macroblock Ordering (FMO) is used to group macroblocks of similar estimated distortion into a slice, with different levels of Reed–Solomon (RS) coding over slices. The method in [2] optimally decides both slicing based on the estimated incurred distortion, and optimal forward error correction (FEC) rate for each packet. Two different RS codes are assigned to video data of high/low priorities in [3].

Joint source and channel coding that trades between source video quality and error resilience of the transmission in a lossy network is a well-studied area. The work in [4] studied a combined source-channel coding problem for transmission in an AWGN channel using RCPC codes. A universal operational distortion-rate characteristic is used for the optimization problem, and the performance of the algorithm approaches the information-theoretic bound. The rate-distortion optimization among source coding, channel coding and error concealment is jointly considered in [5]. A selective packet retransmission mechanism is integrated into the algorithm.

Unequal error protection (UEP) can be jointly used with intra updating which stops error propagation. This option requires more source and channel bits to intra-code the slice, therefore there is a tradeoff among video quality, error correctability, and the ability to stop error propagation in case of uncorrectable errors [6], [7]. UEP for progressively coded images/videos is also extensively studied [8], [9].

Most of the work on UEP discussed above involves either progressive/scalable coding or a change to the source encoder. In our work, we consider nonscalable video streams pre-encoded and stored; the problem is choosing optimal packet protection for the channel conditions at the time of transmission. This was treated in [10], however, they consider a packet erasure channel, and optimal RS erasure coding. Our paper treats the same problem of optimal code rate allocation for already encoded packets, but we consider a bit error channel and RCPC codes. The problem is therefore to unequally protect the packets for a given outgoing channel bit rate and channel SNR. The coding decision is based on both the distortion induced by the packet, as well as the size of the packet. The RCPC code allocation for packets is an integer programming problem. In our prior work [11], [12], we solved this problem using variations on the Branch-and-Bound method. In the proposed work, the optimal FEC code rate allocation search is done efficiently in the dual domain by the method of subgradients, and we also include the options of not coding a packet, or not sending it at all.

This paper is organized as follows. Section II formulates the end-to-end distortion minimization problem, and presents the low-complexity subgradient method in the dual domain to solve it. Experimental results for different videos, GOP structures, and channel conditions are presented in Section III. Section IV concludes the paper.

II. Optimal RCPC Code Rate Allocation

In this section, we introduce the problem formulation of the RCPC code rate allocation. The efficient subgradient method to solve this problem in the dual domain is discussed.

A. Problem Formulation for RCPC Code Allocation

Different packet losses introduce different levels of distortion. In our prior work [11], [12], the metric to define distortion is the packet loss visibility for each packet; this is a perceptual quality metric. Here, we use mean square error (MSE) induced by a packet loss as the distortion. This can be substituted with any metric measuring the importance of each packet.

N is the number of packets in each optimization. We optimize over one GOP at a time. As there are 30 frames in a GOP, and 15 packets in a frame, $N = 30 \times 15 = 450$. The ith packet has size S_i bits, and if lost, will induce a distortion D_i, given knowledge of the error concealment used by the decoder (frame-copy in our case). D_i is the MSE over a GOP between a compressed/reconstructed video with no packet loss, and one with the ith packet lost, because we assume we only have pre-encoded bitstreams available in the server when we perform the UEP. Although in

Manuscript received December 04, 2009; revised February 25, 2010. First published March 08, 2010; current version published March 31, 2010. This work was supported by the National Science Foundation and the Office of Naval Research. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lap-Pui Chau.

The authors are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093-0407 USA (e-mail: tinglan@ucsd.edu; pcosman@ucsd.edu).

Digital Object Identifier 10.1109/LSP.2010.2044938

running the algorithm, that D_i is used, in the final performance evaluation, we use the MSE between the actual decoded video and the original uncompressed video. Note that because the D_i is computed by applying the loss and decoding the *whole* GOP to measure MSE, this is a costly but accurate way of computing D_i, that includes the effect of error propagation. For the ith packet, we seek the optimal RCPC code rate r_i from the candidate set $\{R_1, R_2, \ldots, R_K\}$, so as to minimize the end-to-end expected packet loss distortion, subject to the constraint that the outgoing total bit budget is at most B bits.

The packet error probability P_e depends on channel SNR, packet size, and RCPC code rate selected for the packet. Each packet will be appended with a 16-bit CRC for error detection. We include the CRC bits in the packet size and assume perfect error detection. Whenever there is at least one bit error in the packet after channel decoding, we discard the packet. Therefore, $P_e = 1 - (1 - P_b(\mathrm{SNR}, r_i))^{S_i}$ where $P_b(\mathrm{SNR}, r_i)$ (hereafter denoted P_b) is the bit error probability after channel decoding for code rate r_i. The optimization problem of minimizing the expected packet loss distortion subject to the total bit constraint can be formulated as:

$$\min_{\mathbf{r}} \sum_{i=1}^{N} D_i\{1 - (1 - P_b)^{S_i}\} \text{ subject to } \sum_{i=1}^{N} \frac{S_i}{r_i} \le B$$
$$r_i \in \{R_1, R_2, \ldots, R_K\}, i = 1, 2, \ldots, N \quad (1)$$

where $\mathbf{r} = [r_1, r_2, \ldots, r_N]$. This is a nonlinear integer programming problem. We solved it by the BnB method in [11], [12]. To reduce complexity, heuristic packet groupings were used, and only four RCPC codes were used in [11] and six in [12]. In this paper, we use a low-complexity subgradient search in the dual domain to efficiently find the best code for each individual packet from the full set of 13 RCPC rates.

In (1), one must examine the meaningfulness of summing MSE contributions from different packets. Often in this type of formulation, other researchers take the distortion associated with a lost packet to be the initial MSE, computed over only the macroblocks in that packet. This does not account for error propagation. Also, in general H.264 slices can cover arbitrary pixel area; a formulation based on initial MSE would need weights by pixel area. In (1), however, the MSE associated with one packet is the MSE over the entire GOP induced by that one packet being lost. This accounts for all error propagation. Also, it means we can use packets with arbitrary numbers of macroblocks (although we happen to use packets with a constant pixel area) and we would not need to weight by pixel area, because regardless of the packet's pixel area, the MSE induced over the entire GOP is computed for each packet. A drawback of this GOP-based MSE which accounts for error propagation is that the measured distortion associated with losing both packets i and j is, in general, equal to the sum of the distortions associated with losing each one separately only if the error propagation of each did not affect the other one (for example, if the packets are in B frames, or if one loss is in a B frame and the other is in a much later P frame which is not referenced by the B frame). This noninteraction does not hold in general. However, as will be seen, the results from treating the components additively are good, and the complexity is far lower than it would be to account for interaction effects. For example reference method [10] also uses the additive assumption.

B. Dual of the Problem

We first relax our constrained optimization problem in (1) to an unconstrained problem [13], [14]. By absorbing the constraint into the objective with a Lagrange multiplier $\lambda \in \mathbb{R}^+$, we construct the Lagrangian function $L(\mathbf{r}, \lambda)$:

$$L(\mathbf{r}, \lambda) = \sum_{i=1}^{N} D_i(1 - (1 - P_b)^{S_i}) + \lambda \left(\sum_{i=1}^{N} \frac{S_i}{r_i} - B \right).$$

We form a dual function $d(\lambda)$ by minimizing the Lagrangian function for a given λ:

$$\begin{aligned} d(\lambda) &= \min_{\mathbf{r} \in \mathcal{C}} L(\mathbf{r}, \lambda) \\ &= \min_{\mathbf{r} \in \mathcal{C}} \left\{ \left[\sum_{i=1}^{N} D_i(1 - (1 - P_b)^{S_i}) + \lambda \frac{S_i}{r_i} \right] - \lambda B \right\} \\ &= \left\{ \sum_{i=1}^{N} \min_{\substack{r_i \in R_j \\ j=1,2,\ldots,K}} \left[D_i\{1 - (1 - P_b)^{S_i}\} + \lambda \frac{S_i}{r_i} \right] \right\} - \lambda B \\ &= \left\{ \sum_{i=1}^{N} \min_{\substack{r_i \in R_j \\ j=1,2,\ldots,K}} L_i(r_i, \lambda) \right\} - \lambda B \end{aligned} \quad (2)$$

where $\mathcal{C}$ is the space of all possible combinations of $r_i, i = 1, 2, \ldots, N$ selected from $\{R_1, R_2, \ldots, R_K\}$. This minimization for a given λ can be found by minimizing the sub-Lagrangians $L_i(r_i, \lambda)$ individually; the latter is done by exhaustive search over the discrete set $\{R_1, R_2, \ldots, R_K\}$. The solution space of the minimization of $L(\mathbf{r}, \lambda)$ is K^N, but because we can minimize sub-Lagrangians individually, we can compute $d(\lambda)$ with only NK evaluations of $L_i(r_i, \lambda)$ and comparisons [13].

C. Subgradient Method

We use the subgradient method to search for the best λ in the dual domain. The dual function $d(\lambda)$ is a concave function of λ even when the problem in the primal domain is not convex [13], [14]. Therefore, the optimal λ is found by solving: $\max_{\lambda \in \mathbb{R}^+} d(\lambda)$. Since the dual is a piecewise linear concave function [13], the function may not be differentiable at all points. Nevertheless, subgradients can still be found and used to find the optimal value [13]. It can be shown that the subgradient is a descent direction of the Euclidean distance to the set of the maximum points of the dual function [13]. This property is used in the well-known *subgradient method* for the optimization of a nonsmooth function. The subgradient method is an iterative search algorithm for λ. In each iteration, λ^{k+1} is updated by the subgradient ξ^k of d at λ^k:

$$\lambda^{k+1} = \max(0, \lambda^k + s_k \xi^k / \|\xi^k\|) \quad (3)$$

where s_k is the step size. Based on the derivation in [13], the subgradient ξ^k of the dual function $d(\lambda)$ at λ^k is

$$\xi^k = g(\mathbf{r}^k) - B = \sum_{i=1}^{N} \frac{S_i}{r_i^k} - B \quad (4)$$

where g is the constraint function of the problem, and $\mathbf{r}^k = [r_1^k, r_2^k, \ldots, r_N^k]$ is the solution to $L(\mathbf{r}, \lambda^k)$.

The step size s_k trades off between the speed of convergence and the variance of the optimized value in each iteration [13].

The complexity of this algorithm is low. In the proposed algorithm, the stepsize is scaled ten times smaller whenever there is a sign change in subgradient from the previous iteration. When a certain precision of stepsize is achieved, the algorithm terminates. The precision can be chosen differently by context; we used 10^{-18}. By this heuristic method on the change of the stepsize, our method finds the best λ using 82 iterations on average for each optimization.

D. Discarding Packets

In [10], RS codes are used for channel protection. The N packets in each optimization group are sorted based on their D_i. The algorithm discards the first k_d packets, sends the next k_u packets uncoded, and protects the remaining packets by a single code rate r, the strongest one meeting the bit constraint. The objective function to be minimized is the sum of expected distortion over all packets. The algorithm does not consider multiple code rates or variable packet size. We consider variable-sized packets and a bit error channel, so packet size needs to be considered in choosing protection, including the options of discarding and no protection. The same algorithm can be used; we include *"not-sent"* $(r_i = \infty)$ and *"uncoded"* $(r_i = 1)$ in the RCPC set. The optimization in (1) has the same form. In (1), the P_b corresponding to packets not sent is set to 1; those packets will induce distortion for sure. To signal to the decoder which code is used for each packet, we need 4 bits per packet. We assume these bits are collected together and well protected as part of a GOP header. Since the mean packet size is 1917 bits (including both source and channel bits), the 4-bit overhead is negligible.

III. Experimental Results

We used H.264/AVC JM Version 12.1 with SIF resolution (352×240), GOP structure IPPP and IBBP, frame rate 30 fps, and encoding rate 600 kbps. The error concealment is frame copy, which is one of the options provided in the JM.12.1 decoder. We define a packet (a NAL unit) as a horizontal row of macroblocks. There are 15 packets in a frame. For each GOP structure, we tested two videos: *Foreman* and *Mother-Daughter*. We optimize over one GOP at a time. As there are 30 frames in a GOP, the number of packets in each optimization is $N = 30 \times 15 = 450$. The mother code of the RCPC code has rate $1/4$, with memory $M = 4$. The puncturing period is $P = 8$. In the simulation, soft-decision is used for the Viterbi decoder. The RCPC rates each packet can select from are $\{(8/9), (8/10), (8/12), (8/14), (8/16), (8/18), (8/20), (8/22), (8/24), (8/26), (8/28), (8/30), (8/32)\}$. Therefore, there are $K = 13$ candidate code rates for our dual search algorithm (**Dual13**). When we include the options *"not-sent"* and *"uncoded"*, then $K = 15$ (denoted **Dual15**).

We simulate an AWGN channel, and find P_b given RCPC code rate and channel SNR. The RCPC rate used by Equal Error Protection **EEP** is $(8/14)$, and the budget for the UEP optimization problem is the number of bits used by the EEP in the optimization group. Channel SNR ranges from -2 to 5 dB, corresponding to channel bit error rates from about 10^{-1} to 10^{-3}. The end-to-end lossy video quality is measured by Peak Signal to Noise Ratio (PSNR), calculated by the MSE over all frames between decoded and original videos.

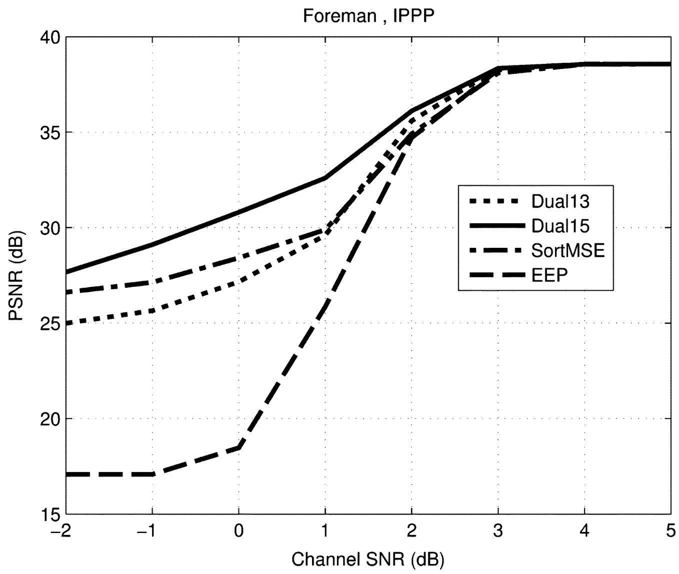
Although [10] is intended for packet erasure channels and uses RS codes, we made a version using RCPC codes intended for bit error channels. Among the N packets sorted on D_i, the first k_d are discarded, the next k_u are sent uncoded, and the remaining are protected with a single code rate r. We denote this **SortMSE**. SortMSE finds $\{k_d, k_u, r\}$ that solves

$$\begin{aligned} \min_{\{k_d,k_u,r\}} & \sum_{\underline{i}=1}^{k_d} D_{\underline{i}} + \sum_{\underline{i}=k_d+1}^{k_d+k_u} D_{\underline{i}}\{1-[(1-P_b(\mathrm{SNR},1))^{S_{\underline{i}}}]\} \\ & + \sum_{\underline{i}=k_d+k_u+1}^{N} D_{\underline{i}}\{1-[(1-P_b(\mathrm{SNR},r))^{S_{\underline{i}}}]\} \\ \text{subject to} & \sum_{\underline{i}=k_d+1}^{k_d+k_u} S_{\underline{i}} + \sum_{\underline{i}=k_d+k_u+1}^{N} \frac{S_{\underline{i}}}{r_{\underline{i}}} \le B \\ & r \in \{R_1, R_2, \ldots, R_K\} \quad (5) \end{aligned}$$

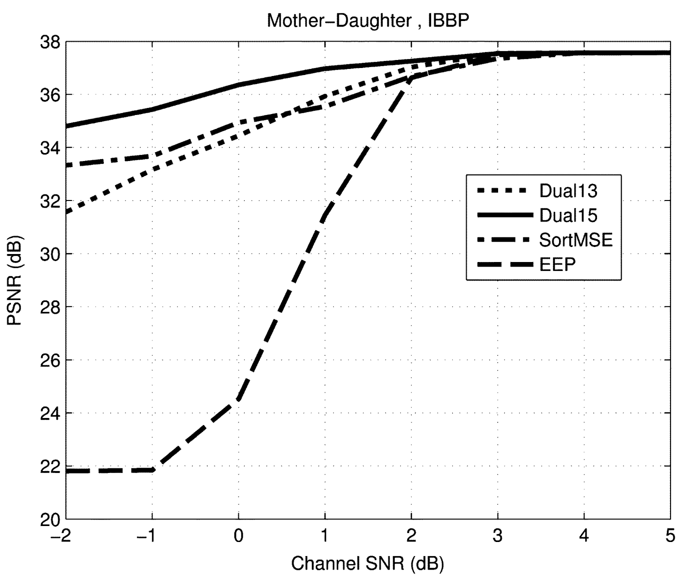
where $\underline{i}$ is the sorted packet index. This problem is solved by the method described in [10]. The proposed Dual method requires evaluating the sub-Lagrangian $L_i(r_i, \lambda)$ $82NK$ times or evaluating the Lagrangian function about $82K$ times. SortMSE, according to [10], requires evaluating the objective function in (5) at most $2N$ times. This objective function has complexity comparable to the Lagrangian function. Therefore the complexity comparison between the proposed Dual method and SortMSE is about $82K : 2N$. For the values we used ($K = 13$ RCPC code rates, $N = 450$ packets in the optimization) the complexities are comparable.

PSNR comparisons among decoded videos for Dual13, Dual15, SortMSE and EEP are performed for *Foreman* and *Mother-Daughter* for IPPP and IBBP GOP structures. All results show similar trends. We present the results for *Foreman* in IPPP in Fig. 1(a) and *Mother-Daughter* in IBBP in Fig. 1(b). We see obvious improvements of Dual15 that allows *"not-sent"* and *"uncoded"* over Dual13 that does not. The advantage of Dual15 takes place in every channel condition, and is more obvious at lower SNR because discarding large or unimportant packets is particularly useful in worse channel conditions. The largest improvements of Dual15 over Dual13 are 3.64 dB, 3.49 dB, 3.59 dB and 3.23 dB for the four comparisons (*Foreman* in IPPP, IBBP, *Mother-Daughter* in IPPP and IBBP). The advantage of discarding also can be observed for SortMSE: for low SNR, SortMSE outperforms Dual13 despite SortMSE allowing only one RCPC code for each optimization group. However for better channels, SortMSE is worse than Dual13 because in better channels, packets are less likely to be discarded, so the discarding option of SortMSE can not compensate for its having only one code rate. At better SNRs, SortMSE performs slightly worse than EEP. One might think that SortMSE should never do worse than EEP. EEP only assigns one code rate to all packets, whereas SortMSE chooses $\{k_d, k_u, r\}$ and should be more flexible. However, by basing the importance on distortion with no consideration of packet size, SortMSE may discard tiny packets that would have cost little to retain, or may heavily protect large packets that are costly to retain, and may do worse than EEP. However in worse channels, the flexibility of being able to discard packets compensates for this disadvantage. Dual15 that features both packet discarding and flexible rate allocation for each packet performs the best for every channel SNR.

The results so far use rate $(8/14)$ for the EEP. The total number of bits after channel coding by this EEP is the constraint for the optimization. It is possible that $(8/14)$ is particularly unsuitable for some channel SNRs. To show that



(a)



(b)

Fig. 1. Average PSNR of decoded video versus channel SNR. Comparison among Dual13, Dual15, SortMSE and EEP over 100 realizations of each AWGN channel. (a) *Foreman* in IPPP GOP structure and (b) *Mother-Daughter* in IBBP GOP structure.

Dual15 performs better than EEP for our entire set of possible EEP rates, we separately channel encode the same pre-encoded video sources using each different EEP (excluding the possibility of discarding everything). This gives rise to 14 different bit constraint totals. For each, we ran Dual15. The average improvement of Dual15 over the corresponding EEP is $\{13.2, 12.3, 9.8, 7.2, 5.6, 4.3, 3.5, 2.4, 1.7, 0.9, 0.5, 0.3, 0.09, 0\}$ dB. Dual15 outperforms EEP in all cases, even with a tiny bit budget (EEP=*"uncoded"*), because Dual15 can discard some packets to free up bits for protecting important packets. The advantage of Dual15 decreases as total bits increase, until finally when each packet is equally protected by the strongest channel code, the improvement of Dual15 over EEP vanishes, because at that high total bit rate Dual15 and EEP both can afford to equally protect every packet with the strongest rate. In summary, Dual15 outperforms all the EEP values except for the extreme cases of maximal protection and total discarding where they are equal.

IV. Conclusion

We propose an efficient RCPC code rate allocation algorithm over error-prone channels. The algorithm searches in the dual domain by the subgradient method for the optimal channel code rate for each packet with different packet size and different induced distortion. The algorithm is of low complexity. We exploit the options of not coding and not sending the packets. For all channel conditions, video clips and GOP structures tested, our dual algorithm significantly outperforms equal error protection as well as a simpler UEP version that considers only options of discarding, not coding, and a single level of protection.

References

[1] N. Thomos *et al.*, "Robust transmission of H.264/AVC video using adaptive slice grouping and unequal error protection," in *IEEE ICME*, Jul. 2006, pp. 593–596.

[2] O. Harmanci and A. M. Tekalp, "Rate-distortion optimal video transport over IP allowing packets with bit errors," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1315–1326, May 2007.

[3] C. Dubuc, D. Boudreau, and F. Patenaude, "The design and simulated performance of a mobile video telephony application for satellite third-generation wireless systems," *IEEE Trans. Multimedia*, vol. 3, no. 4, pp. 424–431, Dec. 2001.

[4] M. Bystrom and J. W. Modestino, "Combined source-channel coding schemes for video transmission over an additive white Gaussian noise channel," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 880–890, Jun. 2000.

[5] F. Zhai, Y. Eisenberg, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Rate-distortion optimized hybrid error control for real-time packetized video transmission," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 40–53, Jan. 2006.

[6] K. Stuhlmuller, 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, Jun. 2000.

[7] Q. Qu, Y. Pei, and J. W. Modestino, "An adaptive motion-based unequal error protection approach for real-time video transport over wireless IP networks," *IEEE Trans. Multimedia*, vol. 8, no. 5, pp. 1033–1044, Oct. 2006.

[8] G. Baruffa, P. Micanti, and F. Frescura, "Error protection and interleaving for wireless transmission of JPEG 2000 images and video," *IEEE Trans. Image Process.*, vol. 18, no. 2, pp. 346–356, Feb. 2009.

[9] A. E. Mohr, E. A. Riskin, and R. E. Ladner, "Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 819–828, Jun. 2000.

[10] Y.-Z. Huang and J. G. Apostolopoulos, "A joint packet selection/omission and FEC system for streaming video," in *IEEE ICASSP*, Apr. 2007, pp. I-845–I-848.

[11] T.-L. Lin and P. Cosman, "Optimal RCPC channel rate allocation in AWGN channel for perceptual video quality using integer programming," in *IEEE Int. Workshop on Quality of Multimedia Experience*, 2009.

[12] T.-L. Lin and P. Cosman, "Perceptual video quality optimization in AWGN channel using low complexity channel code rate allocation," in *Asilomar Conference on Signals, Systems and Computers*, 2009.

[13] D. Li and X. Sun, *Nonlinear Integer Programming*, ser. International Series in Operations Research and Management Science. New York: Springer, 2006.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.