An Adaptive Rate Allocation to Source-Channel Coding for Internet Video

Jae Cheol Kwon[1], Jae-kyoon Kim[2]

[1]Home Networking Research Team, Network Laboratory, Korea Telecom, Yusong-gu, Taejon, 305-811, Korea
[2]Department of Electrical Engineering, KAIST, 373-1, Kusong-dong, Yusong-gu, Taejon, 305-701, Korea
jckwon@kt.co.kr, kimjk@ee.kaist.ac.kr

***Abstract* - A practical method of adaptive rate allocation to source and channel codings for an independent loss channel is proposed for Internet video. It is based on the observations that the values of residual loss probabilities at the optimal code rates for different packet loss probabilities are closely clustered to the average residual loss probability for a transmission frame size *n* in RS(*n,k*) code and for a total bit rate *R*. These observations are then exploited to find the code rate for maximum PSNR. Simulation results demonstrate that the proposed method achieves a near-optimal bit-rate allocation in the joint source-channel coding of H.263 and RS(*n,k*) codings.**

I. INTRODUCTION

When transporting video over the Internet, packets can be dropped or experience excessive delay according to the network status. Video packets that arrive too late at the receiver are of no use and considered lost. Consequently, to enhance the quality of a decoded picture against packet loss, several error control schemes have been proposed [1][2]. Among them, the FEC(forward error correction) technique is useful for real-time applications because it requires less time to recover lost data and does not require an additional backward channel. Since the packet loss probability varies with time in time-varying channels such as Internet, FEC scheme with a fixed code rate either wastes the available channel bandwidth during a period of low loss probability, or is insufficient to completely recover original information during a period of high loss probability. Therefore, allocating the optimal rate to source and channel codings while minimizing end-to-end overall distortion is a key issue of JSCC(joint source-channel coding).

Various previous studies have been focused on finding this solution. Stuhlmuller, *et al.* [3] proposed a model for estimating total distortion of a video transmission system to analyze the performance as a function of the most important system parameters. But their method is not packet-based and needs intensive calculation to identify the system parameters. Zhang, *et al.* [4] proposed an unequal interleaving and fixed length packetization, along with an R-D function-based optimal bit allocation method to minimize the end-to-end distortion. Gallant, *et al.* [5] proposed an algorithm that combines layered coding with transport prioritization in an operational R-D framework that trades off source coding performance for channel coding protection. Kuceren, *et al.* [6] proposed a system employing the combination of passive error recovery together with active forward error correction coding to simultaneously maximize the statistical multiplexing gain while minimizing the end-to-end reconstructed video distortion.

Almost all these methods were focused on solving the optimization problem to identify the optimal bit-rate allocation between the source and channel codings. However, they also involve very intensive computation to obtain the operational R-D functions for all possible packet loss probabilities and all possible source-channel coding rate combinations prior to the actual encoding. As such, none of the above methods are easily applicable in a real situation. Accordingly, we present a simple and practical method to find the near-optimal coding rate allocation to the source and channel codings based on a bounded residual packet loss probability for time-varying packet loss channels.

The remainder of the paper is organized as follows. Section II describes the video transmission system and channel models under consideration. Section III and IV present the observation of joint PSNR *vs.* code rate and the proposed adaptive coding rate allocation method, respectively. The simulation results are discussed in Section V, and the conclusion is given in Section VI.

II. VIDEO TRANSMISSION SYSTEM

Fig. 1 shows a brief sketch of the video transmission system for joint source-channel coding under consideration. It is assumed that the network status information such as the packet loss probability and average burst length (average number of consecutively lost packets) can be made available or estimated by the transmitter side from feedback information, such as SR(sender report)/RR(receiver report) packets of RTCP protocol in Internet. The system consists of a video encoder and decoder, channel encoder and decoder, and the functions related to the estimation of channel characteristics.

Let $D(t)$ be the end-to-end overall distortion of a video frame at time t. Assuming that $D_s(R_s(t))$ and $D_c(R_c(t))$ are uncorrelated, it can be represented by (1).

This work is partially supported by the Korea Science and Engineering Foundation (KOSEF) under grant 98-0101-02-01-3.

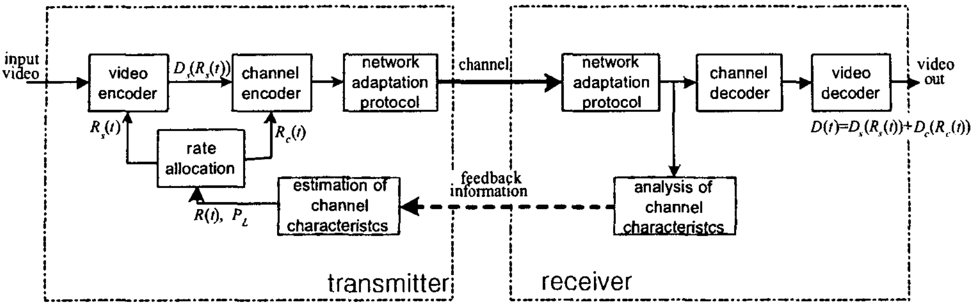


Fig. 1. *A video transmission system for joint source-channel coding*

$$D(t) = D_s(R_s(t)) + D_c(R_c(t)). \tag{1}$$

where $D_s(R_s(t))$ is the source coding distortion for the source rate $R_s(t)$, and $D_c(R_c(t))$ is the distortion introduced by channel errors at the additional channel rate $R_c(t)$ for error correction. $D_c(R_c(t))$ includes the distortion from residual packet error not corrected by an FEC scheme, plus error propagation to successive video frames in both the spatial and temporal directions. The statistical characteristic of $D_s(R_s(t))$ is more or less deterministic, which can be characterized by a quantization process, while that of $D_c(R_c(t))$ is probabilistic.

After source coding, the video bit-stream is packetized into packets of the fixed length of L_p bytes for transmission by a channel encoder. We assume that the channel encoder generates n-k parity packets for every k video packets by an RS (Reed-Solomon) code RS(n,k) with the code rate $r=k/n$, where n refers to the size of a transmission frame. For an RS(n,k) code with the information of packet sequence numbers available, the erroneous packets up to n-k packets in a transmission frame can be corrected.

If more than n-k packets are in error, the whole frame cannot be corrected. But the video packets received correctly among the erroneous frame can be still used for video reconstruction if its position information is available. This can be made, for example, by starting each packet with a picture or slice header information.

For an independent loss channel with packet loss probability P_L, and frame size n, the video packet loss probability $P_E(n,e|P_L)$ for e lost packets out of n transmitted packets is expressed by (2). Therefore, the residual video packet loss probability $P_R(r, P_L)$ can be described as in (3).

$$\begin{aligned} P_E(n,e|P_L) &= \sum_{m=e}^{\min(e+n-k,n)} \binom{k}{m-(n-k)} P_L^{m-(n-k)}(1-P_L)^{n-m} \binom{n-k}{e+(n-k)-m} P_L^{e+(n-k)-m}(1-P_L)^{m-e}\,\frac{m}{n} \\ &= \sum_{m=e}^{\min(e+n-k,n)} \frac{m}{n}\binom{k}{m-(n-k)}\binom{n-k}{e+(n-k)-m} P_L^{e}(1-P_L)^{n-e}. \end{aligned} \tag{2}$$

$$\begin{aligned} P_R(r,P_L) &= P_R(n,k,P_L) \\ &= \sum_{e=n-k+1}^{n} P_E(n,e|P_L), \text{ where } r=k/n. \end{aligned} \tag{3}$$

The average number of consecutively lost packets (L_B) for this model is $1/(1-P_L)$.

Some examples of the residual video packet loss probabilities(P_R) obtained by the above results is shown in Fig. 2, which shows $P_R(r, P_L)$ for several values of n with P_L =1% and 10%. We notice that for the same n, a higher P_L results in a higher P_R. And for the same P_L, a smaller n results in a higher residual loss probability. Since a higher P_R means more distortion, it is expected that a smaller n or higher P_L would bring an inferior average PSNR($=10\log_{10}255^2/D$) performance for the same code rate. If the delay of channel decoding is of no concern, a larger n will be desirable for the same r and P_L.

And we also observe in Fig. 2 that to a value of $\log_{10} P_R$ or to a range of $\log_{10} P_R$ values as indicated by the region of dotted lines, there are several points corresponding to different pairs of (r, P_L). It will be a key observation in finding later on the optimal code rate r_o for a changed P_L after an initial code rate r_o for a given P_L.

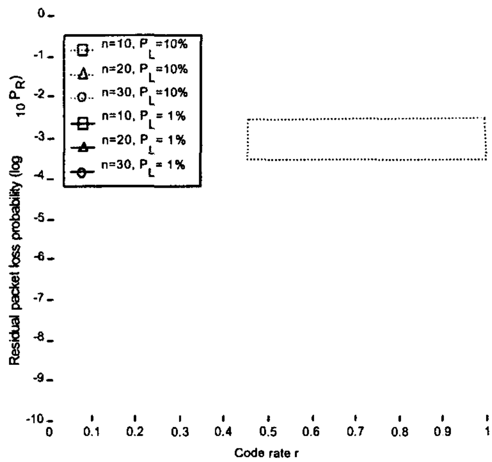


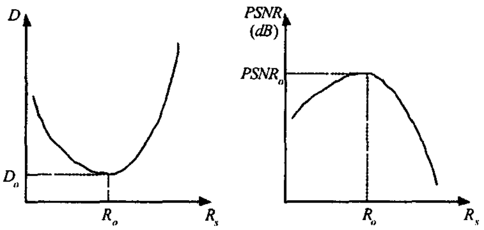
Fig. 2. *Residual packet loss probabilities.*

III. OBSERVATION OF JOINT PSNR VS. CODE RATE

The rate-distortion (R-D) function in a JSCC scheme, including both the source coding distortion and the channel distortion introduced by packet loss, will show the relationship between the end-to-end overall distortion of the video frames and the source coding rates.

When the total bit rate $R=R_s + R_c$ of available channel bandwidth is fixed, the typical shape of a joint distortion and combined PSNR(dB) *vs.* source coding rate for lossy channels will be like Fig. 3. The source coding rate R_o corresponding to the lowest point of the overall distortion D_o is the optimal source coding rate, thereby minimizing the end-to-end overall distortion. The left side of R_o indicates that almost all the lost packets are corrected by an FEC scheme, and hence the overall distortion decreases as the source coding rate increases. Whereas the right side of R_o means that as the source coding rate increases, more video packets are generated, resulting in higher packet losses that are not corrected. Accordingly, the overall distortion increases as the source coding rate increases for a fixed channel bandwidth. The combined PSNR curve is of inverse shape to the joint distortion curve as Fig. 3 (b).

Fig. 4 shows the measured average PSNR for a JSCC of an H.263 source coding [10] with source coding rate $R_s = R\ r$ and of a RS(20, k) channel coding with channel coding rate $R_c = R\ (1-r)$ and the packet size of 300 bytes for the total bit-rate R=1, 2 Mbps.



(a) Joint distortion (b) Joint PSNR

Fig. 3. Typical shape of joint distortion and joint PSNR

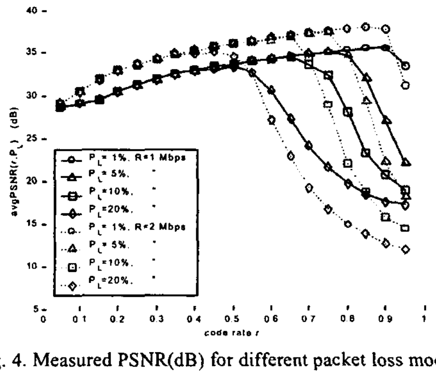


Fig. 4. Measured PSNR(dB) for different packet loss model.

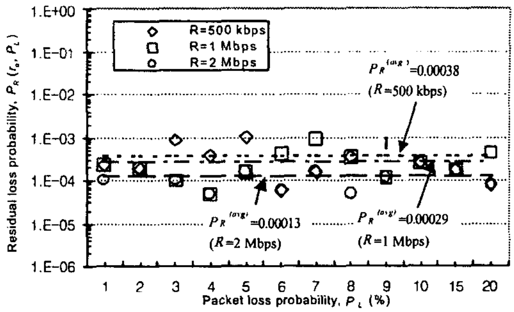


Fig. 5. Bounded residual loss probability at optimal code rate $P_R(r_o, P_L)$

The errorneous image blocks due to residual packet loss after channel decoding are substituted by the average video level of 128, and the measured PSNR(r, P_L) is the average PSNR(r, P_L) for 300 frames of test sequence *Foreman*.

We observe in Fig. 4 that for each curve of given P_L and total bit rate R, there is a point of peak PSNR(r_o, P_L) for an "optimal" code rate r_o. For the code rate beyond r_o, the *avgPSNR* decreases rapidly since the number of lost packets is beyond the correction capability of the FEC coding. We also notice that for a fixed P_L, at the code rates before reaching r_o, the PSNR is higher for higher total bit rate R, but at the code rates beyond r_o, the trend is reversed. Also for a fixed bit rate R, the PSNR decreases with increasing P_L at code rate beyond r_o.

From the above observations, the optimal rate allocation based on the combined PSNR for a given total bit rate R and packet loss probability P_L can be obtained by choosing the code rate r_o corresponding to the highest average PSNR as (4).

$$r_o = \arg\max_r \{ avgPSNR\ (r, P_L, R) |\ P_L, R \} \tag{4}$$

But it is a time consuming task to plot *avgPSNR*(r) curve for a given R and P_L. Yet, as expected from Fig. 2, the values of residual loss probabilities $P_R(r_o, P_L)$ at the optimal code rates for given packet loss probabilities are closely clustered for the given n and total bit rate R as shown in Fig. 5. From Fig. 5, we notice that $P_R(r_o, P_L)$ is concentrated to an average value $P_R^{(avg)}$ that is determined mostly by the video characteristics for given total bit rate R, packet loss model, channel code structure (n,k), but not much by P_L. This is due to the fact that P_R is proportional to both P_L and code rate r_o as shown in Fig. 2. But when P_L is large, r_o is small, and vice versa for the same total bit rate as in Fig. 4.

Hence, $P_R(r_o, P_L)$ stays near to a constant value. This is verified by comparing the (r, P_L) curve for $P_R(r, P_L) \equiv Const = P_R^{(avg)}$ from (3) for an arbitrary large $n(=k/r)$, and the plot of (r_o, P_L) pairs for $\max_r \{avgPSNR(r, P_L)\}$ of Fig. 4 as shown in Fig. 6.

It is interesting to note that the measured (r_o, P_L) pairs are very close to the calculated (r, P_L) curve for a fixed $P_R^{(avg)}$.

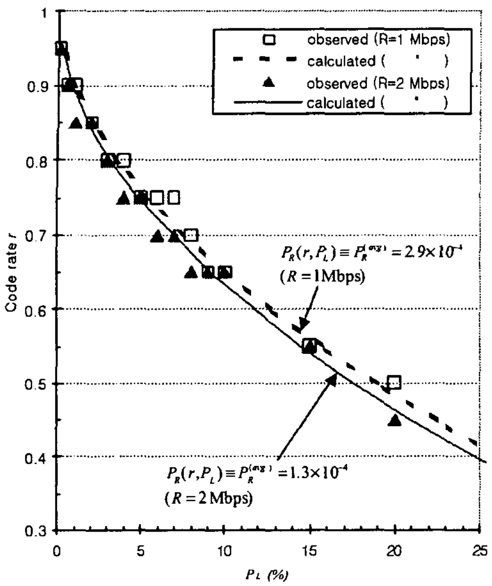


Fig. 6. Comparison of (r,P_L)-curve for $P_R(r,P_L) \equiv P_R^{(avg)}$ with observed (r_o, P_L) pairs for $\max_r\{avgPSNR(r,P_L)\}$.

This leads to the verification that the values of residual loss probabilities at the peak PSNRs are very close to the average value as shown in Fig. 5.

IV. ADAPTIVE CODE RATE DECISION METHOD

We propose a practical method to determine the code rate r for a changed packet loss probability without complete knowledge of the entire PSNR curves or analytic estimation of $D_c(R_c(t))$. For this, we need to exploit the observations described in Section II and III. Fromthe characteristics of combined PSNR function that the values of P_R at the peak PSNR are closely located, finding the approximate average value $P_R^{(avg)}$ of residual loss probability is the key of our approach.

The detailed procedures of adaptive code rate decision for a given total bit rate R and a RS(n,k) channel code is described as follows. However this method requires some pre-calculations to obtain the average residual packet loss probability $P_R^{(avg)}$ for a sequence at the initial stage. The $P_R^{(avg)}$ needs to be updated when there are significant changes in statistical characteristics like scene changes, or in total bit rate.

Initial stage:

Step 1: For a given $P_L^{(1)}$ and for its adjacent arbitrary packet loss probabilities $P_L^{(2)}$ and $P_L^{(3)}$, make test-runs to get $avgPSNR(r,P_L^{(i)})$ curves, and to obtain the matching code rates $r_o^{(i)}$ for maximum PSNR and the average residual packet loss probability $\hat{P}_R^{(avg)}$:

$$r_o^{(i)} = \arg\max_r\{avgPSNR(r,P_L^{(i)})\},\ i = 1,2,3, \tag{5}$$

$$\hat{P}_R^{(avg)} = \frac{1}{3}\sum_{i=1}^{3} P_R(r_o^{(i)},P_L^{(i)}). \tag{6}$$

where $P_R^{(i)}(r_o^{(i)},P_L^{(i)})$ is from (3).

Step 2: The initial code rate for the given $P_L^{(1)}$ is $r_o^{(1)}(0)$, and hence $R_s = r_o^{(1)}\cdot R,\ R_c = R - R_s$.

Update stage:

Step 3: When there is a change in P_L, find the matching code rate r^* for the same $\hat{P}_R^{(avg)}$:

$$r^* = \arg\min_r\left\{P_R(r,P_L) - \hat{P}_R^{(avg)}\ \middle|\ n,P_L\right\}, \tag{7}$$

The proposed method does not require any analytic estimation of channel-induced distortion nor need several system parameters to be updated frequently except $P_R(r,P_L)$.

V. SIMULATION RESULTS

Some simulations are carried out to evaluate the performance of the proposed method using 300 frames from the *Foreman* sequence in a CIF(common intermediate format) format. The assumptions and parameter settings for simulation are as follows.

The total bit rate (available channel bandwidth) is fixed at 1 Mbps. H.263 coding method [7] with the modified TMN-8 rate control algorithm [8] is used to support the time-varying source coding rate. In the RS(n,k) code for channel coding, the transmission frame size is set at $n=20$, and k is varied according to the time-varying loss probability P_L in an integer granularity, *i.e.*, $1 \le k \le 19$. No error concealment algorithm is applied in the video decoder, but a slice structure of three slices/frame and MB(macroblock) based INTRA refreshment at the rate of *1/60* (*frames*$^{-1}$) are used to protect against error propagation due to channel errors. The payload length L_p of a video packet is set at *300 bytes*. No packetization or protocol overhead bits are considered, as they are not significant. It is also assumed that the picture- and slice-header would not experience any loss for comparison at fair conditions.

Fig. 7 compares the optimal code rates obtained for a given P_L over whole sequence by the full search method of (4) and the code rates obtained adaptively by the proposed method. For the proposed method, the average residual packet loss probability $\hat{P}_R^{(avg)}$ for the initial packet loss probabilities P_L=1,2,3% are found to be 1.8×10^{-4}. We note that the code rates obtained by both methods are very similar.

Fig. 9 shows the code rates obtained by the two methods for the time-varying packet loss probability of Fig. 8.

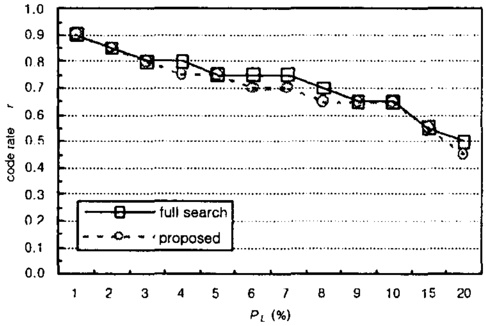


Fig. 7. Optimal code rates r_o *vs.* P_L by full search and by the proposed method.

We notice that the difference in code rate between the two methods is within 0.05 in Fig. 7 and in Fig. 9. Fig. 10 compares the frame-based PSNR for the code rates of Fig. 9 and for P_L of Fig. 8. The results of proposed method is very close to the results of the full search method, confirming the justification of the proposed method. At the fixed code rate of *0.4* and *0.9*, the PSNR performance is very much dependent on the channel loss as expected. We also obtained almost the same results for other test sequence "*Stefan*".

VI. CONCLUSIONS

We have proposed a practical method of adaptive rate allocation to source and channel codings for a given total bit rate and packet loss probability for an independent loss channel. It is based on the observations that the values of residual loss probabilities at the optimal code rates for different packet loss probabilities are closely clustered to the average residual loss probability.

It is shown that the allocated code rates obtained by the proposed method are very close to the optimal code rates by the full search method. Therefore, it can be applied for practical selection of code rate in the joint source-channel coding without complete knowledge on the relationship between joint PSNR and code rate in advance.

We also verified that the proposed method could be applied for a bursty loss channel with modification of (2) only, which is not included here due to the lack of space.

REFERENCES

[1] D. Wu, Y.T. Hou, Y.Q.Zhang, "Transporting Real-time Video over the Internet: Challenges and Approaches", *Proceedings of the IEEE*, vol. 88, No. 12, Dec, 2000.

[2] Y. Wang, Qin-Fan Zhu, "Error Control and Concealment for Video Communication: A review", *Proceedings of the IEEE*, vol. 86, No. 5, May, 1998.

[3] Klaus Stuhlmuller, N. Farber, B. Girod, "Analysis of video transmission over lossy channels", *IEEE Trans. on Jou. of Sel. Areas on Comm.*, vol. 18, No. 6, pp. 1012-1032, June. 2000.

[4] Qian Zhang, Wen Zhu, Ya-Qin Zhang, "Resource allocation for multicast streaming over the Internet", *IEEE Trans. on Multimedia*, vol. 3, No. 3, Sep. 2001.

[5] M. Gallant et al, "Rate-Distortion Optimized Layered coding with Unequal Error Protection for robust Internet Video", *IEEE Trans. on Circuit and Systems for Video Technology*, vol. 11, No. 3, pp. 357, Mar. 2001.

[6] R. Kurceren, J. W. Modestino,"A Joint Source-channel coding approach to Network transport of Digital Video", *in proceeding of IEEE INFOCOM' 2000*, pp. 717. 2000.

[7] ITU-T Recommendation H.263 (02/98)- Video Coding for Low Bit Rate Communications.

[8] ITU-T Video Codec Test Model, Near-Term, version 8 (TMN8), Sep. 1997.

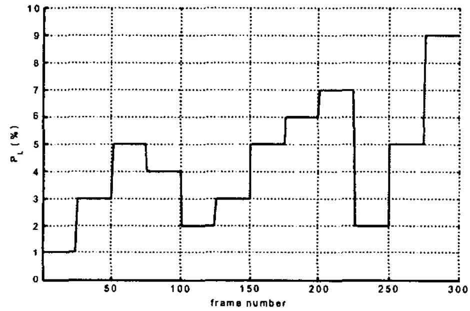


Fig. 8. Time-varying packet loss probabilities for simulations

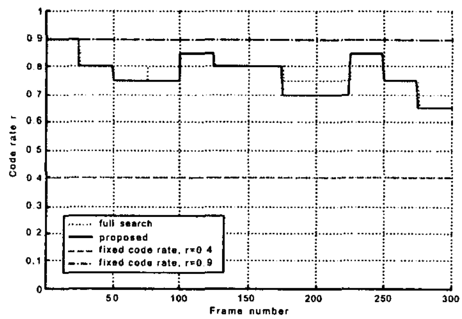


Fig. 9. Allocated code rates for the channel loss of Fig. 8.

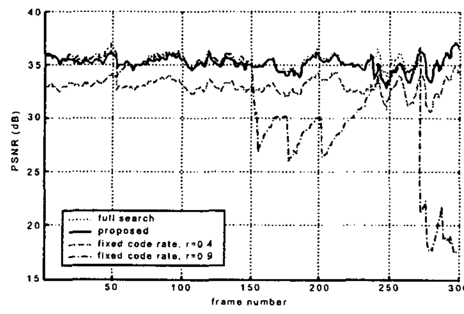


Fig. 10. PSNR performances for the code rate of Fig. 9.