A NEW JOINT SOURCE-CHANNEL MODELING APPROACH FOR ADAPTIVE FEC CODE RATE DECISION

Yo-Won Jeong, Jae Cheol Kwon**, Jae-kyoon Kim*, Kyu-Ho Park**

* Korea Advanced Institute of Science and Technology, Daejon, Republic of Korea
** Advanced Technology Lab., R&D Group, KT, Daejon, Republic of Korea

ABSTRACT

We propose a joint source-channel modeling approach for adaptively determining the quality-optimal FEC code rate. Our objective is to obtain the optimal video quality in the receiver, while taking time-varying packet loss into consideration. First, we define the loss threshold set as the set of packet loss probabilities in which the code rate have to be adjusted to maintain the maximum video quality. Proposed model can estimate the loss threshold set efficiently. Then, whenever there is a change of the packet loss status in the channel, video sender can always find the optimal code rate on-the-fly by using the estimated loss threshold set. Simulation results show that the proposed method can determine the near-optimal code rate in joint source-channel coding.

1. INTRODUCTION

When transporting video over the wireline/wireless channel, packets can be dropped, be corrupted or experience excessive delay by the time-varying channel status. These packets are considered lost, and result in the reconstructed video of unacceptable quality at the receiver. To protect the quality of a decoded picture against packet loss, several error control schemes have been proposed. Among these schemes, the forward error correction (FEC) technique is considered as a useful tool for real-time applications. Since packet loss probability is not fixed in the channel, the FEC scheme with a fixed code rate either wastes the channel bandwidth during the status of low packet loss, or is insufficient to completely recover the original information during the status of high packet loss. Therefore, it becomes a key issue in joint source-channel coding (JSCC) to determine the optimal FEC code rate, or allocate optimal bit rates to source and channel codings for minimizing the end-to-end distortion of the reconstructed video quality.

Several previous studies have focused on this optimal code rate decision. Stuhlmuller *et al.* [1] proposed an analytic model to estimate the end-to-end distortion over error prone channels. Frossard *et al.* [2] also proposed an end-to-end distortion model, in which channel-induced distortion was assumed to be proportional to the number of lost pixels. Even though they showed the accuracy of their model via extensive simulation results, because they considered two models, the source-coding model and the channel-coding model, their modeling equations are complex and have many model parameters. Kwon *et al.* [3] proposed a practical method based on an observation that the residual packet loss probability in the optimal code rate stays near to a constant value. However, actually this residual packet loss probability is not constant any more at the high channel packet loss probabilities.

In this paper, we present a novel joint source-channel modeling approach based on the *loss threshold set* for determining the quality-optimal FEC code rate. Loss threshold set is defined as the set of the packet loss probabilities in which the code rate must be adjusted to maintain the maximum video quality regardless of the packet loss status of the channel. There are two types of loss threshold sets: the *channel loss threshold set* and the *residual loss threshold set*. The proposed model represented as one equation having two model parameters can estimate the loss threshold sets accurately. Then, when there is any change in the channel packet loss status, video sender can find the optimal code rate on-the-fly by using the estimated loss threshold sets.

The remainder of this paper is organized as follows. Section 2 describes the video transmission system and channel models being considered in this paper. Section 3 introduces the two types of loss threshold sets in detail. The proposed joint source-channel model based on these loss threshold sets will be explained in Section 4. The simulation results are discussed in Section 5, and the conclusion is given in Section 6.

2. VIDEO TRANSMISSION SYSTEM AND CHANNEL MODELS

Fig. 1 shows a brief sketch of the video transmission system under consideration. It is assumed that network status information, such as channel packet loss probability (P_L) and total channel rate (R_{TC}), can be estimated at the sender from feedback information of real-time transport protocol (RTCP). This video transmission system consists of a video encoder/decoder, a channel encoder/decoder, and the functional blocks for the estimation of channel characteristics and *rate allocation*. Finding the optimal code rate allocated to the source and channel codings (R_S and R_C), is intimately related to the rate allocation block.

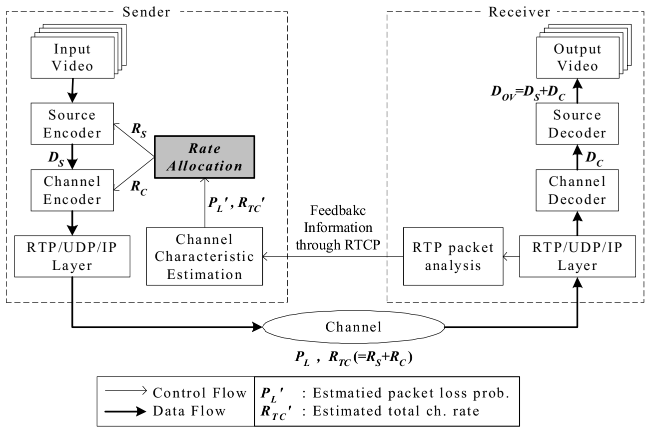

Fig.1. A video transmission system considered in this paper

The overall distortion (D_{OV}) reconstructed video depends upon both the source coding distortion (D_S) and channel-induced distortion (D_C).

In this paper, we make the following assumptions.

- The interval of abrupt changes in a picture content such as scene change is longer than the calculation time of the optimal code rate. Generally, scene changes do not occur frequently in live video communications (e.g., video telephony and distance learning) [4].
- Hybrid motion compensation video coding such as H.261/3/4, MPEG-1/2/4, is used as the source coding.
- The resynchronization-marker (RSM) structure of three markers per frame and macroblock (MB) based INTRA refreshment at the rate of 1/60 (frames^{-1}), are used to prevent spatio-temporal error propagation due to the residual packet loss [4].
- After source coding, the coded bitstream is fragmented into packets with a unit of the RSM, in which each video frame is fragmented into three packets. The underlying protocol overhead (or header) information is not considered.
- A simple error concealment algorithm is used at the decoder side so that the damaged MBs due to the packet loss can be replaced with spatially corresponding MBs in the previous frame.
- The packet loss characteristic of the channel is assumed to be independent and random.
- The first I-frame would not experience any loss for the performance comparison between the optimal and the proposed schemes at fair conditions.

The channel encoder generates n-k parity packets for every k video packets by the Reed-Solomon code or $RS(n,k)$-code with the code rate $r=n/k$. n refers to the number of total packets of each transmission group. We adjust this code rate by making n be fixed and k changed from 1 to n. For an $RS(n,k)$ code with the information of packet sequence numbers available, lost packets up to n-k packets in a transmission group can be completely recovered [4]. If more than n-k packets are lost, this transmission group cannot be recovered in its entirety. We will hereafter refer to the residual packet loss probability after channel decoding as P_R, to distinguish it from P_L, the packet loss probability in the channel.

The residual packet loss probability depends on r and P_L, and can be described as

$$P_R(r,P_L)=\sum_{i=n-k+1}^{n}\binom{n}{k}P_L^i(1-P_L)^{n-i} \quad [1]. \tag{1}$$

We can also express the optimal code rate decision with respect to minimizing the overall distortion for the given n as follows:

$$r_O=\arg\min_r\{D_{OV}(r,P_L)|\, n\} \tag{2}$$

3. LOSS THRESHOLD SETS

It is difficult to adaptively determine the optimal code rate because of the intensive calculation necessary to obtain the operational R-D functions for all possible packet loss probabilities and all possible source-channel coding rate combinations. In this section, we introduce two types of loss threshold sets, the channel loss threshold set and the residual loss threshold set, which is of help to determine the optimal code rate quickly and accurately.

3.1 Concept of channel loss threshold set

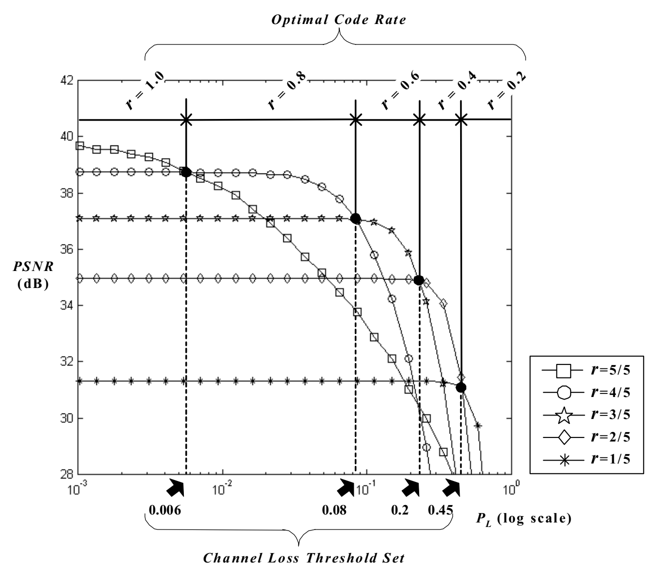

Fig. 2. Relationship between code rate and CLTS. (Test sequence: *Akiyo*)

Fig. 2 shows the curves of measured average PSNR, $PSNR(r,P_L)$ for the test sequence of *Akiyo*. We used an H.263 source encoder [5] with $R_S=R_{TC}r$, and $RS(5,k)$, $k=1,2,\cdots,5$ as a channel code with coding rate of $R_C=R_{TC}(1-$

r) where R_{TC}=100kbps. Since we can adjust r by changing k and fixing n as n_0, r can be redefined as follows.

$$r_k = \frac{k}{n_0}\ ,\ k = 1, 2, \cdots, n_0 \tag{3}$$

In Fig. 2, we can see that though P_L increases, $PSNR(r_k,P_L)$ barely decreases in the ranges of low P_L where almost all lost packets are recovered by the channel decoder. If P_L increases more and more, $PSNR(r_k,P_L)$ start to decrease sharply because the channel decoder cannot recover the lost packets any more and the error propagation in the decoded video occurs. We can observe that for each curve, there is a range of P_L where it has higher PSNR than other curves. For example, $PSNR(r_4,P_L)$ curve (-o- mark) has higher PSNR than other curves in the range of P_L from 0.006 to 0.08. If we adjust the code rate like in Eq. (4) for time-varying P_L, we can always obtain the highest average PNSR for all P_L.

$$\begin{cases} r = r_5,\ P_L \le 0.006 \\ r = r_4,\ 0.006 < P_L \le 0.08 \\ r = r_3,\ 0.08 < P_L \le 0.2 \\ r = r_2,\ 0.2 < P_L \le 0.45 \\ r = r_1,\ 0.45 < P_L \end{cases} \tag{4}$$

Therefore, the values of P_L, which determine the boundaries to change the code rate, are important. We define the set of these values as the *channel loss threshold set (CLTS)*. In the case of Fig. 2, the CLTS is {0.006,0.08, 0.2,0.45}. The number of elements of the CLTS is n_0-1. Since all elements of the CLTS are the crossing points of two curves as shown in Fig. 2, the CLTS is represented as

$$CLTS = \left\{ P_{Lk} \middle| \begin{array}{l} P_L \text{ val. of } PSNR(r_k, P_L) = PSNR(r_{k-1}, P_L), \\ k = 2, 3, \cdots, n_0 \end{array} \right\}. \tag{5}$$

3.2 Concept of residual loss threshold set

Elements in the CLTS have to be lastingly updated according to the channel status during the real-time video communication. However, it is difficult to directly calculate this CLTS because we cannot draw all PSNR curves like in Fig. 2. To get the CLTS quickly and accurately, we have to introduce another loss threshold set, *residual loss threshold set (RLTS)*. We define the RLTS as

$$RLTS = \left\{ P_{Rk} \mid P_{Rk} = P_R(r_k, P_{Lk}),\ k = 2, 3, \cdots, n_0 \right\}, \tag{6}$$

in which each element of RLTS, P_{Rk}, is the value shown in Eq. (1) for the given r_k and P_{Lk}. Note that we can get the CLTS from the RLTS directly if the RLTS is available.

This RLTS can be obtained more easily than the CLTS by using our joint source-channel model explained in the next section.

4. JOINT SOURCE-CHANNEL MODEL FOR RLTS

We know that the optimal code rate for a given P_L can be obtained by the CLTS as in Eq. (4). But, it is very time-consuming task to get CLTS by plotting the PSNR curves like Fig. 2 whenever input video changes. If we define distortion as

$$D(r_k, P_L) = \frac{255^2}{10^\wedge\left(PSNR(r_k, P_L) / 10 \right)}, \tag{7}$$

we can observe that the difference of source distortion between r_k and r_{k-1} is likely equal to the channel distortion at r=r_k and P_L=P_{Lk}. This is due to the fact that $PSNR(r_k,P_L)$ curve abruptly decreases and $PSNR(r_{k-1},P_L)$ curve maintains nearly constant value at their crossing point. We can formulate this observation as follows;

$$D(r_{k-1}, 0) - D(r_k, 0) \approx D_C(r_k, P_{Lk}),\ \ k = 2, 3, \cdots, n_0 \tag{8}$$

$D_C(r_k,P_{Lk})$ means the channel distortion. Fig. 3 shows experimental results for the measured $D_C(r_k,P_{Lk})$ and $D(r_{k-1},0)$-$D(r_k,0)$ for several test sequences. We can see that Eq. (8) is valid for almost all ks.

$D(r_k,0)$ and $D(r_{k-1},0)$ in Eq. (8) can be represented by a source coding model because they are values in the lossless case, and $D_C(r_k,P_{Lk})$ can be represented by a channel distortion model. In this paper, we use an inverse-proportional model as the source coding model [1]. $D(r_k,0)$ can be represented as

$$D(r_k, 0) = \frac{\theta}{R_{TC} r_k - R_0} + D_0\ \ , k = 2, 3, \cdots, n_0, \tag{9}$$

$$\theta, R_0, D_0 : \text{modeling parameters}$$

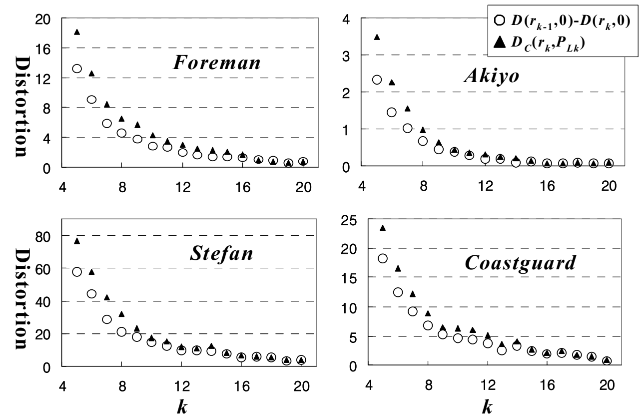

Fig. 3. Experimental results of $D_C(r_k,P_{Lk})$ and $D(r_k,0)$-$D(r_{k-1},0)$ for four test sequences, n_0=20

To represent $D_C(r_k,P_{Lk})$ as a channel coding model, we have simulated the distortion patterns for a wide range of the residual packet loss probability using several test sequences. Fig. 4 shows that the distortion of the received video is in average proportional to the residual packet loss probability regardless of k. The reason is that intuitively, the perceptual distortion is in average proportional to the number of lost MBs [2], and the video packet rate is nearly constant regardless of source coding rate because each video frame is packetized into 3 packets by RSM. Therefore, by using Eq. (1), we can represent $D_C(r_k,P_{Lk})$ as

$$D_C(r_k, P_{Lk}) = D_C(r_k, P_{Rk}) = bP_{Rk}\ , k = 2, 3, \cdots, n_0, \tag{10}$$

$$b : \text{constant}$$

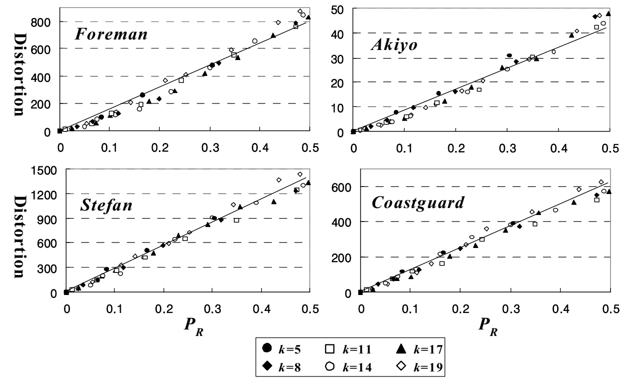

Fig. 4. Video distortion versus residual packet loss probability P_{Rk} for various k where n_0=20

Note that P_{Rk} and P_{Lk} can be one-to-one mapped by Eq. (1). If we substitute Eq. (9) and (10) into Eq. (8), finally, we get P_{Rk} set as

$$P_{Rk} = \frac{\beta}{(k-(\alpha+1))(k-\alpha)}\ ,\ \alpha = \frac{R_0 n_0}{R_{TC}}, \beta = \frac{\theta n_0}{bR_{TC}}, \qquad (11)$$
$$k = 2, 3, \cdots, n_0$$

where α and β are the final model parameters of the joint source-channel model. By Eq. (11), we can get all elements of the RLTS defined as in Eq. (6), and the CLTS can be also obtained.

5. SIMULATION RESULTS

Some simulations are carried out to evaluate the performance of the proposed model for the test sequences, *Coastguard*, *Akiyo*, *Stefan*, and *Foreman* of QCIF format.

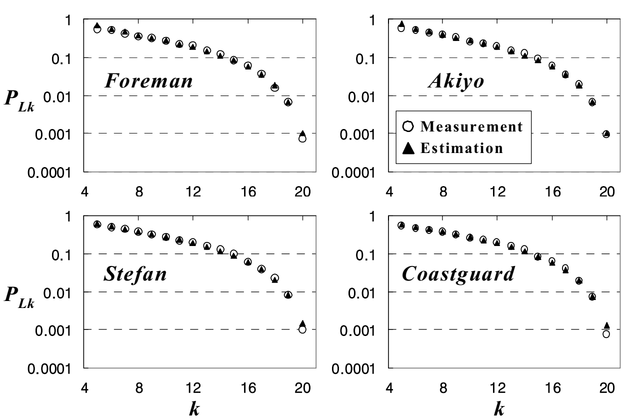

Fig. 5. Elements of CLTS vs. k obtained by full- measurement and using the proposed model for several test sequences. (α , β) is (3.15,0.27) in foreman, (3.6,0.25) in Akiyo, (1.7, 0.48) in Stefan, and (1.04, 0.45) in Coastguard.

All assumptions in Section 2 are applied. The parameter settings for simulation are as follows. The total channel rate (available channel bandwidth) is fixed at 384 kbps. The H.263 codec [5] is modified to support the time-varying source coding rate. In the $RS(n_0,k)$ code for channel coding, n_0=20 and k is varied in an integer granularity, *i.e.*, k=1,2,⋯,n_0.

Fig. 5 compares the elements of CLTS, versus k, obtained by the full-measurement method like in Fig. 2 and estimation using the proposed model of Eq. (11).

Note that the elements of CLTS obtained by both methods are very similar in all cases of test sequences, so that we can always get optimal code rate adaptively in time-varying loss channel without complete knowledge of the relationship between overall distortion and code rate.

6. CONCLUSIONS

We have proposed a new approach for adaptive rate allocation for source and channel codings. This approach is based on the loss threshold set, which is a set of packet loss probabilities in which the FEC code rate have to be changed in order to maximize average end-to-end video quality (or PSNR). The proposed model can estimate the residual loss threshold set by on-line background processing simultaneously with transmitting video packets, and then uses these loss probability elements to find the optimal code rate when there is a change in channel packet loss probability. The proposed method can obtain a near-optimal code rate faster than conventional methods because the proposed method uses the RLTS model equation, which is simple and has only two model parameters.

More work is left to refine the modeling of channel distortion of Eq. (10). A more theoretically based model should be used. At the moment we are working on the online estimation of the model parameters and their influence on the optimal RLTS elements.

7. REFERENCES

[1] K. Stuhlmuller, N. Farber, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE Trans. on Selected Areas in Comm.*, vol. 18, no. 6, pp. 1012-1032, June. 2000

[2] P. Frossard and O. Verscheure, "Joint Source/FEC Rate Selection for Quality-optimal MPEG-2 Video Delivery," *IEEE Trans. on Image Processing*, vol. 10, no. 12, pp.1815-1824, Dec. 2001

[3] J. C. Kwon and J. K. Kim, "Adaptive code rate decision of joint source-channel coding for wireless video," *Electronics Letters,* vol. 38, no. 25, pp. 1752-1753, Dec. 2002

[4] J. C. Kwon and J. K. Kim, "An Adaptive Rate Allocation to Source-Channel Coding for Internet Video," in *Proc. of GLOBECOM 2002*, Nov. 2002

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