OPTIMIZED RATE ALLOCATION FOR UNBALANCED MULTIPLE DESCRIPTION VIDEO CODING OVER UNRELIABLE PACKET NETWORK

Li Bin, Huang Feng, Sun Lifeng and Yang Shiqiang

Department of Computer Science and Technology
Tsinghua University
Beijing, China, 100084

ABSTRACT

Video transmission over unreliable packet network is in general hampered by the packet losses and constraint by stringent playback deadline. With these two key factors in consideration, Multiple Description Coding (MDC),comprising balanced and Unbalanced MDC has been proposed as an error-robust source coding technique. Recently, transmitting multiple descriptions over a single path is interesting due to the unavailability of multiple independent paths. Therefore, in this paper, we investigate the problem of rate allocation for the High-Resolution (HR) and Low-Resolution (LR) descriptions in UMDC transmission over single path. We first propose an approximate but efficient rate allocation model with the aid of two-state Markov link model and a simple distortion model at the sender side. Then we conduct extensive experiments to verity the proposed model and more excitedly the simulation results clearly demonstrate the effectiveness of proposed model.

1. INTRODUCTION

Video transmission over unreliable packet network is hampered by the high error rates and stringent delay constraints. In order to alleviate inevitable channel disturbances, Multiple Description Coding (MDC) has been proposed as an error-robust source coding technique [1]. MDC approach generally consists in generating two or more descriptions (MDC with two descriptions is fairly usual and herein it is illustrated as an example in the following). As long as these descriptions do not encounter packet losses concurrently, an acceptable quality could be obtained.

With the attractive properties of low delay and error robustness, MDC is particularly beneficial for those applications that have very stringent delay constraints or have a relatively long Round Trip Time (RTT). MDC essentially come in two flavors, consisting of balanced and unbalanced MDC. In balanced MDC approaches [2] the two descriptions have the same importance while in unbalanced MDC [3] one description has a higher quality than the other. Since in UMDC, the Low Resolution (LR) description is primarily used as redundancy and exploited to conceal errors in the High Resolution (HR) description, UMDC gives better control on the amount of introduced redundancy than BMDC.

In general, the MDC solutions have been designed to exploit path diversity to decrease the correlation among the descriptions [4][5]. However, it complicates the existing underlay network topology and requires additional link resource in reality. Particularly, for mobile ad hoc networks, what makes the multiple paths extremely difficult is the fact that there is no any guarantee on enough link resources. So UMDC over a single path is explored in [6]. Towards this end, since the LR description is used as the redundancy to relieve the effect of packet loss, a critical research question rises: *how to effectively and efficiently allocate rates for HR and LR descriptions to achieve better visual quality according to the varied network conditions*? Several rate allocation schemes for the UMDC solutions are proposed in recent years. Comas et al. [7] has proposed an UMD system which considers a rate-distortion optimal algorithm for generating the descriptions. However the analysis is specific for their proposed UMDC and not suitable for other UMDCs. Ekmekci et al. [4] also discuss rate allocation for the UMDC. They perform the RD curves generation/analysis in the encoder part and use the results of the analysis guide the rate allocation process in the encoder. However, this method is time-consuming and is not suitable for real-time applications.

In this paper, we mainly investigate the problem of an effective and efficient rate allocation scheme for the general UMDC when only a single physical path is used. First, we use the two-state continuous-time Markov chain to model the packet losses on the network. Second, a simplified distortion model is adopted to reveal the relationship between the bit rate and the distortion. Then we integrate the two models to deduce an approximate rate allocation model in dependence of the loss probability to guide the rate allocation in practice to achieve a better PSNR at the receiver side. Finally, we conduct extensive experiments to verify our approximate rate

This work is supported by the National Natural Science Foundation of China under Grant No. 60503063.

allocation model. The experimental results and the theoretical results match very well.

The rest of this paper is organized as follows. We introduce the packet loss model and distortion model and use these two models to deduce our rate allocation model in Section 2. Then in Section 3, simulation results are presented. Finally we conclude in Section 4 with a summary of our work.

2. THE PROPOSED RATE ALLOCATION MODEL

We introduce the packet loss model and distortion model in this section. Then we use these two models to deduce our approximate rate allocation model to guide the rate allocation between the HR and LR descriptions for real-time applications according to the varied network conditions.

2.1. UMDC scheme

2.1.1. The UMDC Architecture

We use a general UMDC as shown in Fig. 1 and Fig. 2 to discuss our approximate rate allocation model. One standard encoder (H.264 in our experiments) produces the HR coded stream and another produces the LR coded stream. The LR encoder is designed to produce a bitrate significantly lower than that produced by the HR one.

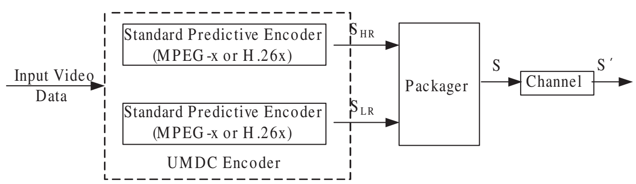

Fig. 1. UMDC video encoder.

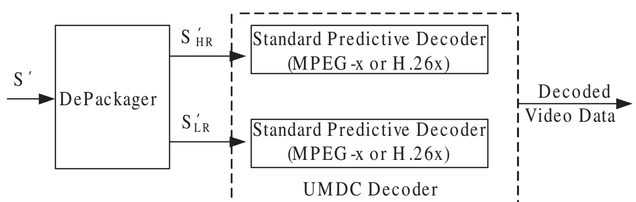

Fig. 2. UMDC video decoder.

2.1.2. The UMDC packager

As mentioned in Section 1, our proposed UMDC scheme uses only one channel. Thus both the HR and LR descriptions need to be transmitted over the same channel. In order to guarantee that the concurrence of packet loss for the two descriptions has a low probability, the corresponding HR and LR packets must not be transmitted simultaneously. Then we packetize the HR and LR data, which belong to different frames respectively, into the same packet to decrease their correlation. The offset of the HR and LR data in the same packet is adjusting according to the channel condition. Commonly, the size of the LR coded bitstream is smaller than the one of the HR coded bitstream. Thus in the same packet, the HR data occupies more proportion. Our packetization method with the offset of k is illustrated in Fig. 3.

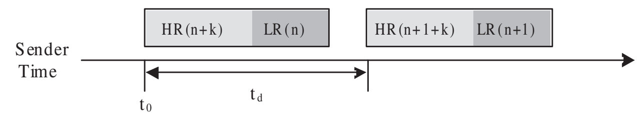

Fig. 3. Our packetization method.

2.2. Packet loss over networks

We adopt the two-state continuous-time Markov chain $\{X_t\}$, where $X_t \in \{0, 1\}$ presented in [8] to model the packet losses on the link. A packet transmitted at time t is lost if $X_t = 1$ and not lost if $X_t = 0$. The infinitesimal generator matrix Q of the chain is denoted by

$$Q = \begin{bmatrix} -\mu_0 & \mu_0 \\ \mu_1 & -\mu_1 \end{bmatrix}. \tag{1}$$

Then the stationary distribution is $\pi = (\pi_0, \pi_1)$, where $\pi_0 = \mu_1/(\mu_0 + \mu_1)$ and $\pi_1 = \mu_0/(\mu_0 + \mu_1)$. And the transition probability between state i to state j with an interval t can be compute as following:

$$\begin{cases} P(X_{t_0+t} = 0|X_{t_0} = 1) = \pi_0 - \pi_0 \exp(-(\mu_0 + \mu_1)t), \\ P(X_{t_0+t} = 1|X_{t_0} = 1) = \pi_1 + \pi_0 \exp(-(\mu_0 + \mu_1)t), \\ P(X_{t_0+t} = 0|X_{t_0} = 0) = \pi_0 + \pi_1 \exp(-(\mu_0 + \mu_1)t), \\ P(X_{t_0+t} = 1|X_{t_0} = 0) = \pi_1 - \pi_1 \exp(-(\mu_0 + \mu_1)t). \end{cases} \tag{2}$$

Given the average packet loss probability P_L, the average burst loss length $\overline{b}$ in packets, and the sending rate λ, the parameters of μ_0 and μ_1 can be deduced as following:

$$\begin{aligned} \mu_0 &= -P_L\lambda \left(1 - \frac{1}{\overline{b}\,(1 - P_L)}\right), \\ \mu_1 &= \mu_0\,(1 - P_L)\,/P_L. \end{aligned} \tag{3}$$

In fact, the values of P_L and $\overline{b}$ can be measured at the receiver and be fed back to the sender. Then the sender is able to calculate and update the parameters of μ_0 and μ_1 timely.

2.3. Distortion Model

For compressed-video applications, several distortion models have been developed to estimate the distortion for a given source rate. However most of them are computationally expensive, thereby reducing the practicality of the proposed rate allocation method. So we seek a relatively easy-to-implement, yet representative model. For distortion estimation, we adopt the following equation proposed in [9], where K is a model parameter, R_i and $D(R_i)$ denotes the rate and the distortion respectively.

$$D(R_i) \approx \frac{K}{R_i}. \tag{4}$$

For offline-encoded video sequences, the value of K can be precomputed. However, for real-time encoding, an initial value is assumed for K, and this value is refined as more frames are encoded.

2.4. Rate Allocation

Since we transmit both HR and LR descriptions over a same path, the total available bit rate R_T is fixed. Therefore, how to allocate an appropriate rate for HR and LR is important. We discuss our rate allocation model in the following. Let D_{HR} (D_{LR}) denote the achieved distortion when HR (LR) arrives correctly and D_C denote the achieved distortion when HR and LR are both lost. And p denotes the probability of HR lost and p_{10} (p_{11}) denotes the conditional probability of LR received (lost) when HR is lost. Then we can write the achieved distortion D at the receiver as

$$D = (1-p) \times D_{HR} + p \times p_{10} \times D_{LR} + p \times p_{11} \times D_C. \tag{5}$$

For our packetization method, the delay between the HR data and LR data which belong to the same frame is kt_d, as shown in Fig. 4. With (2) and (4), we can rewrite (5) as

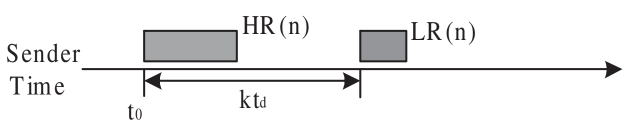

Fig. 4. Delay between HR and LR.

follows

$$\begin{aligned} D \approx\ & (1-p) \times \frac{K}{R_1} \\ & + p \times (\pi_0 - \pi_0 \exp(-(\mu_0+\mu_1)kt_d)) \times \frac{K}{R_2} \\ & + p \times (\pi_1 + \pi_0 \exp(-(\mu_0+\mu_1)kt_d)) \times D_C. \end{aligned} \tag{6}$$

Since the delay should be as small as possible, the value of k cannot be set too large. From Fig. 5, we can see that for all $k \geq 3$, the term $\exp(-(\mu_0+\mu_1)kt_d)$ is close to zero. Therefore we can select an appropriate and practical k, for example $k = 3$, to make

$$\exp(-(\mu_0+\mu_1)kt_d) \approx 0. \tag{7}$$

Subsequently we approximate (6) by

$$D \approx (1-p) \times \frac{K}{R_1} + p \times (1-p) \times \frac{K}{R_2} + p \times p \times D_C. \tag{8}$$

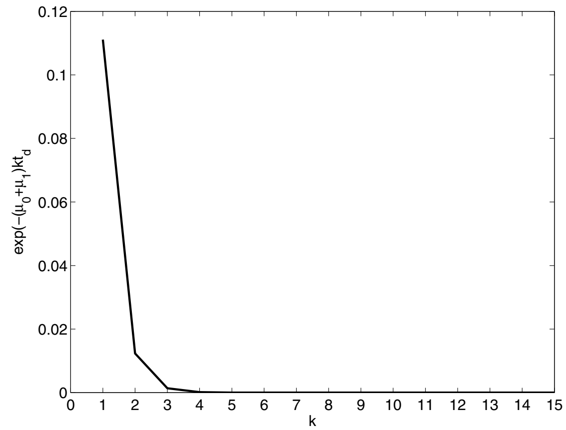

Fig. 5. The curve of $\exp(-(\mu_0+\mu_1)kt_d)$ under the network condition of average burst length $\overline{b}$=1.5, P_L=25%.

The objective is to minimize the above distortion subject to the constraint

$$R_1 + R_2 = R_T. \tag{9}$$

Given p, we can easily know

$$\min(D) \Leftrightarrow \min(\frac{1}{R_1} + \frac{p}{R_2}). \tag{10}$$

Toward this end, obviously we could figure out when $R_2 = \frac{\sqrt{p}}{1+\sqrt{p}} R_T$, $\frac{1}{R_1} + \frac{p}{R_2}$ achieves its minimal value, i.e. D achieves its minimal value.

3. SIMULATIONS

In this section we present the simulation results of the proposed rate allocation model and the comparison with the theoretical results.

We implemented the UMDC system by modifying the JVT Software Version M8.6. We assume the RTP payload format for packetizing the H.264 video stream [10]. We test different video sequences and present the results for two of them (QCIF, 15fps, 150 frames), Foreman and News. We run the simulations over 50 different loss realizations with the same model parameters to get convictive results.

In the experiments, the average burst length is set to 1.5 and 2.0 and the packet loss probability is set to 15%, 20%, 25%, and 30% respectively to simulate different network conditions. We encode the test video sequences at the targeting total bit rate 300kbps. Due to the limited space, we only present the result under the network condition of average burst length $\overline{b}$=1.5, P_L=25%, total bit rate R_T=300kbps in this paper, as shown in Fig. 6 and Fig. 7. The theoretical result from our model is presented in Fig. 6. From Fig. 6 it is easily seen that the distortion is minimized if the LR bit rate is about 90kbps. The simulation results are presented in Fig.

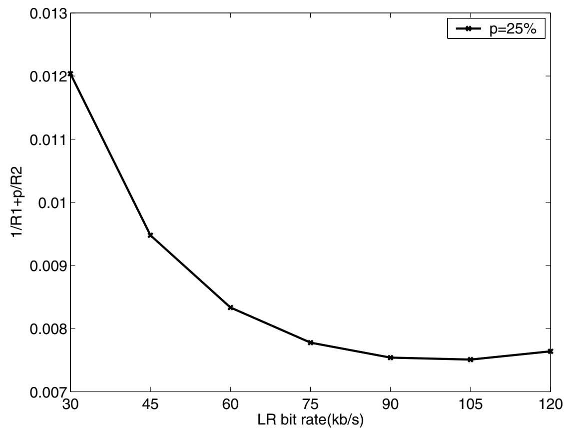

Fig. 6. The theoretical result under the network condition of $\bar{b}$=1.5, P_L=25%, R_T=300kbps.

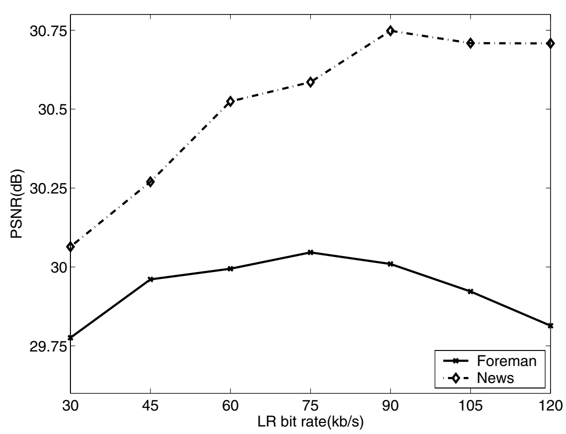

Fig. 7. The simulation result under the network condition of $\bar{b}$=1.5, P_L=25%, R_T=300kbps.

7. We can easily see that the maximal PSNR is reached if the LR bit rate is about 90kbps for news sequence and 75kbps for foreman sequence. It shows that the theoretical result from our approximate model matches the experimental result very well. So our approximate model is efficient to guide the rate allocation to achieve a better quality according to the network condition.

4. CONCLUSION

In this paper, we mainly discuss the issue of rate allocation for HR and LR descriptions in the area of UMDC. Our contributions come in twofold. Firstly, we introduce two-state Markov packet loss model and a simplified distortion model to incubate a novel rate allocation model. Secondly, we conduct experiments to further verify that the theoretical results match the experimental results very well, which demonstrates the proposed rate allocation model is efficient enough to guide the rate allocation while ensuring better visual quality within the horizon of delay deadline according to the varied network conditions. In future, we will

5. REFERENCES

[1] V. K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, pp. 74–93, Sept. 2001.

[2] K. R. Matty and L. P. Kondi, "Balanced multiple description video coding using optimal partitioning of the dct coefficients," *IEEE Transactions on CSVT*, vol. 15, no. 7, pp. 928–934, Jul. 2005.

[3] M. Fumagalli, R. Lancini, and S. Tubaro, "A novel error-concealment algorithm for an unbalanced multiple description coding architecture," *Proceedings of International Packet Video Workshop*, Dec. 2004.

[4] S. Ekmekci and T. Silora, "Unbalanced quantized multiple description video transmission using path diversity," *Proceedings of SPIE's Electronic Imaging 2003*, Jan. 2003.

[5] J. G. Apostolopoulos and S. J. Wee, "Unbalanced multiple description video communication using path diversity," *Proceedings of International Conference on Imaging Processing*, Oct. 2001.

[6] A. Vilei, G. Convertino, S. Oliva, and R. Cuppone, "A novel unbalanced multiple description scheme for video transmission over wlan," *Proceedings of the 3rd ACM international workshop on Wireless mobile applications and services on WLAN hotspots*, Sept. 2005.

[7] D. Comas, R. Singh, A. Ortega, and F. Marques, "Unbalanced multiple-description video coding with rate-distortion optimization," *EURASIP Journal on Applied Signal Processing 2003*, , no. 1, pp. 81–90, Jan. 2003.

[8] Yen-Chi Lee, Y. Altunbasak, and R. M. Mersereau, "Optimal packet scheduling for multiple description coded video transmissions over lossy networks," *Proceedings of IEEE GLOBECOM'03*, vol. 6, pp. 3569–3573, Dec. 2003.

[9] A. C. Begen, Y. Altunbasak, O. Ergun, and M. H. Ammar, "Multi-path selection for multiple description video streaming over overlay networks," *Signal Processing: Image Communication*, vol. 20, no. 1, pp. 39–60, Jan. 2005.

[10] S. Wenger, T. Stockhammer, and M. M. Hannuksela, "Rtp payload format for h.264 video," *Internet Draft*, Mar. 2003.