OPTIMAL RCPC CHANNEL RATE ALLOCATION IN AWGN CHANNEL FOR PERCEPTUAL VIDEO QUALITY USING INTEGER PROGRAMMING

Ting-Lan Lin and Pamela C. Cosman

University of California, San Diego – ECE

ABSTRACT

For videos transmitted in an error-prone network, it is necessary to protect the source bitstream. Based on our packet loss visibility model, we minimize the end-to-end video quality degradation when transmitted in an AWGN channel using Rate-Compatible Punctured Convolutional codes for a given channel rate budget. We transform the original problem into a binary-decision problem, then we solve this integer program by the Branch and Bound method. Experimental results show that our visibility-based unequal error protection can improve the received video quality compared to equal error protection.

Index Terms— Unequal error protection, packet loss visibility model, perceptual quality, AWGN channel, RCPC codes.

1. INTRODUCTION

For predictively compressed video, losses of different packets induce different levels of distortion. Thus, it is reasonable to have unequal levels of error protection for video packets when transmitted in a lossy communication channel. In [1], channel protection bits are assigned unequally among frames in a GOP using the Genetic Algorithm. FMO (Flexible Macroblock Ordering) in H.264 is employed in [2] to group macroblocks of similar estimated distortion in a frame into a slice, with different levels of channel protection over slices. This method is extended in [3] with Converged Motion Estimation, which performs motion estimation for the current frame using mostly the highly-protected MBs in the previous frame as reference. In [4], frames closer to the end of a GOP have less error protection. Intra-updating is used in [5] to aid forward error coding to handle bursty packet losses in a 3G wireless environment. Unequal error protection for joint source and channel coding is considered in [6] using a source and channel distortion function to find the best source and channel rate allocation, where the channel codes used are Rate-Compatible Punctured Convolutional (RCPC) codes. Traditionally, video quality degradation is measured with MSE (mean-squared error). However, MSE is poorly correlated with human perception [7, 8]. Therefore, to improve the performance of channel rate allocation in terms of human visual perception, a metric developed/verified by subjective experiments, instead of MSE, should be used to evaluate quality.

In our past work, we built a Generalized Linear Model from subjective experiments to predict the visibility when a certain packet is lost, based on information extracted from the video encoding process. Packet loss visibility models for MPEG-2 [9] and for H.264 [10] were built from various subjective experiments individually. In [11], we built a more general model analyzing the data from subjective experiments with different codecs, encoding parameters (GOP structure, encoding rate etc.) and decoder error concealment strategies. Also, we included scene-related and reference-frame-related factors, shown to be significant to the visibility. We improved this model in [12], and showed its usefulness for packet prioritization.

In this paper, we propose an unequal error protection (UEP) scheme based on the visibility of each packet to minimize the visual quality degradation when transmitted over an AWGN channel, given a channel rate budget. We use RCPC codes to flexibly change code rates across packets. Integer programming is used to solve this optimization problem. We compare our method with equal error protection (EEP) using Video Quality Metric (VQM) [13, 14], a full reference (FR) metric developed by the National Telecommunication and Information Administration, shown to be well correlated with human perception compared to other FR video quality metrics [15].

The organization of this paper is as follows. In Section 2, we briefly describe the factors of our visibility model. Section 3 formulates the RCPC rate allocation problem as an integer programming problem. In Section 4, the branch and bound method is introduced to solve the optimization problem. Experimental results are shown in section 5. We discuss possible further improvements to this work and conclude in section 6.

2. PACKET LOSS VISIBILITY MODEL

The definition of packet loss visibility is the probability that the end user will observe the packet loss artifact if the packet is lost. In this paper, we use the packet loss visibility model

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

from [12]. This model was built from three subjective experiments with various codecs, encoding rates, GOP structures and error concealment methods. Here we focus on the description of the factors used in the model; this description is taken from [12].

We define the original video frame at time t as $f(t)$, the compressed video frame as $\hat{f}(t)$, and the decoded video frame as $\tilde{f}(t)$ (with possible packet losses). The error is $e(t) = \hat{f}(t) - \tilde{f}(t)$. The factors were classified by accessibility of the original video as a reference when measured. The RR (Reduced-Reference) measurements for a packet can be obtained when a video encoder or video server reliably provides per-MB information based on $e(t)$, $\hat{f}(t)$ and $\tilde{f}(t)$, assuming knowledge of the decoder concealment strategy. It was found that **IMSE**, **ISSIM** (the average MSE and SSIM [16] among all MBs in an initial packet loss) and **MaxIMSE** (the maximum per-MB MSE over all MBs in the initial packet loss) are significant to the packet loss visibility. To measure the motion information (x,y) per MB that is independent of any codec, a forward motion estimation using 16x16 motion blocks from the *uncompressed* signal $f(t)$ was used. (MOTX,MOTY) is the average motion vector, and **ResidEng** is the average residual energy after motion compensation, over MBs in a packet. The boolean variable **HighMOT** is TRUE if $MOTM = \sqrt{MOTX^2 + MOTY^2} > \sqrt{2}$.

Reference-Scene-related factors are shown to be important by exploratory data analysis (EDA) in [11]. A method for detection of quick scene cuts from $\hat{f}(t)$ was presented in [17]. Each packet loss was labeled by the distance in time between the first frame affected by the packet loss and the nearest scene cut, either before or after. This is DistFromSceneCut, and is positive if the packet loss happens after the closest scene cut in display order, and negative otherwise. DistToRef per MB describes the distance between the current frame (with the packet loss) and the reference frame used for concealment. This variable is positive if the frame with the packet loss uses a previous (in display order) frame as reference, and negative otherwise. **FarConceal** is TRUE if MaxDistToRef (maximum of |DistToRef| in a slice) ≥ 3. In this inequality, MaxDistToRef has units of frames. The Boolean variable, **OtherSceneConceal**, is TRUE if |DistFromSceneCut| < |MaxDistToRef|, where the compared variables must be of the same sign (same direction). In this inequality, the compared variables have units of seconds. If the compared variables have different signs, OtherSceneConceal is FALSE. OtherSceneConceal describes whether the packet loss will be concealed by an out-of-scene reference frame which will increase the visibility of packet loss. To account for the depressed visibility immediately *before* a scene cut, **BeforeSceneCut** is TRUE if $-0.4sec < \mathrm{DistFromSceneCut} < 0sec$. Scenes are classified based on four camera-motion types: still, panning, zooming, or complex camera motions. Since significantly fewer viewers see packet loss in still scenes than in panning or zooming scenes [11], the Boolean variable **NotStill** is TRUE if motion type is not still.

Factors	Coeff. for Final Model
Intercept	4.18061
$\log(1 - \mathrm{ISSIM} + 10^{-7})$	0.22871
SXTNT2	-0.41208
SXTNTFrame	-1.47672
Error1Frame	-0.33009
$\log(\mathrm{MaxIMSE} + 10^{-7})$	0.27578
$\log(\mathrm{ResidEng} + 10^{-7})$	-0.61219
HighMOT	0.18290
NotStill	0.73364
BeforeSceneCut	-1.14434
OtherSceneConceal	2.08966
$\log(\mathrm{IMSE} + 10^{-7})$	0.30492
$\log(\mathrm{IMSE} + 10^{-7})$: FarConceal	0.25720

Table 1. Factors of the final model. Note that the colon (:) means "interact with"

NR (No-Reference) factors can be measured from the lossy pixels only (NR-P), lossy bitstream only (NR-B), or both bitstream and pixels (NR-BP). Factors found by these methods describe *exactly* the spatial extent, pattern, location, and temporal duration of the loss. Variants of these factors that are significant to the visibility were defined as follows: **SXTNT2** is TRUE when two consecutive slices are lost, and **SXTNTFrame** means all slices in the frame are lost. **Error1Frame** is TRUE if the packet loss lasts only one frame. Here we do not consider the number of packets referring to the lost packet, or the number of frames for which the error propagates. As in our past work, we adopt the NR-B measurements to directly obtain the above factors. The factors and coefficients of our final model are summarized in Table 1, taken from [12].

3. RCPC RATE ALLOCATION FOR EXPECTED PACKET LOSS VISIBILITY

Using the model described in Section 2, we can compute for each packet the loss visibility. The visibility scores can be regarded as the visual importance of each packet. If we assign a lower channel code rate to the packet, the probability that the end users observe the packet loss from this packet will be lower. However, the lower code rate requires more FEC (Forward Error Correction) bits. Thus it is important to find out the best code rates to be allocated to each packet given a total bandwidth and other characteristics of the packets.

Assume we have N packets, with size S_i and packet loss visibility V_i, $i = 1, 2, ..., N$. We seek the optimal RCPC rate selection r_i for the ith packet from the RCPC candidate set $\{R_1, R_2, ..., R_K\}$, so as to minimize the end-to-end expected packet loss visibility, while the outgoing total rate budget is constrained to be B. The problem can be formulated as:

$$\begin{aligned} \min_{r_i} \quad & \frac{1}{N}\sum_{i=1}^{N} V_i \times PacketErrorRate(SNR, S_i, r_i) \\ & s.t. \ \sum_{i=1}^{N} S_i \times \frac{1}{r_i} \le B \\ & r_i \in \{R_1, R_2, ..., R_K\} \\ & i = 1, 2, ..., N \end{aligned} \tag{1}$$

Here we define a packet to be in error (undecodable by the source decoder) when any of the bits in the packet is incorrect. Therefore,

$$\begin{aligned} & PacketErrorRate(SNR, S_i, r_i) \\ & = 1 - (1 - BER(SNR, r_i))^{S_i} \end{aligned} \tag{2}$$

Also, from [18] and the simulation section, the logarithm of BER (bit error rate after channel decoding) is linearly related to the inverse of the RCPC rate for a given SNR in an AWGN channel. Thus we have

$$\begin{aligned} & BER = 10^{a(SNR)\frac{1}{r} + b(SNR)} \\ & r \in \{R_1, R_2, ..., R_K\} \end{aligned} \tag{3}$$

where the regression coefficients a and b are a function of SNR. From here we fix SNR for discussion simplicity. Substituting the expression of packet error rate and BER into problem (1), we have the following:

$$\begin{aligned} \min_{r_i} \quad & \frac{1}{N}\sum_{i=1}^{N} V_i \times \{1 - (1 - 10^{a\frac{1}{r_i} + b})^{S_i}\} \\ & s.t. \ \sum_{i=1}^{N} S_i \times \frac{1}{r_i} \le B \\ & r_i \in \{R_1, R_2, ..., R_K\} \\ & i = 1, 2, ..., N \end{aligned} \tag{4}$$

This problem can be recognized as a nonlinear integer programming problem, and can be solved by a well-developed method called branch-and-bound (BnB). To use BnB to solve an integer optimization problem, it is preferable to transform the original discrete optimization variables into binary boolean variables [19]. For our problem, the optimization variables and the set $r_i \in \{R_1, R_2, ..., R_K\}$ are transformed into $x_{ij} \in \{0, 1\}$ defined as

$$x_{ij} = \begin{cases} 1 & \text{if packet } i \text{ uses rate } j \\ 0 & \text{otherwise} \end{cases}$$

Since one packet can only use one rate, we have the following linear equality constraint:

$$\sum_{j=1}^{K} x_{ij} = 1, \forall i = 1, 2, ..., N$$

Therefore, our problem in (4) can be written as:

$$\begin{aligned} \min_{x_{ij}} \quad & \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{K} x_{ij} \times V_i \times \{1 - (1 - 10^{a\frac{1}{R_j} + b})^{S_i}\} \\ & s.t. \ \sum_{i=1}^{N}\sum_{j=1}^{K} x_{ij} \times S_i \times \frac{1}{R_j} \le B \\ & \sum_{j=1}^{K} x_{ij} = 1, \forall i = 1, 2, ..., N \\ & x_{ij} \in \{0, 1\} \\ & i = 1, 2, ..., N, j = 1, 2, ..., K \end{aligned} \tag{5}$$

Now we can use BnB, which is described later, to solve the rate allocation problem.

4. BRANCH AND BOUND METHOD

The Branch-and-Bound (BnB) method implicitly enumerates all the possible combinations of the optimization variables, and in principle produces an exactly optimal solution [19]. BnB partitions the original problem into smaller subsets by tree-growing, and eliminates further consideration of the feasible solutions that can not be better than the current one. BnB solves the integer programming problem as follows :

1. The original problem is solved with the integer constraint being relaxed, i.e., to allow $0 \le x_{ij} \le 1$. The lower bound (LOWER) of the optimized value to the original problem is the optimized value of the relaxed problem since it has a larger feasible set. The upper bound (UPPER) of the optimized value to the original problem can be obtained by substituting the rounded solutions (to zero or one so that they are feasible) into the problem. This is an upper bound of the optimal value to the original problem since it is the best feasible value we can find in this stage. The optimal value to the original problem must be less than or equal to UPPER, and greater than or equal to LOWER.

2. For non-zero-or-one entries (which are infeasible) in the solution from the previous stage, the algorithm grows binary subtrees that fix one entry to zero or one, then solves the problem while relaxing other entries. The optimized value is the lower bound to this subproblem. If this lower bound is greater than UPPER, we prune this branch, since the solutions to any feasible combination of the trees growing from this point are not going to be better (less) than UPPER we currently have. Otherwise, we keep the node for further growing. The upper bound of this subproblem can again be yielded by substituting the rounded solution to the problem. If the upper bound to this subproblem is less than UPPER, we update the UPPER, and mark this rounded solution (feasible) as the best solution so far.

3. Step 2 will be repeated until there is no node to be grown.

Details of the Branch-and-bound algorithm can be found in [19].

5. EXPERIMENTAL RESULTS

In this section, we demonstrate the performance of our end-to-end visibility minimization method compared to EEP. The rate allocation problem is preferably performed across all the packets in a GOP using the rate budget available in a GOP, so that the number of variables and budget are larger to improve performance. However, the computational time of BnB grows nonlinearly with the number of variables (N). Therefore we should find a strategy to balance between results and computational efficiency.

The video sequence used in our experiment is encoded by H.264/AVC JM Version 12.1 in SIF resolution (352×240) with GOP structure IPPP, frame rate 30 fps, and encoding rate 600 $kbps$. We define a packet (a NAL unit) as a horizontal row of macroblocks. Therefore, there are 15 packets in a frame. If the optimization is done over one frame, the algorithm is not able to exploit the relative differences in visibility that occur in different frames and redistribute FEC bits efficiently. For example, if we optimize the channel rate allocation over one I frame where the loss visibility of most packets would be high, we are not making use of the less necessary FEC bits used in B frames where there are more packets with lower visibility. However, if the optimization is to be done over a GOP (30 frames), there will be $N = 450$ variables, which is too large for the BnB method.

In this experiment, we consider all packets in a GOP, but we partition the packets into groups. We first sort the packets by their packet loss visibility in ascending order. Each group of packets ($N = 15$) to be optimized by BnB includes 13 packets of low visibility from the head of the sorted packets, and 2 packets of high visibility from the tail of the sorted packets. For example, the first group of packets includes packet number [1:13, 449:450] from the sorted packets, and second [14:26, 447:448], etc, where the Matlab notation is used. In this way, we attempt to distribute FEC bits from more packets of low visibility, to few packets of high visibility.

The convolutional coder to produce the mother code of the RCPC code has rate $\frac{1}{L}$ where $L = 4$, with memory $M = 4$. The puncturing period of the RCPC code is $P = 8$. The channel we simulate for the wireless communication is AWGN. As mentioned previously, the logarithm of the bit error rate at a given SNR can be linearly related to the inverse of the RCPC rates, as shown in Fig 1. The coefficients of models for different SNRs are shown in Table 2. In this simulation, the RCPC rate used by EEP is $\frac{8}{14}$, and the RCPC rates from which our method can select are $\{\frac{8}{12}, \frac{8}{14}, \frac{8}{16}, \frac{8}{18}\}$. The budget for the optimization problem will be the number of bits used by the EEP in the optimization group. The simulated AWGN channel SNR ranges from -2 to 2 dB, corresponding to the channel bit error rate from about 10^{-1} to 10^{-4} when the RCPC rate of EEP ($\frac{8}{14}$) is considered. Instead of using mean square error, the resulting source-decoded videos are evaluated by a full-reference metric VQM (Video Quality Metric) [13] which is much closer to human perception. VQM ranges from 0 (excellent quality) to 1 (poor quality).

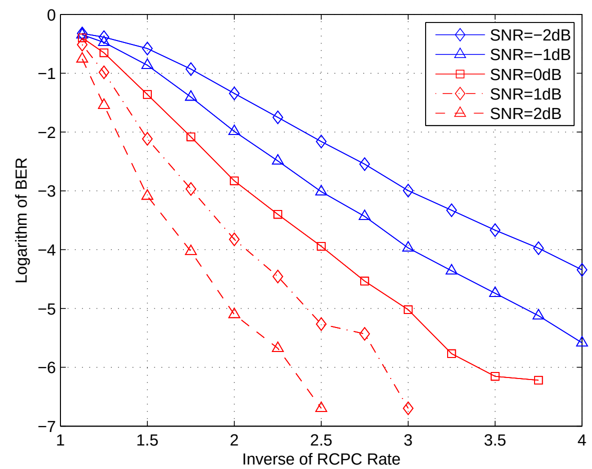

Fig. 1. Linearity of the logarithm of BER to the inverse of RCPC rates for different SNRs of an AWGN channel.

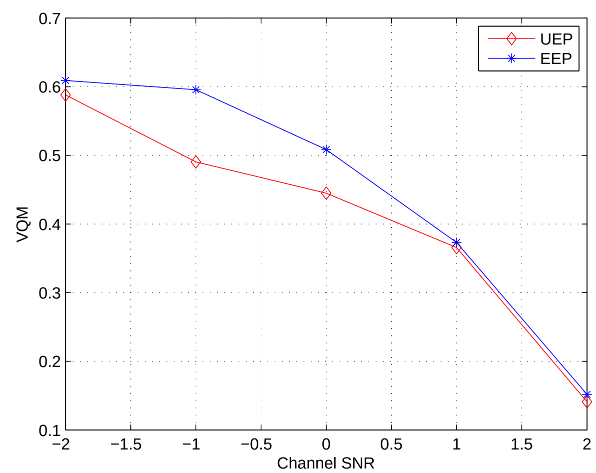

Fig. 2. The average VQM comparison between EEP and UEP over 100 realizations of each AWGN channel.

Figure 2 shows the VQM comparison result between EEP and our proposed UEP. Each data point is obtained by averaging the performance of the corresponding algorithm from 100 realizations of the AWGN channel. We observe that our proposed UEP method consistently performs better (lower VQM scores) than EEP for all SNRs. The largest VQM difference

SNR	-2dB	-1dB	0dB	1dB	2dB
$a(SNR)$	-1.59	-2.15	-2.59	-3.11	-3.43
$b(SNR)$	1.82	2.35	2.46	2.5	2.01

Table 2. The coefficients of the linear model for BER in Equation (3) for different AWGN SNR.

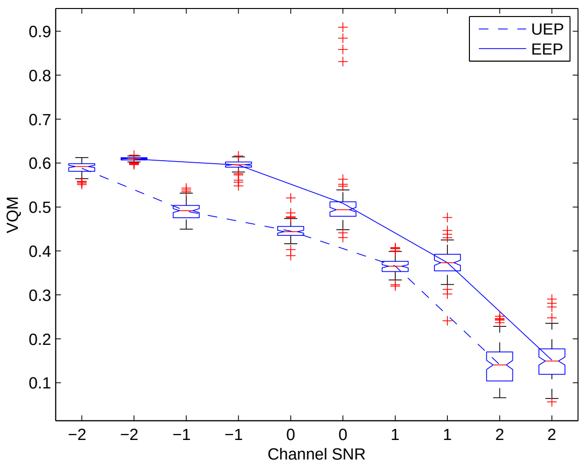

Fig. 3. The boxplot VQM comparison between EEP and UEP for 100 realizations of each AWGN channel. The x-axis shows each SNR twice, each for a specified method.

between the EEP and the proposed UEP is about 0.1, which is considered to be a significant difference in VQM score (see, e.g., [20]).

Besides the mean comparison, we can have more insight to the performance of the two methods by investigating the distributions of their 100 performance realizations. Figure 3 demonstrates the boxplot distributions for each particular method versus SNR. Each box indicates 1st quartile, median and 3rd quartile, and the cross symbols stand for outliers to the distribution. We notice that most of the time the distribution of VQMs produced by EEP is wider than the ones by UEP. In particular when SNR=0dB, the mean of the VQMs for EEP is 0.5, but the VQM in this distribution can be as bad as 0.9, indicating that the performance of EEP is less stable than UEP in most of the cases.

6. DISCUSSION AND CONCLUSIONS

The results by our algorithm presented in previous section are promising, but would improve if we use our algorithm for all packets in a GOP without further partition. However, due to the complexity of the branch and bound method, we choose to optimize over 15 packets at a time; this limits the ability of the algorithm to regulate a larger budget for a larger number of packets. One solution to the complexity problem is to accelerate the branch and bound method by adding more constraints in the optimization problem [19]; this will reduce the variable search range. Of course, the additional constraints should make sense for the desired solution. For example, if the packet sizes are all the same, we can set a linear constraint stating that if packet A has higher loss visibility than packet B, the channel rate assigned to packet A should never be higher than the one to packet B. However this is not a valid constraint for our case since the packets in our problem do not have the same size; larger size can lead to higher packet error rate. Therefore one of our future work is to increase the number optimization variables and find additional valid constraints to reduce the search time of the BnB algorithm.

Compared to EEP where there is only one channel rate, we need to signal which channel code rate is used for each packet; this introduces an overhead to each packet. In our simulation, we have 4 RCPC rates for each packet to select from. Therefore, for each packet, we need an additional 2 bits to tell the channel decoder which RCPC rate is used for this packet. However, this 2-bit overhead is negligible compared to the overall bits for each packet. In our simulation, the mean packet size is 2294 bits (including both source and channel bits). Compared to this, we ignore the 2-bit signalling overhead.

In conclusion, we use the branch and bound method to solve the channel rate RCPC allocation problem to reduce the end-to-end packet loss visibility over an AWGN channel. The result shows that our method consistently achieves better end-to-end video quality in different channel SNRs than Equal Error Protection.

7. REFERENCES

[1] T. Fang and L.-P. Chau, "A novel unequal error protection approach for error resilient video transmission," *IEEE International Symposium on Circuits and Systems*, vol. 4, pp. 4022– 4025, May 2005.

[2] N. Thomos et al., "Robust Transmission of H.264/AVC Video using Adaptive Slice Grouping and Unequal Error Protection," *IEEE International Conference on Multimedia and Expo*, pp. 593–596, July 2006.

[3] J.Y. Shih and W.J. Tsai, "A new unequal error protection scheme based on FMO," *ICIP*, pp. 3068–3071, 2008.

[4] F. Marx and J. Farah, "A novel approach to achieve unequal error protection for video transmission over 3G wireless networks," *Signal Processing: Image Communication*, vol. 19, no. 4, pp. 313 – 323, 2004.

[5] Q. Qu, Y. Pei, and J. W. Modestino, "Robust H.264 video coding and transmission over bursty packet-loss

wireless networks," *IEEE Vehicular Technology Conference*, vol. 5, pp. 3395–3399, October 2003.

[6] C.-L. Huang and S. Liang, "Unequal error protection for MPEG-2 video transmission over wireless channels," *Signal Processing: Image Communication*, vol. 19, pp. 67–79, Jan. 2004.

[7] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Transactions on Communications*, vol. 43, no. 12, pp. 2959–2965, 1995.

[8] B. Girod, *What's wrong with mean-squared error?*, MIT Press, Cambridge, MA, USA, 1993.

[9] S. Kanumuri, P. C. Cosman, A. R. Reibman, and V. Vaishampayan, "Modeling Packet-Loss Visibility in MPEG-2 Video," *IEEE Trans. Multimedia*, vol. 8, pp. 341–355, Apr 2006.

[10] S. Kanumuri et al., "Packet-loss visibility in H.264 videos using a reduced reference method," *IEEE ICIP*, Oct 2006.

[11] A. R. Reibman et al., "Predicting packet-loss visibility using scene characteristics," *Packet Video*, pp. 308–317, Sept 2007.

[12] T.-L. Lin, Y. Zhi, S. Kanumuri, P. Cosman, and A. Reibman, "Perceptual quality based packet dropping for generalized video GOP structures," *ICASSP*, 2009, accepted.

[13] VQM, "http://www.its.bldrdoc.gov/n3/video/," .

[14] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Transactions on broadcasting*, vol. 50, no. 3, pp. 312–322, September 2004.

[15] M. H. Loke, E. P. Ong, W. Lin, Z. Lu, and S. Yao, "Comparison of video quality metrics on multimedia videos," *IEEE ICIP*, October 2006.

[16] Z. Wang et al., "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Im. Proc*, vol. 13, Apr 2004.

[17] A. Hanjalic, "Content-based analysis of digital video," *Kluwer Academic Publishers, Boston*, 2004.

[18] B. Wimmer and T. Stockhammer, "Unequal error protection for H.263 video using RCPC codes," *ITU - Telecommunications Standardization Sector*, vol. Q11-G-12, July 1998.

[19] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 2nd edition.

[20] S. L. P. Yasakethu, W. A. C. Fernando, S. Adedoyin, and A. Kondoz, "A rate control technique for off line H.264/AVC video coding using subjective quality of video," *Consumer Electronics, IEEE Transactions on*, vol. 54, pp. 1465–1472, Aug. 2008.