Transactions Letters

Content-Based Retransmission for 3-D Wavelet Video Streaming Over Lossy Networks

Jianyu Dong, *Member, IEEE*, and Yuan F. Zheng, *Fellow, IEEE*

***Abstract*—In video streaming, time-varying packet loss rate leads both degradation and inconsistency to video quality. In this paper, we propose a novel content-based retransmission scheme to address the above two problems in 3-D wavelet video streaming. The fundamental idea is to consider the contents of the packets and the traffic of the Internet simultaneously to optimize the quality of service (QoS). In the encoding stage, the content information is generated and stored, based on which an intelligent hybrid decision system is designed to perform a two-step optimization in transmission. First, the scheme for retransmitting the lost packets is designed to minimize the distortion of each group of video frames (GOF). Secondly, the bandwidth used for transmitting each GOF is dynamically allocated according to the changing traffic conditions such that the quality of video is consistent. The goal is to enable the streaming system to deliver the best possible visual quality of video under the same channel conditions. Simulation results show that our proposed approach can improve the video quality (more than 2 dB) as well as reduce the quality variation significantly.**

***Index Terms*—Content-based, quality of service (QoS), retransmission, 3-D wavelet, video streaming.**

I. Introduction

IMPROVING the quality of service (QoS) of video streaming over lossy networks is a challenging topic of research. A lot of research efforts have been made to optimize the video quality subject to channel constraints [1]–[8]. Chou and Miao [1]–[3] presented a general rate-distortion optimized streaming framework based on Markov decision and Lagrangian optimization, which provided a valuable guideline for future works in this field. More recently, a great deal of research has been conducted to improve the general rate-distortion streaming framework and to extend its usage to various transmission scenarios with different video coding schemes [9]–[11]. However, most of the existing work on rate-distortion optimized streaming did not consider the inconsistency of video quality caused by time varying packet loss behavior in a realistic transmission channel. As shown in Fig. 1, change of packet loss rate during transmission can cause significant quality difference in adjacent video groups of frames (GOFs). Since human eyes are very sensitive to abrupt quality changes in continuous motion pictures, capability of keeping consistent video quality should be an important factor in the design of a good video streaming system.

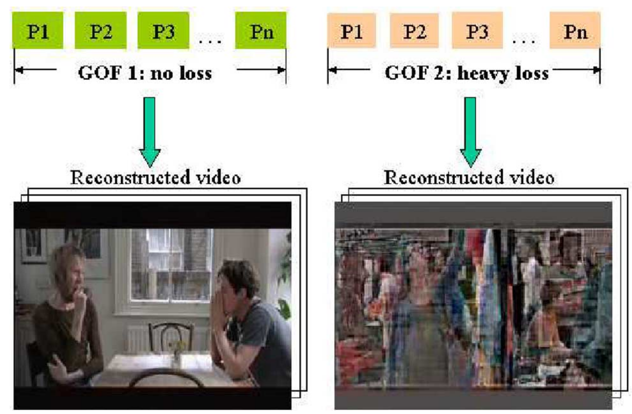


Fig. 1. Quality variation caused by time-varying packet loss rate in two adjacent groups of video frames. The quality of the video on the right is not acceptable because of heavy loss of the packets.

In this paper, we present the design of an optimized transmission system for 3-D wavelet video across time-varying packet loss channels, which takes the quality consistency into account. Hence, our designed system aims to not only maximize the quality of the received video, but also reduce the quality variation of the video sequence. To achieve this optimization goal, a content-based retransmission scheme is proposed. During the encoding, content information about the video is generated and stored; during the transmission, a two-step optimization is performed.

- For each GOF, minimize the distortion caused by the packet loss due to limited network resources (intra-GOF optimization).
- For multiple GOFs, minimize the variance of their distortions (inter-GOF optimization).

Both steps of the proposed algorithm are based on the properties of the GOF-structured 3-D wavelet video. For intra-GOF optimization, the transmission decision is based on the packet content and channel conditions. For inter-GOF optimization, a dynamic bit allocation algorithm is developed based on the rate-distortion characteristic of each GOF delivered by

Manuscript received December 16, 2003; revised April 22, 2006. This paper was recommended by Associate Editor A. Puri.

J. Dong is with the Department of Electrical and Computer Engineering, California State University, Los Angeles, CA 90032 USA (e-mail: jdong2@calstatela.edu).

Y. F. Zheng is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA, and also with the Shanghai Jiao Tong University, Shanghai 200030, China (e-mail: zheng@ece.osu.edu).

Color versions of Figs. 1, 4–6, 9 and 10 are available at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2006.881866

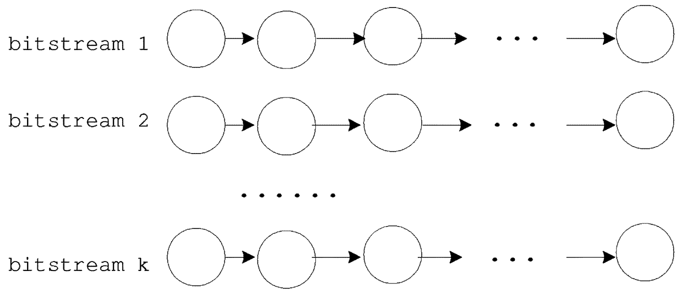


Fig. 2. Inter-packet dependency in one 3-D wavelet video GOF where each circle represents a packet. The arrow points to the packet which depends on the current packet where the arrow starts.

the intra-GOF optimized scheme. An experimental streaming system has been developed based on our proposed algorithm, and simulations have been conducted to verify its effectiveness in achieving good streaming effect in lossy networks.

II. Problem Definition

A. Backgrounds and Motivations

Although not so widely used as popular video coding standards such as MPEG2 and H.264, 3-D wavelet based video codec has its unique advantages. One widely recognized advantage of 3-D wavelet video codec is its great scalability due to the multiresolution nature of wavelet decomposition. In addition, error propagation is reduced in error prone environment since 3-D wavelet video codec utilizes wavelet transforms along the temporal direction, instead of motion compensation, to remove the inter-frame redundancy. Many widely used 3-D wavelet coding methods, such as packetized zerotree wavelet (PZW) [12], [13], take advantage of this property to generate independently-decodable video packets. Thus, the loss of one packet will not affect the reconstruction of the others. Therefore, 3-D wavelet video codec has great potential to be a desirable solution for streaming applications over lossy, heterogeneous networks, and it is meaningful to further explore its properties for developing a best streaming approach for 3-D wavelet video.

- The GOF structure in 3-D wavelet video: GOF can be viewed as a basic coding unit in 3-D wavelet codec. A GOF has to be accumulated before 3-D wavelet transform can be applied. Similarly, inverse wavelet transform cannot proceed until all packets of a GOF are received at the decoder. Therefore, 3-D wavelet video packets of the same GOF have identical transmission deadline.[1]
- The inter-packet dependency in 3-D wavelet video: The inter-packet dependency of 3-D wavelet video is very different from the well-known relationship between I-, P-, and B-frames in MPEG and H.264. Fig. 2 depicts the generic inter-packet relationship of 3-D wavelet video. In the figure, there are k independent bitstreams resulted from error-resilient packetization, but in each bitstream a later packet is dependent on its immediate earlier packet due to the embedded bitstream structure.

[1]Due to the delay constraint, a packet P must be transmitted before time T at the sender. The time T is defined as the transmission deadline for packet P.

To design a best streaming system for 3-D wavelet video, the above properties regarding the transmission deadline and packet dependency should be taken into account.

B. Formulation of the Optimal Retransmission Problem

For convenience, we first partition the packets at the sender into four sets: Ω_{Sr} (set of packets which has been received); Ω_{Su} (set of packets which has been sent but is unknown if received); Ω_R (set of packets which is requested for retransmission); and Ω_U (set of packets which is yet to send).

- **Formulation of the intra-GOF optimization problem**: Assuming that the average data transmission rate constrained by the channel capacity to be R, and the delay constraint imposed on the ith GOF in the application is $T_D(i), i = 1, 2, \ldots$. Note that a GOF must be delivered on time to display the video continuously on the receiver side. Therefore, both rate and delay constraints should be met when designing the retransmission scheme. An intra-GOF retransmission scheme is a policy π to choose a subset $\Omega_T(\pi) \subseteq \Omega_T(\Omega_T = \Omega_R \cup \Omega_U)$ of the current GOF to transmit while meeting the following constraints:

$$\begin{cases} r_s \leq R \\ \sum_{j \in \Omega_T(\pi)} B_j \leq r_s \cdot (T_D(i) - T_D(i-1)) \end{cases} \quad (1)$$

where r_s is the average data sending rate and B_j is the size of the jth packet selected to transmit in the ith GOF under the policy π. An optimal policy Π maximizes the quality of the received video. If the video quality is represented by the distortion, the policy should be selected according to the following formula:

$$\Pi = \arg\min_{\pi} D(\pi) \quad (2)$$

where $D(\pi)$ is the expected total distortion caused by the lost or dropped packets under policy π.

- **Formulation of the inter-GOF optimization problem**: In order to meet the display schedule at the receiver, the delay constraint for each GOF must be satisfied. Assuming that the sending time for the latest packet in GOF i is $T_{\text{latest}}(i)$, and the scheduled transmission deadline for GOF i is $T_D(i)$, the inter-GOF optimization problem can be formulated as follows:

$$\begin{cases} T_{\text{latest}}(i) \leq T_D(i), \quad i = 1, \ldots, N-1 \\ T_{\text{latest}}(N) = T_D(N) \\ \mathcal{T} = \arg\min_{\mathcal{T}_k} (\operatorname{Var}(Q_1, Q_2, \ldots, Q_N)) \end{cases} \quad (3)$$

where $\mathcal{T} = (T_{\text{latest}}(1), T_{\text{latest}}(2), \ldots, T_{\text{latest}}(N))$ is the optimal policy to transmit GOF 1 to GOF N, and $Q_i, i = 1, \ldots, N$ is the quality of each GOF, which can be represented by either distortion [mean square error (MSE)] or peak signal-to-noise ratio (PSNR).

In (3), the first two expressions are the delay constraints, and the last one means that the goal of the optimization policy is to minimize the variation of the video quality among multiple GOFs.

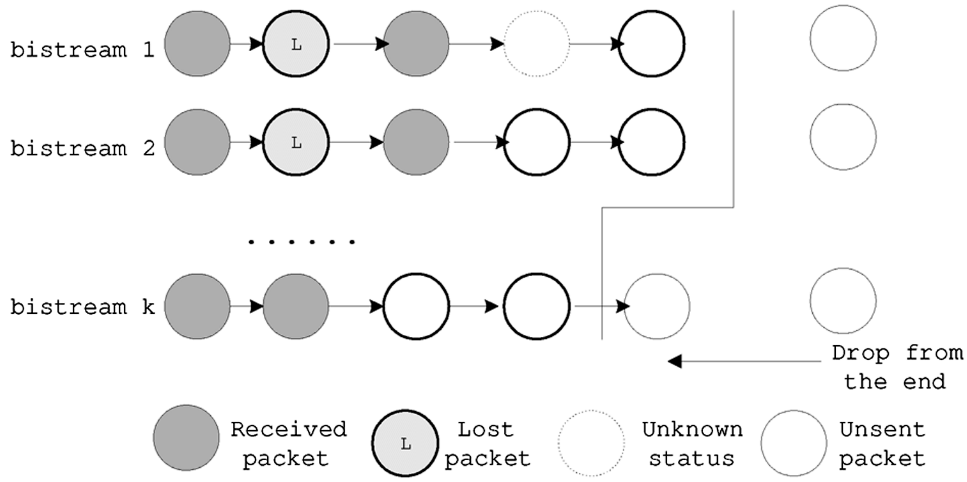


Fig. 3. Transmission scenario for one GOF.

III. Content-Based Retransmission Mechanism

In this section, we will describe the proposed content-based retransmission mechanism to achieve the optimization goals presented in Section II-B, which comprises two steps: intra-GOF optimization and inter-GOF optimization. The basic idea of the former is to ensure the delivery of the most important packets under delay and rate constraints; and the latter will determine the transmission bit budget for each GOF based on the channel condition and video content to achieve a smooth display quality.

A. *Intra-GOF Optimization Algorithm*

Assume that there are N packets in the current GOF and they have to be sent out before the transmission deadline t_D. At a certain transmission instance t_{Cur} when n out of N packets have been sent out, the remaining bit budget of the current GOP is given by

$$C_{\mathrm{Cur}} = R \cdot (t_D - t_{\mathrm{Cur}}). \tag{4}$$

As shown in the transmission scenario depicted in Fig. 3, if the remaining bit budget C_{Cur} is not sufficient to transmit all the packets in Ω_U and Ω_R, packet dropping is unavoidable. Thus, an optimal dropping policy should be developed to minimize the incurred distortion.

A packet dropping policy $\rho = (\rho_1, \rho_2, \ldots, \rho_m)$ is a binary vector, where

$$\rho_i = \begin{cases} 1, & \text{dropped} \\ 0, & \text{otherwise} \end{cases}$$

and m is the total number of packets in Ω_T. To achieve the goal of intra-GOF optimization defined by (2), an optimal packet dropping policy Π_D should be found to minimize the distortion caused by dropping while subject to the bit constraint:

$$\begin{cases} \Pi_D = \arg\min_\rho \Delta D(\rho) \\ \sum_{i=1}^{m} \rho_i B_i \geq B(\Omega_T) - C_{\mathrm{Cur}} \end{cases} \tag{5}$$

where B_i is the size of packet i and $B(\Omega_T)$ is the total size of all packets in Ω_T.

With the selected dropping policy Π_D, the remaining packets (denoted by Ω_{ss}) should contain the most important information and have the greatest potential to maximize the quality of the received video.

To find the best dropping policy for (5), it is important to quantify the distortion caused by dropping one packet. Hereby a content index, namely *drop-distortion*, is defined to measure the impact of packet dropping. It will be shown later that the drop-distortion is an accurate measurement of packet significance when considering dropping policy, and it can be calculated easily for 3-D wavelet video packets.

- **Definition of drop-distortion**: Drop-distortion of packet i is defined to be the distortion increment of the received video when packet i is selected to be dropped by policy ρ

$$\Delta D_i = D_i - D_i^* + \sum_{j \in \Omega_{\mathrm{Dependent}}} (1 - p_{\mathrm{lost}}(j)) \cdot \Delta D_j \tag{6}$$

where D_i is the distortion of the video reconstructed from all the packets except both packet i and its dependent packets,[2] D_i^* is the distortion of the video reconstructed from all the packets except the dependents of packet $i, \Omega_{\mathrm{Dependent}}$ is the set of direct dependent packet of packet i, and $p_{\mathrm{lost}}(j)$ is the loss probability of the direct dependent packet j that is defined as follows:

$$p_{\mathrm{lost}}(j) = \begin{cases} 1, & \text{if } j \text{ is dropped or lost} \\ 0, & \text{if } j \text{ is received} \\ P_e, & \text{if } j \text{ has unknown status} \end{cases} \tag{7}$$

where P_e is the channel packet loss rate.

Drop-distortion can be viewed as the *conditional* expected distortion under certain dropping policy ρ. When dropping of packet i is assumed in ρ, the incurred distortion consists of two parts: 1) the impact of losing packet i itself, which is measured by the first term $(D_i - D_i^*)$ in (6), and 2) the effect of error propagation over its dependent packets, which is quantified by the second term in (6).

[2]If a packet is lost, all its dependent packets become useless in reconstruction.

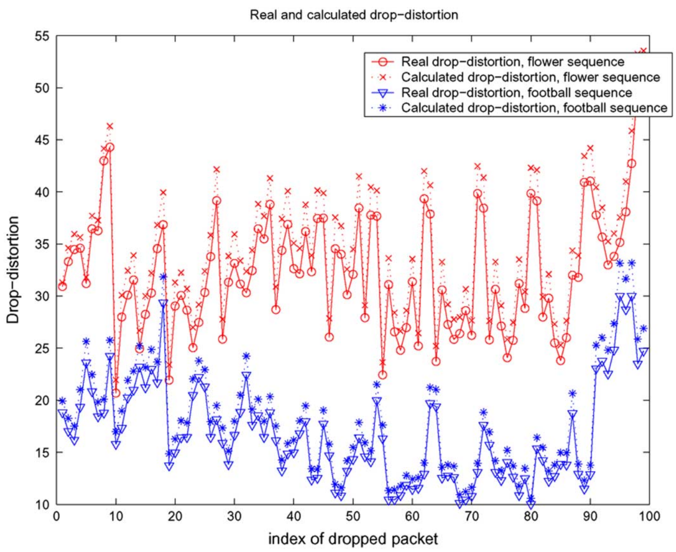


Fig. 4. Comparison between real and calculated drop-distortion using wavelet coefficients. Here both video sequences (*Flower* and *Football*) are compressed using 3-D zerotree based codec with robust packetization. 9-7 wavelet transform is used in spatial domain and Haar wavelet transform is used in temporal domain.

For a valid dropping policy ρ, if both packet i and its dependents belong to Ω_T, the dropping of packet i will not happen until all its dependents are dropped. In this case, there is no error propagation. If some of its dependents have been sent out, dropping of packet i could lead to the failure of reconstruction of its dependents packets, and thus cause additional distortion increment. The definition in (6) provides a generic measure of the dropping effect for all possible situations.

- **Calculation of drop-distortion**: Since the actual value of the distortion caused by packet dropping $(D_i - D_i^*)$ in (6), cannot be known before the video is reconstructed, for 3-D wavelet video packets shown in Fig. 2, drop-distortion of a packet i in bitstream l can be approximated via recursion

$$\Delta D_{l,i} = E_{l,i} + (1 - p_{\text{lost}}(i+1))\Delta D_{l,i+1} \tag{8}$$

where $E_{l,i}$ is the energy in packet i of bitstream l and $\Delta D_{l,N} = E_{l,N}$ is the energy contained in the last packet in bitstream l that has no child. The energy of a 3-D wavelet packet can be estimated using square sum of the wavelet coefficients.

Experimental results have shown that the above approximation is very close to the real drop-distortion defined in (6), as demonstrated by Fig. 4. From the figure, we can see that (8) provides a quick estimation of drop-distortion with sufficient accuracy. To speed up the calculation, the energy contained in 3-D wavelet packets can be pre-calculated during the encoding stage and stored with the video bitstream.

To compare the effect of different dropping policies, it needs to represent the total incurred distortion using the drop-distortion of individual packets. For 3-D wavelet video with packet dependency shown in Fig. 2, it can be shown using (8) that the total distortion, $\Delta D(\rho)$, caused by a valid dropping policy ρ is the linear addition of individual drop-distortion as follows:

$$\Delta D(\rho) = \sum_{i=1}^{m} \rho_i \cdot \Delta D_i \tag{9}$$

where m is the number of packets in Ω_T.

The above linear addictive property of drop-distortion significantly simplifies the computation to search for optimal dropping policy.

- **Optimal packet dropping based on drop-distortion**: Theoretically, the optimal dropping policy can be found using exhaustive search or dynamical programming. However, similar to the conventional backpack problem, the worst case computation complexity for optimal solution is $O(2^n)$. In reality, a typical GOF in 3-D wavelet video can be divided into 50 or more packets. That could cause the computation complexity to the order of 2^{50}, which is too high to handle in real time. Therefore, we propose a heuristic algorithm to find a suboptimal dropping policy with much lower computation complexity. Assume that we want to find the best dropping policy at time t_{Cur} under the transmission scenario shown in Fig. 3. If not all packets can be sent, packets will be dropped as follows:

Initialization: A dropping candidate set will be constructed using the leaf packets that has no child in each bitstream; $\Omega_{ss} = \Omega_T$.

Step 1) Check the packets in the dropping candidate set, if the packet is in Ω_{ss}, calculate its drop-distortion per bit.

Step 2) Based on the calculation of Step 1, drop the one with the smallest drop-distortion per bit, update drop policy ρ and Ω_{ss} accordingly.

Step 3) Check the number of total bits in Ω_{ss}. If $B(\Omega_{ss}) \leq C_{\text{Cur}}$, stop.

Step 4) Suppose the dropped packet is in bitstream l, add its most recent parent that belongs to Ω_{ss} to the dropping candidate set, and then go back to Step 1.

The packets in Ω_{ss} will be sent in the descending order of drop-distortion per bit.

The proposed algorithm has the following advantages. First, it guarantees that the selected dropping policy is a valid one, which means a parent packet will never be dropped before its children. Secondly, it guarantees that the least important packet (measured by drop-distortion per bit) be dropped at every step. Therefore, the remaining packets that will be sent should be the most important ones, which is consistent with our optimization goal.

B. Inter-GOF Optimization Algorithm

Intra-GOF optimization provides a local optimal streaming method for each GOF. Since the packet loss rate changes from time to time, quality of different GOFs may vary a lot. It is necessary to develop an inter-GOF optimization algorithm to keep the video quality as consistent as possible. In fact, the inter-GOF optimization problem as formulated by (3) can be viewed as a bit allocation problem because the vector of transmission interval

$(T_1, \ldots, T_N)$ can be mapped into a vector of bits allocated to each GOF $(W_1, W_2, \ldots, W_N)$ by the following function:

$$\begin{cases} T_i = T_{\text{latest}}(i) - T_{\text{latest}}(i-1) \\ T_i R_i = W_i, i = 1, \ldots, N \end{cases}$$

where R_i is the average data rate during T_i and W_i is the total amount of bits transmitted.

Inter-GOF optimization can be achieved by adjusting the available bits for each GOF dynamically to alleviate the degree of quality fluctuation caused by the variation of the packet loss rate. Assume that there are N GOFs and the maximum number of bits that can be allocated to send the first k GOFs is $W_k^C, k = 1, 2, \ldots, N$. The purpose of dynamical bit allocation is to minimize the variance of the distortions among the N GOFs while satisfying the bit constraints, as shown

$$\begin{cases} \sum_{i=1}^{k} W_i \leq W_k^C, \quad k = 1, 2, \ldots, N-1 \\ \sum_{i=1}^{N} W_i = W_N^C (\text{bit constraints}) \\ \min \text{Var}(\delta D_1, \delta D_2, \ldots, \delta D_N) \end{cases} \tag{10}$$

where $\delta D_i, i = 1, 2, \ldots, N$ denotes the distortion caused by the packet loss in each GOF. Let D_i denote the actual distortion of the reconstructed video in GOF i, and $D0_i$ is the distortion when no packet is lost. One has $\delta D_i = D_i - D0_i$.

To obtain an optimal solution for the above problem is challenging. Using programming methods such as exhaustive search and dynamical programming is not realistic since for each possible allocation policy, an optimal retransmission scheme must be generated to obtain δD_i for each GOF. In the following we present a closed-form allocation formula to speed up the computation.

1) Rate-Distortion Model for Transmission: Different from the one for compression, a rate-distortion model reflects the relationship between the amount of bits received and the distortion of the reconstructed video. Apparently, it varies with different transmission schemes. We have found that the rate-distortion model of the intra-GOF optimization scheme can be represented by an exponential function of W_i

$$\delta D_i = A_i e^{-\lambda_i (1 - p_{\text{lost}}) W_i}, \qquad i = 1, 2, \ldots, N \tag{11}$$

where p_{lost} is the packet loss rate, and A_i and λ_i are model parameters.

This exponential representation has been demonstrated by many experiments, of which Fig. 5(a) and (b) show the rate-distortion curves for transmitting two video sequences, namely *Austin-Powers* and *Table Tennis*, respectively. In the figures, $\delta D_i, i = 1, 2, \ldots, 32$ is obtained from the experimental results of our proposed intra-GOF optimization scheme under different packet loss rates, and the parameters, A_i and λ_i, in the exponential approximation, are calculated using regression. Since δD_i cannot be known before transmission, it can be well represented by $\sum \bar{c}_{xyz}^2$ of the dropped packets in GOF i, where $\bar{c}_{xyz}$ is the quantized coefficient and known after the video compression.

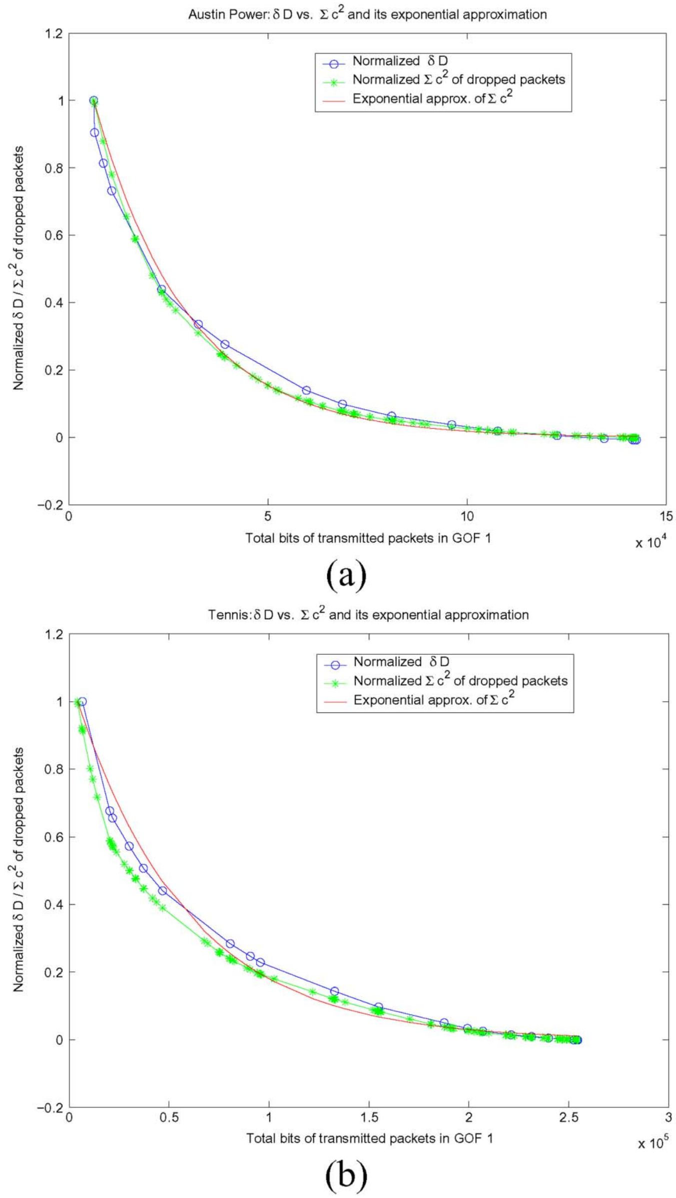


Fig. 5. δD and $\sum \bar{c}_{xyz}^2$ of the dropped packets under Intra-GOF optimization scheme for video sequences. (a) *Austin-Powers*. (b) *Table-Tennis*.

2) Recursive Bit Allocation Algorithm: An ideal solution to (10) is $\delta D_i = \delta D_{i-1}, i = 2, 3, \ldots, N$. Assuming that this is achievable, we can get the following linear equation set:

$$\begin{cases} \ln A_1 - \lambda_1' W_1 = \ln A_2 - \lambda_2' W_2 \\ \ln A_2 - \lambda_2' W_2 = \ln A_3 - \lambda_3' W_3 \\ \cdots \\ \ln A_{N-1} - \lambda_{N-1}' W_{N-1} = \ln A_N - \lambda_N' W_N \\ W_1 + W_2 + \cdots + W_N = W_N^C, \end{cases} \tag{12}$$

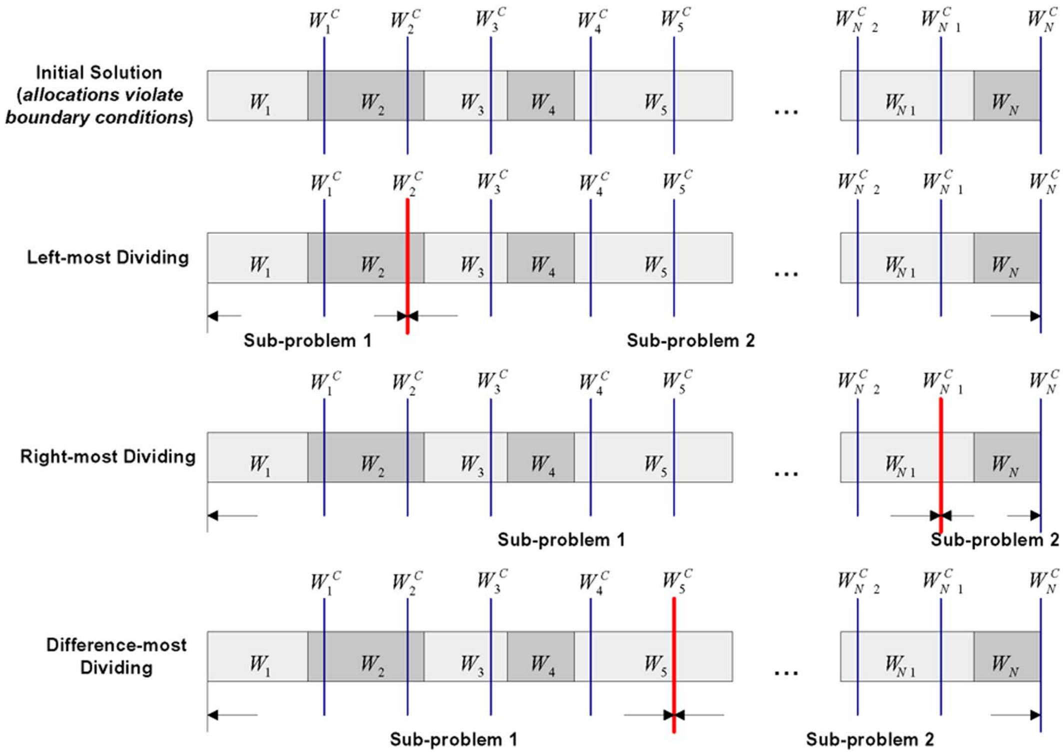


Fig. 6. Principle of divide and conquer: three different dividing methods.

where $\lambda_i' = \lambda_i(1-p_{\text{lost}}), i = 1, \ldots, N$. Apparently, the solution for this linear equation set is $\mathcal{W} = \Lambda^{-1}\mathcal{A}$ when Λ is nonsingular and

$$\Lambda = \begin{pmatrix} \lambda_1' & -\lambda_2' & 0 & \cdots & 0 \\ 0 & \lambda_2' & -\lambda_3' & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}$$

$$\mathcal{W} = \begin{pmatrix} W_1 \\ W_2 \\ \cdots \\ W_N \end{pmatrix}$$

$$\mathcal{A} = \begin{pmatrix} \ln A_1 - \ln A_2 \\ \cdots \\ \ln A_{N-1} - \ln A_N \\ W_N^C \end{pmatrix}.$$

Each element in the solution $\mathcal{W}$ is the bit assignment for one GOF. However, this ideal solution $\mathcal{W}$ may not satisfy the boundary conditions in (10). A simple method to solve this problem is to change any $\sum_i^k W_i, i = 1, \ldots, N-1$ that violates the boundary condition into $W_i^C, i = 1, \ldots, N$. Apparently, this simple method may not yield a good allocation.

In order to achieve a near-optimal allocation effect while satisfying the bit constraint for each GOF, we design a recursive algorithm based on the principle of divide and conquer. The basic idea is to divide the allocation problem into subproblems that are simpler to solve. The performance of divide and conquer algorithm depends on how to divide the original problem. As shown in Fig. 6, there are three possible dividing methods, namely *left-most, right-most*, and *difference-most*. The left-most method divides the problem at the first violation spot (the first allocation that violates the boundary condition), while the right-most one divides at the last violation spot. The difference-most method looks for the bit assignment which exceeds the bit constraint by the largest amount, and divides the problem there. In order to find which method is the best, many experiments were conducted. In the experiments, the three different dividing methods were used to allocate bits to a number of GOFs, and their performances are compared in terms of the STD (standard deviation) of the distortion of the received video. Tables I and II record the simulation results for 6-GOF and 8-GOF allocation, respectively. As the reference of comparison, the performance of the fixed allocation method which sends the packet without inter-GOF optimization is also displayed in the tables.

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT BIT ALLOCATION METHODS FOR 6 GOFS [THE AVERAGE (AVG) AND STANDARD DEVIATION (STD) OF THE DISTORTION ARE USED AS METRICS TO MEASURE PERFORMANCE]

	fixed	left-most	right-most	diff-most
AVG	256.4067	177.6931	177.2754	175.5762
STD	341.1849	131.4545	124.9346	120.8918

TABLE II
PERFORMANCE COMPARISON OF DIFFERENT BIT ALLOCATION METHODS FOR 8 GOFS [THE AVERAGE (AVG) AND STANDARD DEVIATION (STD) OF THE DISTORTION ARE USED AS METRICS TO MEASURE PERFORMANCE]

	fixed	left-most	right-most	diff-most
AVG	264.0386	174.7182	174.2515	171.3767
STD	369.5669	137.0856	130.7151	119.992

From the simulation results, we can see that all three methods of dynamical allocation can achieve much more stable quality

TABLE III
SIMULATION PARAMETERS FOR DIFFERENT TEST VIDEO SEQUENCES

Video Seq.	Frame size	Data rate	Delay constraint
Notting Hill	352×240	64 kbps	4.5 s
Table Tennis	352×240	64 kbps	6.6 s
Grandam	160×120	28 kbps	2.1 s

TABLE IV
AVERAGE PSNRS (IN DECIBELS) WITH LOW PACKET LOSS RATE (5%)

Video Seq.	T1	T2	T3
Notting Hill	30.8261	32.9869	33.2980
Table Tennis	26.6546	26.8784	28.1627
Grandam	31.3454	32.3150	32.4453

than the fixed allocation scheme, and the difference-most method is able to obtain the best allocation scheme. Hence, the difference-most method is adopted in the recursive algorithm.

The basic steps of our proposed recursive bit allocation algorithm can be described as follows:

Step 1) Solve (12) and get the ideal solution $\mathcal{W}$.

Step 2) Search for $\sum_i^l W_i$ that violates the boundary condition and exceeds the constraint by the largest amount. If found, set $\sum_i^l W_i = W_l^C$ and go to step 3; otherwise, stop.

Step 3) Divide the allocation problem into two subproblems. The first is one bit-allocation problem among the first l GOFs and the second is another bit-allocation problem among the remaining GOFs. Repeat the previous steps to solve the subproblems.

In reality, we rarely need to allocate the bits among more than 3 GOFs, since the long-term variation of the traffic condition is hard to predict. For the 2-GOF case, we can even have a close-form solution. Let the allocated bits of GOF 1 (W_1) be mapped into its transmission deadline $T_D^*(1)$. It can be expressed as

$$T_D^*(1) = \frac{\ln A_1 - \ln A_2 + \lambda_1' r_1 T_c + \lambda_2' r_2 T_D(2)}{\lambda_1' r_1 + \lambda_2' r_2}$$

where r_1 and r_2 are the average data rates of GOF 1 and GOF 2, respectively, $T_D(2)$ is the transmission deadline for GOF 2, and T_c is the current time.

IV. PERFORMANCE EVALUATION

Simulations are conducted to evaluate the performance of the proposed content-based retransmission with two-step optimization. 3-D wavelet video bitstreams generated by subband based coding [13] are delivered across a simulated packet erasure channel with independent packet loss rate. The metric of evaluation is the PSNR of the received video. An error-free feedback channel is assumed. The performances of four different transmission schemes are compared: UDP/IP without retransmission (T1), received-based retransmission (T2), content-based retransmission with intra-GOF optimization only (T3), and content-based retransmission with two-step optimization (T4).

TABLE V
AVERAGE PSNRS (IN DECIBELS) WITH HIGH PACKET LOSS RATE (20%)

Video Seq.	T1	T2	T3
Notting Hill	28.5122	29.9875	31.2541
Table Tennis	24.6821	22.7166	26.8667
Grand-mom	28.4871	30.9301	31.1814

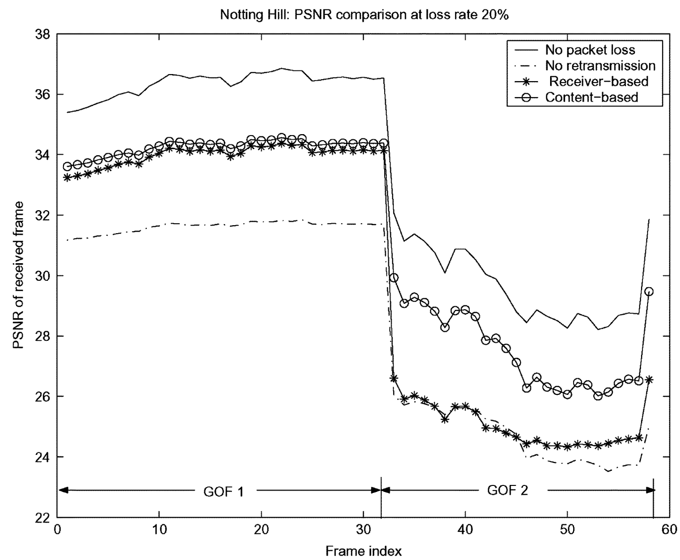


Fig. 7. PSNR comparison of none/receiver-based/content-based retransmission schemes with the packet loss rates of 20%.

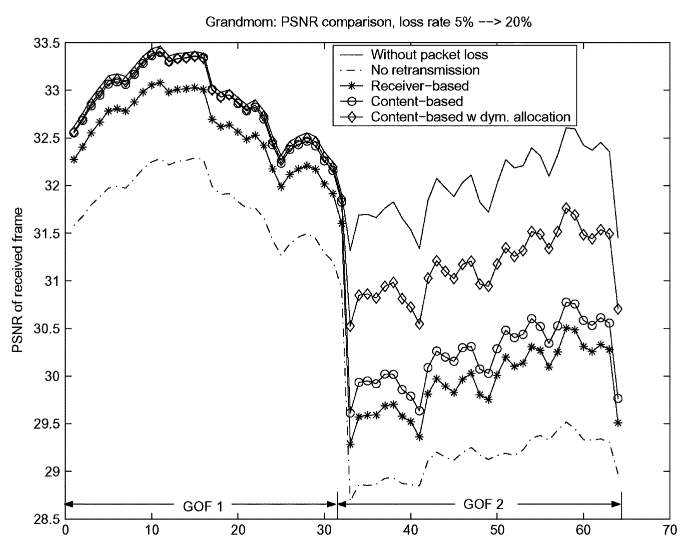


Fig. 8. PSNR comparison of none, receiver-based, content-based retransmission with intra-GOF optimization, and content-based retransmission with both inter-GOF and intra-GOF optimizations schemes with changing packet loss rate.

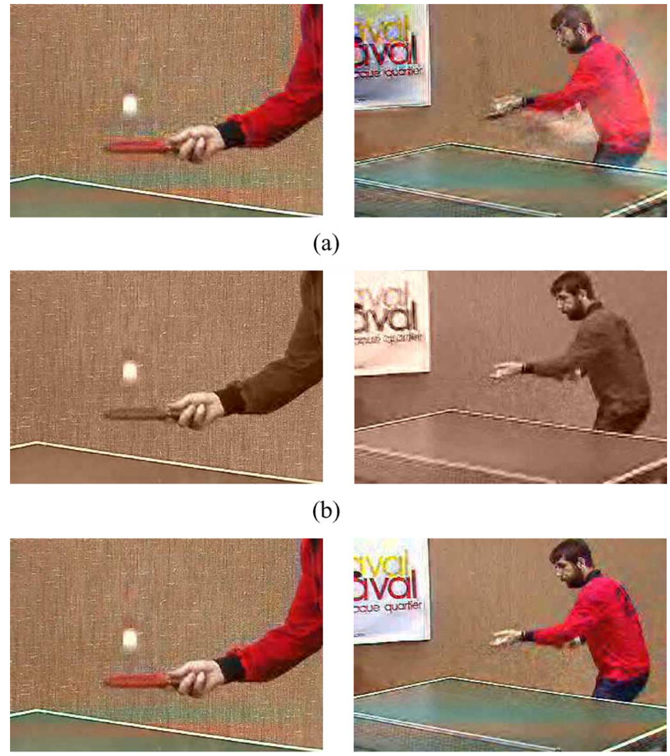
Table III summarizes the simulation parameters for three test video sequences with different characteristics. Among them, the video sequence *Table Tennis* contains fast motions in each GOF, *Grandma* contains slow and smooth motions of the "head and shoulder" type, and in *Notting Hill*, the motions are initially slow, and become much faster with scene changes.

TABLE VI
AVERAGE PSNR (IN DECIBELS) OF THE RECEIVED VIDEO SEQUENCES UNDER DIFFERENT TRANSMISSION SCHEMES (T1, T2, T3, T4) WHEN PACKET LOSS RATE CHANGES FROM 5% TO 20%

Video Seq.	T1	T2	T3	T4
Notting Hill	29.3905	31.2701	32.1777	32.1818
Table Tennis	25.4975	25.3205	26.8865	27.2427
Grandmom	30.4715	31.2432	31.5324	32.0023

(a) (b) (c)

Fig. 9. 1st and 58th reconstructed video frames in *Notting Hill* (packet loss rate 20%). (a) Conventional UDP/IP without retransmission. (b) Receiver-based retransmission. (c) Content-based retransmission.



(a) (b) (c)

Fig. 10. 1st and 63rd reconstructed video frames in *Table Tennis* (packet loss rate 5%–20%). (a) Conventional UDP/IP without retransmission. (b) Receiver-based retransmission. (c) Content-based retransmission.

Simulation results under different traffic conditions are presented as follows:

- **Constant packet loss rate**: Tables IV and V present the simulation results under low (5%) and high (20%) packet loss rates, respectively. It is clear that the content-based retransmission outperforms other transmission schemes in general. In addition, the performance gain is more significant when packet loss rate is higher.

  Fig. 7 plots the frame-by-frame PSNR of the video *Notting Hill* to show the effect of the packet loss in more details when the loss rate is 20%. It can be seen that the second GOF suffers more from the packet loss since it contains larger motions. The simulation results show that the receiver-based retransmission scheme has unstable performance since it does not consider the contents of the packets and extensive retransmission may cause the dropping of the future packets due to delay constraint. If the dropped future packets are more important than the retransmitted ones, the quality of the received video will be degraded, as demonstrated by the experimental results.
- **Changing packet loss rate (from 5% to 20%)**: Table VI compares the simulation results when the packet loss rate changes from 5% to 20% during transmission. Since the packet loss rate varies with transmission time in this scenario, inter-GOF optimization is used to achieve a smooth transmission quality. Again, the frame-by-frame PSNR of video *Grandmom* is plotted in Fig. 8 to show the influence of the changing traffic condition. From Table VI and Fig. 8, it is clear that the content-based retransmission using two-step optimizations has the best performance while the content-based scheme using only intra-GOF ranks the second. Furthermore, inter-GOF optimization is effective in reducing the quality variation among different GOFs.

As an objective measurement, sometimes PSNR cannot represent the visual effect very well. To show the effectiveness of the proposed error control method, some reconstructed video frames are displayed. Fig. 9 displays the reconstructed video frames (the 1st and 58th in the video *Notting Hill*) of the different transmission schemes when the channel loss rate is 20% in packet erasure channel; Fig. 10 displays the reconstructed video frame with varying packet loss rate. Apparently, our proposed approach achieves the best visual quality. Compared to protocol-based error control method, it can significantly enhance the overall quality while reducing the quality variation of the received video.

V. CONCLUSION

In this paper, we have presented a novel content-based retransmission with two-step optimization for 3-D wavelet video streaming. The goal is to minimize the damage caused by packet loss while keeping the video quality consistent. The two steps, intra-GOF and inter-GOF optimizations, are tightly related since the latter is based on the rate-distortion model achieved by the former algorithm. Simulation results show that the proposed method is effective in combating the time-varying packet loss rate in video streaming.

REFERENCES

[1] P. A. Chou and Z. Miao, "Rate-distortion optimized sender-driven streaming over best-effort networks," in *Proc. IEEE Workshop Multimedia Signal Process.*, Oct. 2001, pp. 587–592.

[2] P. A. Chou and Z. Miao, "Rate-distortion optimized streaming of packetized media," *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 390–404, Apr. 2006.

[3] P. A. Chou and Z. Miao, Rate-distortion optimized streaming of packetized media, Microsoft Research Center, Redmond, WA, Feb. 2001, Tech. Rep. MSR-TR-2001-35.

[4] Z. Miao and A. Ortega, "Optimal scheduling for streaming of scalable media," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 2000, pp. 1357–1362.

[5] Z. Miao and A. Ortega, "Expected run-time distortion based scheduling for delivery of scalable media," in *Proc. 12th Int. Packet Video Workshop (PVW 2002)*, Apr. 2002, pp. 687–693.

[6] M. Podolsky, M. Vetterli, and S. McCanne, "Limited retransmission of real-time layered multimedia," in *Proc. IEEE Workshop Multimedia Signal Process.*, Dec. 1998, pp. 591–596.

[7] Z. L. Zhang, S. Nelakuditi, R. Aggarwal, and R. P. Tsang, "Efficient selective frame discard algorithms for stored video delivery across resource constrained networks," in *Proc. IEEE INFOCOM'99*, Mar. 1999, pp. 472–479.

[8] M. Gallant and F. Kossentini, "Rate-distortion optimized layered coding with unequal error protection for robust Internet video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 357–372, Mar. 2001.

[9] C. L. Chang, S. Han, and B. Girod, "Rate-distortion optimized streaming for 3-D wavelet video," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2004, pp. 3141–3144.

[10] C. L. Chang, S. Han, and B. Girod, "Sender-based rate-distortion optimized streaming for 3-d wavelet video with low latency," in *Proc. IEEE 6th Workshop Multimedia Signal Process.*, 2004, pp. 510–517.

[11] J. Chakareski, J. G. Apostolopoulos, E. Wee, W. T. Tan, and B. Girod, "Rate-distortion hint tracks for adaptive video streaming," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 10, pp. 1257–1269, Oct. 2005.

[12] J. Rogers and P. Cosman, "Wavelet zerotree image compression with packetization," *IEEE Signal Process. Lett.*, vol. 5, no. 5, pp. 105–107, May 1998.

[13] J. Dong and Y. F. Zheng, "An innovative framework for 3-D wavelet video streaming," in *Proc. SPIE AreoSense'2002*, Apr. 2002, vol. 4736, pp. 127–138.