MPEG-4 FGS 비디오 스트리밍에 대한 네트워크 차별화 서비스의 성능분석

정회원 신 지 태*, 김 종 원**

Performance Evaluation of Differentiated Services to MPEG-4 FGS Video Streaming

Jitae Shin*, JongWon Kim** *Regular Members*

요 약

ISO/IEC MPEG-4 FGS (finer granular scalable) 비디오 스트림을 패킷손실 측면의 차등서비스 (differentiated services: DiffServ) 네트워크상에서 차별화 전송하는 시스템을 제안하고, 그 성능을 분석한다. 이를 위한 전체 제안시스템의 구조는 크게 다음의 3 부분으로 나눌 수 있다. 즉 1) 선형 근사화한 전송율-왜곡치 (rate-distortion: R-D) 모델를 사용하여 비디오 품질를 일정하게 유지하는 최적의 계층화된 전송율 적응 제어 부분, 2) 각각의 비디오 패킷이 손실될 때 전체품질에 미치는 영향을 고려하는 우선순위 패킷화 (prioritized packetization) 부분, 그리고 3) 이와 같이 우선순위화된 비디오 패킷 스트림을 차등서비스 네트워크 상에서 차별화 전송을 수행하는 부분으로 구분할 수 있다. 따라서 상기한 3 부분들이 효율적으로 연동되어 비디오 전송을 수행할 때, 동일한 네트워크 자원이 주어진 경우 얻을 수 있는 종단간 (end-to-end) 비디오 품질의 향상을 비교 분석하였다.

ABSTRACT

A finer granular scalable (FGS) version of ISO/IEC MPEG-4 video streaming is investigated in this work with the prioritized stream delivery over loss-rate differentiated networks. Our proposed system is focused on the seamless integration of rate adaptation, prioritized packetization, and simplified differentiation for the MPEG-4 FGS video streaming. The proposed system consists of three key components: 1) rate adaptation with scalable source encoding, 2) content-aware prioritized packetization, and 3) loss-based differential forwarding. More specifically, a constant-quality rate adaptation is first achieved by optimally truncating the over-coded FGS stream based on the embedding rate-distortion (R-D) information (obtained from a piecewise linear R-D model). The rate-controlled video stream is then packetized and prioritized according to the loss impact of each packet. Prioritized packets are transmitted over the underlying network, where packets are subject to differentiated dropping and forwarding. By focusing on the end-to-end quality, we establish an effective working conditions for the proposed video streaming and the superior performance is verified by simulated MPEG-4 FGS video streaming.

* 성균관대학교 정보통신공학부 멀티미디어 네트워킹 연구실 (jtshin@ece.skku.ac.kr),
** 광주과학기술원 정보통신공학과 네트워크 미디어 연구실 (jongwon@kjist.ac.kr)
논문번호 : 020116-0313, 접수일자 : 2002년 3월 13일
※ 본 연구는 두뇌한국 21 (BK-21) 사업의 일부 지원으로 수행되었습니다.

I. Introduction

Streaming media applications over IP (Internet protocol) networks are attracting ever-increasing attention as the enabling technology for future multimedia distribution. Streaming offers significant improvements over download-and-play approach for on-/off-line media forwarded from a media server. Users may have different preferences, processing capabilities, diverse network accesses to the streaming media server. Especially for streaming video, this user and network heterogeneity requires both highly scalable video coding and flexible delivery techniques to overcome the challenges imposed by best-effort Internet.

The unpredictable channel variation requires finer granularity than what can be provided with the layering options of conventional MPEG-2 and H.263+ video. The complex dependency geared for coding efficiency poses another bottleneck since it hurts the video robustness under erroneous environment. The fine grain scalability (FGS) of MPEG-4 [1] is one big step towards scalable video solution where the base layer is targeted to provide the basic visual quality to meet the minimal user bandwidth, and the scalable enhancement layer can be arbitrarily truncated to meet the heterogeneous network conditions. With the help of this scalable stream, the video streaming is much simplified since all the transcoding overhead required in the non-scalable codec is bypassed. However, scalable coding only solves the part of the problem, and packet loss is very common with the unpredictable channel condition. To address this problem fully, both efficient scalable coding scheme and flexible delivery technique are needed. The application-oriented people is working by starting from the current best-effort network model to find more innovative streaming scheme to mitigate the effect from this unpredictable packet loss [2,3]. That is, by rate adaptation and error control, the application-layer Quality of Service (QoS) is provided to the end user. Traditionally, rate adaptation and error control are usually investigated separately. However, rate adaptation and error control affect each other, hence, several recent works attempt to integrate these aspects and consider them together. In [4], the rate adaptation is performed to smoothly adjust the sending rate of MPEG4-FGS encoded video based on the estimated network bandwidth, and then each packet is protected unequally. However, this sender-oriented rate adaptation is mainly for unicast video rather than multicast one. Moreover, forward error correction (FEC) level for unequal error protection is decided in a heuristic manner (i.e., without optimization for end-to-end quality). In [5], one fully receiver-driven approach for joint rate adaptation and error control is proposed with pseudo-automatic repeat request (ARQ) layers. The sender injects multiple source/channel layers into network, where the delayed channel layers (relative to the corresponding source layers) serve as packet recovery role. Each receiver performs rate adaptation and error control by subscribing selected number of source and channel layers according to the receiver's available bandwidth and channel condition. However, as shown in [6], the entire receiver-driven approach are subject to several drawbacks including persistent instability in video quality, arbitrary unfairness with other sessions, and difficult receiver synchronization.

The network infrastructure-oriented approach, on the contrary, is promoting more QoS support in network node. Two representative approaches in the Internet engineering task force (IETF) are the integrated services (IntServ) with the resource reservation protocol (RSVP) and the differentiated services (DiffServ or DS) [7]. Those IP-QoS methods are more suitable in accommodating various QoS requirements of different applications than the best-effort model. Between these two main IP-QoS approaches, the DiffServ scheme provides a less complicated and scalable solution since IntServ requires to maintain per-flow state across the whole path for resource reservation. In the DiffServ model, resources are allocated differently for various aggregated traffic flows based on a set of bits (i.e., the DS codepoint bits). Consequently, the DiffServ approach allows different QoS grades to different classes of aggregated traffic flows. Two services are supported; the premium service(PS) that support low loss and delay/jitter, and an assured service (AS) that provides QoS better than best effort but without guarantee. Several previous works have been performed on the non-scalable (i.e., coarse granular) video streaming over QoS provision network [8,9]. Significant gains from the resulting unequal error protection(UEP) are usually claimed. However, the gains are often so conspicuous, especially when the source and network parameters do not match well.

By applying error-resilient scalable source coding, constant quality rate adaptation and packet prioritization, and DiffServ-based prioritized network, we tackle this challenging problem from both application and network viewpoints. The scalable codec is based on the ongoing standardized MPEG-4 FGS coding, whose scalable stream allows to be truncated arbitrary according to the rate budget. The main contribution of this paper is to propose a framework that integrates rate adaptation with scalable MPEG-4 FGS coding into the DiffServ framework through a rate-distortion (R-D) oriented finer granular packet priority and source prioritized packetization. It can be further divided into the followings in detail.

First, an optimal truncation strategy for rate adaptation of enhancement layer (EL) is realized by embedding the minimal R-D information and relying on a piecewise linear R-D model. During encoding process, R-D sample points are generated and embedded for each bitplane in EL. It is then piecewise-linear interpolated to serve as the R-D model of EL. With this assisted R-D model, the rate adaptation can be easily achieved for optimal distortion or its variation. Second, following MPEG-4 packetization principle, fixed size packets are generated and they are prioritized with a basis on their impact to the end-to-end video quality. For EL, from the above piecewise R-D model, the loss impact (i.e. distortion increase) of each EL packet can be easily calculated and the priority of each packet can be determined. MPEG-4 FGS base layer (BL) is known to be less flexible / more error sensitive than EL[4]. At times of severe network loading or provisioning mismatch, even the BL will get lost and render the quality unacceptable. To protect those BL's, error resilience coding [10], source-/channel-level unequal error protection (UEP) [11], or optimal packetization[12] may be employed. In our work, following MPEG-4 packetization principle, fixed-size packets are generated for both BL and EL. The loss impacts of BL packets to the end-to-end quality are measured and priority is assigned accordingly. Similarly, for EL packets, the loss impacts(i.e., distortion increase) are calculated from piecewise R-D model. With the fine granular packet priority, more graceful quality degradation is achieved than [4] that does not differentiate the packets. Third, The proposed framework takes full advantage of the integration of the fine granular scalable video coding into the QoS-enabled DiffServ network. After careful examination from both source and network angle, an appropriate DiffServ service model is selected to efficiently handle this FGS stream. To avoid unpractical performance bias in favor of UEP over equal error protection (EEP), we attempt a fair comparison between UEP and EEP examined for both BL and EL. From the above UEP/EEP comparison, we exploit and evaluate several deployment scenarios. As a result, the differentiated forwarding of FGS video shows sufficient efficiency and flexibility to overcome the short-term network variation and lower-to-middle range packet loss. Thus, by leveraging this, the rate adaptation at the sender is required to match only longer-term network variation (i.e. less frequent rate adaptation).

The rest of this paper is organized as follows. In Section 2, we present the overview of whole framework where the key components such as the scalable video coding and DiffServ network model are identified. In Section 3, we discuss the details of rate adaptation and prioritized packetization for MPEG-4 FGS codec with minimal distortion variation, and the error resilient MPEG-4 FGS coding streams of interaction with DiffServ model. In Section 4, we demonstrate that the video streaming can benefit from the rate adaptation, prioritized packetization as well as the differentiated forwarding. We also analyze the effect of both source and network operations for BL/EL UEP/EEP. Finally, we conclude our work and point out future directions in Section 5.

II. Overview of the Proposed System

The goal of this work is to incorporate the efforts of traffic sources like rate adaptation upon available network resources and the efforts of network providing network QoS. Our proposed system is focused on the integration of rate adaptation with MPEG-4 FGS video streaming, prioritized packetization based on contents, and simple loss DiffServ. The overview of the proposed scalable video streaming system with DiffServ network QoS provisioning is shown in Fig. 1. The system consists of three key components: (1) rate adaptation with scalable source encoding, (2) content-aware prioritized packetization, and (3) differential forwarding. They are briefly described below:

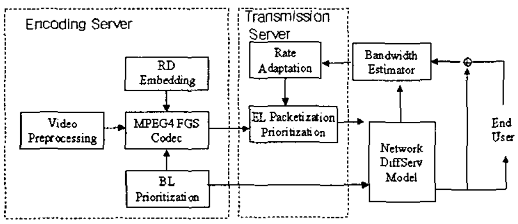


Fig. 1 Overview of the proposed scalable video streaming system with network QoS provisioning

1. Rate adaptation with scalable coding

The source is first encoded by the MPEG-4 FGS codec, where the estimated minimal bandwidth gives the bandwidth constraint for BL. In the coding process, the BL is encoded with a two-pass process, where the first pass focuses on analyzing the sequence, and the second pass performs the real encoding based on the gathered statistics. For the EL, R-D samples are generated for each bitplane and embedded as the VOP user data. For non-real time applications, the over-coded bitstream is prestored in the streaming server. Upon the streaming request, the rate adaptation module takes place to scale the EL stream based on the feedback of the available bandwidth to increase quality by referring to the embedded R-D samples.

2. Prioritized packetization

The rate-adapted stream that contains both BL and EL is then packetized. The fixed-length packetization is adopted for both EL and BL streams as recommended by MPEG-4 [1]. By evaluating the anticipated loss impact of each packet on the end-to-end video quality (i.e. considering the loss impact to itself and depending packets), we assign a priority index, called the relative priority index (RPI), to each packet within the priority range of each layer. The priority assignment of a packet is dynamically determined in the media adaptation module, since not only the concerned packet but also its children packets (which is actually dynamically varying) will affect the assignment.

3. Differentiated forwarding

For streaming video applications, where encoding/decoding is more resilient to packet loss and delay fluctuations, DiffServ AS seems to be a better match. In this paper, we focus on the DiffServ AS, especially the relative service differentiation, for streaming video applications. Note that MPEG-4 FGS originally assumes guaranteed delivery (e.g. DiffServ PS) of the BL and leaving the EL to the mercy of the best-effort Internet. Here, we relax this requirement partly to allow a broader range of applicability and partly to avoid the high cost of DiffServ PS. There exist several ways to realize DiffServ AS, especially the proportional (i.e. relative) DiffServ. For a more detailed description, we refer to [13,14]. This research adopts the proportional DiffServ model described in [13] for network simulation. In terms of packet loss, the proportional DiffServ model demands that loss rates of different DS levels are spaced as

$$\frac{\overline{l_i}}{\overline{l_j}} = \frac{\sigma_i}{\sigma_j}, \quad 1 \le i, j \le N \tag{1}$$

where $\overline{l_i}$ is the average loss rate for DS level i, and σ_i, i = 1, · · ·, N are loss differentiation parameters ordered as $\sigma_1 > \sigma_2 > \ldots > \sigma_N > 0$.

With the assigned priority, these packets are sent to the DiffServ network to receive different forwarding treatment [13]. By mapping these prioritized packets to the different QoS DS levels, packets will experience different packet loss rates with this differential forwarding mechanism. This transport-side prioritization may be accompanied by the application-side prioritization such as FEC and ARQ. Besides prioritized dropping performed by DiffServ routers, traffic policing can be explicitly carried out at intermediate video gateways/filters (e.g. inside the active DiffServ routers or other special network devices) by using packet filtering. Thus, based on the assigned RPI, rate adaptation and error control can be jointly performed in the proposed system.

Ⅲ. Detailed Components of Proposed Systems

1. Rate Adaptation with scalable coding

R-D samples of each bitplane introduces a negligible amount of overhead and demands a very low computational complexity. To obtain the distortion information, there are two major approaches in traditional rate control schemes. One approach as in [15] fully relies on the closed form model, however, it is found inaccurate at low bit rates. To overcome the inaccuracy existing in closed form models, Lin and Ortega utilize a set of RD samples to approximate the complete R-D relationship by a cubic interpolation [16]. Besides the model itself accuracy, the keys of RD sampling based approximation are the low complexity and low overhead. Here we proposed one low overhead, low computation and relatively accurate RD information embedding for MPEG4 FGS EL as follows. The EL is coded bitplane by bitplane, intuitively, the RD characteristic should be uniform with the same bitplane since the distortion reduction is approximately determined by the quantization parameters(QPs) to which the concerned bitplane corresponds. That is, only the RD points at the beginning of each bitplane is needed to be embedded and calculated. Typically, there are only few bitplanes i.e., 6 ~ 7 bitplanes corresponding the wide range QP (e.g., from 1 to 26 or 27 times QP). Moreover, since the DCT is the unit transform which is invariant to the pixel variance, the distortion associated with each bitplane RD points can be directly calculated in the coefficient domain. This bitplane associated RD sample generation incurs negligible overhead and computational complexity without affecting the original encoding process. The generated R-D samples can be either stored in the user data of each VOP or meta data in a separate file.

Given a bit allocation R_{EL}, suppose $R_i < R_{EL} < R_{i+1}$. Here R_i and R_{i+1} corresponds to rate of bitplane i and i + 1. The corresponding distortion of R_{EL} is then equal to

$$c_{g,i} = \frac{R_{i+1} - R_{EL}}{R_{i+1} - R_i} \times (D_i - D_{i+1}) + D_{i+1} \tag{2}$$

In Fig. 2, the piecewise-linear R-D curve obtained via interpolation('Interp' in Fig.2(b)) is compared to the empirical curve('Real' in Fig.2(b)) for the 1st, 15th, and 30th frames(i.e., F1, F2, and F3) of the Foreman CIF sequence.

It is verified that the piecewise-linear R-D model can approximate the empirical R-D curve quite well. With this interpolated R-D model, we adopt a sliding-window approach to perform rate adaptation. Suppose that window W_i includes M_{W_i} frames. Then, the rate allocation to achieve constant quality can be performed within each

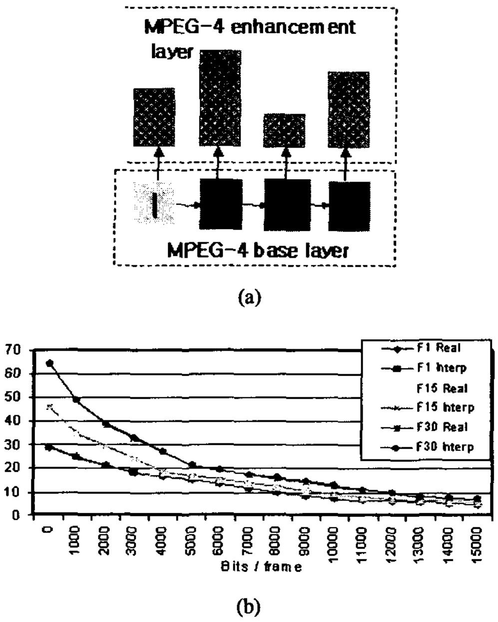


Fig. 2 (a) MPEG-4 FGS scalable structure (b) The comparison of interpolated and real (i.e., empirical) distortion [Y-axis represents mean square error(MSE)]

window independently. In terms of mathematics, the problem can be written as

$$\min \sum_{j \in W_i} \| D_j - D_{j-1} \|. \tag{3}$$

subject to

$$\sum_{j \in W_i} B_j \le \frac{M_{W_i} \cdot R_{W_i}}{FR_{W_i}} - B_{BL}$$

where FR_{Wi} is the encoding frame rate inside the window W_i, R_{Wi} represents the available bit rate (bits/sec) at the start time of window i and B_{BL} is the total bit budget for BL within this window. To reach the optimal solution, binary search is utilized to find the best D_{Wi}, which minimizes the distortion variation among frames. That is,

- Step 1: Take the minimal frame distortion of all BL frames within one sliding window as the initial value of D_{wi} .
- Step 2: Calculate $\sum_j B_j$ based on the given DWi by using the piecewise linear R-D model.
- Step3: If $\sum_j B_j < \left(\frac{M_{W_i} \cdot R_{W_i}}{FR_{W_i}} - B_{BL} \right) - \Delta B - \delta$, then set $D_{W_i} = \frac{D_{W_i} + D_{low}}{2}$, $D_{high} = D_{W_i}$ and return to Step 2. Else, if $\sum_j B_j > \left(\frac{M_{W_i} \cdot R_{W_i}}{FR_{W_i}} - B_{BL} \right) - \Delta B + \delta$, then set $D_{W_i} = \frac{D_{W_i} + D_{high}}{2}$, $D_{low} = D_{W_i}$, and return to Step 2. Note that δ is negligible factor to control rate adaptation accuracy. Else, continue with Step 4.
- Step 4: Move to the next window, update by $\Delta B = B_{real} - \frac{M_{W_i} \cdot R_{W_i}}{FR_{W_i}}$, $R_{w_i+1} = R_{W_i + M_{W_i}}$.
  B_{real} means the bits actually required to transmit the chosen portion of EL stream and thus ΔB is the offset between the real and reference bit allocation. If we are not at the end of sequence, go back to Step 1.

Only a few iterations are needed in the above binary search of R_{Wi} and D_{Wi} . Since a simple interpolation scheme is needed to calculate these values, its complexity is low enough to be performed in real-time. This approach works best for slowly varying channel conditions. The proposed approach is still acceptable for a fast varying channel when differentiated forwarding can mitigate packet losses due to inaccurate rate adaptation.

2. Prioritized packetization

The packetization scheme has a significant effect on the efficiency and error-resiliency of video streaming. Two packetization schemes are used in the context of video streaming, i.e. the variable-length packetization (e.g., GOB packets of H.263+) and the fixed-length packetization (e.g. MPEG-4 video packets), where video packets of a similar length are formed. The packet size is related to efficiency and error-resilency, since a smaller packet size demands a higher overhead but is more resilient to errors. Recently, to improve error resiliency, a discrete optimization problem to minimize the distortion was formulated in the packetization of an embedded stream [12]. In this work, we follow the fixed-length packetization

(FLP) due to the following considerations. First, FLP can avoid the inefficiency of very small size GOBs. More importantly, compared with GOB-based packetization, FLP automatically separates a larger motion region into several packets while grouping several static regions into one single packet. The coding of a motion region into multiple packets is able to spread the loss impact, thus increasing error resiliency.

Each packet is then assigned with certain priority according to its impact to end-to-end visual quality. For different service preferences in terms of loss and delay, the priority can be further divided into the relative loss index (RLI) and the relative delay index (RDI) as given in [8]. If the assigned priority reflects the impact of each packet to end-to-end quality well, graceful quality degradation can be achieved by dropping packets with respect to the priority index. To determine the packet priority with a low complexity is an active research area today. Several features such as the initial error strength (i.e., in MSE by assuming the packet loss concealed), its propagation via motion vectors and the spatial filtering effect were used to develop a corruption model in [15] to determine the packet priority in terms of the loss impact. For BL packets, we adopt the accurate priority rather than its approximation. That is, we empirically measure the overall MSE when the packet of concern gets lost. By doing so, we can analyze the gain from differentiated forwarding. Besides, since BL is normally determined by the minimal bandwidth (which is typically fixed), the priority of BL packets can be calculated off-line. For EL packets, the priority assignment is simplified due to the strict separation of frames along the temporal direction. The packet loss within EL only affects a single frame, and it does not propagate. The incurred distortion from each EL packet can be accurately calculated within each frame. The packet priority can be calculated as

$$\rho_i = \frac{\Delta D_i}{\Delta R_i}\ , \tag{4}$$

where ΔD_i represents the incurred distortion due to the specified loss, and ΔR_i is the rate of the packet of concern. In addition, the packet dependency has to be taken into consideration such that if packets containing more significant bitplane get lost, packets containing the less significant bitplane in the same region should be discarded anyway. Hence, the final packet loss index can be calculated as

$$RLI_i^{(EL)} = \sum_{s \in SD_i} \rho_s + \rho_i \tag{5}$$

where SD_i is the descendent set of packet i. By using the piecewise linear R-D model for each bitplane, the priority of EL packets can be easily calculated on-line during the packetization procedure.

3. Differentiated Forwarding

With each packet assigned with a certain priority, differentiated forwarding can be employed accordingly. We first describe the error resilient coding for both BL and EL in MPEG-4 FGS, and then examine the DiffServ model to perform the differentiated forwarding mechanism to the error-resilient video streams. For non-scalable BL of MPEG-4 FGS, several error resilient tools have been recommended by MPEG-4. They include: video packet, data partitioning (DP), reversible variable length code (VLC), and cyclic/adaptive intra refresh (CIR/AIR), and NewPred (also known as reference picture selection or RPS).

DP and RVLC are mainly for the partial decoding of video packets with bit errors, and it is less relevant to the Internet where the packet loss is dominant. CIR and AIR are encoding choices and AIR improves CIR by using more intelligent refresh based on the motion and so on.

In our approach, both AIR and CIR are applied to the BL and their performance is compared for the UEP as well as the EEP scenarios. Also, intelligent error concealment (EC) scheme shown in Fig. 3(a) is adopted, where lost MBs in the P-frame are interpolated from those of upper and lower MBs based on the motion.

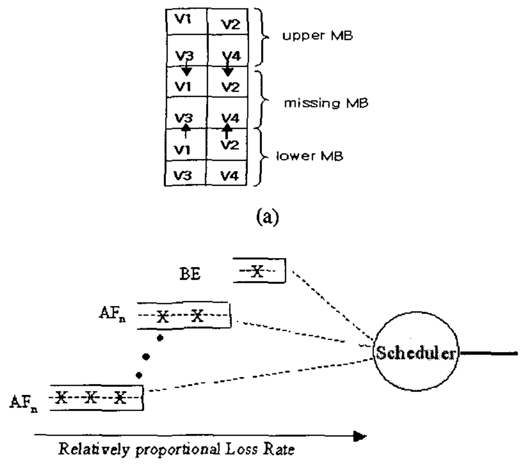


Fig. 3 (a) Error concealment scheme used in BL
(b) The DiffServ node with multiple class-queues.

Finally, as mentioned before, the fixed length packetization is applied to both BL and EL streams. Differentiated forwarding of the non-scalable ITU-T H.263+ stream was discussed in [8], where proportional DS levels were used in prioritized video streaming. For scalable MPEG-4 FGS, it is usually assumed that the available bandwidth is sufficient to cover the BL stream because relatively small bandwidth is required for the BL. Compared to EL packets, the BL packets are very important and should be protected effectively. That is, to secure reliable transmission, the BL stream should be mapped to higher priority DS levels. Even if the available bandwidth goes below the rate demanded by BL incidently, we may assume that some minimal bandwidth (which could be smaller than the BL rate) is still sustained. By prioritizing the BL stream and protecting it accordingly, we can preserve its minimal quality by delivering at least the highest priority portion of BL packets. Thus, as a simple test case, we consider three BL categories for relative proportional DiffServ. For EL packets, another lower priority class queue (i.e. DS level) with two different drop preferences is assigned y. Thus, both BL and EL have three different drop preferences. The multiple class-queue DiffServ node as illustrated in Fig. 3(b) performs the differentiated forwarding policy.

In addition to priority dropping, rate adaptation

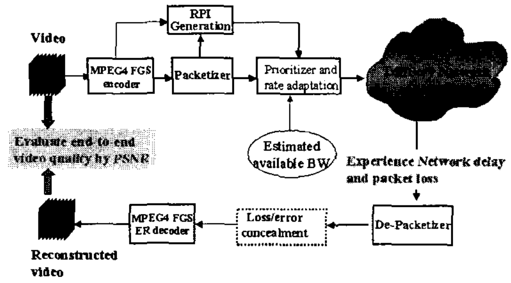


Fig. 4 The diagram of proposed experiment setup.

can significantly reduce network congestion when a good estimate of the available bandwidth is achieved. Rate adaptation can be performed at either the server side or the edge router in the DiffServ network. Packets can be dropped in advance by rate adaptation strictly following the priority order. Compared to priority dropping in differentiated forwarding, rate adaptation provides more graceful quality degradation when the available bandwidth becomes small. However, we have to pay some price to achieve the goal, e.g. the complexity required to estimate the available bandwidth, and the inevitable time delay between bandwidth measure and rate adaptation. Thus, in the proposed system, a compromised solution is suggested. When there is a big change in the available bandwidth, rate adaptation is performed. Otherwise, only DiffServ forwarding is employed.

Ⅳ. Experimental Results

In this section, we will demonstrate the performance of proposed system by enabling or disabling rate adaptation, or enabling or disabling the prioritized transmission. Several typical scenario appearing in video streaming application are identified and evaluated. The gains of prioritized transmission compared with the non-prioritized ones are compared in detailed. The proportional Diffserv with simple three-level loss differentiation and the overall experiment setup is illustrated in Fig. 4.

1. Scenario 1 : Differentiated Forwarding of BL packets

As discussed, the MPEG-4 FGS codec is built on the assumption that the available bandwidth is sufficient to transmit BL packets. However, in case of severe network congestion, it is desirable to degrade the quality of BL gracefully. Here, we first compare the performance of prioritized transmission of BL packets with that of non-prioritized one. The priority transmission adopts the proportional DiffServ model with three DS levels for proportional loss rates. One interesting issue here is how to map the continuously prioritized packets to three discrete DS levels. One possible solution is the mapping strategy proposed in [8], where packets are mapped to a limited number of DS levels, with a goal to minimize quality degradation under a pricing mechanism. However, for simplicity, we adopt a simpler QOS mapping policy in this simulation by adopting a direct mapping from RLI to DS levels.

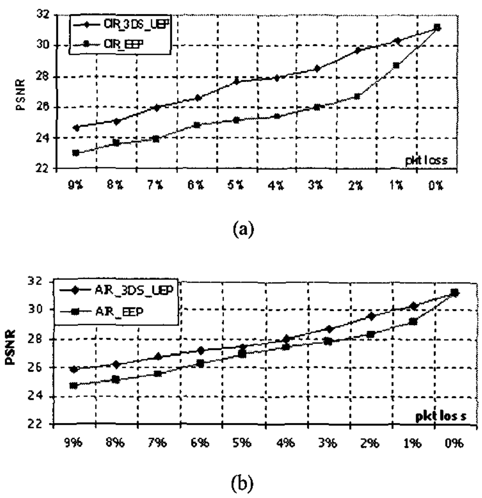


Fig. 5 The PSNR comparison of EEP/UEP for BL packets with 3 different drop-level DiffServ (a) case 1: Gain 0.97dB, (b) Case 2: Gain 2.13dB.

All packets are clustered into three groups, each of which has a similar number of packets, and each group of packets is mapped to one DS level.

For the simulation setup in Fig. 4, we have to set various parameters in the encoding and the packetization modules. Selected parameters are shown in Table 1. The BL packets are encoded by using the MPEG-4 FGS codec with MPEG-2 TM5 rate control at 128 kbps. The CIF sequence is encoded at 10 fps with a leading I-frame followed by all P frames. Since packet loss from the initial I-frame is too catastrophic, we limit the packet loss only to P-frames. Both AIR and CIR simulations are performed and compared. The performance gain from those two modes are shown in Fig. 5(a) and 5(b). The DiffServ transmission has a clear gain in terms of PSNR under the same bit budget and the overall packet loss ratio. It is interesting to note that gains are varying significantly in two cases as shown in Table 1. The less gain is obtained from Case 1. The RLI distribution of BL packets under Case 1 is illustrated in Fig. 6(a).

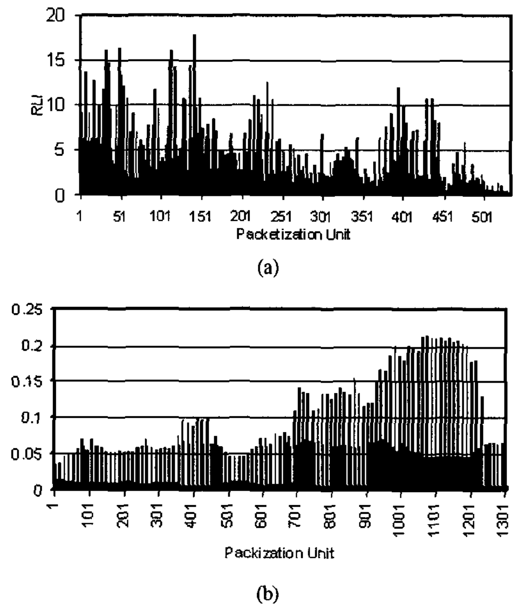


Fig. 6 The RLI distribution for packets in the 'Foreman' sequence: (a) BL packets under Case 1 and (b) EL packets under 384kbps, where y axis represents 'RLI'.

2. Scenario 2: Differentiated Forwarding of EL Packets

BL and EL packets are usually protected by

Table 1. Simulation Setup for BL of MPEG-4 FGS codec

Case	Bitrate	Frame rate	Rate control	GOP mode	ER options	Packet size
1	128kbps	10fps	TM5	IPPPP....	VP, AIR	400 bytes
2	128kbps	10fps	TM5	IPPPP....	VP, CIR	400 bytes

two error protection levels. In [4], different bitplanes within EL are unequally protected. However, even the same bitplanes in different frames may have different contributions to the end-to-end visual quality so that they may be protected unequally. For example, EL packets of a low-quality BL frame typically has a higher impact than that of a high-quality BL frame. Thus, we propose to apply differentiated forwarding to EL packets. By utilizing the R-D sample derived priority, we prioritize each EL packet and perform differentiated forwarding accordingly. The RLI distribution of EL packets at a rate of 384kbps are given in Fig. 6(b).

We show the performance advantage of priority dropping in DiffServ over uniform dropping at 512 kbps, 384 kbps, 256kbps, and 160 kbps of EL packets in Figs. 7(a)-(d), respectively. As shown in these figures, prioritized transmission has a clear gain in PSNR under the same bit budget and the same packet loss ratio. We also present the ideal performance of rate adaptation when an accurate estimate of the bandwidth is available. More graceful quality degradation can be achieved since rate adaptation is efficient in dropping packets in a strict order of priority. Another observation is that the advantage of rate adaptation over UEP and the advantage of UEP over EEP are varying under different packet-loss ratios and different EL bit rates.

Under a high packet loss ratio (i.e. 25 percent or more), rate adaptation has a significant gain over UEP. On the other hand, the gain of rate adaptation over UEP becomes smaller under a

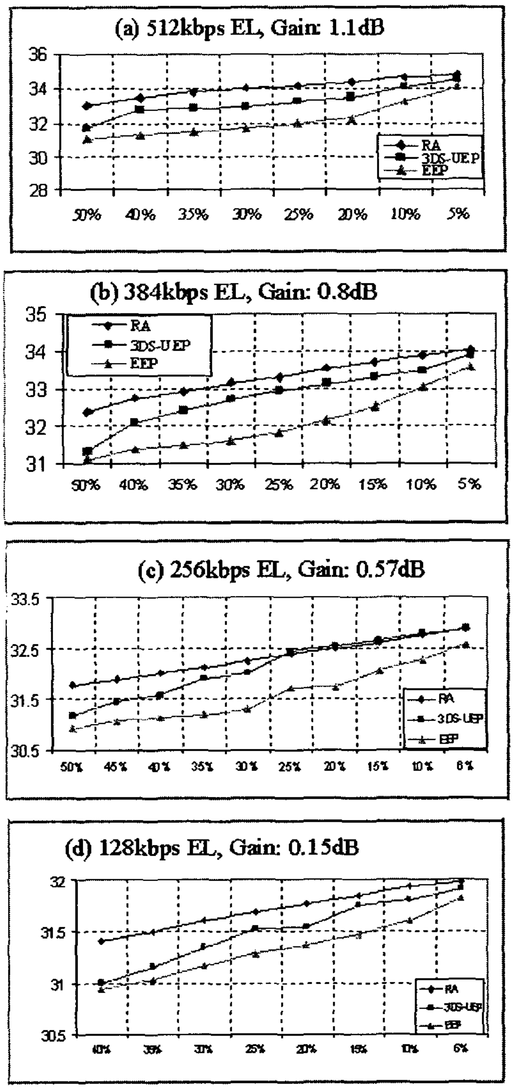


Fig. 7 The PSNR(y-axis) comparison of rate adaptation, three-level DiffServ UEP, and EEP for the Foreman sequence in terms of different packet loss rate(x-axis) under different EL rates: (a) EL at 512 kbps, (b) EL at 384 kbps, (c) EL at 256 kbps, and (d) EL at 160 kbps, where the respective no-loss PSNR is 34.99 dB, 34.19 dB, 32.97 dB and 32.07 dB (Gain is for UEP over EEP).

low packet loss. If there is a substantial amount of bandwidth variation (i.e. corresponding to a high packet loss ratio), a higher gain can be achieved from rate adaptation at the cost of additional complexity to estimate the available bandwidth. Under a small or medium range of bandwidth fluctuation, UEP without rate adaptation

Table 2. Priority distribution of BL and EL packets under different encoding and packetization parameters.

BL Case 1			BL Case 2			128kbps EL		
P_{avg}	σ_P	*Gain*	P_{avg}	σ_P	*Gain*	P_{avg}	σ_P	*Gain*
3.04	2.45	0.97dB	3.35	3.09	2.13dB	0.023	0.026	0.15dB
256kbps EL			**384kbps EL**			**512kbps EL**		
P_{avg}	σ_P	*Gain*	P_{avg}	σ_P	*Gain*	P_{avg}	σ_P	*Gain*
0.024	0.031	0.57dB	0.026	0.035	0.80dB	0.027	0.0038	1.10dB

works well without losing much efficiency. Finally, different EL bit rates also affect the gain of UEP over EEP. The smaller the EL bit rate is, the smaller gain of UEP over EEP is achieved.

3. Impact of Packet Priority Distribution on the UEP Gain

Coding and packetization parameters have an impact on the gain of UEP over EEP in both BL and EL packets. The AIR coding choice and the fixed-length packetization narrow down the performance gap between UEP and EEP (in reference to CIR and GOB packetization). This phenomenon can be explained by the packet priority distribution as shown in Table 2. As an extreme case, when the packet priority is set the same, the gain of UEP over EEP is zero. By modifying the coding and packetization modes, we modify the priority distribution as well as the performance gap. Most work done in UEP attempts to spread the priority distribution over a wide region so that the gain from UEP is highlighted. In conclusion, for the UEP approach, a more widely spread priority index provides more graceful quality degradation. Thus, the performance gap between UEP and EEP is largest when the priorities are spread across a wide range. On the other hand, when packet priorities are clustered into a small region, it narrows down the performance gap.

V. Conclusion and future work

A framework of rate adaptation, prioritized packetization, and differentiated packet forwarding is proposed for MPEG4 FGS video streaming. It performs by embedding the RD information within each bitplane, and relies on piecewise linear model to obtain the real distortion. Then one differentiated forwarding framework of error resilient MPEG-4 FGS video is investigated with the fine granular BL and EL packet priority. Starting from the real distortion of each packet, we show the gains of priority dropping over the uniform dropping under different encoding and packetization parameters. We generalize that the gain gap of UEP over EEP can be illustrated in the different distribution of packet priority. By integrating the rate adaptation with the proposed DiffServ framework, even more gains can be achieved.

A couple of issues should be elaborated further. First, the mapping of both BL and EL packets to DS level is very heuristic. We believe the both the distribution of packet priority and the price mechanism associated with the DS level should plays a role in this mapping. Second, how to exploit the maximal gain by mapping packets from different streams to different DS levels if multiple MPEG-4 FGS packets are multiplexed. Third, the current service model should be polished more to cover the rate adaptation, packet filtering and differentiated forwarding in more realistic scenario.

References

[1] W. Li, "Overview of fine granular scalability in MPEG-4 video standard", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp.301-317, Mar. 2001.

[2] D.Wu, Y.T.Hou, W.Zhu, Y.-Q.Zhang, and J.M. Peha, "Streaming video over the Internet: approaches and directions", *IEEE Trans. on Circuits and Systems for Video Technology*, vol.

11, no. 1, pp. 1-20, Feb. 2001.

[3] J. Kim, Y.-G. Kim, H. Song, T.-Y. Kuo, Y.J. Chung, and C.-C. J. Kuo, "TCP-friendly Internet video streaming employing variable frame-rate encoding and interpolation", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, no. 7, pp. 1164-1177, Oct. 2000.

[4] Q. Zhang, G. Wang, W. Zhu, and Y.-Q.Zhang, "Robust scalable video streaming over Internet with network adaptive congestion control and unequal loss protection", *in Proc. Packet Video Workshop 2001*, Kyungju, Korea, May2001.

[5] P. A. Chou, A. E. Mohr, A. Wang, and S. Mehrotra, "Error control for receiver-driven layered multicast of audio and video", *IEEE Trans. Multimedia*, vol. 3, no. 1, pp. 108-122, March 2001.

[6] R. Gopalakrishnan and J. Griffioen, "A simple loss differentiation Approach to Layered multicast", *in Proc. IEEE INFOCOM 2000*, Tel-Aviv, Israel, Mar. 2000.

[7] S. Blake, D. Black, M. Carlson, E. Davies,Z. Wang, and W. Weiss, "An architecture for differentiated services", *RFC 2475*, IETF, Dec. 1998.

[8] J. Shin, J. Kim, and C.-C. J. Kuo, "Quality-of-Service mapping mechanism for packet video in differentiated services network", *IEEE Trans. on Multimedia*, vol. 3, no. 2, pp. 219-231, June 2001.

[9] J. Shin, J.-G. Kim, J. Kim, and C.-C. J.Kuo, "Dynamic QoS mapping control for streaming video in relative service differentiation networks", *European Transactions on Telecommunications*, vol. 12, no. 3, pp. 217-230, May-June 2001.

[10] Y. Wang and Q. Zhu, "Error control and concealment for video communication: a review", *in Proceedings of IEEE*, pp. 974-997, May, 1998.

[11] G. Cheung and A. Zakhor, "Bit allocation for joint-source channel coding of scalable video", *IEEE Trans. Image Processing*, vol. 9, no. 3, pp. 340-356, Mar. 2000.

[12] X. Wu, S. Cheng, and Z. Xiong, "On packetization of embedded multimedia bitstreams", *IEEE Trans. on Multimedia*, vol. 3, no. 1, pp. 132-140, March 2001.

[13] C. Dovrolis, D. Stiliadis, and P.Ramanathan, "Proportional differential services: Delay differentiation and packet scheduling", *in Proc. SIGCOMM*, Boston, MA, Sept. 1999.

[14] C. Dovrolis and P. Ramanathan, "Proportional differential services - PartII: Loss rate differentiation and packet dropping", *in Proc. International Workshop on Quality of Service*, Pittsburgh, PA, June 2000.

[15] J.-G. Kim, J. Kim, J. Shin, and C.-C. J.Kuo, "Coordinated packet level protection employing corruption model for robust video transmission", *in SPIE Proc. VCIP*, San Jose, CA, Jan. 2001.

[16] J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics", *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, no. 4, pp. 446-459, Aug. 1998.

신 지 태(Jitae Shin) 정회원

1986년 2월 : 서울대학교 전기 공학과 학사

1988년 2월 : 한국과학기술원 원자력 공학과 석사

1988년 3월~1996년 12월: 원자력연구소 선임연구원

1998년 12월 (2001년 5월) : 미국 Univ. of Southern California 전자공학과 석사 및 (공학박사)

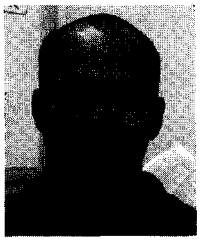
2001년 9월~2002년 2월 : 경희대학교 정보통신대학원 교수

2002년 3월~현재 : 성균관대학교 정보통신공학부 교수

<주관심 분야> 유무선 멀티미디어 네트워크 및 통신

: http://icc.skku.ac.kr/~jtshin

김 종 원(Jongwon Kim) 정회원



1987년 2월 : 서울대학교 제어계측공학과 학사
1989년 2월 : 서울대학교 제어계측공학과 석사
1994년 2월: 서울대학교 제어계측공학과 박사

2001년 9월~현재 : 광주과학기술원 정보통신공학과 부교수
2000년 7월~2001년 6월: 미국 InterVideo Inc., Fremont, CA, 개발자문
1998년 12월~2001년 7월: 미국 Univ. of Southern California, Los Angeles, CA, EE-Systems Department 연구조교수
1994년 3월~1999년 7월: 공주대학교 전자공학과 조교수
<주관심 분야> Networked Media Systems and Protocols focusing "Reliable and Flexible Delivery for Integrated Media over Wired/Wireless Networks(네트워크미디어: http://netmedia.kjist.ac.kr)