國 立 交 通 大 學

資訊科學與工程研究所

博 士 論 文

應用於雜訊網路通道具容錯能力的可調式視訊傳輸之研究

Efficient Error-Tolerant Design for Scalable Video Transmission over Error-Prone Channels

研 究 生：何健鵬

指導教授：李素瑛 教授

中 華 民 國 一０一年 二 月

應用於雜訊網路通道具容錯能力的可調式視訊傳輸之研究
Efficient Error-Tolerant Design for Scalable Video Transmission over Error-Prone Channels

研 究 生：何健鵬　　Student：Chien-Peng Ho

指導教授：李素瑛　　Advisor：Suh-Yin Lee

國 立 交 通 大 學

資 訊 科 學 與 工 程 研 究 所

博 士 論 文

A Dissertation

Submitted to Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Computer Science

February 2012

Hsinchu, Taiwan, Republic of China

應用於雜訊網路通道具容錯能力的可調式視訊傳輸之研究

研究生：何健鵬　　　　　　　　　　　指導教授：李素瑛 博士

國立交通大學資訊科學與工程研究所

摘要

近年來因為寬頻服務的日趨成熟，多媒體串流技術和點對點(P2P)分散式網路架構的隨選視訊點播服務的傳輸與流通量也越來越大，在現有的網路傳輸系統中，存在封包遺失累積問題，因此穩定可靠的多媒體數據傳輸方法正成為越來越重要的研究主題。其中，在典型的多媒體通信應用中高效率容錯和錯誤恢復能力的演算法透過分析和數值方法來實現多目標優化性能指標的有效技術。與傳統數據通信的特點相較，多媒體串流有四個方面的主要差別：(1)、封包丟失是視訊串流品質不確定性的主要原因之一，並可能會導致許多有價值的視訊資訊缺失，進而造成影像品質嚴重下降。(2)、總頻寬的需求仍遠遠超過了現有的通信網絡基礎設施(如：支援高品質視訊隨選點播服務)。(3)、雖然在客戶端的緩衝器設計可以吸收傳輸速率的變化，但這不足以保證多媒體串流的服務品質(如：網路電視和網路電話）。(4)、雖然大多數的多媒體數據壓縮傳送允許應用在漏失封包和多錯誤的網路環境，並且由於人類視覺感知系統對於某些影像的變化不夠敏感，因此能容忍有某些程度的影像品質下降。進而，在視頻傳輸可調式編碼的基礎上使用錯誤控制和非均等錯誤保護碼可以是一個廣為被接受的方法來改善接收到的視訊品質。

在本論文中我們針對視訊串流在不斷變化的網路條件中，提出可擴展性、具錯誤容忍度和有效的錯誤保護的演算法，來達成系統優化性能之有效視訊傳輸技術。我們針對在漏失封包和多錯誤的網路環境下提出一個三維小波可調式視訊編碼演算法為基礎的內容自適應之錯誤保護

技術，利用細緻可調式的方式去增強資料錯誤的保護能力。所提出的細緻可調式之錯誤保護演算法可在視訊編碼率和影像失真之間選擇一個適當的折衷，以適切決定可調式視訊編碼率和通道編碼-前向錯誤更正碼的保護安排。實驗結果顯示，確實能達成在多媒體數據傳輸應用中重建視訊品質和系統容錯能力之間的平衡。在P2P分散式網路架構的視訊串流之研究中，我們結合非結構性對等網路運算與隨選視訊技術，發展一套以視訊編碼之失真模型為基礎的複本容錯機制，希望藉由有效的佈建高視訊品質影響及高點播率之複本於點播視訊流中高失真區與非連續區域來提升整體服務成效。根據模擬實驗證明我們所提的方法，相對於其它的容錯機制，在點對點網路環境中廣泛存在的搭便車 (free-rider) 現象出現時亦可有效的降低終端使用者間影像品質損失，並在不同點對點網路條件下能有較穩定的視訊撥放品質。整體而言本方法可以有效降低隨選視訊服務伺服器的工作負載。

關鍵詞：多媒體通信、錯誤保護、隨選視訊、點對點同儕式網路、容錯設計。

Efficient Error-Tolerant Design for Scalable Video Transmission over Error-Prone Channels

Student: *Chien-Peng Ho* Advisor: Dr. *Suh-Yin Lee*

Institute of Computer Science and Information Engineering
National Chiao Tung University

Abstract

Due to the growing maturity of broadband services, multimedia streaming systems and peer-to-peer on demand service are gaining vast popularity in recent times. Stable and reliable transmission of multimedia data is becoming increasingly important for multimedia communication over networks subject to packet erasures. To achieve stability and reliability, efficient fault-tolerant and error-resilience methods for multimedia communication are typical analytical and numerical approaches, so as to attain multi-objective performance metrics. The characteristics of video traffic differ substantially from those of traditional data traffic in four ways. First, packet loss is the major cause of nondeterministic distortion on the Internet and may have significant impact on perceptual quality of the streaming video. Second, aggregate bandwidth requirements, supporting video-on-demand services, are still far in excess of the existing communication network infrastructure. Third, although buffering on the client side can provide an opportunity to absorb variations in transmission rates, it is not sufficient to guarantee the service quality of multimedia streams like IPTV (Internet Protocol television) and VoIP (Voice over IP). Finally, most compressed media data are transmitted over lossy and error-prone networks, and a certain degree of quality degradation is tolerable due to noise in some regions which are below the threshold of human visual perception. Thus, video transmission based on scalable coding and unequal error protection codes can be one of the approaches of maintaining acceptable media quality in a network. Error control for video communication and resource allocation in peer-to-peer multimedia systems remain open issues and are the

focus of this work.

In this thesis, we built scalable, error-resilient, and high-performance multimedia frameworks to adapt to changing network conditions. We developed a framework of fine-level packetization schemes for the streaming of 3D wavelet-based video content over lossy packet networks. An adaptive fine-granularity unequal error protection algorithm was proposed to allow a tradeoff between rate and distortion, and jointly adopt scalable source coding rates and the level of FEC protection. Experimental results show that the proposed framework strikes a fine balance between the reconstructed video quality and the level of error protection under time-varying lossy channels. In the study of P2P video streaming, we developed a replication strategy to optimize resource allocation based on the video-distortion technique for unstructured P2P overlay networks. Failure recovery action can be accomplished by distributing high quality impact and popular replicas to regions of low peers density or discontinuous areas. The results demonstrated the efficiency and robustness of the proposed method at compensating for network-induced errors, and the framework can be applied at a range of different scales of free-riding peers. Moreover, the proposed algorithm was able to handle the load imposed on the system efficiently and improved average visual quality in the overall system.

Keywords: *multimedia communication, error protection, video-on-demand, peer-to-peer, fault-tolerance.*

Acknowledgments

I would like to express my sincere gratitude and appreciation to my advisor, Professor Suh-Yin Lee, for her constant advice and guidance in conducting this research work. I have really appreciated the candor and frankness of her comments, and I have learned tremendously from her. It is impossible for me to complete this dissertation without her expert guidance. I would also like to thank my dissertation committee members, Professor Ja-Ling Wu, Kuo-Chin Fan, Long-Wen Chang, Hong-Yuan Mark Liao, Hsueh-Ming Hang, and Hsu-Feng Hsiao for their helpful suggestions, time and advice.

I would like to give a special acknowledgment to Professor Chia-Ling Hsieh and Sheau-Ling Hsieh for encouraging me to finish this work no matter what difficulties I have met. Sincere appreciation extends to all my friends and the members of the Information System Laboratory for sharing their research experiences.

Finally, my deepest thanks go to my family for their encouragement and support. This dissertation is dedicated to them.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Motivation - the Impacts of Network-based Multimedia Services

Internet based multimedia applications, such as video conferencing, Internet telephony, video-on-demand (VoD), and peer-to-peer (P2P) applications, are a very important trend for future communication and entertainment systems [1]. Streaming media delivery over packet networks is centered on a network layer capable of dynamically selecting a path from the source node to the destination node, with no guarantees on network conditions. In a realistic network, the traffic conditions are dynamically changing, a large variation in bandwidth and end-to-end delays are major problems for supporting multimedia traffic in the network. Hence, multimedia delivery over lossy networks is a challenging task due to bandwidth variation, delay jitter and packet losses [2]. This raises numerous issues that researchers must deal with in order to maintain smooth presentation of video, with visual quality as constant as possible. For smooth streaming presentation, a media source node should track packet losses and adjust media source rates and channel protection levels accordingly. Scalable video coding techniques (such as H.26X, MPEG-4 or wavelet standards) have been standardized on new digital technologies for applications related to streaming media. With scalable video content, a variable bitrate (VBR) source bitstream can be composed on the fly to match the channel bandwidth (assuming the bandwidth can be predicted via some model) and smoothly transmitted to the receivers for presentation. However, if the video quality of the composed stream varies too much it becomes visually unpleasant. Another source of visual degradation comes from packet losses [3]. Unlike distortion from source coding, distortion due to channel loss is more difficult to quantify since the value of a single missing packet depends on the coded data it contains. Such source and channel distortion removal has not been sufficiently explored until now and needs to be further investigated. Therefore, a more fault tolerant design (with duplicates and redundancies from the given video) may be required, and one should find a way to formulate the distortion and rate impact of packet losses for constructing an rate-distortion (R-D) optimized streaming system that produces smooth video quality streams.

1.1.1 Packet Loss

Packet loss is one of the main factors affecting quality of video in video streaming applications [4]. For instance, it is reasonable to consider a sequence of video-packets transmitting over a route with a packet loss rate varying from 0 to 10 percent [5]. Network congestion is the dominant reason for packet losses, and that causes packet losses in the network router queue buffer [6]. Any packet loss not caused by network congestion is a non-congestive loss, and such segment losses are often assumed to be negligible. Boyce showed that based on MPEG compressed video using the Real-time Transport Protocol (RTP) [7] and User Datagram Protocol (UDP) [8] transport protocols the experimental Internet packet loss rates were quite varied, ranging from 0 to 8.5 percent of the total packets [9].

Over the past half century, the work on loss or error modeling has also paved the way for the distribution of bit errors on communication channels for network related research. One model for packet losses is the Markov or multi-state model. The first generalized model of a burst error channel was presented by Gilbert [10], which is a two-state Markov model. Elliott [11] extended the Gilbert model to the multiple-state model. The parameters of the two models can be calculated from the measured error trace by using the mean burst length (the error run) and the mean gap length (the error-free run). In addition, both burst length and gap length are geometrically distributed. Cain et al. [12] studied the distribution of burst lengths over a Gilbert channel. Fritchman [13] enhanced the Gilbert-Elliott model to a partitioned Markov model with partitioning of the state space to several error-free and error states, where the gap length distribution uniquely specifies the model. Nguyen et al. [14] presented a two-state model with segmented exponential distributed burst lengths and gap lengths. McCullough introduced more channel states (each state represents a specific channel model) to obtain an M-state Markov model, each with a different error rate [15]. Sadeghi et al. introduced and discussed detailed aspects of the finite-state Markov channel and its applications to fading channels [16].

1.1.2 Bandwidth

Table 1-1: Digital video bandwidths.

Compressed Method	Format	Data Rate
Uncompressed Video	SDTV (480i)	165.9 – 270 megabits per second (Mbps)
	EDTV (480p/576p)	540 Mbps
	HDTV(1080i/720p)	1.485 gigabits per second (Gbps)
	HDTV (1080p)	2.970 Gbps
MPEG-2 Video	SDTV Broadcast (480i)	3 – 6 Mbps
	HDTV Broadcast (1080i)	12 – 20 Mbps
	SDTV Production (4:2:2 I-frame only)	18 – 50 Mbps
	HDTV Production (4:4:4 I-frame 10-bit colour depth)	140 – 500 Mbps
MPEG-4 AVC / H.264	SDTV Broadcast (480i)	1.5 – 3 Mbps
	HDTV Broadcast (1080i)	6 – 9 Mbps

Video bitstreams are typically encoded at a predetermined bitrate for a particular application. The audio and video encoded bitstreams consume large amounts of network resources - primarily bandwidth. Moreover, there are many multimedia services which bring a rising tide of increased bandwidth demand over the Internet, including Voice over IP (VoIP), Internet Protocol television (IPTV), High Definition TV streams (HDTV), Standard-definition television (SDTV, 480 lines interlaced), Enhanced definition Television (EDTV, 480 progressive scan) and VoD. Various statistics derived from MPEG-2 and MPEG-4 AVC compressed methods are presented in Table 1-1. Due to high bandwidth needs, there are several technologies available which serve as enabling technologies for satisfying the huge bandwidth demand. We summarize the related network interface technologies and speeds in Table 1-2 [17-19]. For IPTV services, a single SD MPEG-2 stream at the encoding rate of 3.75 Mbps (digital MPEG-2 programs are typical streams over cable television networks) whereas MPEG-4 AVC SD needs approximately 2 Mbps

Table 1-2: Relative network interface speeds.

Connection Type	Connection Speed
Dedicated PPP/SLIP via modem	28.8 kilobits per second (Kbps)
fast modem	56 Kbps
Integrated Services Digital Network	128 Kbps
Typical DSL	640 Kbps
DS1/T1	1.536 megabits per second (Mbps)
10-megabit Ethernet	8 Mbps (best case)
IEEE 802.11b	11 Mbps (maximum)
IEEE 802.16e	15 Mbps
DS3/T3	44.736 Mbps
Optical Carrier#OC1	51.84 Mbps
IEEE 802.11g	54 Mbps (maximum)
IEEE 802.16a	75 Mbps
100-megabit Ethernet	80 Mbps (best case)
Optical Carrier#OC3	155.52 Mbps
IEEE 802.11n	300 Mbps (maximum)
Optical Carrier#OC12	622.08 Mbps
1-gigabit/sec Ethernet	800 Mbps (best case)
Optical Carrier#OC24	1244.16 Mbps
Optical Carrier#OC48	2488.32 Mbps
Optical Carrier#OC192	9953.28 Mbps
Optical Carrier#OC768	39813.12 Mbps

for the same quality video program. The video streaming platforms implemented use mainly a family of transport protocols, namely, UDP, RTP, Real-Time Control Protocol (RTCP), and Real Time Streaming Protocol (RTSP). UDP is a best-effort transport protocol over IP networks while RTP and RTCP are the media transport protocols used to transmit and control media data. RTSP is an application-level protocol for initiating and directing delivery of streaming with real-time properties. It could be thought of as a remote network control for multimedia servers (a detail of RTSP is available through RFC 2326 [20]). Current streaming systems for the Internet are based on a best-effort delivery service in the form of UDP, which does not guarantee reliability or the order delivery of the network packets. The most commonly encountered issues of multimedia transmission and streaming applications are unreliable Internet connections and heterogeneous bandwidth of the different end users. When the network bandwidth fluctuates, the coded bit-rate may not match the real bandwidth [21]. Hence, scalable video coding techniques are often used to provide real-time quality adaptation for streaming systems.

1.1.3 Client-side Buffer

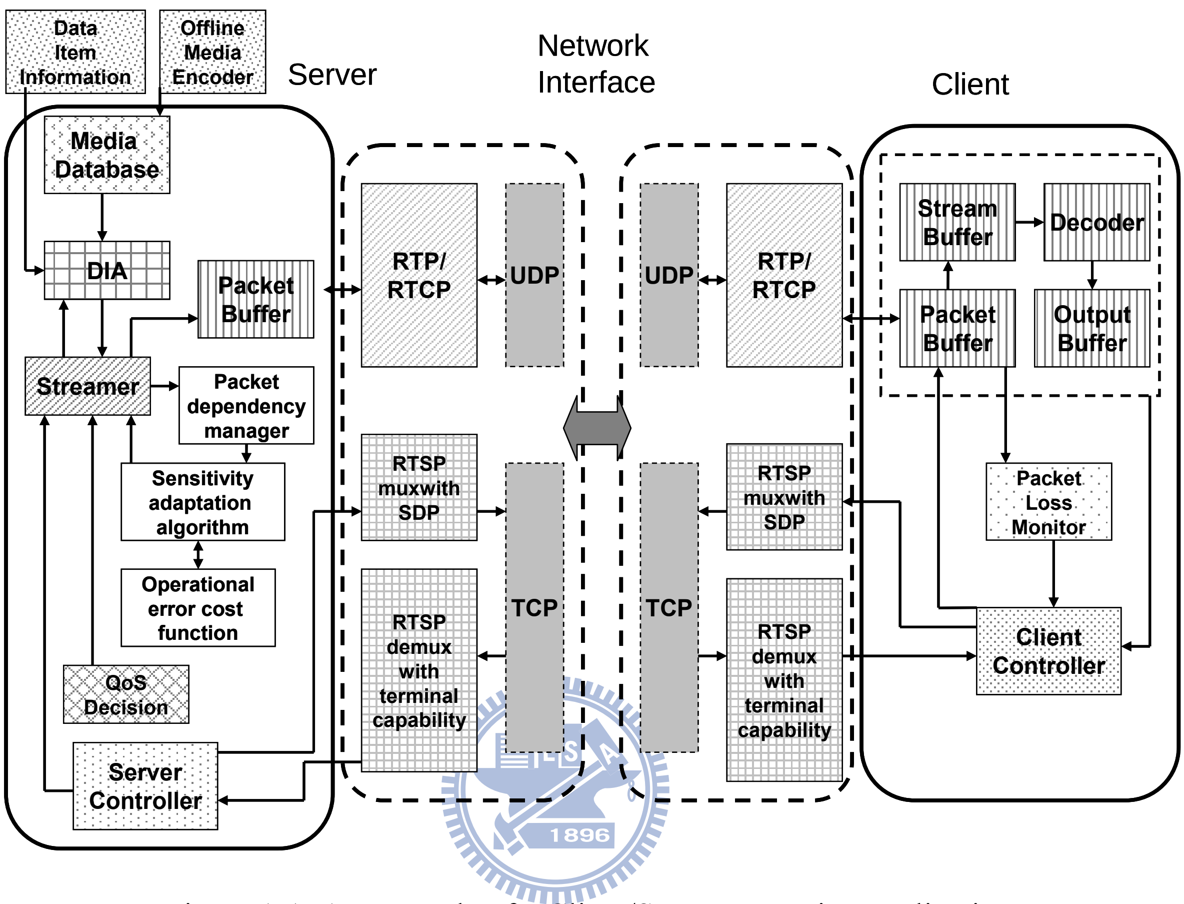


Figure 1-1: An example of a Client/Server streaming application.

A video streaming application should allow for short-term network fluctuation. Packet buffers are commonly located in the sending/receiving device. Incoming packets are queued in the packet buffer at the streaming client to overcome the network latency. The structure of a client/server type of streaming application is depicted in Fig. 1-1. We can compute the distortion and the streaming rate that result from a streaming policy, which are implemented in modules including the packet dependency manager, the sensitivity adaptation algorithm, and the operational error cost function. The quality of service (QoS) decision function decides which of the admission control or equivalent bandwidth strategies should be employed. Based on the former policy and updating the policy given the new value, the streamer component splits bit-streams into packets and stores them in the packet buffer memory for transmission to the server side. The Digital item adaptation (DIA) module can be used to describe the usage environment and content format features. The arriving packet stream is buffered at the destination packet buffer, which temporarily stores the receiving data before providing it to the stream buffer for decoding at the client side (packet losses are monitored for quality assurance purposes). The media decoder fetches coded units from the stream buffer and stores decoded data in the output buffer, then the decoded content is provided to a screen for playback. In particular, the packet buffer can be a shared resource at network nodes, especially for P2P overlay networks (including video sharing applications and P2P video-on-demand applications) [22].

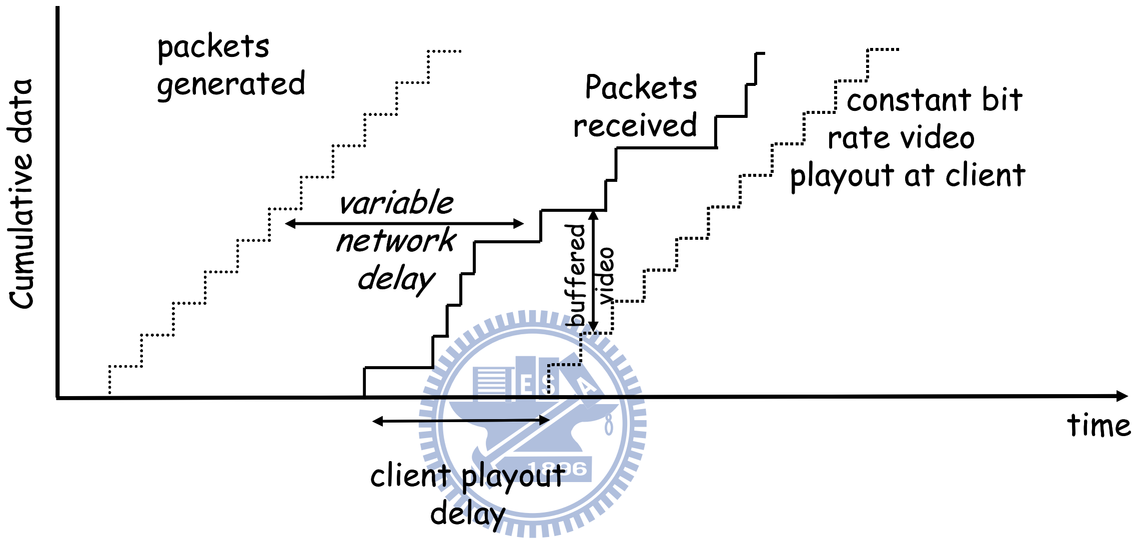


Figure 1-2: Client-side buffering.

As shown in Fig. 1-2, the packet inter-arrival time is random [23] (inter-arrival time of media sessions and file-transfer sessions may follow a Poisson distribution [24]), or the packet arrival time is fixed time but with explosive, random error bursts. Adaptive media playout (AMP) techniques play one frame at a slower pace, and extend the deadline for the arrival of subsequent frames. Network delay contains transmission delay, propagation delay, processing delay, and queuing delay. It is one of the solutions to reduce the impact of network delays and control the rate of the playout without involvement of the server. Thus, in order to ensure continuous playback of media streams, the client device can adaptively change playout speed to prevent buffer overflow and underflow. Thus, AMP techniques allow the client-device to buffer less data, the playback rate of the streaming media player is varied according to the state of its playout buffer [25]. When designing or redesigning network-based multimedia systems, minimizing playout delay and keeping a low loss rate are design considerations when deciding which multimedia-service solution is most applicable.

1.1.4 Packet Dependency Control

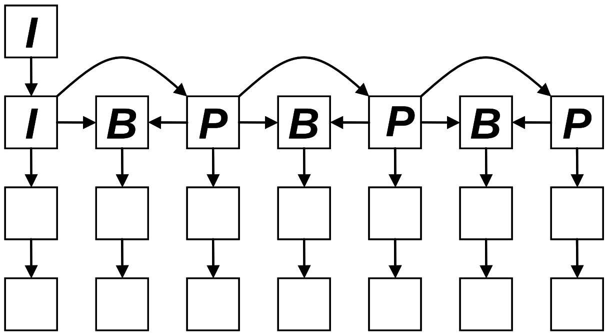


Figure 1-3: Directed acyclic dependency graph.

A certain degree of quality degradation is tolerable for most multimedia applications, and distortion tolerance can be designed to fit human visual perception. Information loss is quantified based on the distortion tolerance and thus the best reconstructed video quality can be obtained at a given bit rate. Packet dependency control (PDC) [26] is designed to provide increased distortion tolerance and elimination of video packet retransmission [27]. Video packets in typical multimedia streaming have potentially strong dependencies between the media packets. If one of the video packets is lost during transmission, then all subsequent packets are dependencies with the lost packet and the follow-up frames may be affected by the error while the decoding procedure is performed at the client. Figure 1-3 shows a dependency of a directed acyclic graph (DAG), each node representing a packetized data unit. Each packetized data unit will be labeled as follows: the actual size of data unit n, distortion reduction if data unit n can be normally decoded at the receiving side, and the decoding deadline of data unit n. The PDC scheme should be used to improve error-resistance capabilities and reduce latency error during the scalable media-streaming phase.

1.2 Research Objectives and Contributions

The major research objectives and goals of this study include the development of error-resilient video communication to mitigate the effects of packet loss and inter-frame error propagation. The study of how active replication [28] can assist fault tolerance and smooth video playback in a P2P overlay network. The study of how perceptual quality-based progressive packet-level transmission on a VBR channel can be used to increase the robustness of video communication systems. The results of this research could help open the way of finding the appropriate solutions for sufficient streaming quality. In summary, there are the following objectives:

- Determining a fine-level FEC (forward error correction) packetization protection framework and evaluating the adaptive fine-granularity of FEC protection scheme.
- Studying the properties of the parameterized R-D model which is incorporated into the fine-level FEC framework to facilitate multiple adaptations.
- Constraining the analysis of video content complexity with a rate-distortion optimized scalable video streaming system
- Characterizing the underlying correlations between peer-network characteristics and video quality impacts over unstructured P2P VoD overlay networks.
- Using a distortion-based active replication method for enhancing system scalability and tolerance of peer failures in a P2P VoD system.
- Specifying the agent-peers of the overlay P2P VoD system as sources and sinks of internet advertising under robust online marketing communication channels.

Scalable video delivery using fault-tolerant and error-resilient techniques has been a subject of study for many years. Nonetheless, although some of the existing error control techniques may establish a different degree of protection based on the degree of importance of the content, unequal error protection is limited to varying degrees in coarsely structured bit-streams, since the error control scheme is based on either a single-layer video coding model or a coarse-granularity layered scalable-video-coding (SVC) mode. The main contribution of this work is the development of adaptive fine-granularity methodologies to support error-protection in video streaming services with an effective streaming policy. This streaming approach takes into account the R-D information of each block (it is only a chunk of data in a coding block) which is derived doing the encoding process, and searches for an optimal operating point between the scalable source coding rate and the FEC protection level which is dependent on the video content entropy. The content adaptation of the FEC protection has a fine granularity. Therefore, adaption conditions between video content entropy and run-time packet loss rate can be accurately established.

P2P services have been exhibiting a rapid growth in the world. For instance, forecasts from Cisco Visual Networking Index (Cisco VNI) reveal that P2P traffic in 2014 will be more than double the amount of data P2P traffic generated in 2009 (around 7 peta bytes per month). The implementation and validation of a P2P VoD system involves carrying out a performance evaluation of the P2P replication mechanism. P2P multimedia applications claim lower costs and more efficient content distribution than traditional client-server applications. The searching area of P2P services are controlled by the time-to-live (TTL) value, which is decreased each time P2P-related commands are forwarded (to limit the maximum number of intermediate peers) until the command is accepted or the TTL value is zero. Whenever the degree of peer sharing resources drops (such as due to an upload bandwidth outage) below the marginal value at the TTL value in a P2P multimedia system, especially for VoD services, each of the peers is automatically trying to reconnect to the media server for service operations. Heavy traffic leads to high waiting times and a lowered quality of service, which could significantly degrade operational performance and service. In particular, our proposed method provides a reliable set of replica data that can accommodate a large number of free-riding nodes with less jerky video playback.

1.3 Synopsis of the Dissertation

This dissertation is organized as follows. Chapter 2 reviews some related works regarding the techniques of significant effect on video streaming and P2P streaming-based video performance. Section 2.1 gives an introduction to wavelet-based video coding system.

The works related to the error resilience tools is given in Section 2.2, an overview of P2P streaming systems involved in Section 2.3.

Chapter 3 studies error control techniques and fine-level packetization scheme for streaming of 3D wavelet-based video over a lossy packet network. We propose an adaptive fine-granularity unequal error protection algorithm in which the foundation of strategy decision making is based on the tradeoff between rate and distortion, and jointly adapting scalable source coding rate and level of forward error correction (FEC) protection. We also explore the concept of multiple adaptations of video bitstream and fine-level FEC protection in a parameterized R-D model-based approach. The fine-level FEC protection to the granularity of coding pass level is discussed in this chapter.

Chapter 4 studies an approach to integrate smooth quality constraint into a rate-distortion optimized scalable video streaming system over a variable bandwidth channel. Accordingly, a solution is needed to preserve smooth video quality in communication systems in which the available network bandwidth and video data rate may vary. Hence, the scheme takes into account the degree of motion in each frame to an R-D optimized framework for determining the packet scheduling policy.

Chapter 5 further studies the problem of optimizing the media-data replication strategies over unstructured P2P VoD networks in the presence of peer failures or the regions of low resource area. By taking video popularity, uploading bandwidth, supportive buffer into considerations, the proposed mechanism provide a solution on reducing the required video server operations, smoothing playback of videos, and improving the averaged visual quality for overall performance in P2P VoD systems. Finally, Chapter 6 concludes the dissertation and discusses some possible directions for future research.

Chapter 2 Background and Survey of Related Work

In this chapter, we briefly summarize related work on the design of error-resilient video communication. This work is classified into two categories. Section 2.1 gives a brief overview of wavelet-based video coding system. Section 2.2 describes error correction approaches which improve the signal to noise ratio of the transmission link to efficiently support multimedia data transmission over lossy packet networks. Section 2.3 gives an overview of P2P video-on-demand services on network usage.

2.1 Wavelet-based Video Coding System

The fundamentals of 3D wavelet coding which enable temporal, spatial and quality salability are discussed in this section.

2.1.1 Overview of the 3D Wavelet Codec

The block diagram of a wavelet-based video coding system is shown in Fig. 2-1. In a *t*+2D wavelet coder, an input video sequence is temporally decomposed first using motion-compensated temporal filtering (MCTF) [29]. The output of MCTF is then further decomposed by a 2D spatial wavelet transform on a frame-by-frame basis. For example, two-level temporal decomposition results in three temporal subbands, namely, P(H_t, YUV), P(LH_t, YUV), and P(LL_t, YUV). When the group of pictures (GOP) size is eight, a typical set of transformed subband data produced by the *t*+2D wavelet coder has four P(H_t,YUV) frames, two P(LH_t, YUV) frames, and two P(LL_t, YUV) frames. Each frame contains one luminance component (Y) and two chrominance components (U and V). The coefficients of different subbands are logically segmented into coding blocks, based on the structure of Fig. 2-2, and each coding block is independently coded by an entropy coder. The analysis is presented using the wavelet-based video coder for a CIF resolution sequence. CIF defines a video sequence with a resolution of 352 × 288. For 2-level temporal transform encoding and GOP=8, a coding block size in Fig. 2-2 has block depth 2 (the number of frames, i.e. two frames), block height 36(=288/2^3), and block width 44(=352/2^3).

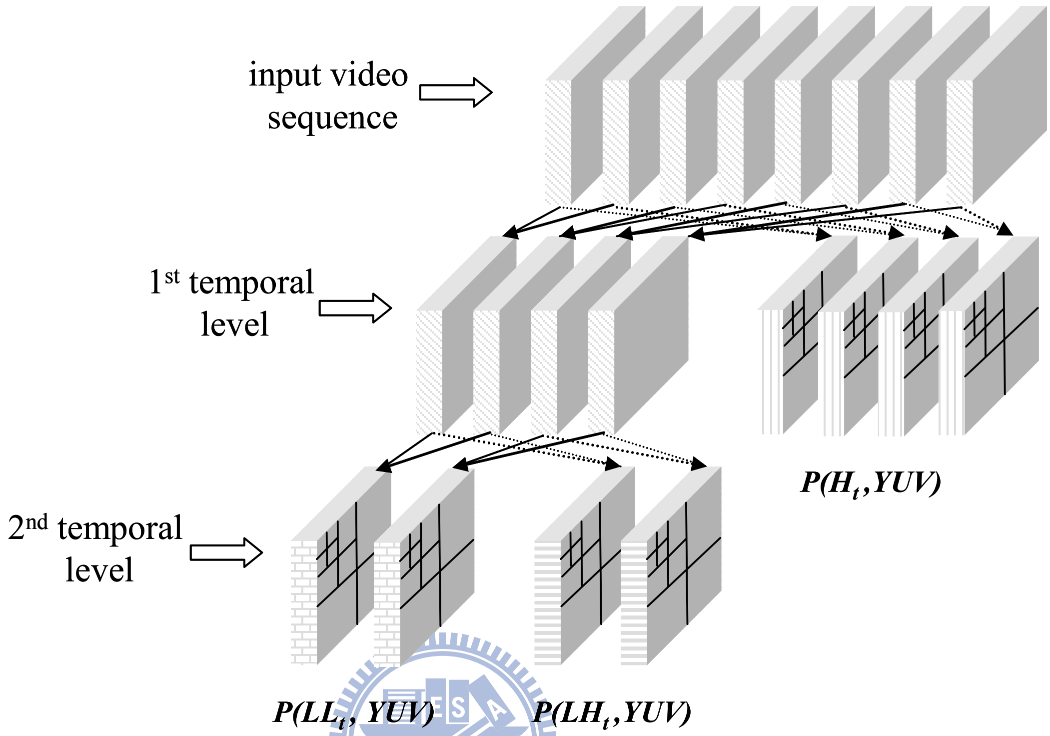


Figure 2-1: Wavelet video coding block diagram

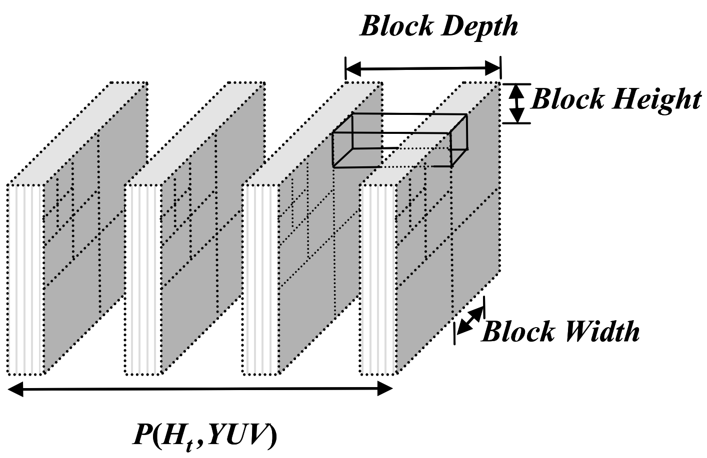


Figure 2-2: Examples of coding block in wavelet video coding

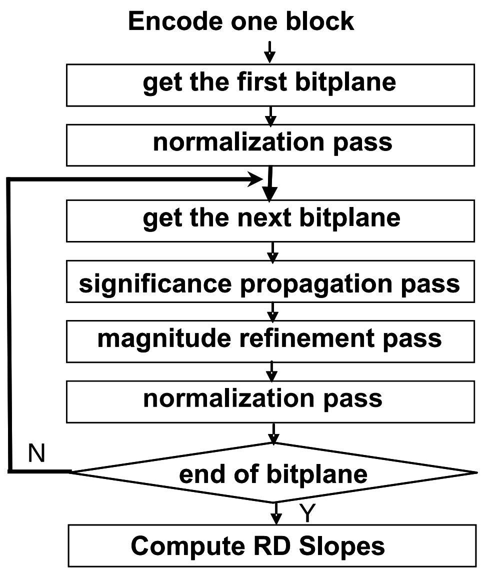


Figure 2-3: 3D ESCOT entropy coding.

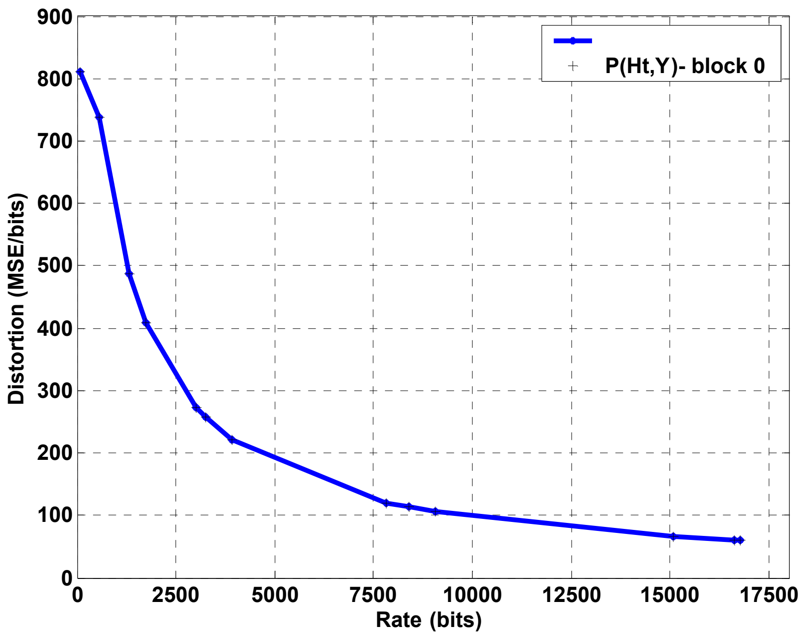


Figure 2-4: The R-D curve of coding block 0 of subband $P(H_t,Y)$of STEFAN.

Common entropy coding techniques for wavelet video are 3D embedded subband coding with optimized truncation (3D-ESCOT) [30] and 3D set partitioning in hierarchical trees (3D-SPIHT) [31]. The 3D-ESCOT algorithm has higher compression efficiency and better scalability than the 3D-SPIHT algorithm. During the 3D-ESCOT entropy coding process, the entropy coder (fractional bit plane coding and context-based arithmetic coding) operates one coding block at a time, and each coding block consists of N total bit planes.

Three encoding operations of the context-based arithmetic coding (zero coding, sign coding, and magnitude refinement) are used to characterize the significance of coefficients in a bit plane. Following the 3D context modeling, fractional bit plane coding ensures that the bit stream is arranged with fine granularity of signal-to-noise ratio scalability for each coding block. As shown in Fig. 2-3, the fractional bit plane coding procedure consists of three distinct passes: the significant propagation pass, the magnitude refinement pass, and the normalization pass. Since the first bit plane of a coding block can only be processed with the normalization pass, a coding block contains $3N-2$ coding passes. After entropy coding, candidate truncation points of a coding block are associated with R-D slopes. Any truncation points that are not on the convex hull are eliminated, and the R-D slopes are λ_0, λ_1,..., λ_{3N-2},where $|\lambda_0|>|\lambda_1|>...>|\lambda_{3N-2}|$. All coding blocks have R-D curves similar to the example shown in Fig. 2-4. A single bit error of a top coding pass may cause severe degradation in video quality. Therefore, higher level of protection is required for top coding passes. In addition, the proposed scheme of this dissertation is based on 3D-ESCOT coding technique.

2.1.2 Bitstream and R-D Information

A typical coded video sequence consists of a variable number of GOPs which is composed of a series of consecutive frames. Video encoders can be setting in a static or dynamic GOP length to encode video sequences, and bit-streams are generated based on the selected GOP structure. Each GOP bit-stream comprises four major parts: coded motion vectors, transformed and scaled luminance (Y) and chrominance (UV) data, and the header information (layer header, component header, and block bitplane selection information (e.g. inclusion flag, bit depth information, delta bytes, delta slope)). The R-D information of coding blocks contains rates, distortions, and the Lagrange multiplier values (refer to as the R-D slope λ value). The Lagrange multiplier is used to find the optimal point for different video quality. The video encoded using a 3D wavelet codec [32, 33] (Microsoft Research Asia - VidWav reference software) is organized in the format shown in Fig. 2-5, where m is the total number of temporal and spatial subbands and n_i is the number of coding blocks in subband *i*. Figure 2-6 depicts an example data of the first GOP of Stefan by using the VidWav codec, the peak signal-to-noise ratio (PSNR) is 41.13 db. The coded motion vectors is 24,218 bytes, Y component is 476,073 bytes, U component is 26,870 bytes, V component is 25,667 bytes, and the header information is 565,278 bytes where the total bytes of the GOP is 1,118,106 bytes, where the header information and Y component are the major amount of data.

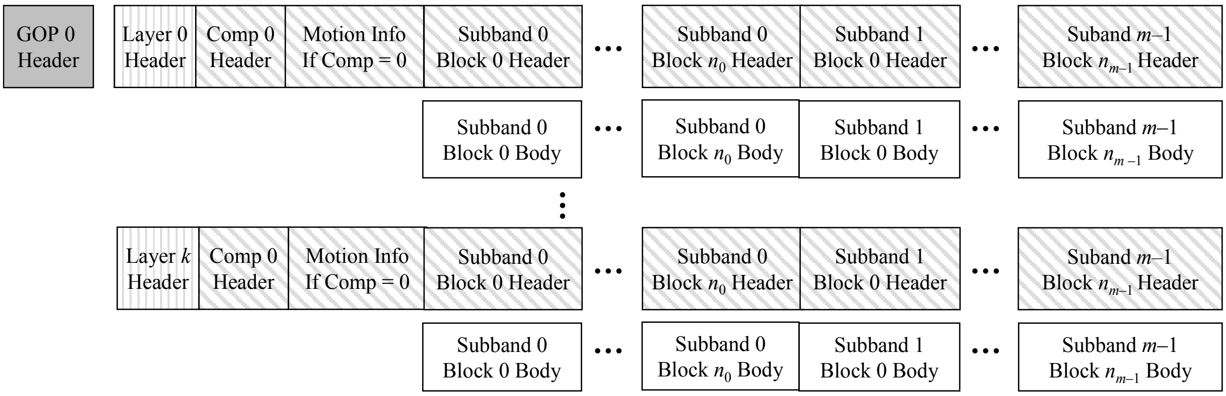


Figure 2-5: MSRA Wavelet bitstream format (please note that there is no need to enforce layer structure for MCTF-based wavelet bitstreams)

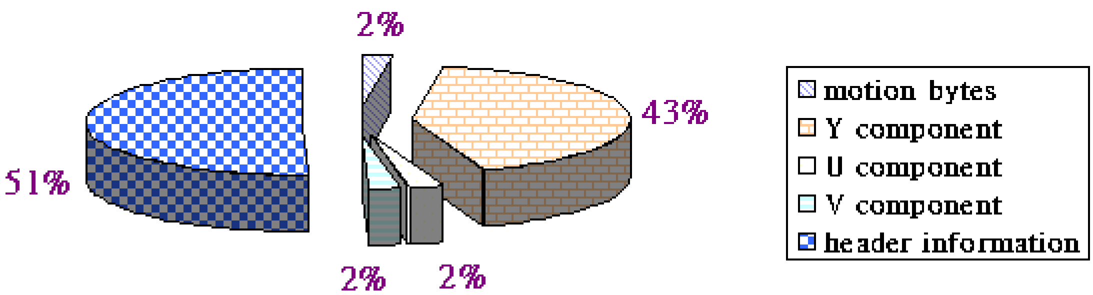


Figure 2-6: An example data of the first GOP of Stefan (2048 kbps, CIF, PSNR: 41.13 db, GOP size: 64)

We modeled the data loss ratio in case of partial failures in a bit-stream file. Figure 2-7 shows the different data loss scenarios of block 1 under the same amount of data loss ratio. The lost data is replaced by zero value. As shown in Fig. 2-7(a), the reconstructed video quality on the video will be severely degraded, if the most visually important data is affected by errors. On the other hand, if the less important data is corrupted or lost during transmission, the video frame can be partially reconstructed, as shown in Fig. 2-7(b). A slightly degraded video-quality but uncorrupted video stream is less annoying to the audiences than a corrupted stream. Figure 2-7(c) shows that the reconstructed frame has less visual quality decrement comparing with the original frame, if the lost data is not critical. This leads to the need of unequal error protection on compressed video bit-stream, in which the more important compressed bit-stream is delivered with a higher degree of error protection [34].

(a) The most important 10% data of block 1 is dropped

(b) The less important 10% data of block 1 is dropped

(c) The least important 10% data of block 1 is dropped

Figure 2-7: 10% of data loss of blocks 1 on the $P(H_t,Y)$ subband (Stefan, CIF, 150 frames, 15 fps, GOP = 64, target bitrates: 2048k).

2.2 Robust Internet Video Transmission based on Scalable Coding Streams

When audio and video streams are transmitted over an unreliable IP network, it is not only the dependency relationship among packets but also the importance of information carried that should be taken into account in the selection of the best video-quality under variable network conditions. Therefore, many research efforts have been made towards improving performance of multimedia transmission systems in general, and designing partial reliability techniques for media content distribution over packet erasure channels which are usually considered on both the forward and feedback links. Efficient transmission over packet erasure channels can be realized by FEC, Automatic Retransmission reQuest (ARQ), error resilience, and error concealment [35]. FEC adds redundancy to the transmitted compressed bit-streams so that the receiver may correct errors and protect the video quality from degradation caused by packet loss and jitter. In ARQ, retransmission is requested when the receiver detects errors, and the most common schemes are: Stop-and-Wait ARQ, Go-Back-N ARQ, and Selective Repeat ARQ. In addition, a number of algorithms have been proposed in the literature for Hybrid ARQ (HARQ) which is essentially a combination of FEC with ARQ in an optimal strategy.

2.1.1 Two Basic Approaches to Error Correction

Automatic Retransmission reQuest (ARQ) is one of the commonly used channel coding approaches for error protection, which is particularly effective against burst channel errors. From a practical viewpoint, the main disadvantages of ARQ are the retransmission costs in terms of huge delays, and some receivers may receive duplicate packets which do not contribute to the quality of the stream due to the long delay propagation. Moreover, ARQ techniques require data stores for incoming packets and a feedback channel. The feedback channel is required to supply for the acknowledgment of error-free packet delivery or for packet retransmission request to the data sender.

FEC is an important low-delay mechanism which can be used as an open-loop error control technique without the need of a feedback channel to reduce or eliminate packet losses in a video communication system. In a FEC scheme, the amount of FEC is pre-determined and based on the probability distribution of the packet loss. However, a larger amount of FEC or redundancy helps to recover from random packet loss event, but FEC codes are not designed to deal with packet burst loss. A burst error may cause substantial degradation to the transmitted video quality [36]. Hence, FEC with adaptive code rate to the importance of bits or frames would be more efficient for controlling errors in data transmission over unreliable channels.

2.1.2 Rate-Distortion Optimized Packet Scheduling

Some recent approaches [37-39] facilitate the selection of optimal channel-adaptive R-D strategies to minimize the expected loss-distortion for different multimedia applications with various quality-of-service requirements and system constraints. An R-D based framework is developed to describe discrete-time Markov chain models in characterizing the random packet-loss process associated with packet network transport system which is then used to analytically investigate the overall efficacy of packet-level protection in improving the end-to-end media quality. One of R-D optimized (RDO) streaming framework was proposed by Chou et al. [26]. RDO can be used to transmit a group of interdependent data unit before a delivery deadline by using feedback information and retransmission strategies. The RDO transmission policies can be obtained by minimizing the Lagrangian cost function of expected rate and distortion.

As shown in Fig. 2-8, the slope is the tangent to the convex hull of the R-D curve. The RDO policy that minimizes the overall R-D Lagrangian cost: $J(\pi)=D(\pi) + \lambda R(\pi)$. The parameter λ can control the operation of the basic conditions, such as distortion sensitive or rate sensitive. The expected distortion $D(\pi)$ depends on the error probabilities, and the expected rate $R(\pi)$ depends on the size of the data units.

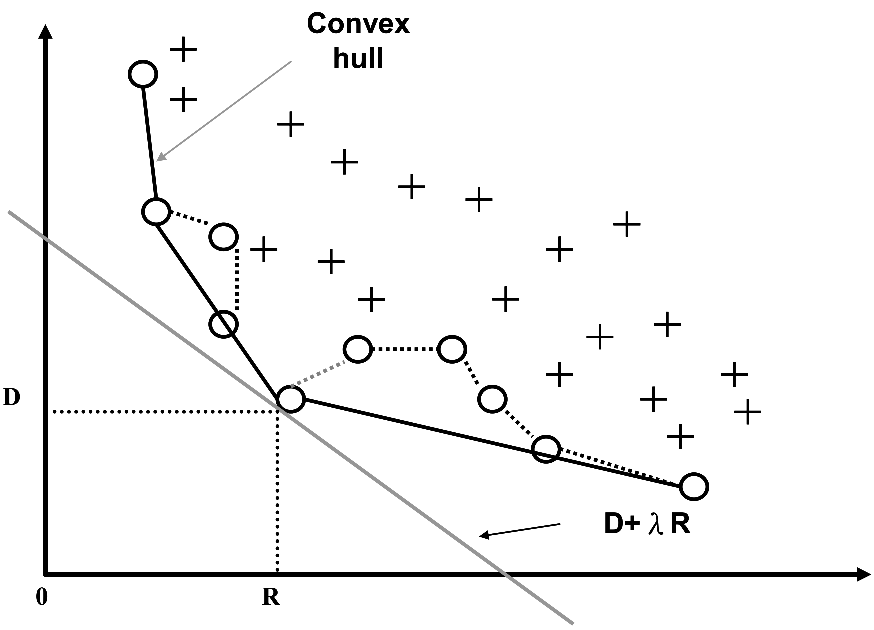


Figure 2-8: The set of R-D pairs, its lower convex hull, and an achievable pair (R, D)

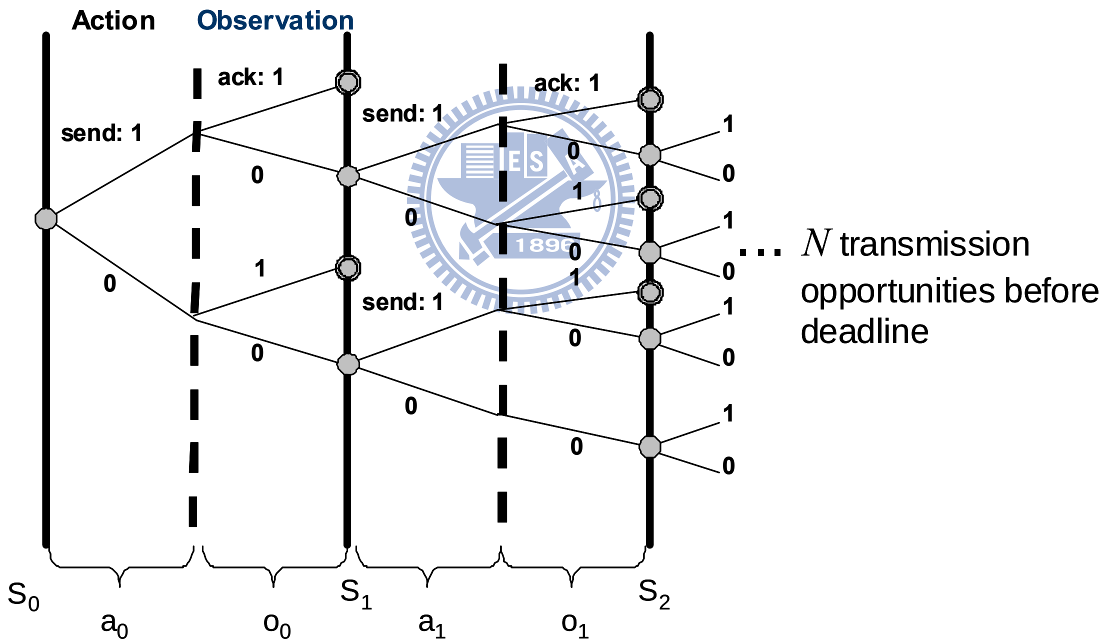


Figure 2-9: Trellis for a Markov decision process. Final state is indicated with double circles.

An example of RDO framework is illustrated in Fig. 2-9. In the figure, we assume that the deadline for the delivery (delivery deadline) prior to each discrete interval is the opportunity to send packets in each transmission opportunities S_0, S_1, ..., S_{n-1}. The sender will be based on the best strategies to select those data to be transmitted to the client. In the action phrase, the sender selects the data unit, then take action a_0 = 1, or not to send data unit, then take action $a_0 = 0$. In the observation phrase, if the data unit receives feedback information $o_0 = 1$, or if the unit has not received any information about the feedback message $o_0 = 0$. In the RDO control streaming system, to decide which opportunities can be sent in each packet should be transmitted to the client is based on the deadline of this packet, the transmission histories, the channel statistics, feedback information, the interdependencies among packets, and the distortion reduction. The policy can indicate whether the video packet should be transmitted at each transmission opportunity. However, the scheme does not address the issue of reducing video quality variation over loss channels. Furthermore, the RDO mechanism maps probability of packet losses into rate increment of redundant packet transmission. The drawback of this approach is that redundant packet transmission will make the resulting R-D curve impractical.

2.2 An Overview on Peer-to-peer Video Systems with Scalable Coding Streams

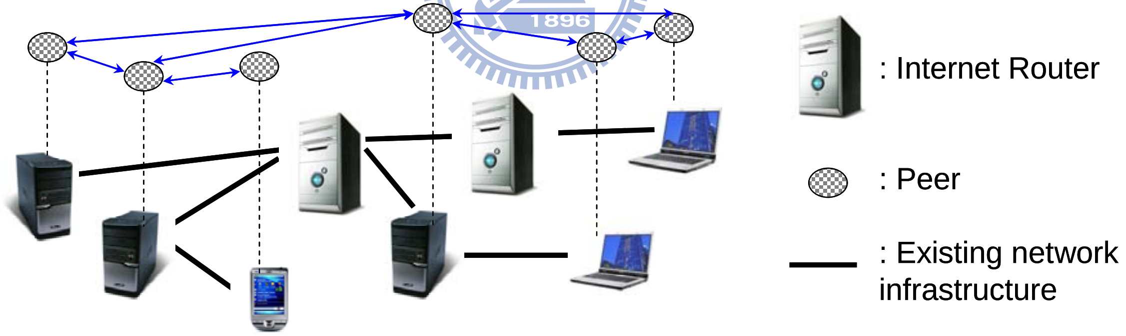


Figure 2-10: An example of P2P overlay network.

The P2P overlay network consists of all the participating peers which cooperate to achieve desired functions on top of the physical network topology, as shown in Fig. 2-10. A wave of innovations on P2P multimedia services has been exhibiting a rapid growth in recent years, ranging from easy and efficient operations of home-to-home user to useful and widespread applications for large-scale user communities [40]. The media-resource sharing P2P systems require a large amount of media repositories [41] and contribute bandwidth to serve other peers. Hence, the P2P multimedia services have emerged as a virtual society that shares and provides access to the resource of peers among P2P community members. Therefore, there has been an increase in research focusing on how to maximize the file availability and minimize the total storage requirements in P2P networks.

2.2.1 Comparison between IP Multicast and Peer-to-Peer Networks

A unicast video stream is dedicated to a single client. In some internet applications, data delivered from a sender to multiple receivers is required, such as video broadcasts. IP multicast [42-44] provides a network-layer solution, which the data packets are delivered to all nodes with respect to a given scope. The IP multicast service can offer scalable one-point to multipoint delivery necessary in the multicast group communications. However, new set of protocols and mechanisms are needed to build multicast services at the network layer, and the technology is still years away from a large commercial deployment by Internet service providers (ISPs). On the other hand, P2P technologies have gained immense popularity as a form of content distribution activity all over the world. P2P services can be fundamentally more scalable than existing multicast methods, particularly when the bandwidth availability is high at the ISP backbone. Several content providers are using legal P2P networks to reduce operational expenses, such as BBC (British Broadcasting Corporation) iPlayer service [45], Babelgum [46], Jaman [47], and GridNetworks [48]. The primary shortcomings of multicast transmissions are summarized below:

- **Configuring routers for multicast:** Routers perform multicast routing and multicast forwarding functions in the IP multicast.
- **Maintenance of the corresponding multicast group:** Nodes of a multicast group need to exchange Internet Group Management Protocols (IGMP) [49, 50] among routers.
- **High complexity and serious scaling constraints at the IP layer:** It is almost impossible for clients to receive the same video quality during different multicast sessions. The majority of high complexity requirements are caused by potentially

large numbers of clients and sessions [51].

- **Not widely deployed:** The implementation of multicast protocols needs special commercial routers which need to be PIM-compatible (Protocol Independent Multicast [52, 53]).
- **Error, flow and congestion control is more complex than in unicast applications:** Advanced features in reliability, congestion control, flow control, and security have been shown to be more difficult than in the unicast case [54, 55].

2.2.2 Difference between Streaming-based and File-based P2P Media Systems

An appropriate fault-tolerant design for P2P VoD systems can help moderate performance degradation during peer failure [56]. This is important especially because continuous operation is an essential feature of P2P VoD systems (e.g. video playback). Another challenging aspect of P2P VoD systems is fault-tolerant design to replicate multimedia files in proper quantities. Replication can hold the greater share of media repositories during high service demand [57]. Thus, numerous P2P replication schemes have been developed for various performance objectives, such as improved startup time, media-file availability, and response time. As shown in Fig. 2-11, P2P replication schemes can be classified into two major types: file-based and media streaming-based data replication. File-based data replication systems are designed for file sharing through downloads, with a typical focus on maximizing file availability or hit rate [58]. There are usually no time-critical constraints associated with transmission (e.g. BitTorrent (BT) [59]). In contrast, media streaming-based data replication systems should determine which data has to be received in time for continuously streaming delivery. In addition, a gracefully degradable quality of video and audio play-back is acceptable, so the transmission of the video and audio stream usually allows some packets to be dropped. For this reason, the BT-like model is less suitable for media streaming-based systems. The remaining differences between file-based and media streaming-based data replications are described as follows.

- **Point-to-Point connections:** The UDP is typically the default network transport protocol used by streaming-based applications. In order to guarantee delivery, the Transmission Control Protocol (TCP) is often used for file-based P2P systems (pull operation) applications.
- **Resolution, frame rate, and bitrate:** For streaming-based P2P systems, the content provider can use flexible approaches to adapt the dynamic QoS requirements of fine granularity scalability by controlling resolution, frame rate, bitrate and video dimension. In contrast, file-based applications are less able to dynamically adapt after the encoding procedure.
- **Quality of streaming session:** In case of severe bandwidth fluctuations or packet losses in the network, TCP-based media delivery may lead to a large number of TCP retransmissions that lead to a higher average delay and jitter. Streaming-based P2P systems are better suited for supporting the playback quality of a streaming session.

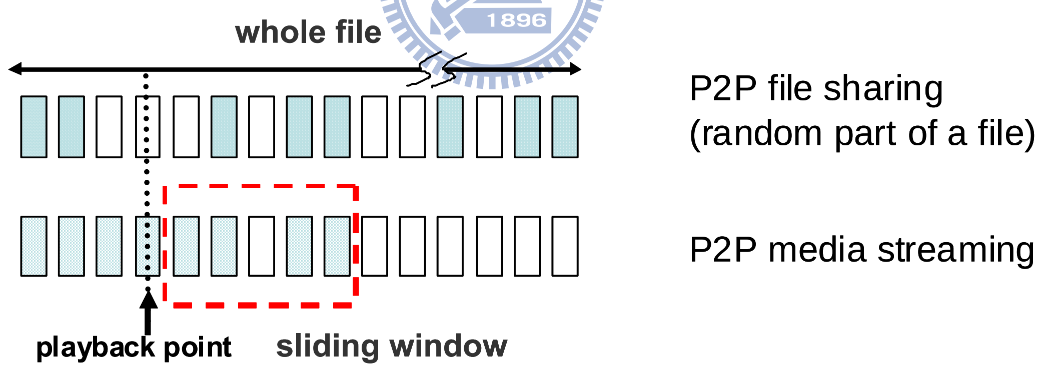


Figure 2-11: A Comparison of the P2P file sharing and P2P streaming scheme.

The problem of streaming-based data replication has been addressed in the past. Ghandeharizadeh et al. [60] developed a home-to-home (H2O) optimized replication framework in a P2P environment. The system minimizes startup-latency and replica-blocks of clips among H2O devices in a hybrid way that replicates partial urgently-blocks and perfecting. A key issue in an H2O-based replication framework is that all video segments are treated equally without discrimination. Such a framework typically relies on the use of non-scalable bit-streams. On the other hand, a scalable bit-stream allows graceful degradation when the smaller distortion of the bit-stream gets lost. Each scalable bit-stream contains a base layer and one or more enhancement layers. Moreover, the base layer is encoded at a lower bit-rate, which is critical in perceptual errors of media data. Hence, the base layer shall have stronger protection against packet losses than the enhancement-layer.

Wan et al. [61] employed local active rendezvous servers (LARS) to coordinate external other LARS and control availability of clips for the P2P streaming in a hierarchical overlay network. Additionally, each LARS is responsible to provide good coverage in its own subgroup. Here network bandwidth and peer capabilities are major factors when generating replicas. The scalability of peers and movies is limited by the special applications. Also a request-rate minimization policy was proposed by Poon et al. [62]. In particular, they studied different replication strategies and showed that the better approach for replication is not only maximizing file availability but also adapting to bandwidth constraints of peers. Ye and Chiu [63] study decentralized and cooperative resource allocation problems with preferential replication in unstructured P2P systems. It is shown that the availability and importance of files are both significant factors in the management of replication service systems. The above types of data replication schemes show that the overlay path length, heterogeneous peer availability and packet loss rates are major factors of perceptual errors of media data in P2P systems. In addition, the replication design of video clips in P2P systems should be postulated based on the popularities and quality impact of video.

Chapter 3 Fine-level Packetization and Streaming of Wavelet Video over IP Networks

This chapter presents a framework of fine-level packetization scheme for streaming of 3D wavelet-based video content over lossy IP networks. The tradeoff between rate and distortion are controlled by jointly adapting scalable source coding rate and level of FEC protection. A content-dependent packetization mechanism with data-interleaving and Reed-Solomon protection for wavelet-based video codecs is proposed to provide unequal error protection. This chapter also tries to answer an important question for scalable video streaming systems. Given extra bandwidth on the systems, should one increase the level of channel protection for the most important packets, or transmit more scalable source data? Experimental results show that the proposed framework achieves good balance between quality of the received video and level of error protection under bandwidth-varying lossy IP networks.

The rest of this chapter is organized as follows. Section 3.1 gives the introduction. Section 3.2 presents a detail analysis on the wavelet compressed video bit stream and the bit stream architecture for fine-level protection. The details of the proposed packetization scheme and streaming framework are described in Section 3.3. Some experimental results of the proposed system are shown in Section 3.4. Finally, some conclusions and discussions are given in Section 3.5.

3.1 Introduction

There is a growing demand for video transmission over heterogeneous networks for communication and entertainment applications. SVC techniques are often proposed for such systems since, ideally, a video sequence can be encoded once and adapted on the fly to different frame rate, bitrate, and resolution for different applications. Although scalable video is an interesting concept, it takes end-to-end system design to show the advantage of SVC over single-layer coding techniques. With single-layer coding, techniques like bitstream switching can be used to achieve video adaptations. However, it is easier to achieve good rate versus source-and-channel distortion tradeoff with scalable coding techniques.

The mainstream video compression techniques are based on hybrid motion compensated transform coding approach, where the transform algorithms are typically either Discrete Cosine Transform (DCT) or 3D wavelet transform. So far, DCT-based SVC approaches have demonstrated better coding efficiency than wavelet-based SVC techniques [64], especially for low bitrate applications. However, a wavelet-based SVC framework can provide fine-granularity bitrate scalability with less system complexity than that of an FGS-based (Fine Granularity Scalability) DCT framework. In addition, many on-going efforts show that wavelet-based SVC approaches still have room for improvement [65]. Therefore, in this chapter, wavelet-based SVC is used as the core codec for the development of a scalable video streaming framework.

The most challenging problem for scalable video streaming over IP networks is about how to optimally adapt source data rate and degree of packet loss protection to real-time network conditions. Video packetization and scheduling algorithms are mostly responsible for mitigating the effects of bandwidth variation and packet losses in the network. The packetization and scheduling algorithms are mainly based on resource-versus-distortion optimization [66-68], where resource can be available computation power, rate, delay, and so forth. A general resource allocation treatment for streaming systems is presented in [69]. Some researches try to apply the rate-distortion optimization (RDO) principle [70] of source coding theories to video streaming over lossy networks [66]. For a streaming system, the distortion is a result from both source coding and channel losses. A key issue in an RDO-based streaming system is that the distortion due to packet losses is much more difficult to quantify than the distortion due to lossy source coding.

Several frameworks for 3D wavelet based video streaming system have been proposed recently. Chu and Xiong [71] introduced a combined packetized wavelet video coding and FEC approach for video streaming and multicasting. The packetized wavelet video coder marks the truncation points of the bit stream at the nearest packet boundaries instead of the end of each fractional bit plane. The FEC-based error protection scheme applies Reed-Solomon (RS) coding to produce parity packets. And then the scheme broadcasts all source packets to one multicast group and parity packets to different multicast groups. Hence, for each client, the optimal number of layers and error protection can be determined by the packet loss ratio and the available channel bandwidth. However, data interleaving is not used in this work, which makes the system less robust to burst errors. Dong and Zheng [72] proposed a content-based retransmission framework for wavelet video streaming. The compression module adopts dynamic grouping and bounded coding scheme for improving compression efficiency and removing unnecessary dependency to each subband. In the transmission module, a packet includes one or more subbands, and a content-based retransmission is used to provide robustness against transmission errors. The content-based retransmission scheme is based on the importance of packet content which is computed by the square sum of coefficients for each wavelet subband. Later, Zhao et al. [73] incorporated an error concealment scheme into this content-based retransmission framework to increase its error resilience capability. Nevertheless, retransmission-based error control requires larger jitter buffer and may consume too much extra bandwidth in high error rate channels [74, 75].

Chou and Miao [26] developed a framework for RDO streaming of packetized media. The RDO framework is flexible to extend the optimizing packet transmission scheduling to a wide range of receiver/sender/proxy driven streaming systems [76]. However, the scheme maps probability of packet losses into rate increment of redundant-packets transmission (ARQ can be avoided in this approach). However, although redundant-packets transmission makes the RDO system simpler for analysis, it is not cost-effective for practical systems. R-D performance can be greatly improved if FEC is used instead. Zhu et al. [68] proposed a congestion-distortion optimized scheme. Zhai et al. [69] presented an integrated joint source-channel coding framework for video streaming. Wang et al. [77] proposed a cost-distortion optimization framework. Chang et al. also proposed sender-based [78] and receiver-based [79] RDO frameworks for 3D wavelet video streaming, which basically follow the framework introduced by Chou [66]. The proposed system uses source rate-distortion profiles to optimize for playout latency and bandwidth allocation among a group of data packets in a way that minimizes distortion in the reconstructed frames.

There are many error control schemes for video streaming, including FEC [80-83], unequal error protection (UEP) [84-86], and ARQ [87]. Until recently, error control schemes for streaming systems are designed independent of rate control schemes. Joint design of error and rate control is important to a variable bandwidth lossy network. For example, when the channel bandwidth increases during runtime, should more bits be allocated to send extra (enhancement) source data, or to increase the level of protection of crucial (also known as base layer) source data? Based on the RDO principle, one should pick whichever approach that reduces more distortion. However, this is not trivial since distortions from channel losses are nondeterministic. Another issue is that not all source data bits carry equal amount of information (i.e. entropy). Although some of the error control techniques try to put different degree of protection based on the degree of importance of the content, unequal error protection is done coarsely since the error control scheme is based on either single-layer video coding model or coarse-granularity layered scalable video coding mode.

In this chapter, a fine-level packetization scheme for wavelet-based streaming video is proposed. The mechanism is based on detailed analysis of the mainstream wavelet-based video codec [88]. Due to its fine-granularity SNR scalability feature, the proposed packetization scheme can apply various error correction capabilities of Reed-Solomon (RS) codes on interleaved video subband data so that the streaming video is robust over IP networks. In addition, the chapter proposes to map the distortion caused by packet loss to distortion caused by source data rate reduction due to extra FEC protection for error-free transmission. Since measuring operational video distortion from packet loss is very difficult while measuring source coding distortion is much simpler, the proposed mechanism can be applied to practical systems. In summary, the main features of the proposed system are highlighted as follows.

- The streaming algorithm searches along the R-D curve for an optimal operating point between the scalable source coding rate and the FEC protection level.

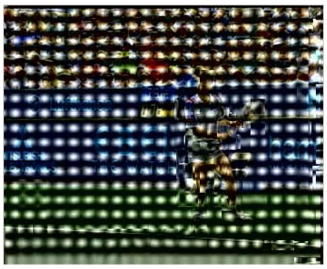
- The FEC protection level is also influenced by runtime channel feedback from the client. Therefore, the proposed system is adaptive to both the video content entropy and the run-time packet loss rate.

- The rate-distortion tradeoff of the system takes into account both distortion due to source data rate reduction and distortion due to packet losses (predicted by FEC protection bits required for error-free transmission).

3.2 Investigation of Wavelet Video Bit Streams with Data Losses

For streaming applications, the quality of video is affected by packet losses. One of the most difficult problems for RDO streaming is about how to measure the distortion caused by packet losses. The distortion depends heavily on the source coding method. In this section, the wavelet video coding schemes presented in [88] are investigated in detail. In particular, some experiments are conducted to exhibit the impact of different wavelet subband data losses on the reconstructed video quality.

In order to gain better insight into the significance of different bit stream segments across different temporal subbands, some experiments are conducted. For example, using a four-level MCTF temporal decomposition, a group of frames is temporally decomposed into the *LLLL*, *LLLH*, *LLH*, *LH*, and *H* subbands. In addition, each temporal subband may further be spatially decomposed. For an encoded video with four-level temporal and three-level spatial decompositions, each temporal subband (TSB) is split into nineteen spatial subbands (SSB) indexed from 0 to 18. The distortion impact of the first coding block within a higher spatiotemporal subband (e.g. Figures 3-1(b), 3-1(c), 3-1(d)) is indeed more sensitive than that of the last coding block within a lower spatiotemporal subband as shown in Fig. 3-1(e).

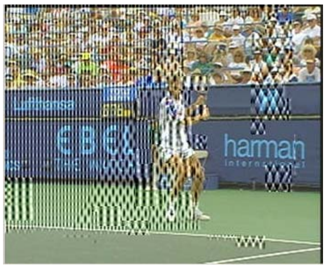


(a) $P(LLLL_t,Y)$

(b) $P(LLLH_t,Y)$

(c) $P(LLH_t,Y)$

(d) $P(LH_t,Y)$



(e) $P(LLLL_t,Y)$

(f) $P(H_t,Y)$

Figure 3-1: Reconstructed video when a chunk of TSB data is lost. The loss occurs in coding block 0 of SSB 0 for the TSB in (a)–(d), and coding block 0 of SSB 18 for the TSB in (e)-(f).

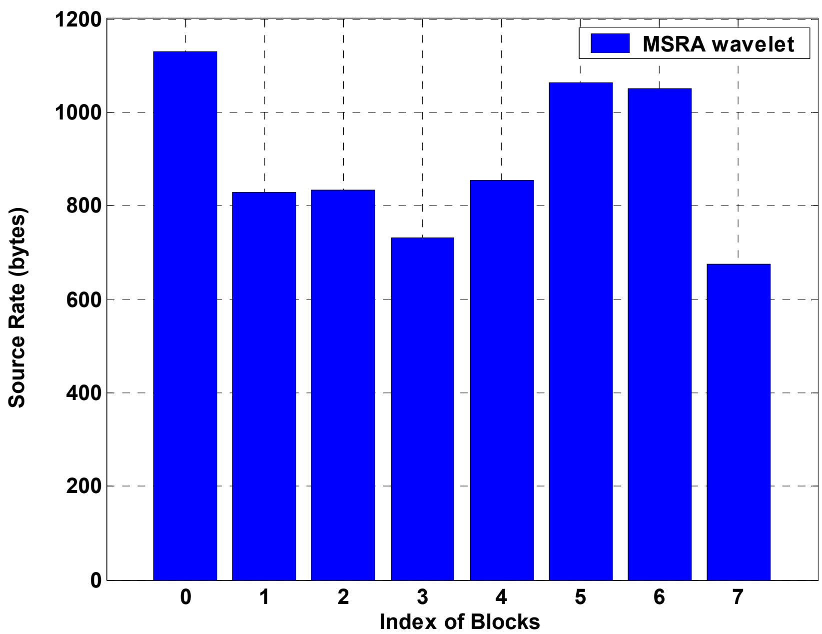


Figure 3-2: Source data rate in SSB 0 of subband $P(H_t, Y)$ of STEFAN.

In practice, given an estimated packet loss rate, different amount of error protection should be applied to different portions of a coding block based on their influence on visual quality. Therefore, further "rate" versus "channel-distortion" analyses of wavelet

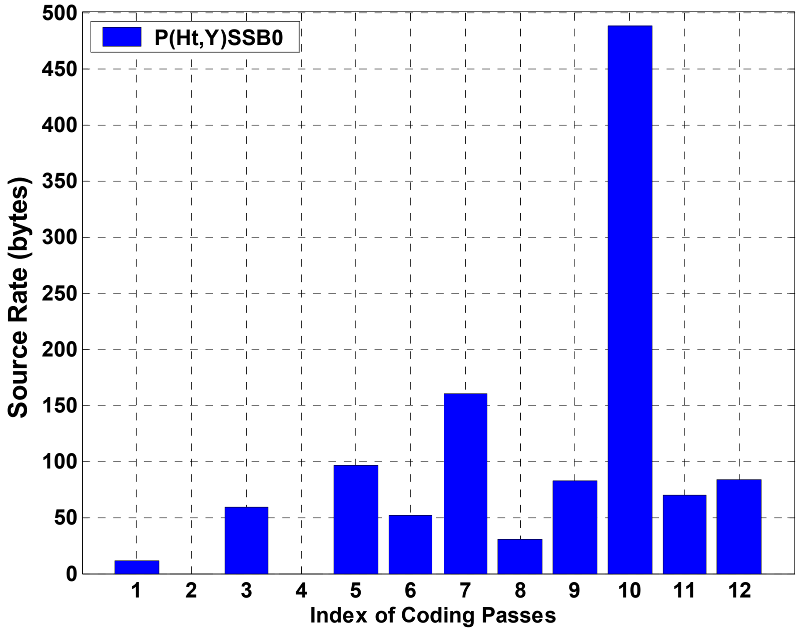


Figure 3-3: Source data rate of coding passes on the convex hull in the block 0 of STEFAN.

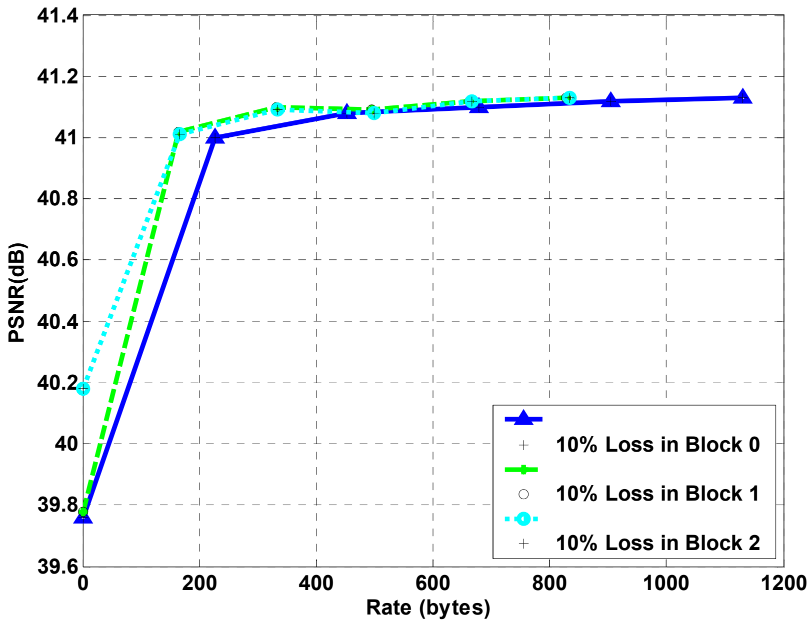


Figure 3-4: R-D curves of STEFAN with 10% loss of coding passes in SSB 0 of the TSB $P(H_t, Y)$.

subband data are conducted as follows. Since the size of different coding blocks varies (see

Figure 3-2), it is not suitable to use coding block as the data interleaving unit for FEC protection. A coding block should be split into several smaller units for data interleaving. Within each coding block, the bit stream size of the first coding pass is usually small (see Figure 3-3), but it has major impact on video quality (see Figure 3-4). To evaluate the effect of degradation from burst data loss, a 10% burst loss of bits is placed in different portions of a coding block as shown in Fig. 3-5. When the burst data loss is located at the beginning of a coding block, it usually causes large degradation of visual quality. Hence, the error protection level for different portions of a coding block should be different.

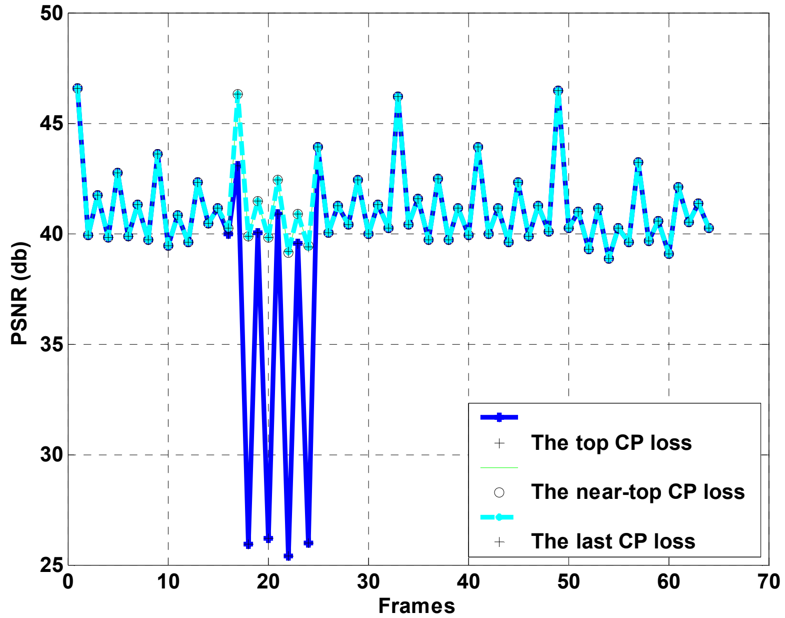


Figure 3-5: PSNR of STEFAN with 10% loss of coding passes (CP) in block 0 of SSB 0 of the TSB $P(H_t, Y)$.

Packet loss is the major cause of nondeterministic distortion for video streaming applications. The main reasons for packet losses are mostly because of network congestion, which causes packet losses in the network router queue buffer [90]. As Fang et al. [89] and Biersack [90] pointed out, FEC protection scheme is effective to recover packet loss with minimum transmission overhead for multimedia streaming. Hence, in this chapter, a fine-level FEC protection scheme for scalable streaming systems is proposed based on previous investigation of channel distortion impact on wavelet video.

The basic concept of our fine-level FEC streaming scheme is to add different FEC protection level to different wavelet subband data based on the R-D slope values of the data set (or, equivalently, the distortion-reduction rates) and the predicted packet loss rate. Figure 3-6 illustrates this concept with some examples of real data. The fine-level FEC protection is applied to the coding block 0 of temporal subband P(H_t, Y) and spatial subband 0 of the STEFAN sequence. In this plot, the y-axis is the distortion reduction rate and the x-axis is the bitrate including source data bits and FEC protection bits. The dashed line is the original subband data without any protection, while the solid line with circle markers is the FEC protected data given 3% estimated packet loss rate and the solid line with "plus" markers is the protected data given 8% estimated packet loss rate. The lower the rate point, the higher the protection level. The equation used to compute the protection level will be described in a moment. Note that the function in Fig. 3-6 can be used for operational RDO streaming decision since the proposed scheme exhibits rate versus source-and-channel distortion tradeoff.

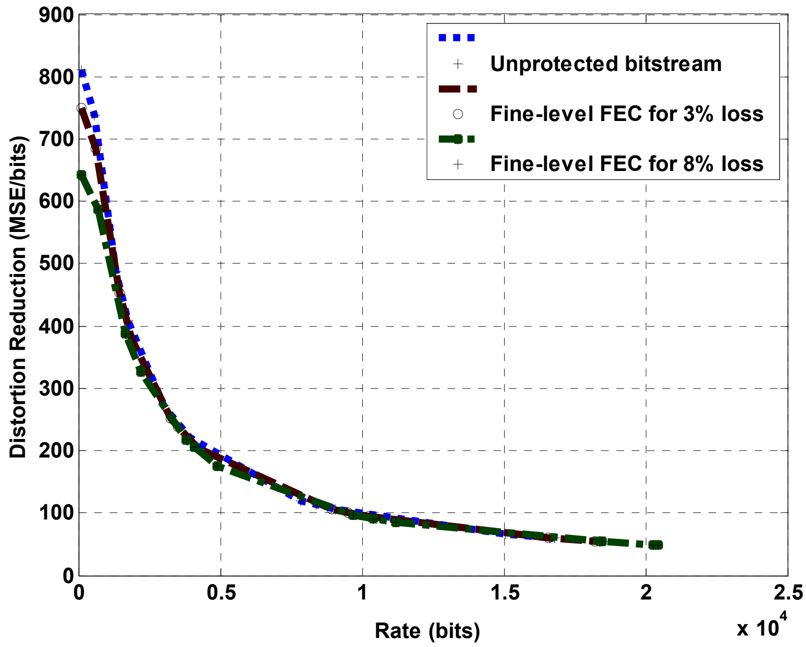


Figure 3-6: Example overhead of fine-level FEC protection for different rate points within a coding block.

In the proposed framework, for each GOP of video bit-streams, an (n, k) Reed-Solomon (RS) code-based FEC is applied to add resiliency to the data. In Fig. 3-7, n is the code word length of the RS encoder, k is the number of video data symbols, and s is the number of correctable symbols. The number of parity symbols is $2s$, where $2s = n - k$. If burst errors occur during transmission, then the RS decoder can correct up to s errors and detect up to $2s$ errors per code word.

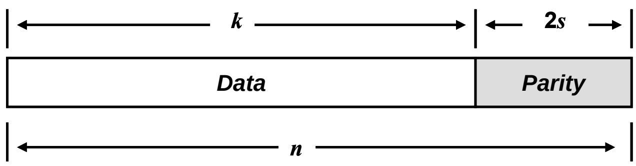


Figure 3-7: An (n, k) RS codeword with k symbols of video data and 2s symbols of parity.

3.2.1 Fine-level adaptive FEC Protection of Wavelet Coefficients

For 3D-ESCOT, each coding block j has temporal level index ω_j, component index v_j, and spatial subband index τ_j. Assuming that the bitstream of a coding block is divided into l codeword, the importance of a coding block can be expressed as in Eq. (3-1):

$$c_j(x, y) = \left\lfloor \exp\left(\frac{\alpha}{y}\cdot\sum_{n=0}^{x}\left(\frac{(T-\omega_j)\cdot U_1}{T} + \frac{U_2}{(Y-v_j)} + \frac{1}{(B-\tau_j)}\right)\right)\right\rfloor, \tag{3-1}$$

where x is the index of the sub-codword and $x = 0, 1, \ldots, l-1$, y is the R-D slope of the first coding pass in block j, α is a scale factor, T is the maximal temporal level index, Y is the maximal component index, B is the maximal spatial subband index, and U_1 and U_2 are weighting factors. Note that the value of $c_j(x, y)$ is defined to be $0 \le c_j(x, y) \le n/2$. The protection level of the fine-level FEC scheme is determined based on the characteristics of the coding block $c_j(x, y)$ given in Eq. (3-1) subject to the network conditions. A coding block is composed of several coding passes. Since the coding passes of a coding block are roughly ordered based on their importance to visual quality, therefore, the protection level applied to different coding passes (indexed by x) of block j is proposed to be $s_{j,x}$, as defined in (3-2):

$$\hat{s}_{j,x} = \left\lfloor \exp\left(\frac{|\lambda_{j,0}|}{\beta}\right) \cdot n_{pl} \right\rfloor - c_j\left(x, |\lambda_{j,0}|\right),$$

$$s_{j,x} = \hat{s}_{j,x} + o, \;\; o = \begin{cases} 0 & \text{if } \hat{s}_{j,x} \text{ is even} \\ 1 & \text{if } \hat{s}_{j,x} \text{ is odd} \end{cases}, \tag{3-2}$$

where $0 \leq s_{j,x} \leq n/2$, $\lambda_{j,0}$ is the R-D slope of the first coding pass in block j, n_{pl} denotes the estimated packet losses given current bandwidth R_{BW}, average packet size P_s, and packet loss rate $\varepsilon_{\mathrm{pl}}$, and β is a scale factor determined empirically. In Eq. (3-2), the level of protection decreases with coding passes order, and $s_{j,0} \geq s_{j,1} \geq \cdots \geq s_{j,l-1}$. Note that $n_{\mathrm{pl}} = \lfloor \varepsilon_{\mathrm{pl}} \times R_{\mathrm{BW}}/P_s \rfloor$, where the operator $\lfloor \cdot \rfloor$ is the floor operation.

3.2.2 Multiple-adaptation and Fine-level FEC Using R-λ Curve Fitting Function

Multimedia distribution over heterogeneous networks and devices has become the mainstream enabling technology for new generations of services. For distribution and playback of a video content on various devices under different network conditions, scalable video coding schemes are usually used. A typical approach for scalable coding is to use a layered coding approach such as MPEG-4 Simple Scalable Profile or fine-grained scalable video coding scheme. In these approaches, the bitstream quality of the encoded video is optimized subject to certain bitrate conditions. Adaptation of the encoded video file to a different target bitrate usually results in sub-optimal bitstreams.

A different approach from the layered coding schemes is to design a scalable codec that produces embedded scalable bitstreams without inherent layered structures. The wavelets-based video codecs belong to this category. After the encoding procedure, there is no inherent layer structure for wavelet video bitstreams, video parameters such as resolution, frame rate, and bitrate can be dynamically adapted with fine granularity. If the R-D tradeoff information is embedded in the bitstream, the adaptation process can produce an R-D optimal bitstream at run-time for the target application. One major advantage of wavelet codecs over coarse-granularity layer-based codecs is that wavelet bistreams can facilitate multiple adaptations. For example, in Fig. 3-8, the video server transmits dynamically adapted scalable bitstreams to two different devices, namely the notebook and the cellular phone. Upon reception of the embedded bitstreams, the notebook plays the high quality bitstream on its screen. In addition, the multiple adaptation application of the notebook truncates (adapts) the received bitstream further and send the new bitstream to another device (the PDA) with tighter channel and device constraints. For the other distribution chain in Fig. 3-8, the cellular phone first receives an adapted bitstream from the server and plays it on its internal large screen. Later, when the user decides to watch the video on the small external screen to conserve power, the video decoder can extract and decode only part of the received bitstream and display a smaller resolution video.

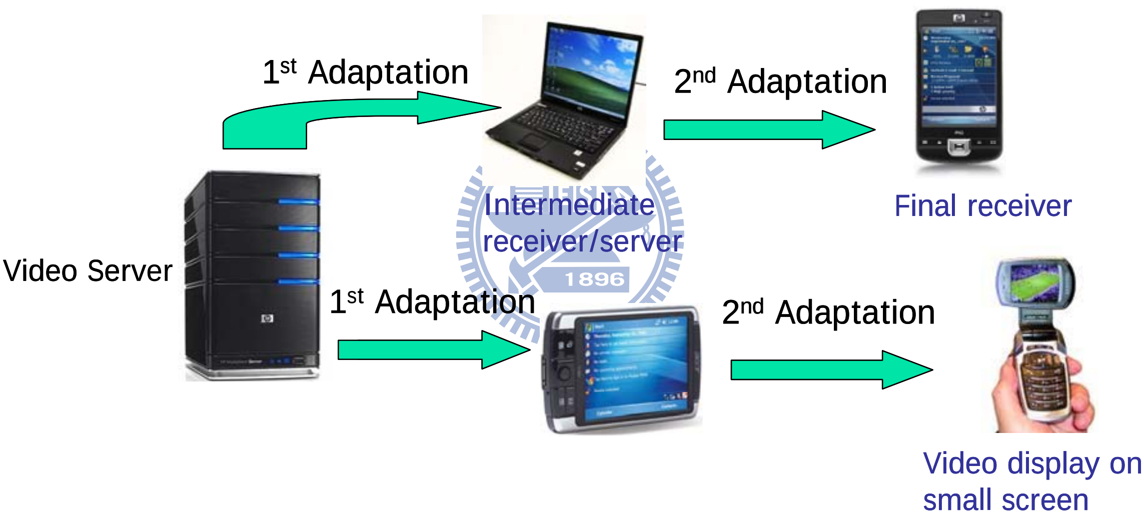


Figure 3-8: Two examples of multiple adaptation applications where the same video content is adapted several times down the distribution chains

The R-D information of the wavelet codec is the discrete rate-lambda (R-λ) pairs for all coding passes. In addition to multiple adaptations of video data, R-D information is also very useful for fine-level FEC protection of video data. Several frameworks for wavelet based video streaming have been proposed in the literature recently [71, 72, 78, 79]. However, none of existing works allows multiple-adaptation of fine-level of FEC protection data. The system can be configured in multiple adaptations of video data. R-D information is also very useful for fine-level FEC protection of video data. In addition, fixed amount of FEC protection consumes considerable overhead, if channel errors are less than expected. Section 3.2.1 proposed a fine-level FEC protection/packetization mechanism of wavelet video data, but multiple adaptations of FEC

Table 3-1: The amount of R-D information.

Sequence Name	*Resolution/ frame rate/bit rate(kbps)*	*R-D information (bvtes, % in bitstream)*		*Saving Ratio*
		MSRA	*Proposed Method*	
Mobile	CIF / 15 / 4096	898,872 (2.14%)	307,120 (0.73%)	65.89%
	CIF / 15 / 2048	557,560 (2.66%)	245,696 (1.71%)	56.02%
	CIF / 15 / 512	163,944 (3.13%)	76,780 (1.47%)	53.04%
	CIF / 30 / 2048	600,648 (2.86%)	297,500 (1.42%)	50.35%
	CIF / 30 / 512	164,864 (3.15%)	77,734 (1.48%)	53.02%
Foreman	CIF / 15 / 1152	433,112 (3.67%)	227,360 (1.93%)	47.41%
	CIF / 15 / 640	290,000 (4.43%)	140,963 (2.15%)	51.47%
	CIF / 15 / 256	127,120 (4.86%)	58,581 (2.24%)	53.91%
	CIF / 30 / 1152	505,224 (4.28%)	245,089 (2.08%)	51.40%
	CIF / 30 / 640	308,480 (4.71%)	142,710 (2.18%)	53.72%
Stefan	CIF / 15 / 4096	805,856(1.92%)	278,960 (0.67%)	65.10%
	CIF / 15 / 3072	711,888(2.26%)	312,435 (0.99%)	56.19%
	CIF / 15 / 1024	350,584 (3.34%)	156,217 (1.49%)	55.39%
	CIF / 15 / 512	218,112 (4.16%)	103,215 (1.97%)	52.64%
	CIF / 30 / 3072	883,672 (2.81%)	380,304 (1.21%)	56.94%
	CIF / 30 / 1024	404,744 (3.86%)	190,152 (1.81%)	53.11%
Average		3.39%	1.60%	54.73%

Table 3-2: R-D information overhead vs. GOP size

Sequence Name	*Resolution/ GOP size/bit rate(kbps)*	*R-D information (bytes, % in bitstream)*		*Saving Ratio*
		MSRA	*Proposed Method*	
Mobile	CIF / 64 / 4096	898,872 (2.14%)	307,120 (0.73%)	65.89%
	CIF / 32 / 4096	1,011,720 (2.41%)	307,440 (0.73%)	69.71%
	CIF / 16 / 4096	890,696 (2.12%)	306,720 (0.73%)	65.57%
Foreman	CIF / 64 / 1152	433,112 (3.67%)	227,360 (1.93%)	47.41%
	CIF / 32 / 1152	552,744 (4.69%)	284,200 (1.95%)	58.85%
	CIF / 16 / 1152	431,088 (3.66%)	227,200 (1.93%)	47.27%
Stefan	CIF / 64 / 4096	805,856 (1.92%)	278,960 (0.67%)	65.1%
	CIF / 32 / 4096	1,005,192 (2.40%)	278,720 (0.66%)	72.5%
	CIF / 16 / 4096	759,384 (1.81%)	278,560 (0.66%)	63.54%
Average		2.76%	1.11%	61.76%

codes are not considered. Transmission of the R-D information was a non-negligible fraction of the data rate. The R-D information can be used to calculate the degree of importance of the wavelet coefficients given estimated packet loss rate of the channel. The granularity of the protection level can be fine-tuned for different wavelet video coefficient coding passes. Although the proposed technique performs very well in practice, it does not allow for multiple adaptations since the R-D information will be discarded after packetization due to its nontrivial overhead. The amount of R-D information overhead in different bit rates and frame rates is illustrated in Table 3-1. And the amount of R-D information overhead in various GOP sizes is shown in Table 3-2. The average R-D information percentage in the bitstream is ranging from 1.9% to 4.9%.

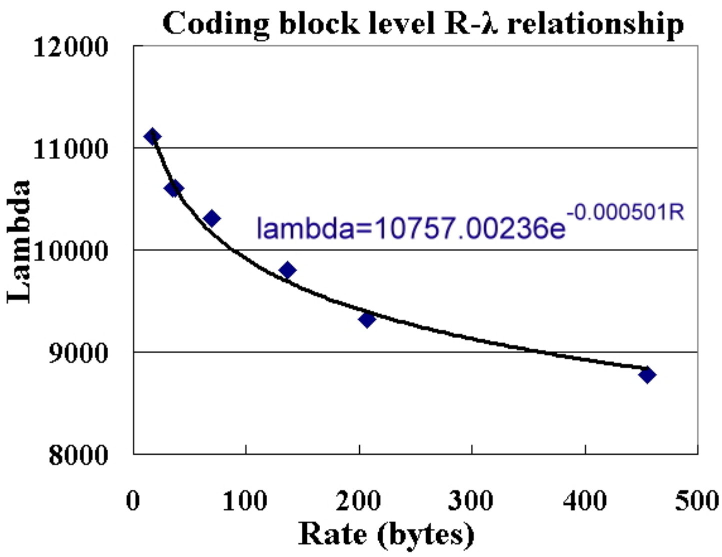


Figure 3-9: R-λ curve fitting function of STEFAN in block 0 of SSB 0 of the TSB $P(H_t, Y)$

The proposed model translates the discrete R-λ pairs into the R-λ function, where λ is R-D slopes. In addition, the R-λ function is incorporated into the fine-level FEC framework to facilitate multiple adaptations. For this purpose, a R-λ curve fitting function for multiple adaptations was proposed in our previous study [91]. The R-λ curve fitting function is an approximation of the functional relationship between rate R and the Lagrange parameter λ. We assume that the R-λ curve fitting function: $\lambda = \alpha e^{\beta R}$ in coding block level can be rewritten as $\ln\lambda = \ln\alpha + \beta \cdot R$, as shown in Fig. 3-9. Therefore, we have an over-determined system of Eq. (3-3):

$$\begin{pmatrix} \ln\lambda_0 \\ \ln\lambda_1 \\ \vdots \\ \ln\lambda_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & R_0 \\ 1 & R_1 \\ 1 & \vdots \\ 1 & R_{n-1} \end{pmatrix} \begin{pmatrix} \ln\alpha \\ \beta \end{pmatrix}. \tag{3-3}$$

where $n > 2$. A least squares fitting technique used for solving the parameter α and β. Note that for fine-level FEC protection, the degree of protection level s should be based on the importance of the video data. The importance of the coefficients within a coding block in a subband can be ranked based on the R-D information of the coding block. After wavelet decomposition, the subbands can be arranged and indexed from low to high frequencies. The smaller the index is, the lower the frequency is. Therefore, each coding block in subband i has a temporal subband index ω_i and a spatial subband index τ_i. The importance of the coefficients in a coding pass is first determined by the importance of the coding block it is located in. The importance of a coding block is in turn determined by the subband it is located in. The importance factor, W_i, of a coding block is computed as in Eq. (3-4):

$$W_i = \exp\left[(-1)\cdot\left(\frac{(T-\omega_i)\cdot U_1}{\mathrm{T}} + \frac{1}{(Z-\tau_i)}\right)\right], \tag{3-4}$$

where T is the maximal temporal level index, Z is the maximal spatial subband index, and U_1 is a weighting factor.

The level of FEC protection is defined by the value s, the number of correctable symbols. Without loss of generality, assume that the bitstream of a coding block j is divided into m codeword. The protection level $s_{j,x}$ of the coefficients in coding pass x of coding block j is computed as in Eq. (3-5):

$$\hat{s}_{j,x} = \left\lfloor \left(\frac{\alpha_j \cdot \exp\left(\beta_j \cdot \sum_{k=0}^{x} R_{j,k}\right)}{\omega} \right) \cdot n_{pl} \cdot W_j \right\rfloor, \tag{3-5}$$

$$s_{j,x} = \hat{s}_{j,x} + o,\quad o = \begin{cases} 0 & \text{if } \hat{s}_{j,x} \text{ is even} \\ 1 & \text{if } \hat{s}_{j,x} \text{ is odd} \end{cases}$$

where x = 0, 1,…, m-1, the parameters α_i and β_i are the close-form R-λ model: $\lambda = \alpha e^{\beta R}$ parameters for the coding block j, $R_{j,x}$ is the length of the x_{th} RS codeword in coding block j, n_{pl} denotes the estimated number of packet losses per second, and ω is a scale factor determined empirically. In Eq. (3-5), the level of protection decreases following fractional bitplane coding pass order, and $s_{i,0} \geq s_{i,1} \geq \ldots \geq s_{i,m\text{-}1}$.

For some multiple-adaptation applications, the second (and above) adaptations may be due to the change of device capabilities instead of channel conditions. For such case, there is no need to re-compute the FEC codes since the level of protection does not change. However, repacketization may still be necessary for efficient transmission of the re-adapted data.

3.3 The Proposed Packetization Scheme and Streaming Framework

In the following discussions, we use the terminology "block bit stream segment" "(BBSS) to describe a portion of bit stream of a coding block across spatiotemporal subbands (see Figure 3-2). A BBSS is composed of one or more coding passes. The packaging of the scalable bit streams into UDP packets is accomplished following both rate control and error control constraints. These constraints try to fulfill the following goals.

- Error protection level of a BBSS should depend on its entropy. The higher the entropy has the higher the protection level. Note that since a BBSS is only a small chunk of data in a coding block, the granularity of content adaptation of the FEC protection is at a very fine scale.
- The streaming packet rate of the system should stay as low as possible. UDP packet size should be smaller than the MTU (maximum transmission unit) allowed by the network links. The typical size of MTU is around 1500 bytes for wired networks. MTUs ranging from 250 to 750 bytes commonly have better throughput under no bit-error-rate circumstances for mobile ad hoc networks [92].

On the other hand, processing a lot of small packets causes very high overhead to the streaming system, especially on the client side. Therefore, a reasonable packet size is slightly smaller than the MTU.

- Although interleaving technique with FEC works well for handling packet losses, it does introduce extra delay to the transmission of video data. Therefore, the selection of interleaving size must take into account the end-to-end delay of the whole system. In general, for broadcast video streaming, overall delay should be less than 20 seconds [93].

3.3.1 Packetization of FEC-protected Data

As mentioned in the previous section, a systematic Reed-Solomon (RS) code word comprising of data and parity is used for fine-level FEC protection. RS coding used for the protection of the BBSSs is depicted in Fig. 3-10. Assume that the total number of coding block is L, $j = 0, \ldots, L-1$, for each coding block j, bit stream can be divided into m-BBSSs, each unit j begins with the first BBSS $C_{j,0}$ and continues through $C_{j,1}$, $C_{j,2}$,... to $C_{j,m}$. An $(n, s_{j,l})$, $l = 0, \ldots, m$, RS code is then applied to add resiliency to the m-BBSSs. Since the BBSSs have varied in size, one must pack variable number of BBSSs into a data unit to reduce packet overhead. In addition, different levels of protection allocate to different portions of the coding block, $s_{j,m} \geq s_{j,m-1} \geq \ldots \geq s_{j,0}$. Furthermore, the BBSSs gathered at the front end of the data unit, and the RS symbols are located at the end of the data unit. For each data unit, the header describes the protection level of the data unit which is also protected by RS coding.

Since we are dealing with a lossy channel, not a bit error channel, a byte-wise data-interleaving scheme is used to shuffle the RS coded data among several data packets before transmission. As illustrated in Fig. 3-11, a BBSS is spread across h packets (each packet is composed of the group of data in dashed lines in Fig. 3-10). For each packet, in addition to video data payload, we also have to transmit the highest protection level, temporal subband index, component index, spatial subband index, and block index in order to properly deinterleave the data. When interleaving is used, the interleaving depth must

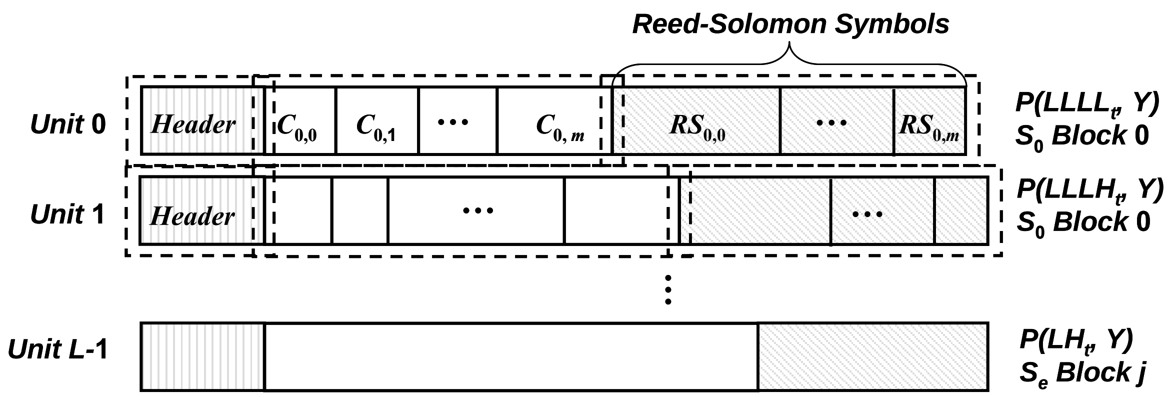


Figure 3-10: Packetization for one GOP of video data.

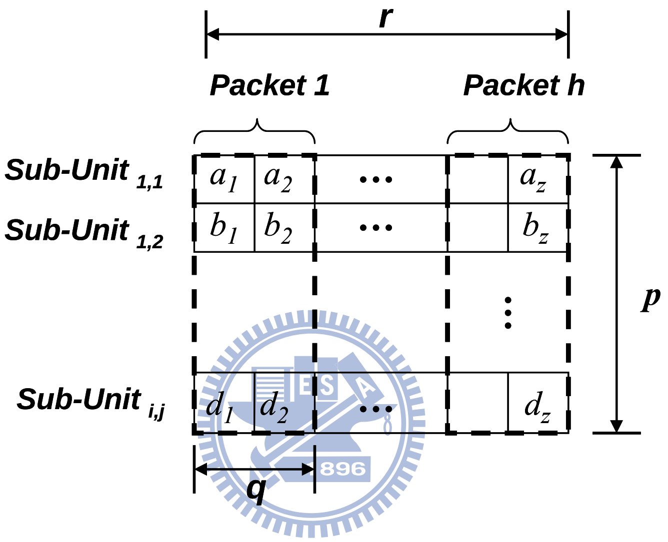


Figure 3-11: Data-interleaving scheme for one GOP of video data.

match the worst case of channel conditions against burst errors. In addition, a large interleaving depth will have impact on the packet buffer size of the client and the end-to-end delay of packet transmissions. The interleaving depth should be appropriately chosen to handle the worst-case error bursts of the networks. As mentioned in Section 5.2, the number of parity symbols is $2s$, where s means the number of correctable errors by an RS decoder. A data unit can be split into several r equal length sub-units and each interleaved packet is composed of q data symbols from each subunit. Hence, q is limited by the number of parity symbols s, and p is limited by the maximum end-to-end delay.

3.3.2 Streaming Policy

The proposed framework will adapt to the fast varying channel conditions by using the real-time network statistics feedbacks from the client side. Through standard RTCP receiver reports, the server can obtain the statistics such as round-trip time (RTT), jitter, short-term packet losses, and accumulative packet losses. The packet loss rate is used to compute the fine-level FEC-protected level as described in Section 5.2. In addition, the server can compute the effective channel bandwidth through the last packet sequence number received by the client and loss rate. Based on the estimated channel bandwidth and the R-D information, the system performs a dynamic rate allocation at discrete transmission time to enhance the perceived quality whenever the network has extra bandwidth for perceptible quality improvement.

For the correction of errors, parity packets are employed to recover from lost data packets. But some parity packets may be lost or corrupted when transmitting packets over the networks based on the UDP protocol. For enhancing the system performance, error recovery mechanisms such as retransmission or error correction can be applied to handle un-correctable errors. Instead of using retransmission to all parity packets, the proposed system delivers more redundancy parity packets to those packets carrying important portion of blocks and fewer to other packets, as shown in Fig. 3-12.

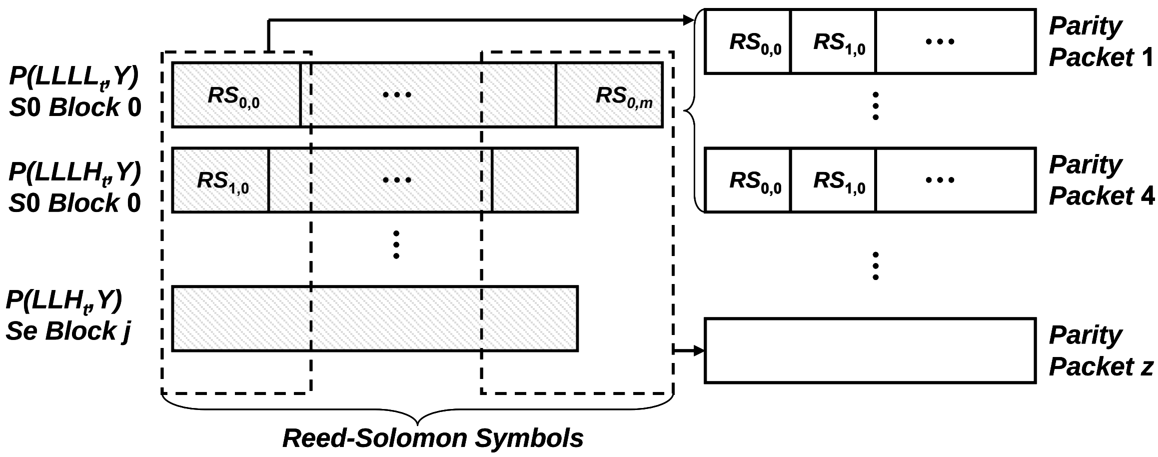


Figure 3-12: Redundancy packets for protect source packets one group of video data.

3.4 Experiments

This section presents the experimental results of the proposed video streaming system. The block diagram of the proposed streaming system is shown in Fig. 3-13. The system is based on the MPEG-21 test bed for resource delivery [94]. The test bed includes an IP transmission channel emulator (based on the NIST net [95]) that allows real-time emulation of various network conditions. We have added Reed-Solomon coding modules, a data interleaving module, and a data deinterleaving module to the original test bed.

The CIF version of the standard MPEG test sequences STEFAN, MOBILE, TABLE TENNIS, FOREMAN, and COASTGUARD is used for the experiments. Those sequences are encoded using MSRA 3D wavelet video coding software [96] at 15 frames per second and a GOP is composed of 64 frames. Four levels temporal decomposition and three levels spatial decomposition are used for subband coding. The number of luminance (Y) blocks is around 1024 BBSSs, and the number of chrominance (U and V) blocks is around 608 BBSSs.

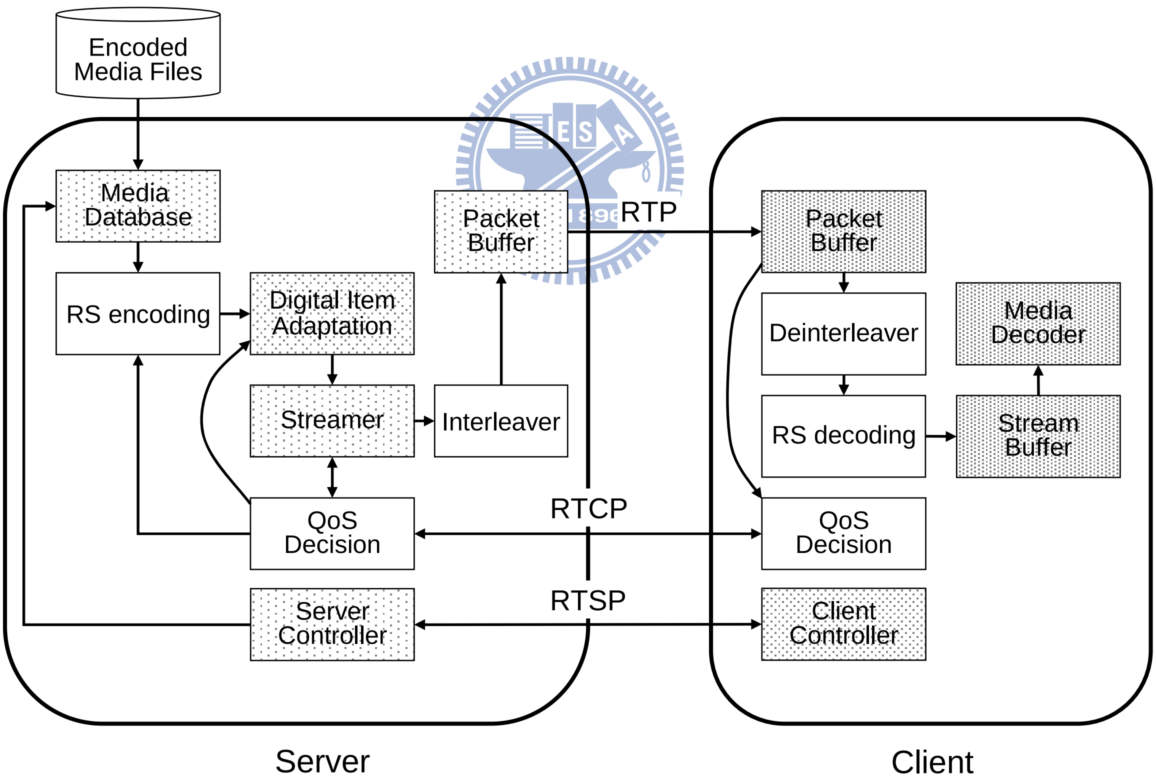


Figure 3-13: Architecture of the proposed video streaming system.

To evaluate the performance of the proposed system, reasonable range of packet loss rates should be used. Over wired links, studies showed that based on MPEG compressed video using the RTP and UDP transport protocols reported the average packet loss rates, ranging from 3.0 to 13.5 percent [97]. Over wireless links, Lai et al. [98] reported the characteristics of the MosquitoNet wireless network. The packet loss rates were 25.6% when packets were sent from a mobile host to a router, and 3.6% when packets are sent from a router to a mobile host. Risueño et al. [99] did a comprehensive study of the handover mechanisms during the disruption time in the wireless network. They reported that the packet loss caused by the handover mechanism was below 0.3%. Based on these published studies, we have set the packet loss rates of our experiments to 5%.

3.4.1 Fine-level FEC Protection Experiments

The proposed fine-level FEC protection framework is compared against a fixed-level FEC protection scheme for video streaming over a 4% packet loss channel. The R-D curves of the luma channel of the reconstructed video sequences are shown in Fig. 3-14 through 3-18. The level of protection for different segment of video data with the fine-level FEC scheme is computed using Eq. (3-2), while the level of protection for video data protected using the fixed-level FEC is determined by the (predicted) average number of packet losses per second. In either case, the maximal packet loss protection level can only recover up to 4% packet losses on average. It is important to point out that the overall number of bits used for FEC protection is the same for both the fine-level scheme and the fixed -level scheme. However, for fine-level protection, more protection bits are applied to more important data (based on Eq. (3-2)). Note that the peak signal-to-noise ratio (PSNR) of the reconstructed video does not increase with the bitrate for the fixed-level FEC protection mechanism. The reason is that if the small set of crucial subband data is corrupted, the PSNR will stay low even if more (less important) data is transmitted. As one can see from the figures, the fine-level FEC protection scheme works much better than the fixed-level protection scheme. The R-D curves of unprotected bitstreams are not shown in the figures because packet losses can severely corrupt an unprotected wavelet video bitstream. Take the STEFAN sequence for example, when the first few coding passes of coding block 0 of $P(LLLL_t, Y)$ are lost, the PSNR is usually less than 10 dB, no matter how high the bitrate is.

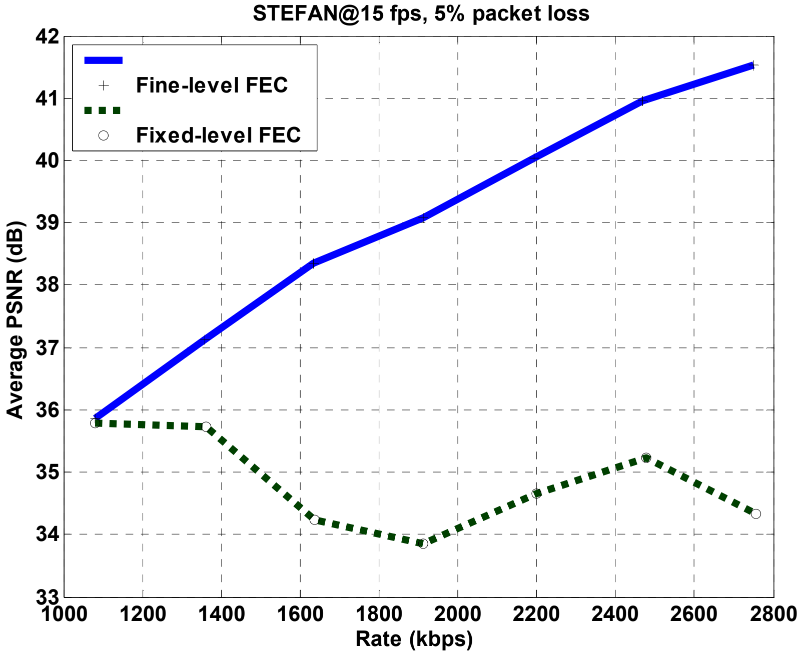


Figure 3-14: A comparison of the R-D curves for the fixed FEC and fine-level FEC protection (both protection levels are for 4% packet loss) using the STEFAN sequence.

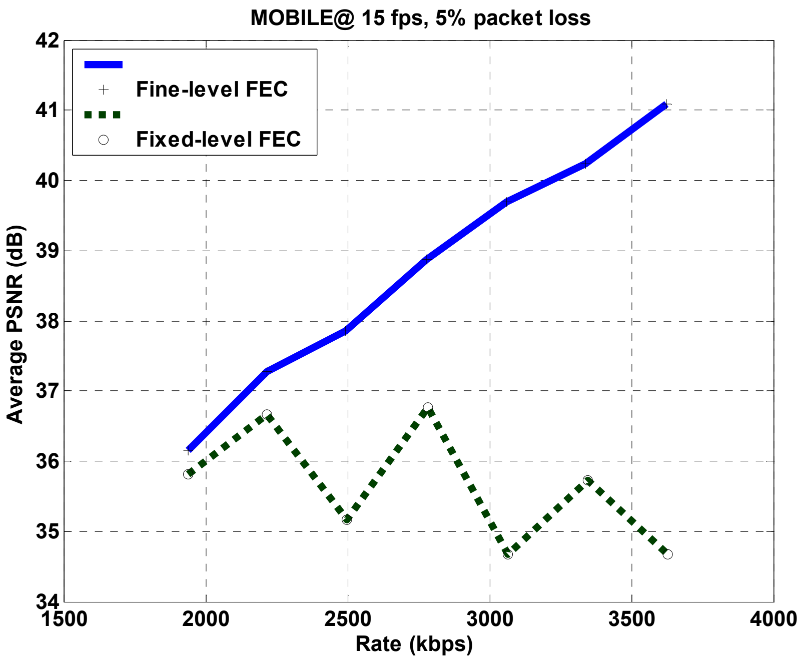


Figure 3-15: A comparison of the R-D curves for the fixed FEC and fine-level FEC protection (both protection levels are for 4% packet loss) using the MOBILE sequence.

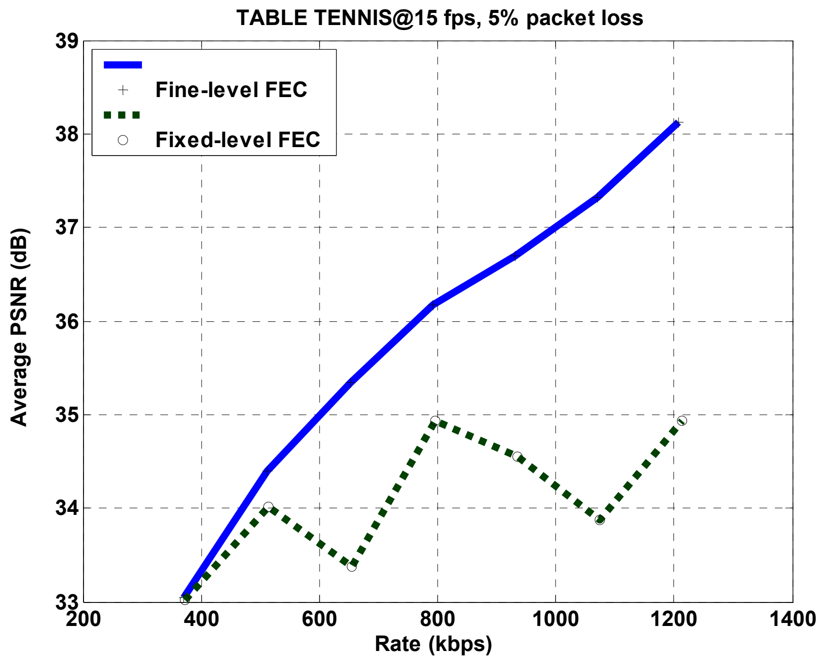


Figure 3-16: A comparison of the R-D curves for the fixed FEC and fine-level FEC protection (both protection levels are for 4% packet loss) using the TABLE TENNIS sequence.

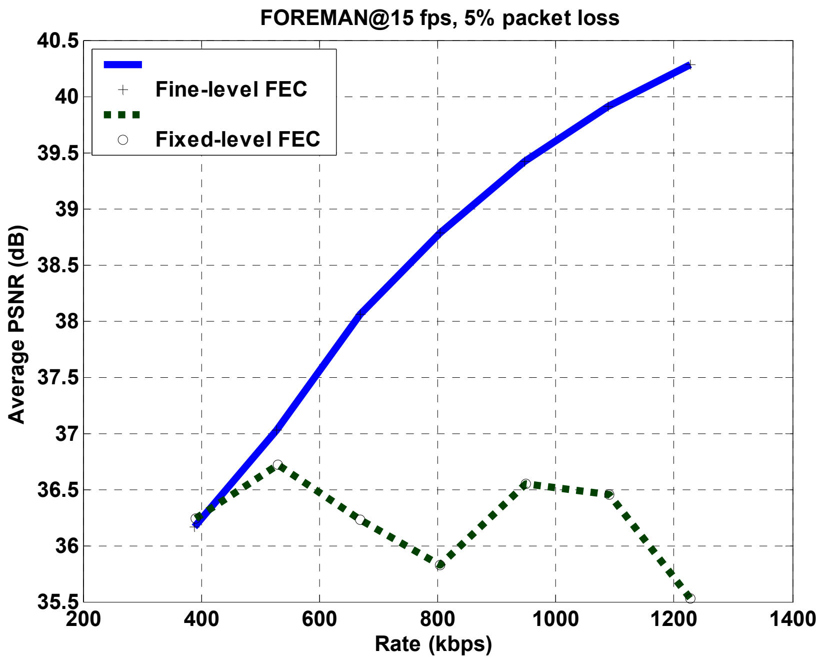


Figure 3-17: A comparison of the R-D curves for the fixed FEC and fine-level FEC protection (both protection levels are for 4% packet loss) using the FOREMAN sequence.

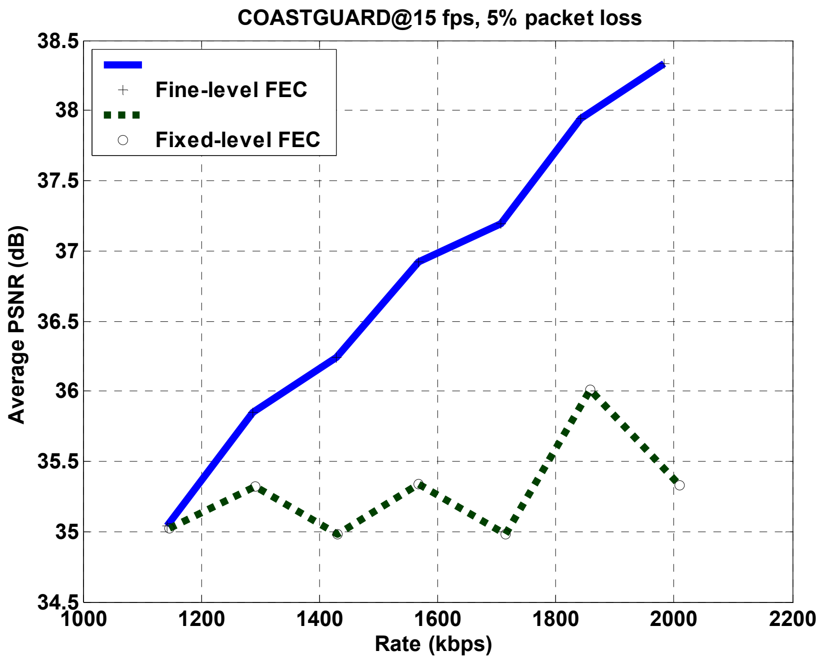


Figure 3-18: A comparison of the R-D curves for the fixed FEC and fine-level FEC protection (both protection levels are for 4% packet loss) using the COASTGUARD sequence.

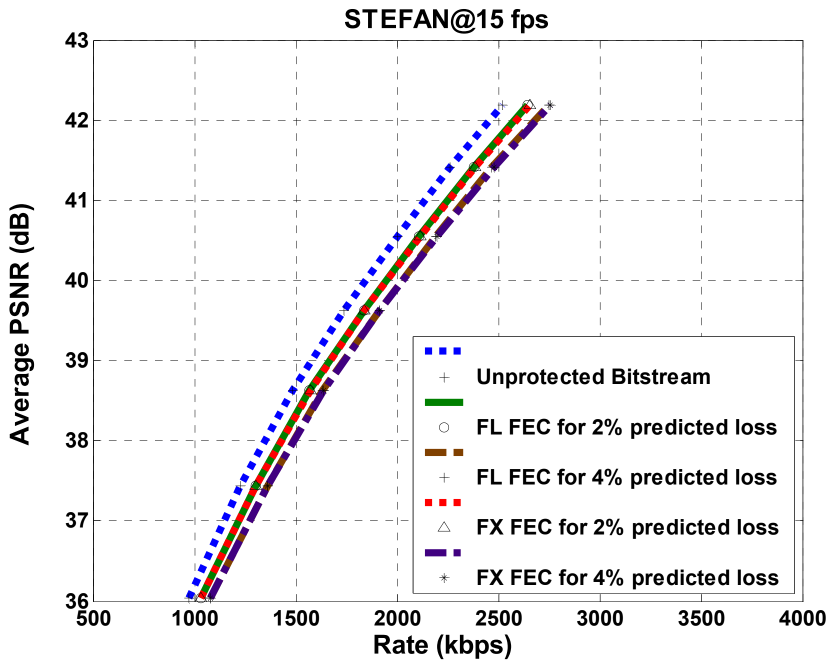


Figure 3-19: Comparison of R-D curves of STEFAN without and with different FEC protections in an error-free environment (FL: Fine-level, FX: Fixed-level).

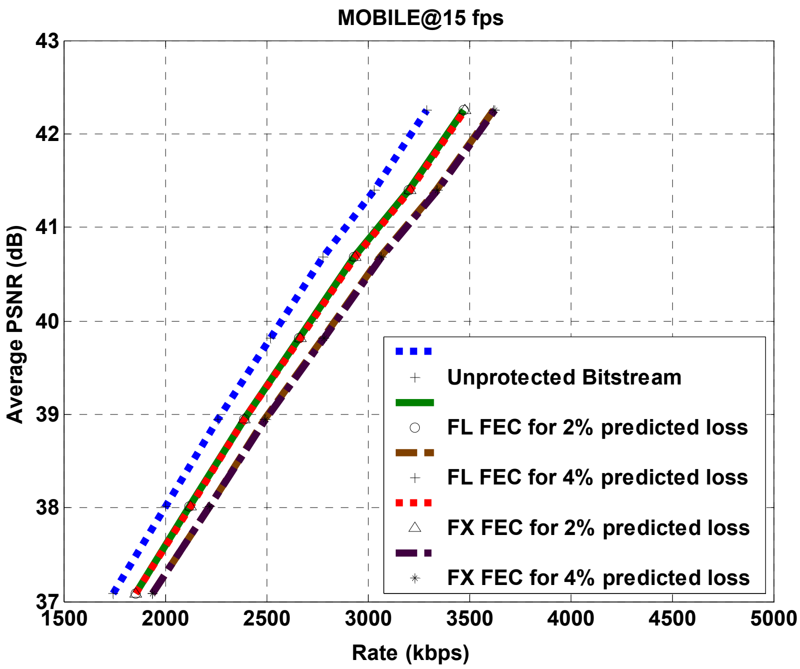


Figure 3-20: Comparison of R-D curves of MOBILE without and with different FEC protections in an error-free environment (FL: Fine-level, FX: Fixed-level).

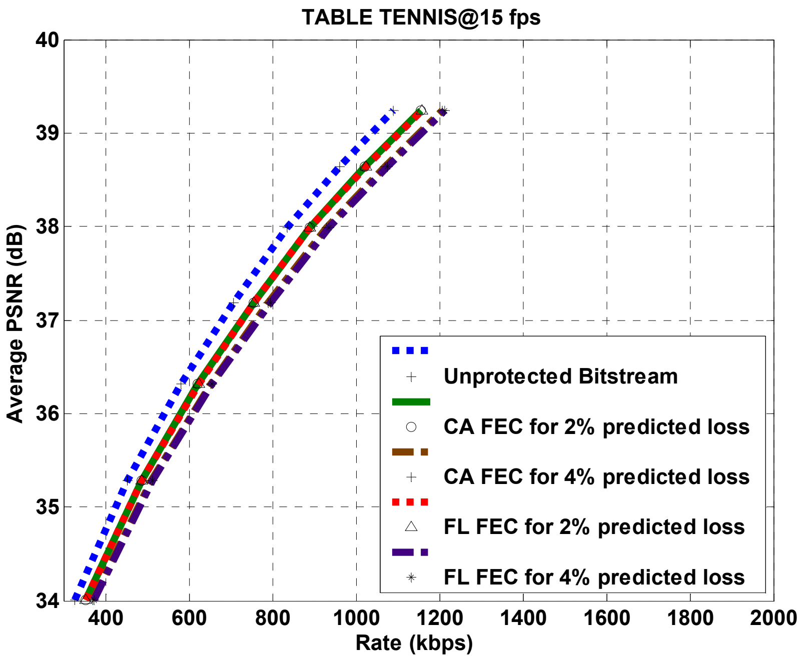


Figure 3-21: Comparison of R-D curves of TABLE TENNIS without and with different FEC protections in an error-free environment (FL: Fine-level, FX: Fixed-level).

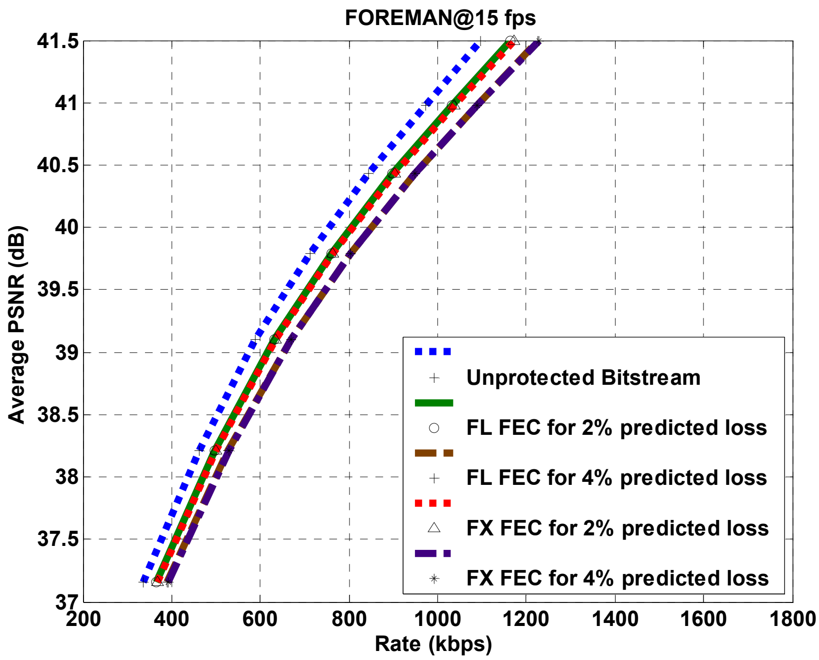


Figure 3-22: Comparison of R-D curves of FOREMAN without and with different FEC protections in an error-free environment (FL: Fine-level, FX: Fixed-level).

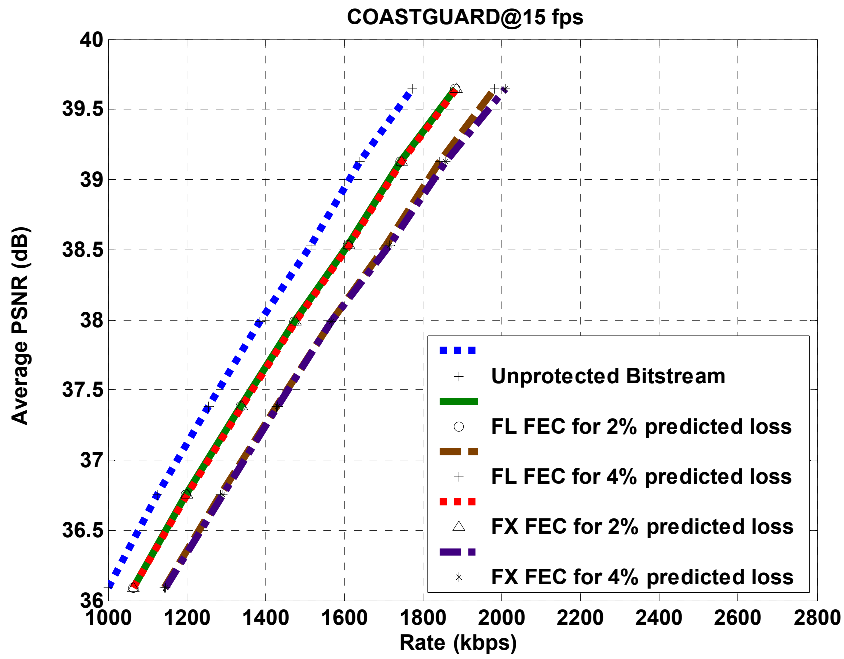


Figure 3-23: Comparison of R-D curves of COASTGUARD without and with different FEC protections in an error-free environment (FL: Fine-level, FX: Fixed-level).

To demonstrate the bitrate overhead of the fine-level FEC protection scheme, the error-free R-D curves of the video bitstreams with and without FEC protection are shown in

Fig. 3-19 through Fig. 3-23. For the bitstreams that are protected using FEC schemes, the level of protection is computed based on an assumption that the channel has an estimated packet loss rate of 2% and 4%. As one can see from these figures, the overhead of the proposed fine-level FEC protection is quite reasonable (about 0.2 to 0.5 dB quality drop across a wide range of bitrates for 2% packet loss protection).

3.4.2 Fine-level FEC Protection Experiments based on R-λ Curve Fitting Function

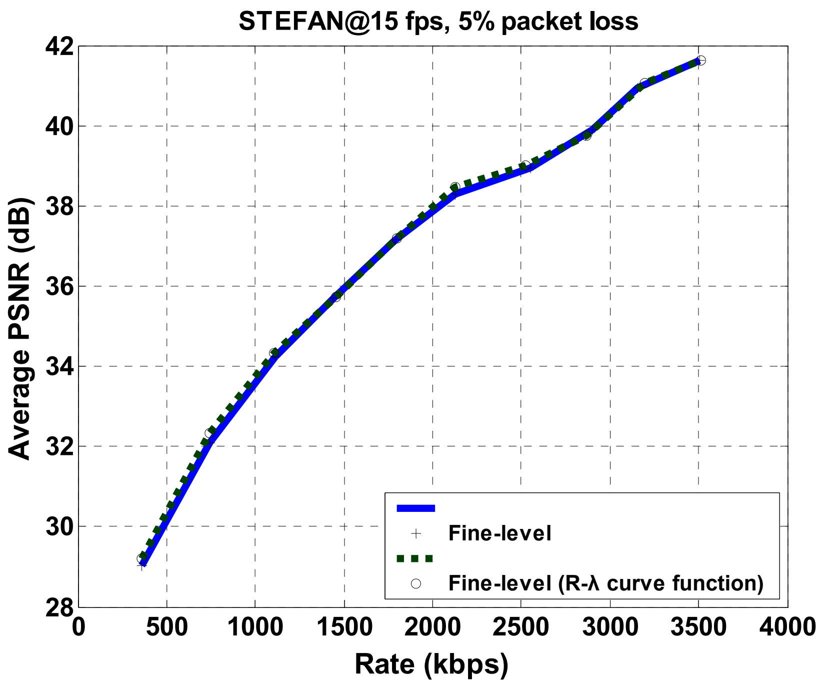


Figure 3-24: Fine-level FEC test for the STEFAN sequence

We have applied 5% packet loss rate to the IP packets in order to evaluate the performance of the proposed fine-level FEC protection system. Adaptive FEC protection using the proposed R-D information is compared against that using the original MSRA R-D information. The PSNR of the luma channel of the reconstructed video sequences are shown from Fig. 3-24 to Fig. 3-28. In either case, the maximal packet loss protection level can only recover up to 4% packet losses on average so that we can evaluate the differences in quality degradation using different R-D information. As one can see from the figures, the proposed R-D information using R-λ curve fitting function (based on Eq. (3-5)) is as efficient as the original R-D information using discrete R-D data points for fine-level FEC protection.

Figure 3-25: Fine-level FEC test for the MOBILE sequence

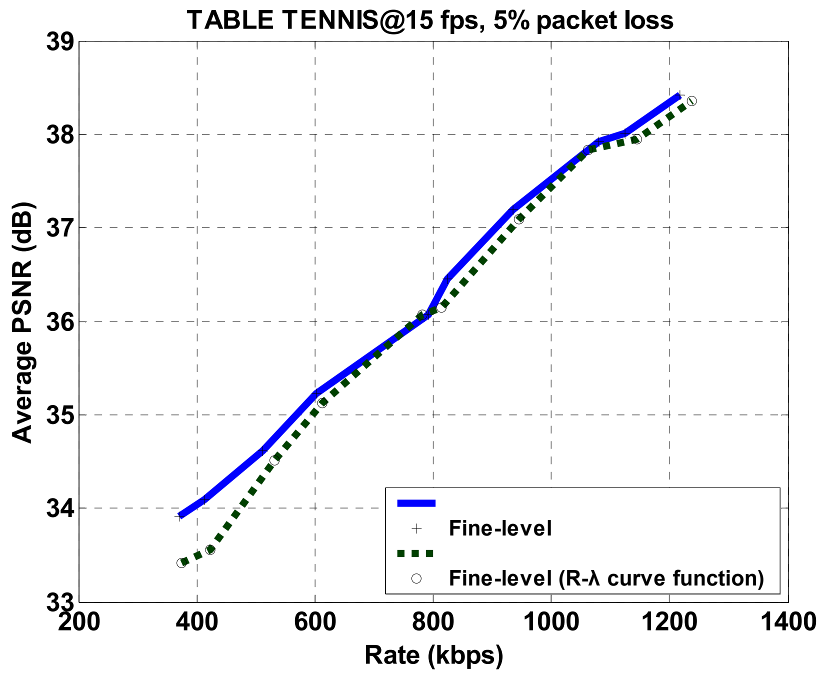


Figure 3-26: Fine-level FEC test for the TABLE TENNIS sequence

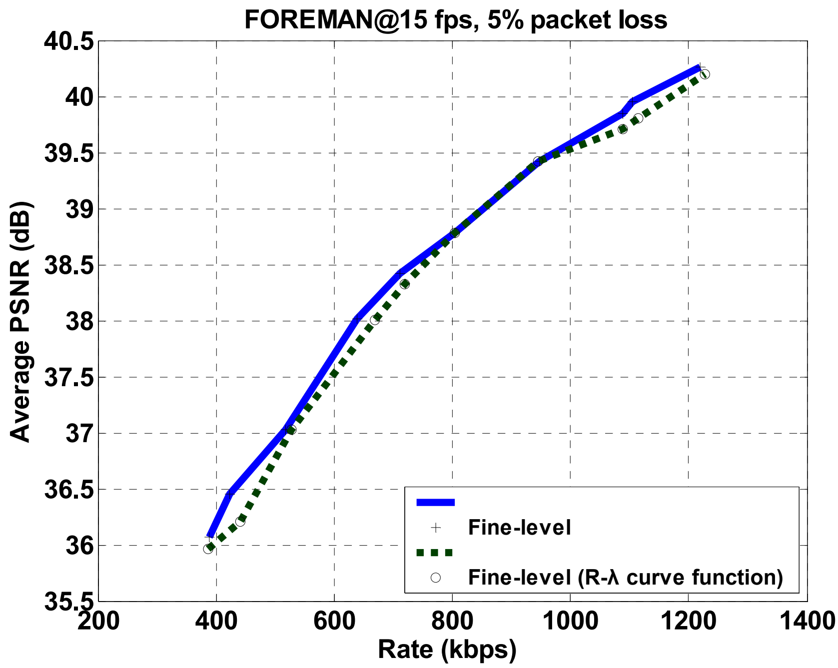


Figure 3-27: Fine-level FEC test for the FOREMAN sequence

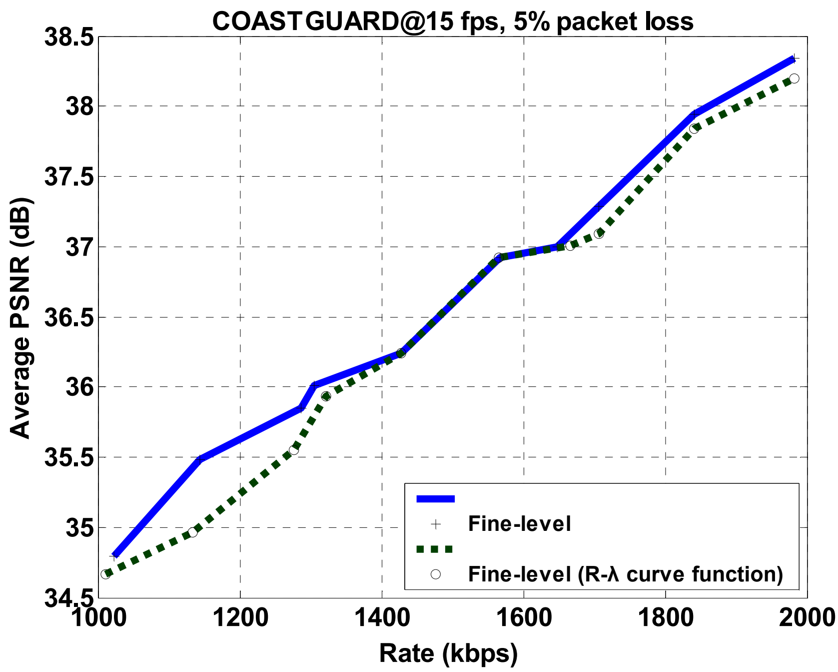


Figure 3-28: Fine-level FEC test for the COASTGUARD sequence

3.5 Summary

In this chapter, a fine-level FEC protection and packetization framework for wavelet video streaming is proposed. The adaptive packet loss protection scheme using Reed-Solomon coding and data-interleaving is based on detailed analysis of rate-distortion tradeoff of wavelet subband data. We have proposed a fine-granularity unequal error protection mechanism for wavelet-based video. The mechanism uses the original R-D slopes information to fine-tune the protection level of coefficients of different fractional bitplanes. The approach maximize the use of protection bit budget to achieves better performances than existing approaches of unequal error protection based on different syntax element types. The experimental results show that with an adaptive fine-granularity FEC protection level packetization scheme, one can achieve much better quality than a fixed-level FEC protection scheme.

The proposed multiple-adaptation framework enables multiple adaptations of fine-level FEC protection scheme for more flexible error resilient transmission of bitstreams. Experimental results show that a higher efficiency can be obtained by applying the proposed scheme on multiple-adaptation architectures. For future work, a runtime operational R-D optimized streaming policy with joint optimization for minimal source coding distortion and packet-loss distortion will be investigated. More rigorous derivation of the FEC protection level function is under investigation.

Chapter 4 Rate-Distortion Optimized Video Streaming with Smooth Quality Constraint

In this chapter, we propose an approach to integrate smooth quality constraint into an R-D optimized video streaming system over variable bandwidth networks. The constraint is based on the analysis of video content complexity. In order to preserve smooth video quality when the network bandwidth and video data rate vary, the scheme takes into account degree of motion in each frame in an R-D optimized framework to determine the packet scheduling policy. An implementation based on MPEG-4 (Moving Picture Experts Group) FGS [100] codec is presented in this chapter. Initial experiments show that the proposed approach can outperform the traditional scalable streaming systems under harsh network conditions.

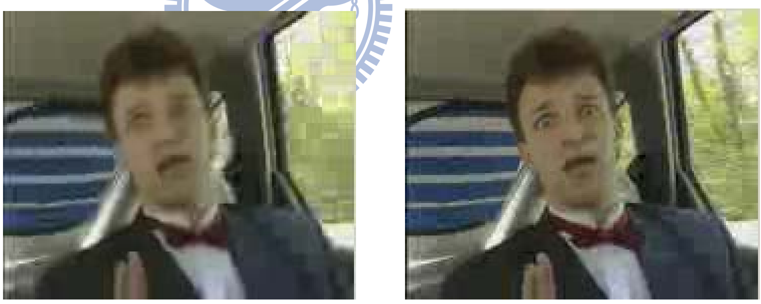
The chapter is organized as follows. Section 4.1 gives the introduction. In Section 4.2, an overview of the proposed system architecture is described. Section 4.3 presents the details of the proposed scheme with R-D optimized streaming framework. Section 4.4 shows some experimental results and compares the proposed system with an FGS streaming system [101]. Section 4.5 gives some conclusions and discusses future research directions.

4.1 Introduction

When delivering video over an unreliable IP network, the importance of each packet varies according to the information it carries. There is also a dependency relationship among packets. Furthermore, if the video content is coded using scalable techniques, a subset of the packets can be transmitted alone to allow smooth (yet lower quality) presentation without re-buffering when the network bandwidth is low. A streaming system should take into account these factors and determine a transmission policy on-the-fly to achieve the best quality under variable network conditions.

A RDO streaming framework was proposed by Chou et al. [66] which is based on the importance and error probabilities of data units to compute transmission policies. The policy indicates whether the video packet should be transmitted at each transmission opportunity. Furthermore, the mechanism maps probability of packet losses into rate increment of redundant packet transmission. Each frame can be represented as a group of data units. Computationally, the system optimizes time and bandwidth resources allocation among a group of data units in a way that minimizes a Lagrangian cost function of expected rate and distortion. Other researchers have proposed techniques to enhance the RDO framework [76]. However, the scheme does not address the issue of reducing video quality variation over lossy channels.

One drawback with most scalable streaming systems is that quality variation across time domain is not explicitly modeled in the R-D optimization framework. In this case, when the network bandwidth varies drastically, video quality can change frequently such that the presentation becomes visually annoying, as shown in Fig. 4-1 (a).



(a) without smooth quality control (b) the proposed method

Figure 4-1: An example of video quality change when the network bandwidth varies drastically (Carphone sequence).

In this chapter, an R-D optimized streaming framework with explicit modeling of temporal quality-smoothness factor is proposed. The proposed system is based on the R-D optimized control model of packet transmission with extra consideration on the degree of motion in each video segment. The system achieves smooth video presentation through two schemes. First, it will compute target bit-allocation for each frame based on the degree of motion. Secondly, it uses an R-D optimized control mechanism to determine the scheduling policy.

4.2 System Overview

In this chapter, a framework of R-D optimized video streaming with smooth quality constraint is proposed. The main features of the proposed system are highlighted as follows:

- The streaming algorithm searches along the R-D curve for an optimal operating point between rate and distortion via (pre-calculated) redundant transmissions of important data units. The impact of packet losses is absorbed by the level of redundant transmissions.
- Both video playback smoothness and visual quality consistency constraints are elegantly incorporated into the R-D optimization framework.

The proposed mechanism is scalable for different bandwidths and resolutions as needed by the network participants. The detailed design is presented from system architecture to source and channel model as described below.

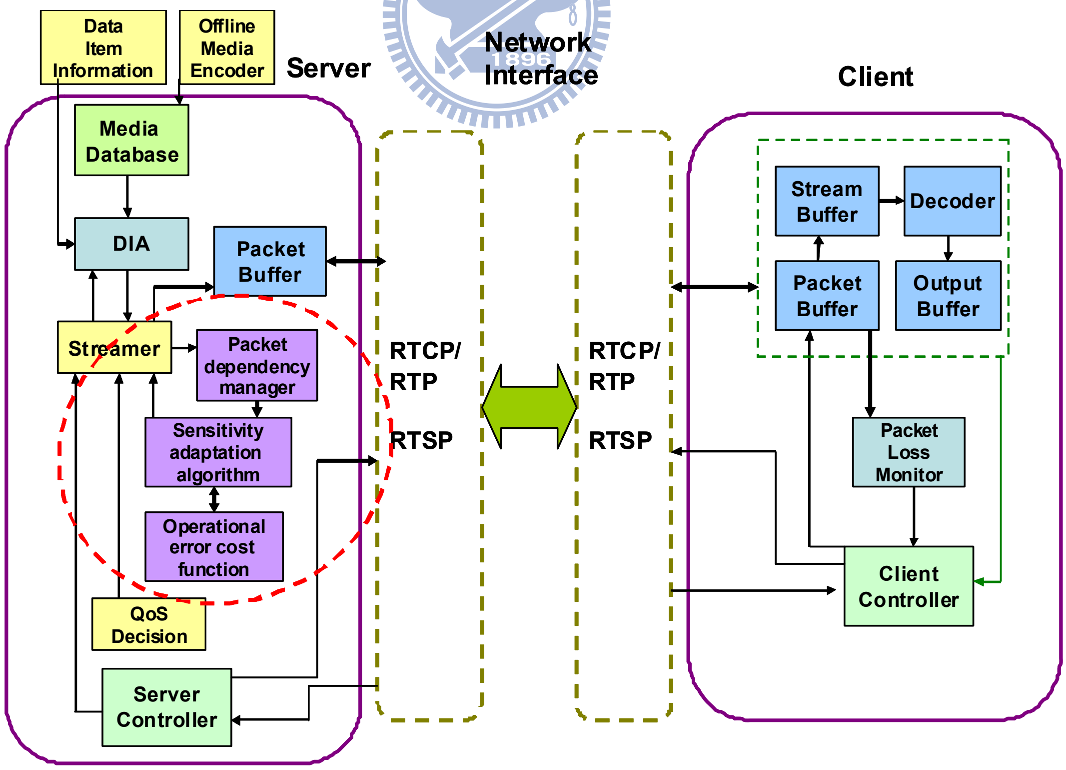


Figure 4-2: System architecture of the proposed streaming system.

4.2.1 System Architecture

A block diagram of the proposed streaming system with the R-D model is shown in Fig. 4-2. The system is based on the MPEG-21 Test Bed for Resource Delivery [94]. The test bed is an FGS-based scalable streaming system with a network emulator that allows real-time simulation of various network conditions. The main differences between the test bed and the proposed system are the following three functional blocks in Fig. 4-2: "Packet Dependency Manager," "Sensitivity Adaptation Algorithm," and "Operational Error Cost Function." The details of these modules will be described in Section 4.3.

4.2.2 Source Model

Packetized multimedia data can be modeled as a DAG [76]. In the DAG, the nodes represent the data units and the arcs represent the dependency relations between data units. Fig. 4-3 illustrates the coding dependencies of the base layer and the enhancement layer of FGS video. In FGS, the enhancement layer is coded using bit-plane coding of the DCT coefficients of the residual images. Each bit-plane is packetized as one data unit.

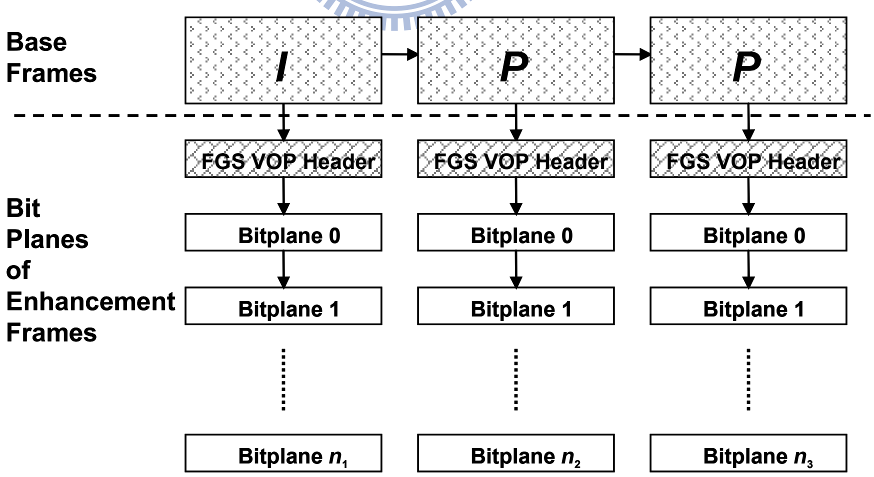


Figure 4-3: Directed acyclic dependence graphs of FGS video.

In a group of frames, the base layer frames are given the highest priority. Each data unit, DU_i, contains the following parameters: playout deadline $T_{d,i}$, frame type F_i, bit-plane coding level BL_i, bit-plane size S_i, and motion degree factor M_i. The playout deadline $T_{d,i}$ is the latest time at which DU_i should arrive at the client. Otherwise, DU_i is too late for decoding and presentation. Sum of absolute differences (SAD) is used to compute M_i and determine the variation of relative motion.

4.2.3 Channel Model

The base-layer bitstream of a scalable sequence contains the most essential video content [102]. When the base-layer bitstream cannot be recovered from transmission errors, the corresponding enhancement layer is useless even if they arrive at the client side successfully [103]. Consequently, the base-layer should be strongly protected against packet loss or corruption [104]. Furthermore, a base layer bitstream free from packet losses (through retransmission protection) makes R-D calculation deterministic and therefore an R-D optimal streaming system can be practically constructed.

The data transport model is shown in Fig. 4-4. An RTP session is created for base-layer packets transmission and the enhancement-layer packets is transmitted in a separate RTP session. Base-layer frames are protected with reliable retransmission mechanism. For enhancement layer, video header and the most significant bit-planes are transmitted with higher priority and are protected by retransmission as well. The channel for each RTP session is assumed to be injected with time-invariant packet losses and delay. Furthermore, the delay is modeled by a shifted Gamma distribution.

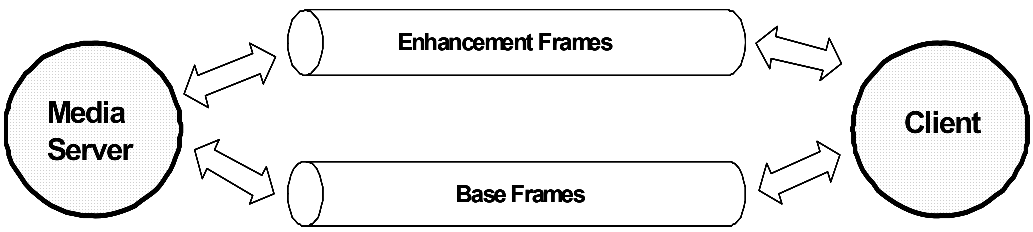


Figure 4-4: Channel model of the proposed scheme.

4.3 Integrated R-D Optimized Policy Selection

In this extension to the R-D optimized streaming framework, the scheme tries to maintain a smooth presentation quality when the network condition varies. In conventional scalable streaming systems, the number of enhancement-layer data units transmitted is constrained by network bandwidth. However, if network bandwidth is the only constraint, a content-rich video frame delivered during low bandwidth period will causes significant drop in video quality since most of the enhancement-layer data units will be dropped. In another words, to ensure smooth quality, the number of enhancement-layer data units transmitted should be based on the channel bandwidth as well as the richness of the content. For a motion-compensation based video codec, the amount of content in a video frame can be estimated using the magnitude of error-residual images or, equivalently, the sum-of-the-absolute-differences (SADs). The proposed scheme will schedule more data units of the enhancement frames when the SAD is larger.

4.3.1 Integration of Quality Smoothness Constraint into the R-D Framework

In the original R-D streaming framework, the tradeoff between rate and distortion is done via (pre-calculated) redundant transmissions of important data units. The argument was that redundant transmission of important packets can reduce distortion over lossy channels at the cost of higher bandwidth (rate) requirement. The importance of a data unit will be described later in Section 4.3.2. In the proposed framework, the number of enhancement-layer data units that will be considered in the R-D framework as a candidate for delivery is determined by the degree of motion of the corresponding frame.

For each frame f_i, i = 1, …, N in a GOP, we need to find an appropriate number of bit-planes that should meet the bit budget. First, the bit budget for each frame f_i is defined as Eq. (4-1):

$$\eta(i) = \left(BW - \sum_{j=1}^{i-1} BS_j \right) \cdot \frac{BS_i}{\sum_{k=i}^{N} BS_k}, \tag{4-1}$$

where BW is the available bandwidth of the transmitting period for the GOP for adapting the varying network conditions, and BS_i is the size of the enhancement layer of frame f_i, therefore $\sum_{j=1}^{i-1} BS_j$ computes the consumed bandwidth before frame i. Hence the number of enhancement layer bit-planes of frame i, L_i, to be transmitted can be determined based on the bit budget for frame $\eta(i)$.

A piecewise-linear model and a piecewise-exponential model can be used for modeling the R-D curve of enhancement-layer bitstream [105]. And the quality variation can be reduced by using the models in finding an optimal constant quality constrained rate allocation from two neighboring sampling points [106]. However, to propose an optimal R-D solution, the increase of computation time is a drawback for real-time applications.

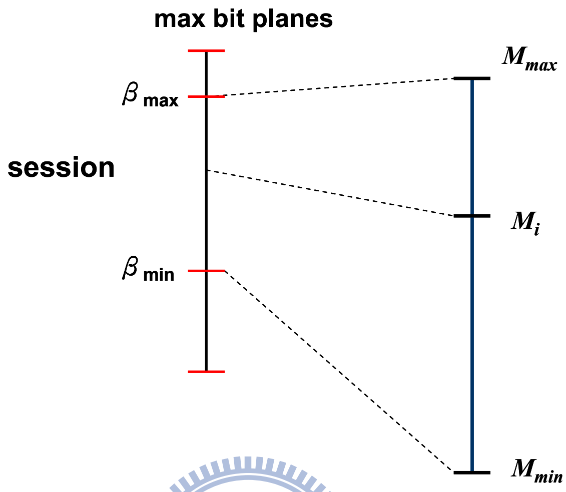


Figure 4-5: An illustrating example to determine how many layers of each frame need to be transmitted.

To reduce the computational cost of the rate allocation, each bit plane is treated as a transmission unit in the enhancement layer. For providing smooth quality across the GOP, we utilize the degree of motion of the corresponding frame to adjust the number of transmitted bit-planes L_i to reduce quality variation, as shown in Fig. 4-5. As mentioned before, each data unit in the proposed FGS-based streaming framework contains a single bit-plane of an enhancement layer frame. The number of enhancement-layer data units for frame i can be adjusted as Eq. (4-2):

$$\xi(i) = \left\lfloor \frac{\beta_{\max} - \beta_{\min}}{M_{\max} - M_{\min} + 1} \right\rfloor \cdot (M_i - M_{\min}) + \beta_{\min}, \qquad (4\text{-}2)$$

where β_{max} is $x = \max(L_i)$, $i = 1, \ldots N$., β_{min} is $x = \max(L_i)$, $i = 1, \ldots N$, and M_i is the motion degree factor for frame f_i.

The motion degree factor for I-frame is 0. The M_i factor for P-frame can be computed as Eq. (4-3):

$$M_i = \sum_{0}^{98} \left(\sum_{p=0}^{15} \sum_{q=0}^{15} \left| g(x+p, y+q) - h((x+r)+p, (y+s)+q) \right| \right), \tag{4-3}$$

where $g(x,y)$ is the current frame pixel at (x,y), $h(x,y)$ is the reference frame pixel at (x,y), and (r,s) is the motion vector. Note that the maximal and minimal values of motion degree are M_{max} and M_{min}, respectively.

4.3.2 Importance Function

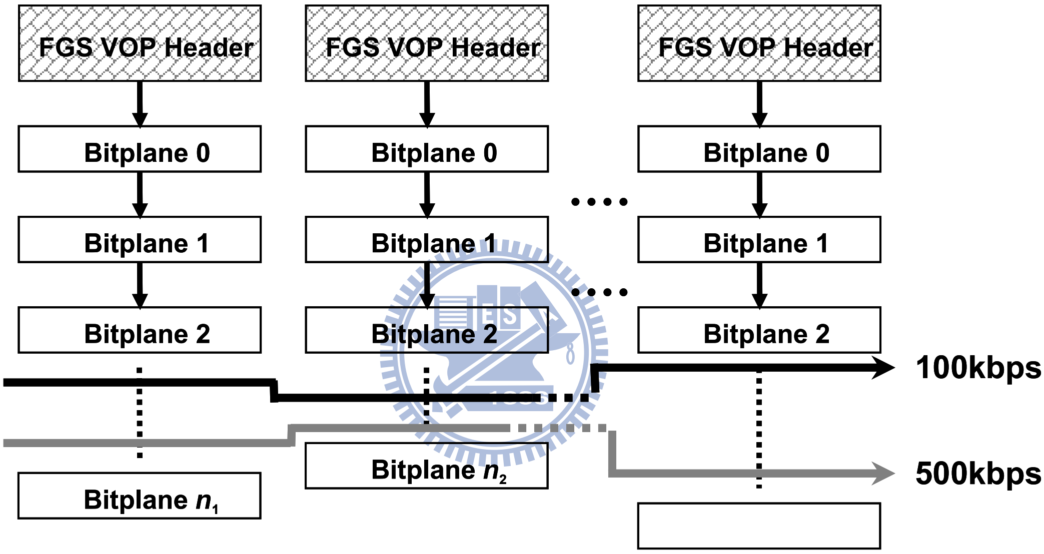


Figure 4-6: An example with a truncated group of data units.

The importance function of a data unit of frame i is composed of the frame type F_i, the bit-plane layer BL_i of the data unit, and the frame partial order FP_i. In other words, the importance of a data unit is the amount of distortion reduction if the data unit is decoded at the client. The importance function ΔD_i can be computed as Eq. (4-4):

$$\Delta D_i = F_i \cdot BL_i + FP_i, \tag{4-4}$$

Note that the frame type F_i is defined so that more important frame type of frames, such as Intra frames, has higher number (in our implementation, I-frame is 3 and P-frame is 1). BL_i ranges from 10 to 0. And finally, FP_i is defined as the number of data units that depends on data unit i. With these assumptions, the status of the optimal frame-level bit allocation for a

GOP can be determined within a fixed time, as shown in Fig. 4-6.

4.3.3 Transmitting a GOP of Data Units

Similar to the mechanism for transmitting a group of data units used in [66], we define different optimization windows for the data units of the enhancement layer. Let the transmission policy be π, the data units will be transmitted at discrete transmission time. The algorithm minimizes the expected Lagrangian for finding the optimal policy π at the transmission time, as defined in Eq. (4-5):

$$J(\pi) = D(\pi) + \lambda R(\pi), \tag{4-5}$$

where $D(\pi)$ is the expected distortion for the data units of the group and $R(\pi)$ is the expected transmission cost. And the parameter λ is a Lagrangian multiplier. The rate constraint and the expected distortion for a group of K data units can be computed by Eq. (4-6) and Eq. (4-7):

$$R(\pi) = \sum_{j=0}^{K-1} \rho_j(\pi) S_j, \quad \text{and} \tag{4-6}$$

$$D(\pi) = D_0 - \sum_{j=0}^{K-1} \left(\Delta D_j \prod_{j'} \left(1 - \varepsilon_{j'}\left(\pi_{j'}\right)\right) \right), \tag{4-7}$$

where $\rho_j(\pi)$ is the expected number of transmissions for data unit j, S_j is the data unit size, D_0 is the distortion if the whole group of data units of the enhancement layer are lost during transmission, and $\varepsilon_{j'}(\pi_{j'})$ is the error probability that the data unit j is not delivered to the client before its deadline. Note that the index j' refers to all the data units that has dependency relationship with data unit j.

4.4 Experimental Results

This section presents the experimental results of the R-D optimized video streaming scheme. The R-D optimized transmission policy will be computed at every transmission opportunity. Each packet will be marked as “transmit” or “not transmit” after the computation. Since λ is fixed in our current implementation, the dropping of data units is only controlled by the quality smoothness constraint. We are working on a derivation of a joint constraint function that express λ as a function of motion degree and available bandwidth. In the experiments, the QCIF version of the FOREMAN sequence is used. The sequence is encoded using MPEG-4 FGS reference software at 10 frames per second.

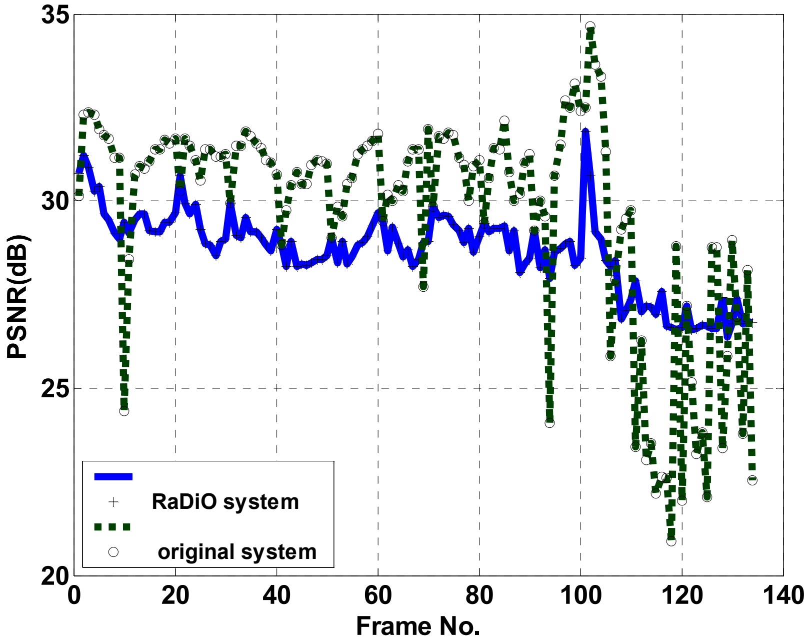


Figure 4-7: PSNR versus each frame at bandwidth of 300kbps.

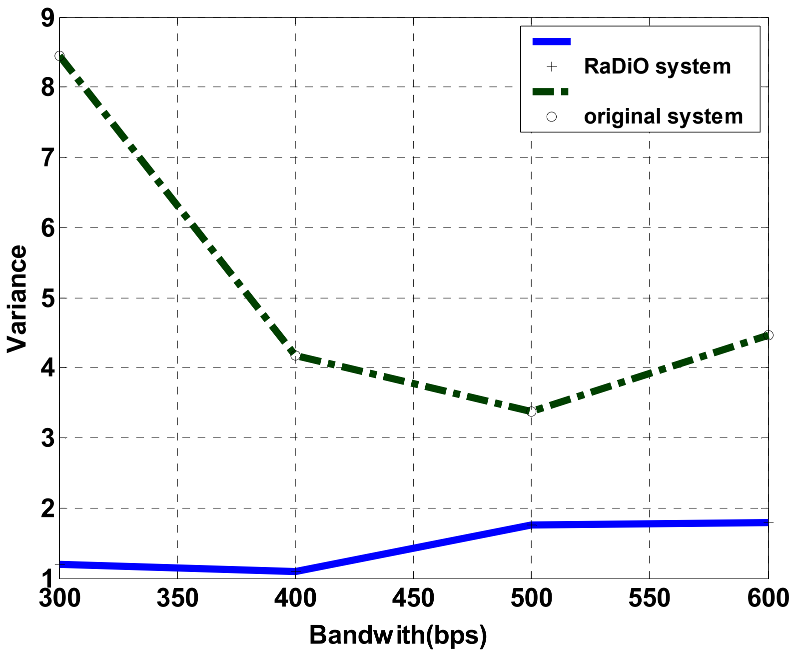


Figure 4-8: Variance for different rate-constraint.

The first experiment simulates video streaming under different constant-rate channel conditions. The PSNR corresponding to each frame at a network bandwidth of 300 kbps as shown in Fig. 4-7. Figure 4-8 shows the PSNR variances corresponding to different bandwidth conditions. The proposed scheme has a smaller variance than the original ARQ scheme of the test bed. Figure 4-9 shows the frame size variation during a streaming session. In FGS streaming, the frame size is directly proportional to the quality of the video. Any significant increase in the motion degree factor would require a significant increase in frame size. Obviously, the proposed system has a better capability to handle the variation than the original system. Furthermore, we observe that the proposed scheme consistently achieves smoother quality compared to the ARQ scheme of the Test Bed.

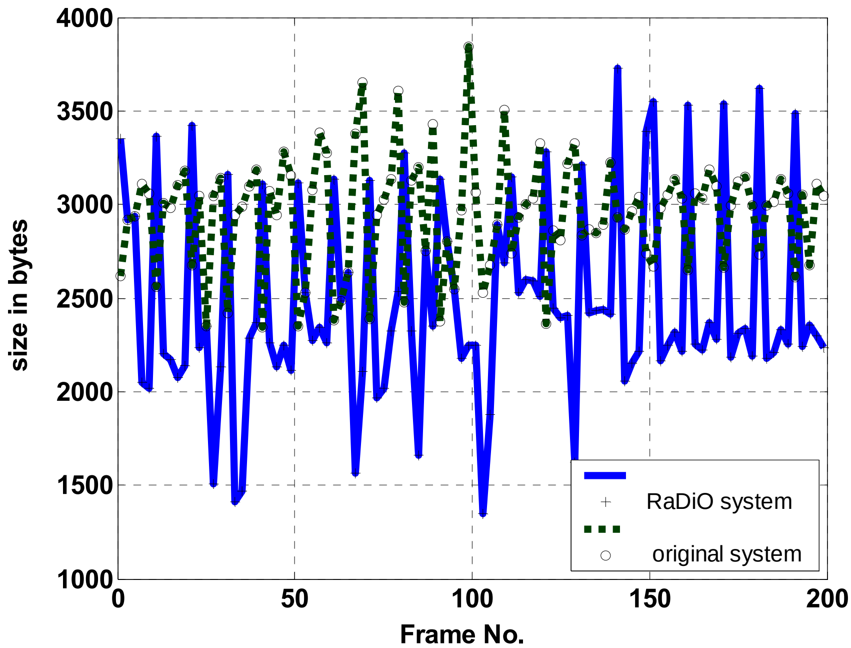


Figure 4-9: Delivered frame size vs. frame number during streaming

Figure 4-10 summarizes the results of streaming under variable bandwidth conditions, ranging from 128kbps to 320kbps. Figure 4-11 shows the frame size variation during a streaming session. The proposed scheme not only has smoother quality but also can adapt to the variable bandwidth better. Overall, from the results of the above two experiments, we observe that the PSNR variance of our proposed scheme is smaller than that of the original ARQ scheme for over half of the R-D data points. This shows that the new scheme is effective in achieving smooth video quality.

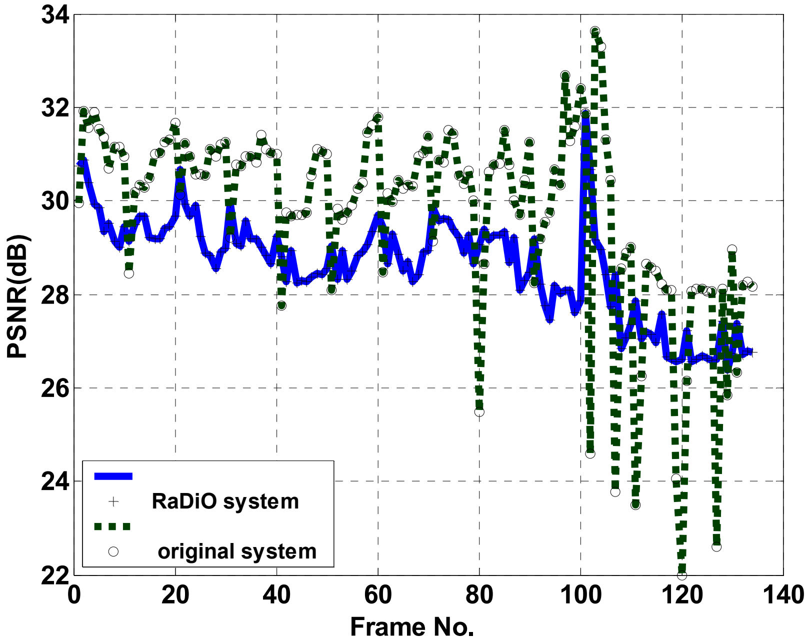


Figure 4-10: The system performance of the variable transmission rate.

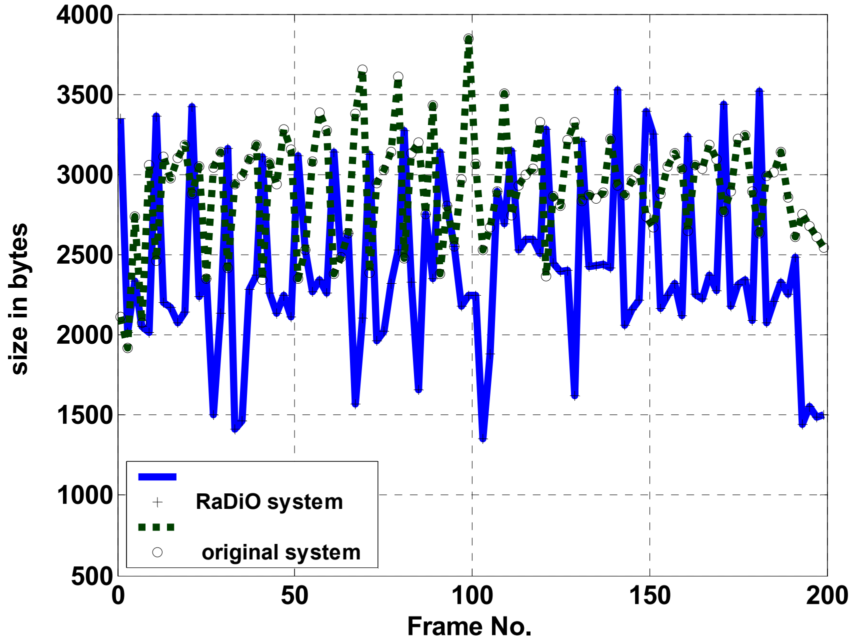


Figure 4-11: The frame size variation during a streaming session.

4.5 Summary

In this chapter, we have proposed an algorithm that integrates the smooth video quality constraint into an R-D optimized scalable streaming system over varying channels. The scheme will flexibly adapt different transmitting strategies when high degree of complex motion is detected. Experiments show that the proposed solution can achieve smoother video quality under variable network conditions, and the system is very promising for practical applications. For future work, efforts will be put on real-time statistics feedback for channel model update in the framework.

Chapter 5 Efficient Data Replication for the Delivery of High-quality Video Content over P2P VoD Advertising Networks

Recent advances in modern television systems have had profound consequences for the scalability, stability, and quality of transmitted digital data signals. This is of particular significance for P2P VoD related platforms, faced with an immediate and growing demand for reliable service delivery. In response to demands for high-quality video, the key objectives in the construction of the proposed framework were user satisfaction with perceived video quality and the effective utilization of available resources on P2P VoD networks. In this chapter, we develop a peer-based promoter to support online advertising in P2P VoD networks based on an estimation of video distortion prior to the replication of data stream chunks. The proposed technology enables the recovery of lost video using replicated stream chunks in real time. Load balance is achieved by adjusting the replication level of each candidate group according to the degree of distortion, thereby enabling a significant reduction in server load and increased scalability in the P2P VoD system. This approach also promotes the use of advertising as an efficient tool for commercial promotion. Results indicate that the proposed system efficiently satisfies the given fault tolerances.

The organization of the chapter is as follows. Section 5.1 gives the introduction. Section 5.2 introduces the problems associated with P2P VoD advertising services. Section 5.3 discusses the operational attributes of the proposed P2P VoD advertising application presented in Section 5.4. The simulation results are shown in Section 5.5. Finally, summarizes this chapter in Section 5.6.

5.1 Introduction

Recent advances in online advertising and P2P VoD networks, enabling peers to watch or download internet video clips on demand, have created considerable interest in the construction of integrated frameworks. Online advertising channels [107], such as online newspapers/magazines, keyword trigger tools, and e-mail, have gained wide public acceptance and considerable importance as advertising media [108]. However, an increasing number of internet content providers, such as Blinkx BBTV [109], Joost [110], and LiveStation [111] are incorporating legal P2P technologies into their delivery platform to reduce operational expenses. In P2P VoD applications, user preferences can be automatically derived from media usage data without the need for direct user input, making them an excellent system for the collection of customer information. This enables advertisers to bid on video clips relevant to their target market. For instance, a toy or a snack advertisement might link to cartoon videos. Hence, a concomitant need has arisen for the delivery of marketing messages to attract customers to the P2P VoD environments [112]. A P2P VoD computing environment can be an ideal platform on which to display advertisements. Perceptions of high video quality and a robust environment are essential for the delivery of online advertising in P2P VoD networks; therefore, this chapter attempts to make a system tolerant of network errors in terms of video-quality enhancement and online P2P advertising availability.

There are many ways to enable efficient and scalable on-demand video distribution over networks, including IP multicast, content delivery network (CDN), and P2P networking. Although IP multicast is an efficient approach for a number of channels with high popularity rankings [113], it has several drawbacks. First, IP multicast has not been widely deployed on the internet [114]. Second, core network routers must process a considerable number of forwarding entries when many active multicast groups are used, resulting in increased memory requirements and slower forwarding processing. Third, IP multicast flow aggregation is not well suited to less popular video channels [113]. In contrast, P2P VoD technologies have gained immense popularity throughout the world

[115-118]. P2P video on-demand services are fundamentally more scalable than existing IP multicast methods when bandwidth availability exists at the ISP backbone. The advantages of using P2P VoD technologies for content distribution over CDN or IP multicast are listed below:

- **Exploitation of the underutilized resources of peers:** Some resource owners (resourceful peers) can become providers by making their underutilized resources available. Peers can be frequently switched or reconnected to resourceful peers, and shared data and services are accessible to other peers.
- **System deployability:** A set of incrementally deployable and extensible solutions bring existing P2P systems closer to commercial production, and many have been introduced in recent years.
- **Hardware economics:** Traditional CDN combine the infrastructure for content-delivery, request-routing, distribution, and accounting to provide an intermediate layer of infrastructure to rapidly deliver content from providers to end users. The disadvantages include the need for a large number of CDN servers (Content Foundry), high costs, and a lack of scalability to accommodate a large audience. Infrastructure management is expensive, and according to Jupiter Research, one-hour video-clips streamed to an audience of 1,000 costs content providers 0.5 cents per megabyte [119].
- **High scalability of P2P services:** The high scalability of P2P systems relies on an aggregate of resource contributions by individual peers. Peers do not need a global view of the overall system, which makes publishing, sending, or downloading shared media easy, quick, and scalable.

In existing P2P advertising systems, delivery services are accomplished through a combination of P2P file sharing and an advertising service, such as ZapShares [120], MediaDefender [121], P2Pads [122], or P2Pwords [123]. A P2P web search engine [124] (e.g. Mininova [125]) is defined as a P2P retrieval service, providing the Uniform Resource Locators of multiple trackers and integrity metadata in answer to a search request by a peer. As shown in Fig. 5-1, when a keyword or comprehensive query is submitted to the P2P web portal server, the results are revealed to enable the selection of content by users. The search results, including commercial-advertisement files, can be downloaded from other peers and shown to participants. The commercial-advertisement video may be an interactive commercial video that is viewed by the target audience, enabling them to interact and make immediate purchasing decisions. P2P advertising platforms enable advertisers to efficiently track through-clicks and historical data. Compared to non-P2P online advertising (e.g. contextual ads on search engine, banners, advertising on social network, and e-mail advertising), P2P networks are more socially-aware and service-oriented because they are self-organizing and decentralized communication [126]. P2P advertising enables the better utilization of all peer resources and the more effective promotion of advertisements.

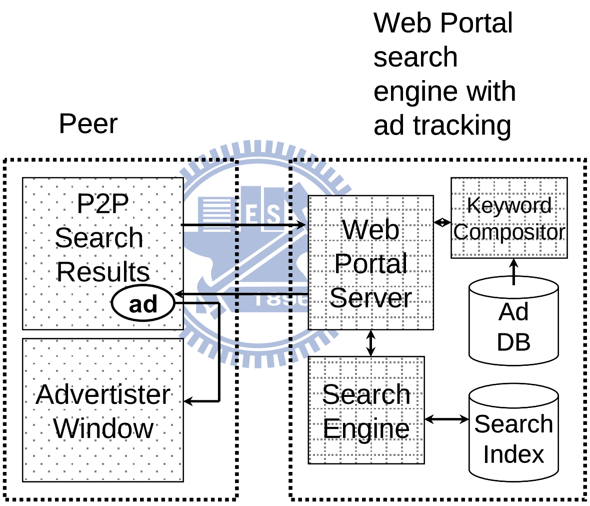


Figure 5-1: Existing advertising mechanisms

The divergent behavior of peers influences the availability of resources in a P2P network. Therefore, it is essential for system designers to determine an appropriate policy for sharing resources to deal with video distortion resulting from packet loss. The aim of this study was to develop a framework in which to integrate advertisements and manage the sharing of resources, according to content and network conditions via video-distortion estimation in P2P VoD networks. P2P systems are commonly classified into three classes: unstructured, structured, and hybrid [127]. The empirical goals of this chapter are twofold: The first is to achieve a high degree of perceived video quality and effectively utilize the resources available on P2P VoD services. Video quality can be improved through the management of replication operations in video sessions involving the estimation of video distortion. The second is to develop a peer-based promoter for delivering online advertising in P2P VoD networks. Online advertising strategies for the proposed P2P VoD framework must consider the display function as well as the stability, efficiency, and robustness required for continuous operations in online marketing communication channels [128]. The proposed framework enables the distribution of advertisements via peer-sharing to make them publicly available in a way that is rarely possible in other media (e.g. banner advertisements on web sites). The proposed framework differs from existing P2P advertising systems in two respects. First, the video title and video description fields may provide useful information about the self-interest of individual peers, making them a useful tool for the promotion of advertisements. Second, advertising can be obtained not only in the initial stages of searching, but also in the subsequent video sessions. In summary, the main contributions of the proposed system are as follows:

(1) This method decomposes the network into separate sub-networks to enable the dynamic replication of data to enhance performance without global knowledge of all peers in the overlap network.

(2) The network conditions of individual peers are integrated with the platform to maximize visual quality in the P2P VoD through the replication of video chunks to maximize the video quality.

(3) The distribution of ads relies on peers and a centralized collection point (a web portal server), making more scalability and flexibility of the P2P ad service.

(4) The proposed framework is evaluated through the simulation of the proposed distortion-based video-chunk replication solution to reduce server load and significantly increase the scalability of P2P VoD systems.

This chapter proposes a P2P VoD advertising framework based on the estimation of video distortion prior to the of replication data stream-chunks, as shown in Fig 5-2. The major achievement of the proposed framework is the reduction of server load and the optimization of overall video quality under given network conditions. We also propose an online video advertising mechanism based on on-demand videos.

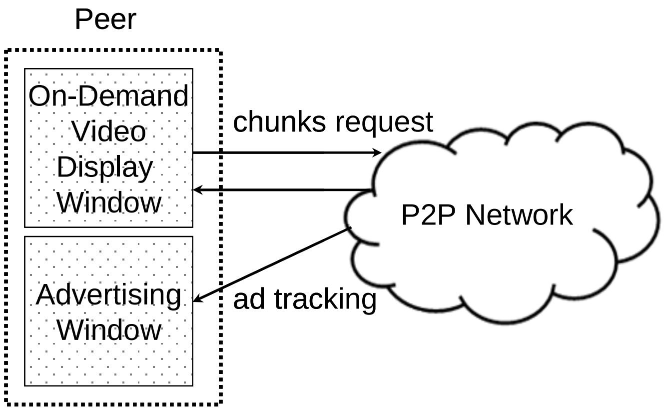


Figure 5-2: P2P VoD advertising scheme

5.2 Previous Studies

Several studies have lent support to the claim that P2P advertising services can effectively facilitate the spread of advertisements and promotions. Research on the effect of P2P network on advertising services is still in its infancy, and even less has been conducted on the effect of P2P advertising services integrated with VoD systems.

Ad-Share [129] provides a P2P distributed advertising scheme to distribute advertisements among a group of participating peers. A large number of free riders contribute nothing or few resources to other peers which may seriously influence system performance. For instance, nearly 20-40% of Napster [130] and 85% of Gnutella [131] are free riders. The impact of free riding is one of the most commonly discussed problems in P2P networks. Hence, the approach of Ad-Share integrates reputation scheme within an incentive-based model to cope with the problem of free riding to improve scalability and efficiency. Chen et al. [132] proposed a location-aware solution for the instantaneous dissemination of advertisements to a target audience within an area of interest over mobile P2P networks. The location-aware solution considers important physical constraints of networks such as a high advertisement delivery rate, low advertisement delivery time, and a flood of advertisement messages. However, the instant advertising method is suitable for limited or specific spatial/locational groups rather than broad audiences.

Previous observations of peer behavior in the P2P overlay networks were the motivation for this chapter, which emphasizes if the users are located close to the advertising promoter. They will obtain the relevant advertising media with high delivery rate and short delay. The chapter differs from related work in two significant respects: (1) we attempt to reduce video degradation through the proposed distortion-based data-replication scheme; and (2) a framework is proposed to enhance the scalability of the propagation space in the P2P advertising network. Poor video quality and frequent interruption of internet services on the user side are typically caused by delivery failures, long packet delays, or packet losses in the P2P network. A pool of common resources can be effective when applied in resource poor sessions to maintain stable video quality. To prevent the overconsumption of common resources, it is necessary to create a promotion strategy in which available resources can be shared among peers. The advertising-supported video scheme is a strategy for agents to earn revenue [133] by delivering relevant advertisements and sharing resources.

The high degree of integration between advertisements and P2P VoD networks has brought new challenges to the design of systems. P2P advertising services should send promotional messages to their preferred audience by associating advertisements with a set of keywords and network conditions. The proposed approach provides a distortion-based replication mechanism to support video on-demand services in a dynamic environment, and promotes the advertisement through resourceful peers. The high visibility of advertisements with a rapid delivery is due to strong coupling between the advertisement service and P2P VoD systems. Using this integrated design also provides flexibility in advertisement timing and placement.

Collaborative caching among peers in the P2P VoD network can be an effective way to accomplish resources sharing. For instance, in a group-management based VoD system [134], the incorporation of optimized bandwidth utilization, including the upload bandwidth, cache content, and cache capacity of each peer is used to support the playback of the entire video. All peers are clustered into groups of various sizes according to the playback point of peers, and local information is collected by the head peers of the groups. Thus, the managed range of cached chunks can be determined in individual groups, and this collaborative caching mechanism compensates for a lack of chunks in nearby groups.

However, a high number of free riders may generate considerable group dynamics, which severely degrades video quality.

In [135], we proposed a technique for the detection of peer-level bottlenecks and density-based clustering as a basis for regional replication and advertising in an unstructured P2P advertising VoD service. In this chapter, we propose a method of estimating video distortion for data stream-chunk replication in the P2P VoD advertising network. The algorithm is capable of balancing the supply and demand of video chunks under non-uniform segment popularity distribution. In addition, because the distribution of high quality video chunks is based on the estimation of video distortion, it is more likely that a client will find resources required to continue the playback and receive video of better quality. The proposed framework provides a method for advertising using the supporting-peers to distribute advertisement videos for potential customers. In addition, this approach also relies on a centralized web portal server for the delivery of advertisement messages to the target audience. It depends even more on dynamic sharing-peers delivering advertisements during video sessions. We evaluate the performance of the proposed algorithm through simulation.

5.3 System Design

The main attributes of P2P video applications can be classified into two categories: video-chunk attributes and peer attributes. Video-chunk attributes include the importance of the video and aspects of video compression (e.g. motion bytes and header information in video streams). Each peer in P2P video systems has several peer attributes [136], including location, uplink bandwidth, and communication latency, which indicate whether the video chunks can (1) be replicated to compensate for video loss; and (2) support interactive VoD. In addition, peer attributes have demonstrated value in assessing the distribution of video content in P2P television systems [137].

5.3.1 Peer-Attributes Related to Data-sharing

This chapter focuses mainly on maintaining smooth playback in a P2P VoD advertising network, configured as an unstructured streaming-based sharing system. Networks have highly unpredictable behavior because peers join and leave at any point in time. When a peer fails, one of the neighbors of the failed peer or VoD server will take over. Compared to traditional client-server unicast services, in which media files are stored on a centralized media server, the media files on P2P channels are stored across peer-to-peer networks in a uniquely decentralized manner. On the other hand, departure misses are major cause of video quality degradation in a P2P VoD system [115]. To ensure smooth video playback, the proposed mechanism is based on distributing important replicas in areas in which video-distortion is expected. All peers are encouraged to contribute resources to a global pool as data-sharing peers (supporting peers). The main peer-level attributes of the proposed system are: (1) the possibility of hiding communication latencies and the extent of distortion among peers and (2) service capacity of supporting-peers (uplink bandwidth).

5.3.1.1 Channel Model of Peers

The proposed framework employs a packet erasure network, in which the probability of packet erasure is estimated according to the estimated communication latency between each receiver-peer and the source-peer. Large communication latency between peers implies longer network RTT and a higher probability of dropping packets. The average round-trip time between peers in a P2P network can be used to indicate the probability of packet loss and end-to-end throughput. In addition, video-chunk delivery in P2P networks typically employs TCP or UDP as the underlying transport control protocol. The proposed system is based on the best-effort delivery service in the form of UDP, which does not guarantee reliability or the delivery order of network packets [138]. Nevertheless, noise (e.g. blocking, blurring, frame freezing, packet loss) due to transmission loss or switching peers in the P2P network can sometimes be an important factor influencing overall performance [139]. Hence, we assume that packet loss or corrupt files are random occurrences, meaning that peers may need to reconnect to other peers to locate required content stored on the overlay network. Video sessions with lower round-trip times imply that both the server load and service time will be reduced.

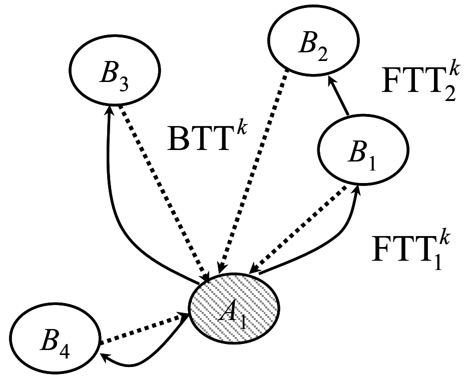


Figure 5-3: Communication latency experienced on path k within the time-to-live: 2.

Figure 5-3 shows an example set of peers during video playback of the same movie clip x_1. We assume that the TTL value is 2. In the above case, peer A_1 uses a ping-pong mechanism (solid line is the PING command; dotted line is the PONG command) to compute the round-trip time between a pair of peers. Peer A_1 sends a PING command to all of its neighbor-peers $\{B_1, B_2, B_3, B_4\}$. When a neighbor-peer receives a ping command, it immediately replies with a PONG message containing information about the neighbor-peer. Thus, we can derive a good approximation of the RTT as the measure of end-to-end latency, and the forward and backward path using an independent time-invariant packet erasure channel with random delay. The RTT between a pair of peers is used to compute the average values of RTT. In the P2P VoD environment, a long average round-trip time implies that data sharing ability is limited, and video quality varies greatly. There exists at least one forward path and a backward path for each peer in the channel. The round trip time is, by definition, the sum of the FTT (forward trip times) and BTT (backward trip times). Let FTT_1^k, FTT_2^k,…, FTT_T^k be the communication latency experienced on path k within the time-to-live T of that packet, as shown in Fig. 5-3. Therefore, the round-trip time can be computed as Eq. (5-1):

$$RTT^k = \sum_{i=1}^{T} FTT_i^k + BTT^k. \tag{5-1}$$

The probabilities of packet loss on the forward and backward channel are denoted by μ_f and μ_b, respectively. For instance, if peer A_1 sends a PING packet on the forward channel at time t, the probability of packet loss will be μ_f. Conversely, if the packet is received at its neighbor-peer B_1 at sender time t', the FFT will be $FTT_1^k = t' - t$, where FTT_1^k is distributed according to the probability density function d_f. Likewise, d_b is the probability density of the transmission delay in the back channel. According to Mukherjee [140], when the network status is stable or changes slowly, the delay over a path satisfies a shifted gamma distribution. The distribution shape depends mainly on the non-network delay (e.g. schedule and interrupt processing). The distribution center mainly depends on the network delay. In addition, the distribution center is shifted to the network traffic and queuing delay changing. Hence, we assume that the probability distribution of the packet loss and the packet delay are combined into a single probability space, and "∞" means the packet is lost or damaged. The packet delays d_f and d_b are approximated by a shifted gamma distribution. The probability of a peer with a PING packet (time t) not receiving a PONG packet by time $t+\tau$ is defined as Eq. (5-2):

$$P(RTT^k > \tau) = \sum_{i=1}^{T}\left(\mu_f^i + (1-\mu_f^i)\int_\tau^\infty d_f^i(t)dt\right) + \mu_b + (1-\mu_b)\int_\tau^\infty d_b(t)dt. \qquad (5\text{-}2)$$

5.3.1.2 Channel Sharing Ability of Peers

The ability of peers to share channels is implemented on the basis of available uplink bandwidth using time-dependent coefficients. Constraints are taken into account in the VoD service framework in which P2P video on-demand streaming could use up the available uplink bandwidth of each peer. Audio and video encoded bit-streams consume significant network resources (primarily bandwidth). The most commonly encountered issues related to P2P streaming applications are unreliable internet connections among various end users [141]. When the network bandwidth fluctuates, the coded bit-rate does not necessarily match the real bandwidth. Hence, scalable video coding techniques are often used to provide real-time quality adaptation for streaming systems. Therefore, we assume that the number of peers and the quality of video delivered to the audience-peers is constrained by the outgoing channel (uplink bandwidth) capacity of sharing peers.

The problem of free-riding, in which peers cannot or will not contribute their resources, is an important issue when designing a P2P video on-demand system. The existence of a large fraction of free riders has been demonstrated to degrade overall performance and cooperative behavior in P2P systems. Nonetheless, incentive schemes [59] or the proposed active distortion-based replication strategy can substantially enhance performance when free-riders are present in video sessions. This study incorporates the factor of free-riders into our design. Let $N(t)$ be the set of peer q connections at time t in the P2P network. Consider a communication channel with an uplink bandwidth of U_q bps. Let Ψ be the maximum uplink bandwidth in the network. When a request for video x arrives at time t, the requested peer may send a response and accept the connection j request to the requesting peer at time t_{ja}. Then, the time t_{jd} is the disconnect time from the requested peer. The connection time of the complete video stream of video x on the channel is $t_{jc} = t_{jd} - t_{ja}$ where $t_{jd} > t_{ja}$. The bandwidth allocated to the connection j of peer q at time t, is defined as Eq. (5-3)

$$\sum_{j \in N(t)} \eta_j(t) \le U_q, \quad t_{ja} \le t < t_{jd}. \tag{5-3}$$

That is, we can define the channel-sharing ability function as in Eq. (5-4).

$$\hat{\eta}_j^q(t) = \begin{cases} 0 & \text{, free riders} \\ \dfrac{1}{\Psi} \cdot \left(U_q - \displaystyle\sum_{j \in N(t)} \left(\dfrac{\int_{t_{ja}}^{t} \eta_j(t)dt}{t_{jc}} \right) \right) & , t_{ja} \le t < t_{jd} \\ \dfrac{U_q}{\Psi} & , t < t_{ja} \text{ or } t \ge t_{jd} \end{cases} \tag{5-4}$$

The channel-sharing ability of free-riding peers is zero. The establishment of all connections arriving and departing depends on the available uplink bandwidth. The remaining uplink bandwidth is equal to the total uplink bandwidth minus the mean allocated bandwidth, while some connections reside with peer q.

5.3.2 The Distortion Estimation in the Packet Bit-stream

The distortion estimation presented in this section is based on a 3D wavelet-coding technique. The SVC extension of the H.264/MPEG-4 (Part 10) Advanced Video Coding (AVC) is the latest video codec based on the DCT of ITU-T and ISO/IEC [142]. Although H.264 has many technical advantages, it also has some shortcomings [143]. Full scalability of H.264 is not well supported due to the usage of hierarchical B-pictures [144]. An alternative technique for video coding is wavelet-based coding, which has some advantages over current H.264 [145]. In addition, the method of interframe wavelet coding overcomes this drawback through the use of MCTF to achieve scalability without additional system-related overhead. In addition, the structure of open-loop prediction in interframe wavelet coding provides greater flexibility in bitstream extraction and robustness against transmission impairment when no feedback is available [146]. Hence, we adopted wavelet-based coding to make our system more robust and widely applicable.

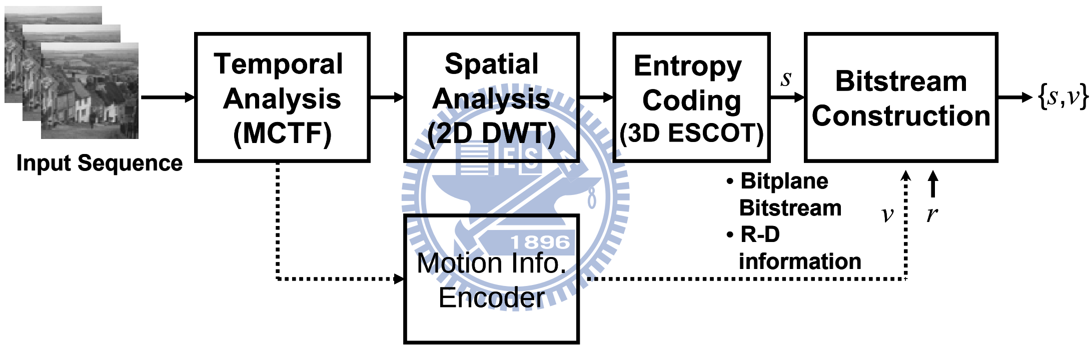


Figure 5-4: The t+2D coding structure of a wavelet encoder.

A general R-D model [147] for an embedded wavelet coder with a square-error distortion measure was used for video texture coding $R(\mathrm{D}) = \varphi \ln(\omega / D)$. φ and ω are source-dependent parameters of the logarithmic R-D model. Note that ω is related to the signal variance of the source. Although this model fits the R-D characteristics of a single coding block, it requires additional computation for source dependent parameters [91]. However, the R-D slopes provide an explicit way to quantify the distortion of texture videos. To obtain accurate distortion information, we coded all of the R-D slope-values from code blocks. As shown in Fig. 5-4, multi-level MCTF is used to decompose the video frames into several temporal subbands, including highpass and lowpass subbands. A two-dimensional discrete wavelet transform (2D-DWT) is then performed in each temporal

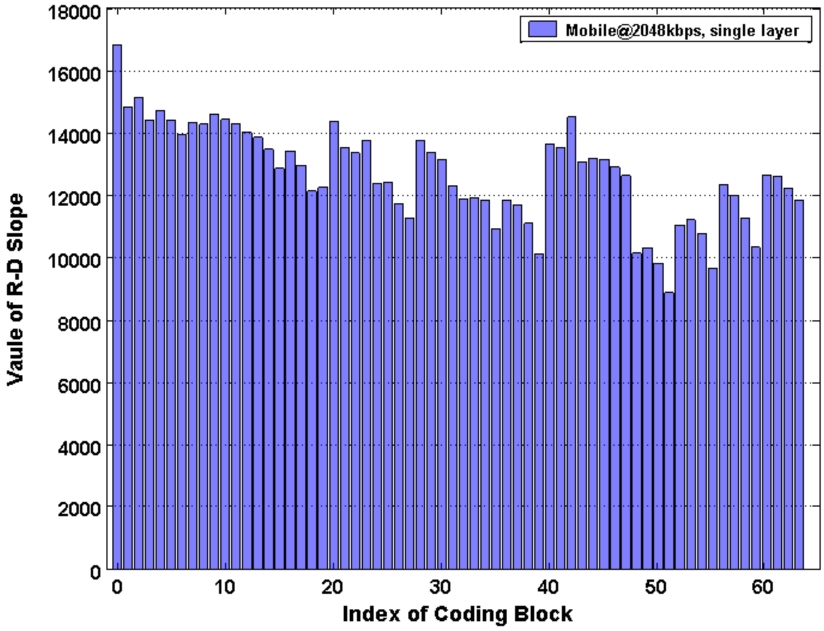


Figure 5-5: Distribution of R-D slopes.

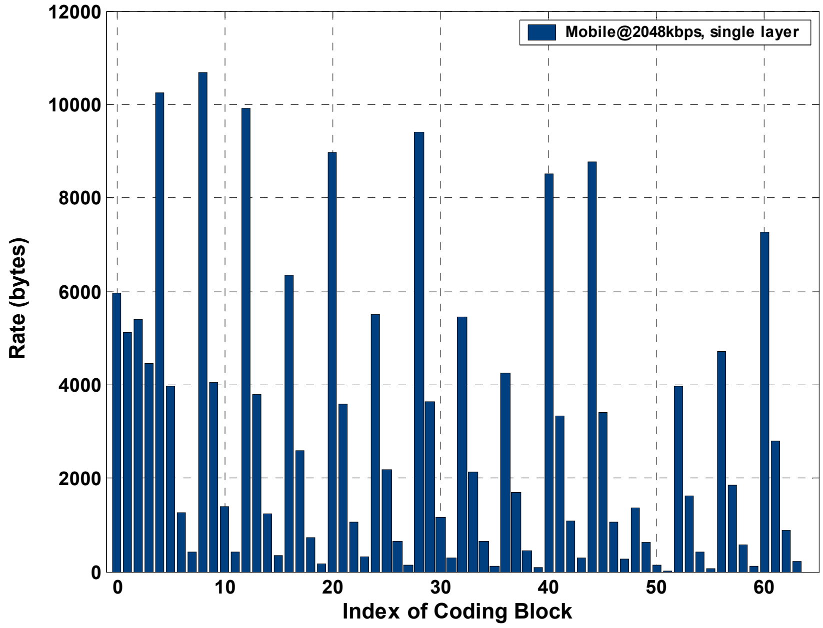


Figure 5-6: The distribution of block data rates.

subband to decompose the frames spatially. The solid line shows the data paths of the texture data, and the dashed line indicates motion information. Through the entropy coding stage 3D-ESCOT, an embedded compressed bit-stream can be generated for each subband of the 3D wavelet transform. In addition, candidate truncation points of each subband are related to R-D slopes, such that all points on the convex hull can be obtained. Finally, a bit-stream construction algorithm optimizes the trade-off between rate and distortion to further truncate each coding pass in the embedded bit-stream to form an output bit-stream. For instance, the distribution of R-D slopes and block data rates of the LLLL subband of MOBILE sequence is shown in Fig. 5-5 and Fig. 5-6. A major video distortion as well as video quality impact can be discriminated on the basis of the R-D slope values.

Based on the above observations, we assume that the amount of video distortion from packet loss is related to R-D slope information of each coding unit. In addition, a packet comprises a header or trailer and a payload which may include one or more coding units, as shown in Fig. 5-7. Thus, the expected amount of distortion reduction in GOPs due to channel conditions can be estimated by the quantity of the received video chunks in a set of resource-sharing peers. We further assume that the maximum value of R-D slopes for any given packet is therefore an approximation for the importance of that packet to the reconstruction of the video. The coding units in a GOP are divided into Y packets, and then there exists a set of coding units $\chi=\{c_1,c_2,\ldots,c_\chi\}$ in a packet. In case no packets are received within the time window (GOP), the expected reconstruction error denoted as D_0 and can be computed in Eq. (5-5):

$$D_0 = \sum_{i=1}^{Y} \max\left(\lambda_{c_1}, \lambda_{c_2}, \ldots, \lambda_{c_\chi}\right)^i . \tag{5-5}$$

The scalable bit-stream is composed of header and texture data. The header contains sensitive data such as GOP size, temporal band index, and motion information, which is variably length coded. One coding block can be coded in one or several Network Adaptation Layer Units (NALUs), and each NALU can be packed into one or several transport packets. In addition, each NALU varies in importance regarding the reconstruction of video frames. The header data of the video bit-stream is particularly important to the quality of the decoded video.

The formation of the bit-stream using the wavelet codec is explained in Fig. 5-7. Using four-level temporal and three-level spatial subband decompositions, a group of frames is decomposed into LLLL, LLLH, LLH, LH, and H subbands, and each subband is divided into a collection of coding-blocks. In addition, each subband consists of luminance (Y: gray-scale) blocks and chrominance (U and V: color) blocks. The luminance signal is the equivalent of a black and white TV signal, and has a significant effect on visual quality. The proposed replication mechanism in this work focuses on the luminance signal.

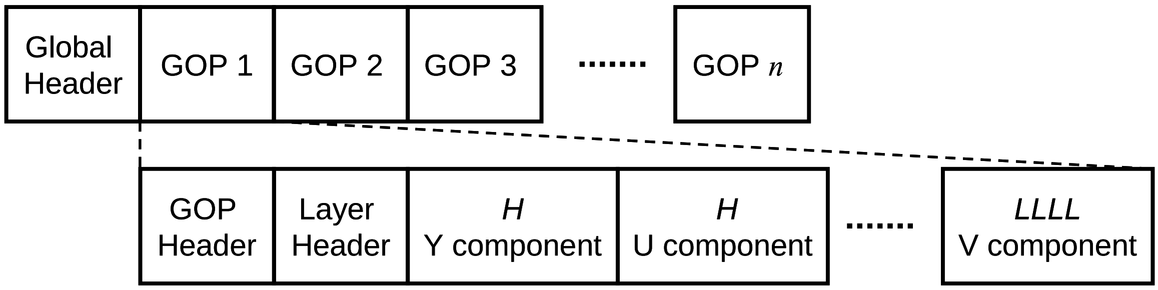


Figure 5-7: Wavelet Bitstream Format.

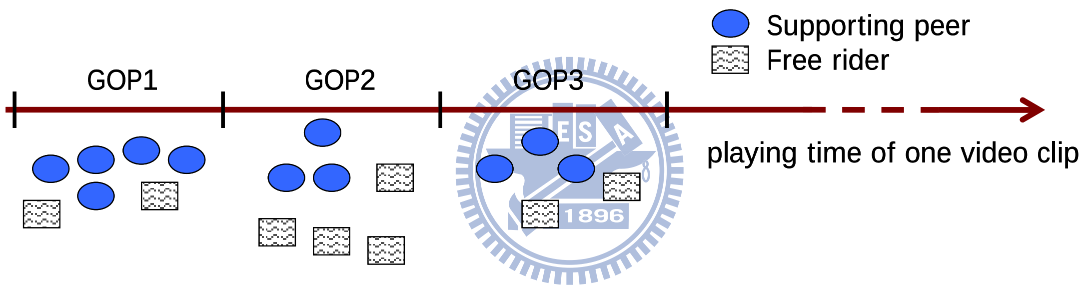


Figure 5-8: Peers at GOP boundaries.

The proposed replication strategy depends on the factor of distortion to select an appropriate mechanism for replication and the degree of replication required. Estimated distortion values are used as indicators related to the severity of video degradation in a particular GOP instance. Our method generalizes the ideas of [148] and [149], by exploiting time-varying P2P channel conditions and maximizing the video quality of the received sequence under the constraint of varying bandwidth resource allocation. For each GOP length of on-demand video x, we can estimate the distortion values of the GOP for peers active on the channel over a particular period of time, as shown in Fig. 5-8. Let $G_x^w(t)$ be the set of peers present within a GOP w of video x at time t in the P2P network, and each peer registers its own stored video-chunks on set $R(t)$. The number of participating peers within the partition is ρ. As discussed in Section 1.1.4, the packetized data unit γ' and the packetized data unit γ have a corresponding relation to the decoder. To construct data unit γ, the encoder requires that the data unit γ' also be decoded. The expected amount of distortion reduction of peer α based on fully resource sharing at time t is defined as Eq. (5-6):

$$\varpi_{\alpha}(t)=\sum_{\gamma\in R(t)}\left(\Delta D_{r}\cdot\prod_{\gamma'\prec\gamma}\left(1-\mu_{r'}(t)\right)\right). \qquad (5\text{-}6)$$

where ΔD_{γ} is the expected reduction in reconstruction error if video chunk γ is decoded on time, and μ is the probability that video chunk γ is not received on time. After the estimated distortion reduction is obtained, and then we adopt the bandwidth sharing properties of each peer. The expected distortion reduction in the GOP w of video x at time t can be computed as Eq. (5-7):

$$\varphi_{x}^{w}(t)=\frac{1}{\rho_{x}^{w}}\sum_{i\in G_{x}^{w}(t)}\left(\hat{\eta}_{j}^{i}(t)\cdot\frac{\varpi_{i}(t)}{D_{0}^{i}}\right). \qquad (5\text{-}7)$$

5.3.3 Advertising Strategies on the P2P VoD network

The concept of textual relevance matching is useful for targeted advertising. Typical examples include keyword-targeted (e.g. AdWords of Google) and content-targeted advertising (e.g. AdSense of Google). Hence, a customized advertisement can be associated with one or more keywords, which can be manually selected by advertisers. Language is a medium of communication, and the target audience often relies on the presentation in native language. Language is a useful criterion for segmenting advertising markets, and advertisers should be able to include this in schemes to customize their own advertising plans without wasting network or processing resources. Such schemes can include launch date, advertising language, and keywords for different audience-peers.

A multitude of advertisement payment models, including cost-per-action, cost-per-click, cost-per-impression, cost-per-download, and cost-per-visitor, can be implemented. Two major categories of internet video advertisements are in-page and in-stream. In-page advertisements are video advertisements embedded in a search-engine result page, containing search results and the retrieval of advertisement tracking. In-stream advertisements can be within streaming video content or played in the advertisement window. In the proposed framework, an internet video advertisement can be placed before, during, and/or after the demanded video content and played within the advertisement window of the application.

Advertisement publishing rules can be created to match advertisements with similar keywords in the VoD clips to describe which advertisements should be associated with clips. The delivery of advertisements is based on the movie clip keywords found on the time line of the audience-peer group that is attracting advertisements, and sharing peers sending them to audience peers in the P2P online marketing communication channel. In this manner, commercial advertisements can be delivered through P2P VoD advertising platforms, and the targeted messages can be delivered to the target audience. P2P VoD advertising services have expanded the horizons of advertising by quickly distinguishing the audience using a video catalog, tightly integrating the video-content and advertisements, and increasing the visibility of advertisements in a scalable manner.

5.4 Proposed Advertising P2P VoD Framework for Wavelet Bit-streams

In this section, we present the proposed distortion-based replication scheme and advertising approach introduced in Section 5.3 for P2P VoD applications using a wavelet codec. The main operating characteristics of the proposed P2P VoD advertising framework includes: (a) an on-demand video repository server, (b) a web portal service, (c) trackers, (d) audience-peers (a set of free-riding peers), and (e) supporting-peers. The on-demand video repository server stores a complete copy of encoded video clips, and serves a number of the requests that arrive in the queue of the server. The web portal service provides audience-peers with online video information and delivers advertisements to each audience-peer who has sent QUERY messages. Trackers help newly joined peers to bootstrap nodes and coordinate the replication of significant chunks through the proposed distortion-based

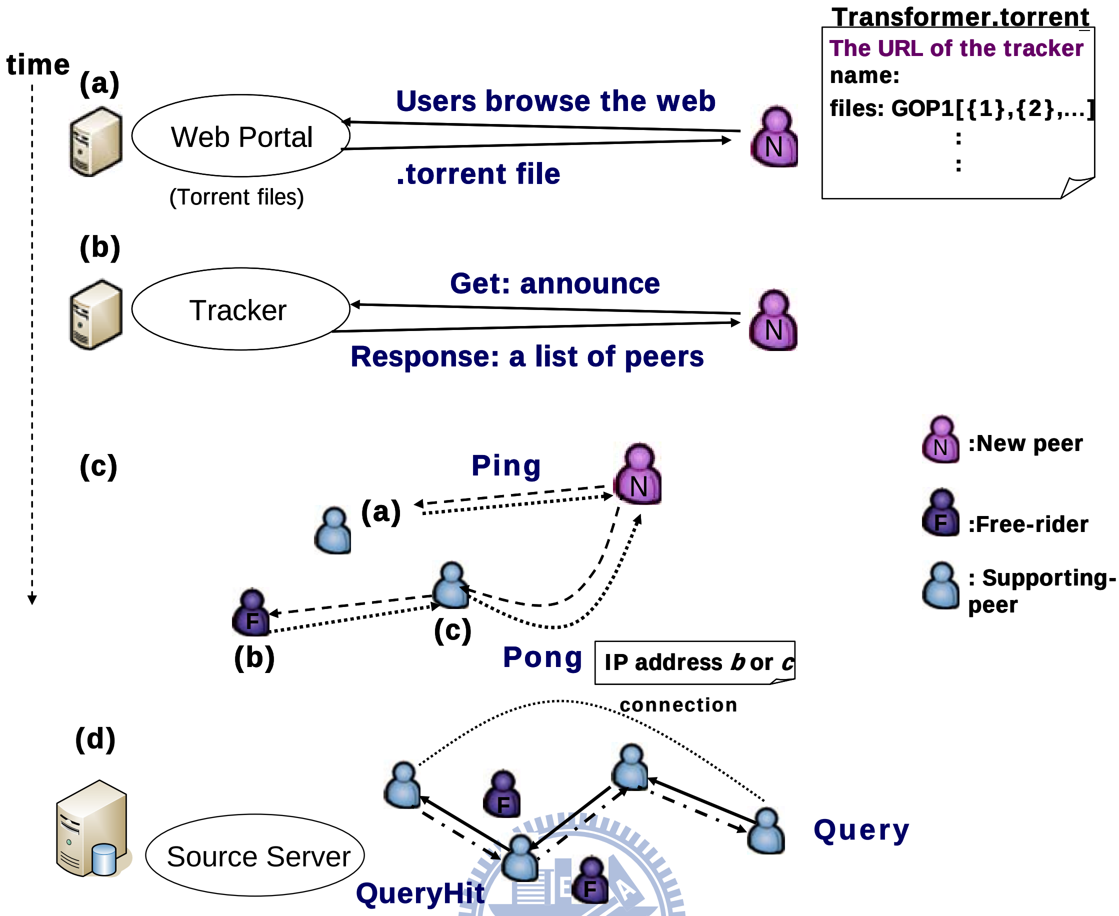


Figure 5-9: The operational scenario of a newly joining peer.

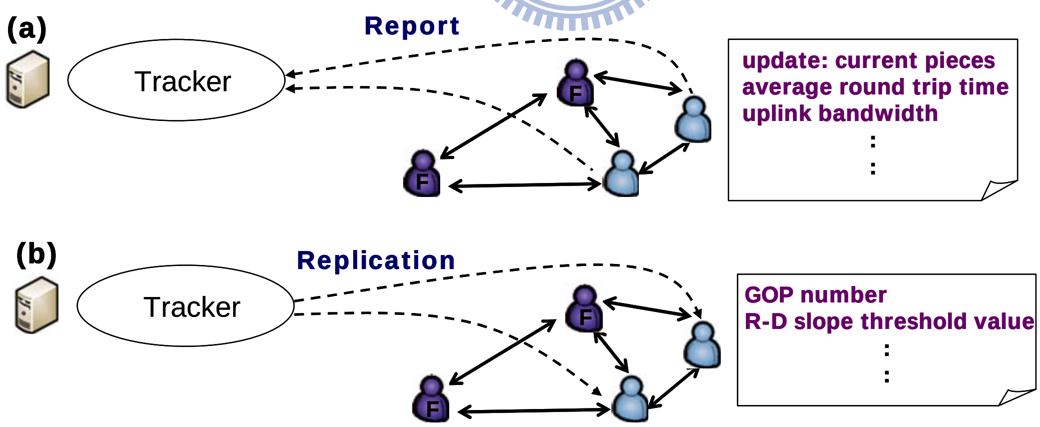


Figure 5-10: The operational scenario of a Tracker.

strategy. Finally, supporting-peers (idle or resourceful peers) fetch chunks from the server or other peers, and deliver advertisements to each supported audience-peer. The supporting peer can also be an audience peer. In the proposed framework, we use a hybrid method that combines Bittorrent and Gnutella protocols [131] to maintain connections with other peers in the network. The operational scenario of the proposed system is shown in Figure 5-9. A newly joining peer connects to the web portal server for enabling a peer to interface with a service configuration. In addition, the new peer will get the main information of video-clips available on the P2P network. When the new peer starts the step of joining, it will try to connect the tracker to get peer list and then send "PING" message to other peers in peer list. A peer receiving a ping message may respond with a pong message. The PONG message is comprised of an IP address and port number of a potentially active peer. To look for video data, a peer submits a QUERY message to its neighboring peers. The neighboring peer will stop relaying the query if the TTL becomes zero. A neighboring peer who receives a QUERY message will respond with a QUERYHIT message if the neighboring peer finds one or several matches in its local data store. The tracker will be responsible for collecting relevant information from the peers, as shown in Figure 5-10. Furthermore, the tracker has to coordinate replication activity in a dynamic P2P environment.

One common difficulty encountered in P2P VoD systems is a severe being lack of resources to allocate to individual peers. An appropriate fault-tolerance design for P2P VoD system can help moderate performance degradation in the presence of peer failures and/or bandwidth degradation. This is particularly important for continuous operations such as video playback are essential in P2P VoD systems. Another challenging aspect of P2P VoD systems is the use of fault-tolerant design in replicating multimedia files in appropriate quantities. Replication enables the holding of a greater share of media repositories during high service demand. Thus, numerous P2P replication schemes have developed for various performance objectives, such as improved startup time, media-file availability, and response time. P2P replication schemes can be classified into two major types: active and passive. Passive replication systems are commonly designed for file sharing through download, with a focus on maximizing data-holder value to improve overall file availability or hit rate. However, the video quality of P2P multimedia applications is greatly affected by variations in bandwidth, delay jitter, and packet loss. Proper active replication in the P2P VoD system is necessary to continuously stream video playback of acceptable quality. Constructing P2P VoD advertising mechanisms involves four key issues associated with packet loss during video transmission over P2P networks. The first is the requirement of timely and continuous streaming to meet the playout deadline at the audience site. The second issue is that bandwidth requirements for the P2P VoD networks are increasing at a rapid rate (from 200-300 Kbit/s to 1-5 Mbit/s [136]). Hence, improving access time and efficient bandwidth utilization over P2P channels is a challenge. The third issue is that the perceived degradation of video quality is often negligible when packet dropping is within acceptable limits. An appropriate data replication scheme should be used to protect video content from network errors, since higher priority packets have to be received on time. The last issue is what we call flash crowd: a sudden increase in peer arrivals on the P2P overlay networks.

Our proposed method indicates replication locations, according to the proposed distortion estimation method of GOP. Supporting peers are designated by the tracker to compensate for loss or damage arising from unexpected neighbor-peer or network failures. Moreover, the popularity index of clips changes dynamically with time. We organize peers in an unstructured P2P network into an undirected graph topology. $G(t)=(Q, E)$ is defined as the undirected graph comprising a set of participating peers and a set of overlay links at time t. Then, Q is a finite set of peers and E is a set of unordered pair $\{u, v\}$ of distinct peers in the P2P streaming overlay network, where the population size $|Q|$ is larger than 2. In the P2P overlay network, each peer may download or upload streaming content from multiple peers. The number of replica is proportional to the number of supporting-peers and the level of replication. In addition, the level of replication is chosen depending on the desired video quality required. We assume that the error probabilities are independent of each other. The proposed algorithm is summarized as follows. Note that replication process is constantly adjusted to maximize the recovery of video quality and operational efficiency.

Algorithm: Distortion-based Replication Strategy

1. Input: Graph $G(t)$, the set of on-demand videos $V(t)$, with sort by video popularity distributions, supporting-peers ζ, the desired level of video quality σ
2. Let v get one video from the set of candidates $V(t)$.
3. For each candidate peers in v from graph $G(t)$

   Obtain the round trip time values within the connections of the candidate peers through ping-pong mechanism and TTL constraint.

Obtain the channel-sharing ability using Eq. (5-4).

End for

4. Calculate the error probabilities of the video chunks using Eq. (5-2).

5. Estimate the expected video distortion at each GOP in the *v*:

a) Calculate the expected reconstruction error denoted as D_0 using Eq. (5-5).

b)For each peer within the GOP *i* of *v*, find the estimated distortion reduction using Eq. (5-7).

c) The expected distortion of the GOP *i* is approximated by the expected distortion reduction in (b)

d) Increase the index *i* to move downstream.

e) Iteratively perform steps (b)-(d) until reaching the end of video clips.

6. Remove v from the set of $V(t)$. Repeat the step 2 for the next on-demand video from $V(t)$ until the set is empty or there are no available supporting peers.

7. The dynamic policy of video-chunk replication:

In step (1), video popularity is dynamic and changes over time. On-demand videos can be divided into equivalence classes by means of periodic partitioning at several levels according to video popularity distributions. For each level, important video-chunks are replicated (based on the step 5 and the desired level of video quality σ) by supporting peers ζ with the aim of providing fault tolerance to the system. Figure 5-11 presents a much more detailed explanation of the proposed replication algorithm. The tracker will collect relevant information of peers from the P2P overlay network, such as video chunks in the buffer, the average RTT among the inter-communicating nodes, and the available uplink bandwidth. The level of important of the video-chunks was determined using an R-D slope threshold value. The resource allocation policy is implemented by two phases of data replication methods. The aim of the first phase is to keep important data available for users. The least important data may be removed from the buffer or disabled from a downloading mode if the available space is limited within the current GOP. The second phase is the data replication support by the neighboring GOP. Each GOP has a certain storage space and uplink bandwidth reserved for participating peers. The spare uplink bandwidth and space can freely assign to connections in order to increase overall throughput. Hence, the uplink bandwidth and storage of peers of the neighboring GOP the can be effectively utilized for maximizing data replication efficiency. The number of replicas for each supporting peer is based on the half of spare uplink bandwidth and space.

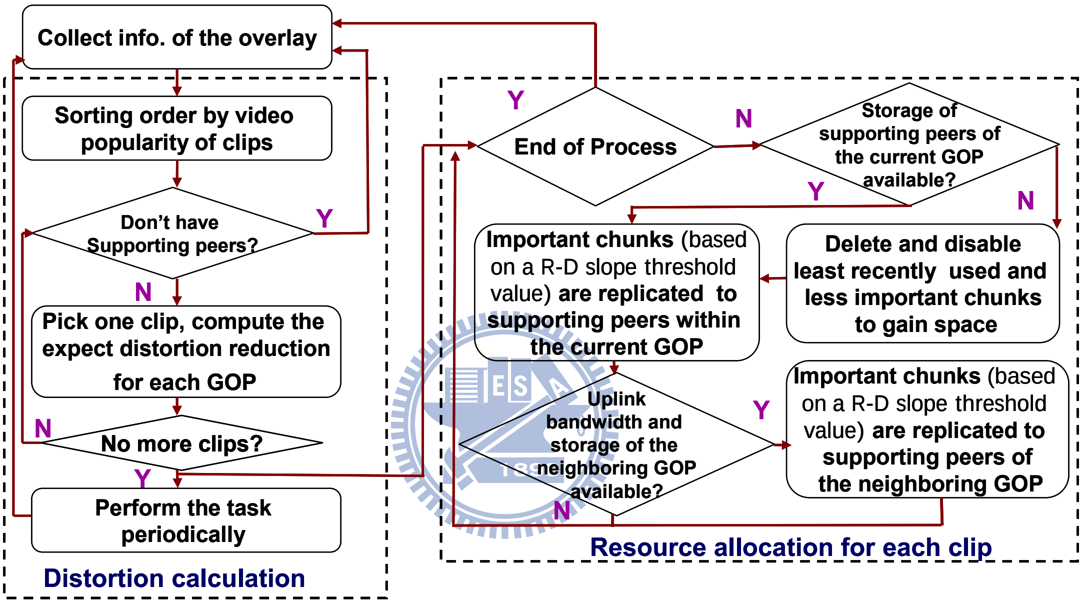


Figure 5-11: The proposed replication mechanism.

The proposed algorithm constructs a fault tolerance approach to the P2P VoD advertising service, providing groups of participants in a video session with video loss recovery using replicated chunks of importance.

5.5 Simulation Experiments

This section presents the results of simulation experiments conducted to evaluate the performance of the proposed P2P VoD advertising framework. The experiments compare the performance of the proposed scheme with a random replication strategy (distributing the replicas in a random order) (RR), the bottleneck-based replication method (BR) [135], the collaborative caching method (CC) [134] (described in Section 5.2), and the proposed method. We evaluate the proposed strategy through four ways: analysis of server load, analysis of advertisement-delivery rate, testing for departure misses, and the impact of free-riding peers. The collaborative caching method competes well with the proposed approach because caching is effective for managing replicas of data in P2P systems. To ensure a fair comparison, the collaborative caching method was modified to introduce advertisements broadcast to group head peers, using the same number of advertisements. Nonetheless, we assumed that the replica files in the four strategies occupy the same amount of storage space. A simulator was developed for analyzing the behavior of a P2P VoD advertising network under multiple design factors.

5.5.1 Experimental Setup

To evaluate the performance of the proposed system, the wavelet coder was used to produce packetized non-layered bit-streams, and all video sequences (STEFAN, MOBILE, FOREMAN, COASTGUARD, and TABLE) were stored as CIF versions (352 × 288) of the standard MPEG format. There were two kinds of videos on the P2P VoD advertising network: on-demand videos and advertisement videos. The length of the advertisement videos was generally between 17 seconds and 4 minutes. For in-page advertisements (video advertisements embedded in a search-engine results page), the most common lengths of play were 30 seconds and 15 seconds [148, 150]. In the simulation, we choose the length of advertisement-clips to be 40 seconds (15kb~20kb, 100 advertisement-clips). The length of the on-demand video followed a normal distribution, ranged from 1 minute to 60 minutes, and was encoded using VidWav reference software [151] at 15 frames per second, with a GOP comprising 64 frames. Moreover, the length of most on-demand videos was assumed to be between 5 minutes to 15 minutes [152]. The number of on-demand videos was 200 (around 200MB). 20% of the total space used for all replicas (around 20MB). We focused on a *t*+2D decomposition scheme (first temporal decomposition followed by spatial decomposition) and a four-level temporal transform employing three spatial decomposition operations. To achieve acceptable PSNR value (> 40dB), the desired video quality (PSNR) is set to 40dB.

In the P2P VoD system, each peer was equipped with asymmetric uplink and downlink bandwidth in the overlay network, such as through an asymmetric digital subscriber line. The performance of the system was particularly influenced by the limited uplink bandwidth of participating peers. Based on the study [153], Table 5-1 provides the uplink bandwidth distribution of peers, all of whom were collected from the internet. In many current operating systems, users are permitted to execute multiple applications simultaneously on a single device. Many internet applications, such as e-mail, internet telephony, and web browsing consume substantial amounts of network bandwidth. Hence, after subtracting non-P2P traffic from the total bandwidth, peers can provide shareable bandwidth in the range from several kbps to around 1,000 kbps for shared resources in overlay networks. We assumed that the distribution of bandwidth could be determined by following Table 5-1.

Table 5-1: Peer uplink bandwidth distribution (kbps).

Percentage	Bandwidth (kbps)	Shareable Bandwidth (kbps)
10.0	256	150
14.3	320	250
8.6	384	300
12.5	448	350
2.2	512	400
1.4	640	500
6.6	768	600
28.1	1,024	800
16.3	>1,500	1,000

The query operation is based on the flooding search algorithm with a TTL constraint. We set TTL=6 and the network size was 100,000 peers with a default setting of 40% free riding peers with peers randomly assigned to view movie-clips to provide uniform probability. The path length of 90% of the internet maps was under 20 hops, and 99% were no more than 25 hops in length [154]. The communication latency between peers was in the range of 20 ms to 300 ms [155]. In this simulation, the distribution of the communication latency closely approximated a normal distribution. As recommended in [156], a Poisson-like distribution was adopted for user arrival rates. The distribution of video popularity followed a Zipf-like distribution as recommended in [152, 156], and all copies of the compressed video-clips were stored on the server. Average PSNR was used as a quantitative video quality metric for evaluating the algorithms.

5.5.2 Server Load Analysis

We conducted two simulations to determine whether the influence of server variability could be reduced using the proposed method. The server load imposed by a large numbers of peers, particularly free riders with simultaneous requests, can result in problems related to performance. Thus, the server load imposed under dynamic network conditions with the proportion of peer number was investigated using 50% free-riding peers. We assume that users who subscribe the same video quality level. Figure 5-12 shows that our proposed method outperforms RR, BR, and CC schemes in terms of reduced server bandwidth usage and peer waiting time. There are two reasons for the poor performance of the CC scheme. The first reason is that the limited cache capacity of peers has a remarkable impact on resource-sharing performance. The second is that the influence of free riders on the groups makes it difficult to cache many time-varying data. There were no available resources for collaboration, which increased server stress. On the other hand, supporting-peers with important video-chunks were able to satisfy the request of peers, and compensate for the possibility of failure in zones with low resources. The average estimated video-distortion proved a good indicator of potential long-term significant improvements in server throughput. Trackers can allocate a large fraction of the important video chunks (carried by supporting-peers) to locations suffering high distortion. Enhancing the average resource rate can help avoid burst traffic and reduce server load in the overall system. The reliance of overhead on the centralized on-demand VoD server limits the scalability of the system in the presence of large and dynamic peers. Hence, the proposed approach enhances the scalability of the network, and achieves significant cooperative gains in obtaining high performance in large P2P VoD systems.

Flash crowd traffic is generated with a small set of popular videos from a large number of peer requests over the internet and server load or end-to-end network bandwidth may suffer large fluctuations due to the flash crowd traffic. Media-segment delivery rate is

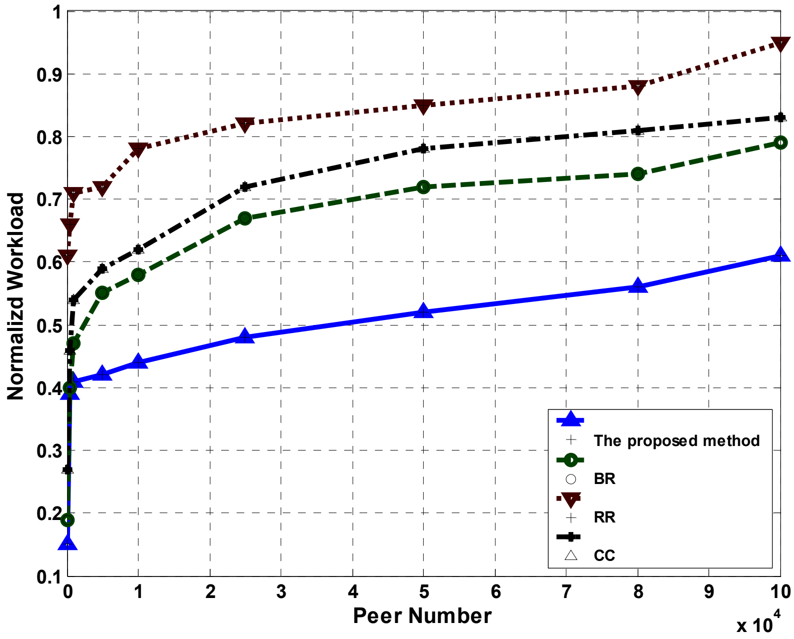


Figure 5-12: Workload imposed on the server.

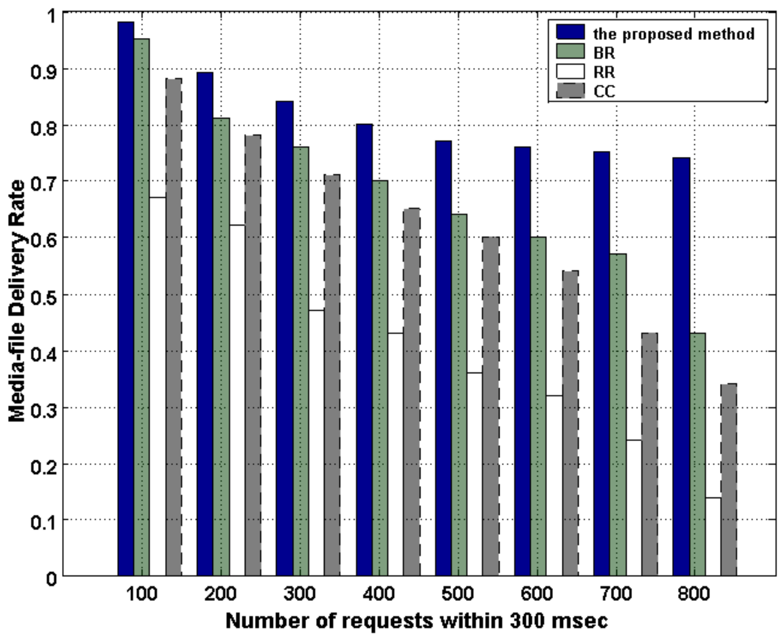


Figure 5-13: Percentage of requested file received.

defined as the percentage of audience-peers that successfully receive requested video-chunks. The simulation was conducted by performing simultaneous requests for access to a particular resource (the same group of video-chunks) within a short time interval (300 ms) with 60% free-riding peers. All peers communicate in an unstructured way. As shown in

Fig. 5-13, the proposed scheme has the potential to reduce server load and a large number of requests between peers can be properly matched by supporting peers. Therefore, the proposed approach is capable of significantly reducing peak demand at the media server.

5.5.3 Advertising-delivery Rate Analysis

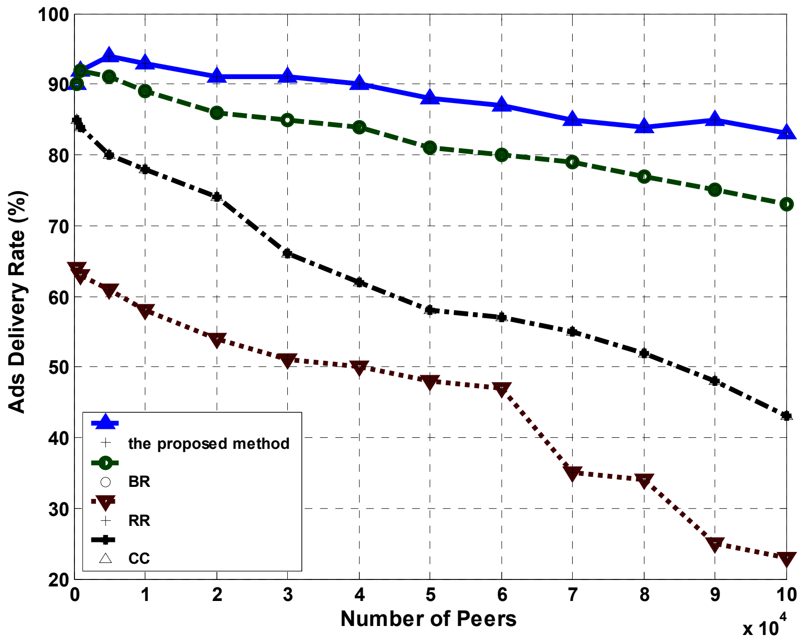


Figure 5-14: Advertisement delivery rate for networks of various sizes.

The advertising service is carried by the supporting peers, with commercial advertisements delivered to the appropriate online audience based on partial matching of the advertising plans and the keywords associated with the clips. The proposed scheme significantly increases the chance that an advertisement will be promoted on the online marketing channel. The advertisement delivery rate is defined as the percentage of audience peers who receive the advertisement within a fixed time period (1 min). Each advertising video is inserted into a packet buffer of the peer's advertising window, from which it is played. As shown in Fig. 5-14, our advertisement delivery rate is higher than those of the RR, BR, and CC methods with the network size varying from 500 to 100,000 peers. This set of curves indicates that as the managed group range of the CC method increases in size, and then it becomes inefficient for a head peer to maintain close bonds with all group members. The advertising can be effectively pushed into the P2P environment due to the ease of adding additional advertising messages.

5.5.4 Analysis of Peer Departure Misses

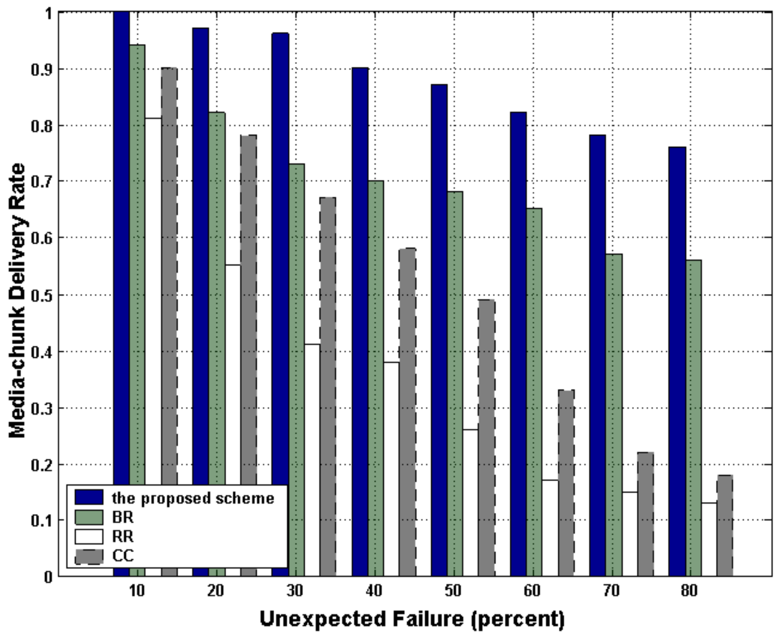


Figure 5-15: Comparison of unexpected failure of peer departure misses.

In the P2P network, each peer can decide whether and when it wishes to join or leave the VoD session. This simulation examines the behavior of peer departure misses. During simulation sessions, a default value of 40% free-riding peers was used. Figure 5-15 shows that in the BR case, the maximum was approximately 20% higher in a limited resource setting. The proposed distortion-based replication method can decrease departure misses and thereby further reduce server load. Some bits in compressed data contain a large amount of information that is very sensitive to errors, and channel-induced distortion leads to a perceived degradation in quality related to the reconstructed frames at the decoder. With the proposed method, the multimedia video-chunks are protected from a potential loss of quality in the data traffic well. In the RR scheme, the allocation of replicas is often less than the number of peer requests for a resource; therefore, the performance is much worse in comparison to the BR or proposed methods. In the case of the CC method, it has been shown that a decrease in the cooperative share rate has a dynamic impact on the delivery rate of media-chunks, particularly with high departure rate for head peers.

5.5.5 Impact of Free-riding Peers

As specified in Section 5.2, free riders make up the majority of peers in P2P overlay networks. In this subsection, we examine the impact of various percentages of free riders. We assume that users who subscribe the same video quality level. Figure 5-16 illustrates that the workload of the VoD server can be decreased by more than 20 % compared to that of the BR under the maximum number of participating peers. The reason is that the range of video-chunk replicas in BR only considers the network conditions according to the number of playbacks. When the free-riding peers aggregate within continuous groups, the failure rate increased in the presence of resource leaks. Although the performance of the proposed method is slightly worse than in the CC scheme when the number of free riders is less than 20%, an increase in the number of free riders results in a considerable improvement in performance. This is mainly due to the fact that resource sharing depends on a willingness to cooperate and contribute to the cached data. A lack of sufficient network cooperation for the same video session leads to a competitive

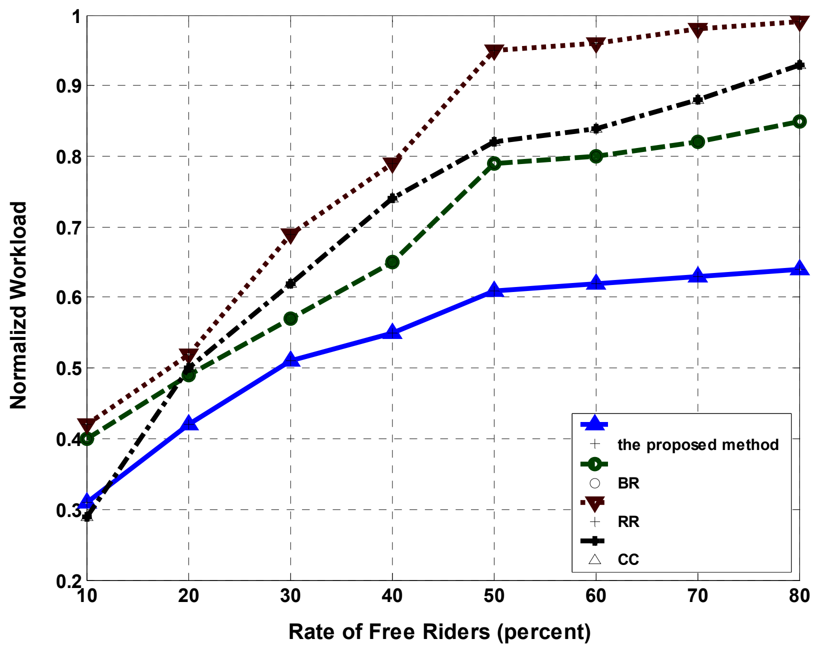


Figure 5-16: Workload imposed on the server with respect to the rate of free riders.

disadvantage and a lack of resource sharing. The simulation illustrates that the proposed scheme adapts to the requirements of the P2P VoD system, enabling the supporting peers to share P2P multimedia stream-sharing workloads from the VoD server. The replication algorithm avoids sub-clip misses and reduces the risk of request implosion caused by free-riding peers.

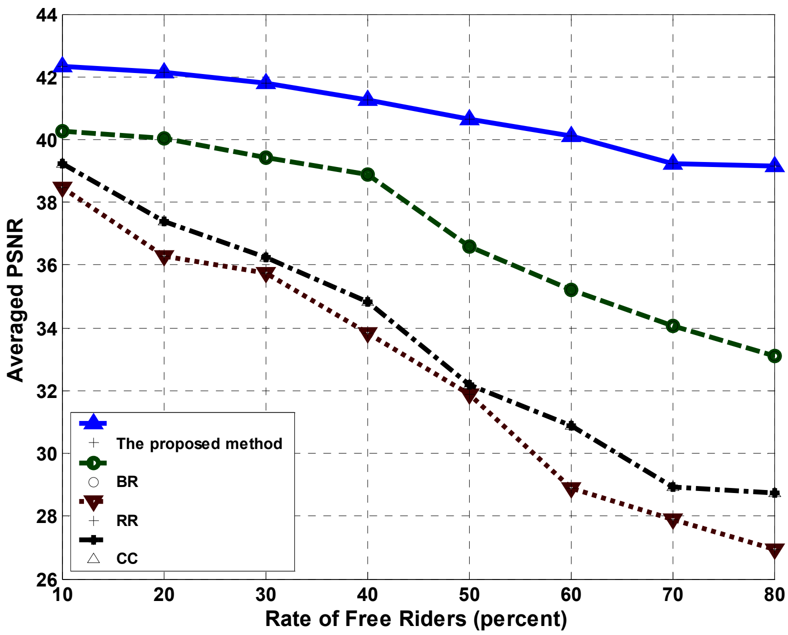


Figure 5-17: Comparison of the proportion of free riders with respect to the averaged PSNR.

The averaged PSNR is an averaged value for the test video sequences. Figure 5-17 shows the comparison of averaged PSNR for the first 256 frames of the video sequences Mobile and Stefan, using different approaches. When the number of replica is the same, the results demonstrate that the proposed method is more robust. When free-riders are a majority in the groups, the proposed scheme has a stronger preference for resource aggregation, particularly when the playing session includes video clips of high popularity and video chunks with a strong impact on distortion. Thus, a given bit budget for supporting peers can be distributed by trackers to improve video quality. By contrast, the RR and BR methods enable downloading of video-chunks in the packet buffer by online peers, but those replication strategies do not accurately represent current resource conditions in the P2P VoD network. In the CC method, the lack of priority given to cached video-content for cooperation can reduce video quality. In addition, it is expected that packets (caused by the limited sharing resources that must be accessed serially) would have a longer round trip time, implying high packet loss rates, causing the decoded video quality to degrade rapidly.

5.6 Summary

The importance of video quality enhancement has become increasingly obvious in recent research on P2P technologies and modern television applications. Internet users, who watch online P2P on-demand programs, require an easy-to-use integrated platform to share resources. In a P2P VoD service, the video program needs to be divided into chunks and packetized by source peers, whereupon the original video program is decoded by the receiving peers from chunks within a playback deadline. However the video program is not necessarily reconstructed perfectly due to packet loss, delays, heavy server load, or lack of shared content. Hence, we demonstrate that the proposed distortion-based video-chunk replication is to ensure the continuous delivery of video streaming in high-quality video coding applications. The effects of P2P advertising on society are relevant in an online marketing environment. We conclude that an effective online marketing communication channel is the key to success in terms of both advertising effectiveness and VoD service satisfaction using P2P networks. Simulation results support the hypothesis that video distortion estimation prior to data stream-chunk replication is an efficient tool for balancing load among peers, reducing latency to audience-peers, and improving the overall visual quality for the end-users. Our future study will extend the proposed framework by creating an integrated interactive advertising platform with increased audience-peers interaction, customer retention, and P2P community-driven advertising.

Chapter 6 Concluding Remarks and Future Works

6.1 Concluding Remarks

This dissertation studies the issues regarding video streaming and P2P VoD applications, which lies within the scalable coding streams research framework. It considers time-varying channel conditions and resource discovery for dynamic network users simultaneously so as to improve the video-quality and reliability of video communication systems. In this chapter, we provide some comments on the efficiency of adaptive fine-granularity error protection methods and media-data replication strategies in P2P environments. In Section 6.2, potential problems and ideas for future work are also discussed.

6.1.1 Comments on Fine-level Packetization Design

Reducing video quality variation over loss channels studied in this dissertation provides a possible way to resolve the rate and distortion impact of packet losses. One solution to this problem is a framework of R-D optimized video streaming over lossy IP networks. The tradeoff between rate and distortion is controlled by changing the scalable source coding rate and the level of channel protection. The channel protection level is achieved using an adaptive FEC scheme using Reed-Solomon coding. Moreover, the system performs dynamic rate allocation to either enhance perceived quality by sending more enhancement-layer data or adding more FEC protection to the base-layer under various network conditions. The R-D impact of packet losses is mapped to various levels of FEC protection using data-interleaving and Reed-Solomon coding for base-layer bitstream. For the enhancement layer, the system performs a dynamic rate allocation at discrete transmission times to enhance the perceived quality when the network bandwidth is good enough for perceptible and consistent quality improvement. As shown in the experimental results and analysis, the proposed system can meet smooth presentation requirements under different network conditions and achieve high bandwidth utilization. Moreover, the proposed system can achieve smooth presentation with reduced video-quality variation between adjacent frames.

For fine-level adaptive FEC protection of video coefficients, the R-D information of packetized video streams during encoding can provide an adequate way of describing the protection-level of coefficients of different fractional bitplanes under given estimated packet loss rates of the error-prone channels. Along this line of consideration, we propose two frameworks that fine-level FEC protection algorithms to achieve wavelet video streaming and multiple adaptation functionalities, respectively. In the wavelet video streaming framework, a fine-level FEC protection/packetization mechanism of wavelet video data is presented. In the wavelet video multiple adaptation framework, the proposed frameworks enables multiple adaptations of fine-level FEC protection scheme for more flexible error-resilient transmission of bitstreams. The adaptive packet loss protection scheme using Reed-Solomon coding and data interleaving is based on detailed analysis of rate-distortion tradeoffs of wavelet subband data. Experimental results show that with an adaptive fine granularity FEC protection scheme, the proposed mechanism can achieve much better video quality than with a fixed-level FEC protection scheme.

6.1.2 Comments on Fault-tolerance P2P VoD System Design

The voluntary resources of P2P services are highly unreliable (join/leave the system frequently) and unpredictable. Therefore, data replication approaches have been used widely in designing P2P video systems, especially for media-on-demand systems to allow data redundancy to achieve service quality. We propose a video-distortion based replication strategy for P2P VoD services over varying overlay channels. This research investigates an efficient control protocol to facilitate continuous and stable P2P VoD playback. For the purpose of media-data replication, all peers are encouraged to contribute resources to a global pool as supporting peers (especially idle or powerful peers). The proposed distortion-based replication mechanism is based on distributing replicas of popular and high video-distortion chunks to supporting-peers. In addition, the proposed method also overcomes free-riding behavior in P2P systems. Simulation results demonstrated that video-distortion based replication can realize efficient VoD services in large-scale unstructured P2P overlay networks. In summary, the main features of the proposed system are as follows: (1) Distributes highly popular and quality impact replicas to the regions of high distortion subgroups or discontinuous areas to act fast response and localized failure recovery (fail randomly). (2) Effectively handles asynchronous requests from peers so as to minimize space and time consumption in the overall system. (3) Minimizes video-quality degradation due to packet losses. The scheme flexibly adapts the popularity and quality impact of video to build a replication strategy. Simulations also show that the proposed solution can achieve better video quality in large-scale unstructured P2P overlay networks even within a high level of free-riding nodes. Moreover, the workload of the VoD server can be decreased, and average visual quality can be improved.

6.2 Future Work

Based on our work, ideas worth further investigation are listed below:

1. *The impact of user mobility on video streaming*: In this dissertation, we did not consider network prediction and user mobility in more realistic scenarios. Further work on supporting such functionality based on our error-protection mechanisms will be pursued.
2. *Channel switching*: Enhancing channel changing time (the time required to join P2P overlay of a channel, fast channel switching is necessary in IPTV world) is still an open question. Thus, traffic localization and transport protocol optimization are issues in reducing re-buffering time demands in a P2P streaming environment. Further study on integrating our work with P2P streaming protocols will be conducted.
3. *Event history analysis of peers*: A primary conducted for using data mining in P2P systems is to extract user preferences and global information from the analysis of traffic streams. This is helpful with the effective operation of VCR-like (video cassette recording) controls. For instance, overall traffic bandwidth, supporting-peers

occupancy, and individual peer behavior can be used to allocate appropriate resources for the duration of playback sessions (including VCR-like operations).

4. *Incentive mechanisms in P2P VoD systems*: An effective incentive mechanism has the potential to improve long-term system performance. Therefore, in addition to providing effective incentives, it is critical for video-on-demand systems to be robust in the presence of cooperative and social enterprises. There will be further work on increasing the cooperative availability and harmonizing the use of the overall resources for VoD services.

References

[1] B. Wang, J. Kurose, P. Shenoy, and D. Towsley, "Peer-to-peer Streaming of Scalable Video in Future Internet Applications," *IEEE Communications Magazine*, vol. 49, no. 3, pp. 128–135, Mar. 2011.

[2] M. Kolberg, M. Merabti, and S. Moyer, "Trends in Consumer Communications: More Services and Media, Less Wires," *IEEE Communications Magazine*, vol. 48, no. 6, pp. 253–258, Jun. 2010.

[3] D. Jurca, P. Frossard, and A. Jovanovic, "Forward Error Correction for Multipath Media Streaming," *IEEE Trans. on Circuits and Systems for Video Technology,* vol. 19, no. 9, pp. 1315–1326, Sept. 2009.

[4] E. Maani and A. K. Katsaggelos, "Unequal Error Protection for Robust Streaming of Scalable Video Over Packet Lossy Networks," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 20, no. 3, pp. 407–416, Mar. 2010.

[5] J.-Y. Chen, C.-W. Chiu, G.-L. Li, and M.-J. Chen, "Burst-Aware Dynamic Rate Control for H.264/AVC Video Streaming," *IEEE Trans. on Broadcasting*, vol. 57, no. 1, pp. 89–93, Mar. 2011.

[6] X. Zhu, R. Pan, M. S. Prabhu, N. Dukkipati, V. Subramanian, and F. Bonomi, "Layered Internet Video Adaptation (LIVA): Network-Assisted Bandwidth Sharing and Transient Loss Protection for Video Streaming," *IEEE Trans. on Multimedia*, vol. 13, no. 4, pp. 720–732, Aug. 2011.

[7] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", RFC 3550, Internet Engineering Task Force, Jul. 2003, http://www.ietf.org/rfc/rfc3550.txt.

[8] J. Postel, "User Datagram Protocol," RFC 768, USC/Information Sciences Institute, Aug. 1980, http://www.ietf.org/rfc/rfc768.txt.

[9] J. M. Boyce, "Packet Loss Resilient Transmission of MPEG Video over the Internet," *Signal Processing: Image Communication*, vol. 15, no. 1-2, pp. 7–24, Sept. 1999.

[10] E. N. Gilbert, "Capacity of a Burst-noise Channel," *The Bell System Technical Journal* vol. 39, pp. 1253–1265, Sept. 1960.

[11] E. O. Elliott, "Estimates of Error Rates for Codes on Burst-Noise Channels," *The Bell Systems Technical Journal*, vol. 42, 1963.

[12] J. Cain and R. Simpson, "The Distribution of Burst Lengths on a Gilbert Channel," *IEEE Trans. on Information Theory*, vol. 15, no. 5, pp. 624–627, Sept. 1969.

[13] B. D. Fritchman, "A Binary Channel Characterization Using Partitioned Markov Chains," *IEEE Trans. on Information Theory*, vol. 13, no. 2, pp. 221–227, 1967.

[14] G. T. Nguyen and B. Noble, "A Trace-Based Approach for Modeling Wireless Channel Behavior," in *Proc. the 1996 Winter Simulation Conf.*, pp. 597–604, 1996.

[15] R. H. McCullough, "The Binary Regenerative Channel," *Bell Syst. Tech. J.*, vol. 47, pp. 1713–1735, Oct. 1968.

[16] P. Sadeghi, Rodney A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-State Markov Modeling of Fading Channels," *IEEE Signal Processing Magazine*, vol. 25, no. 5, pp. 57–80, Sept. 2008.

[17] Estimating Bandwidth Requirements and Connection Speed, http://technet.microsoft.com/en-us/library/cc785130(WS.10).aspx.

[18] SONET/SDH Tech. Summary, http://www.techfest.com/networking/wan/sonet.htm.

[19] IEEE 802.11: WIRELESS LOCAL AREA NETWORKS (LANs), http://standards.ieee.org/about/get/802/802.11.html.

[20] H. Schulzrinne, A. Rao, and R. Lanphier, "Real Time Streaming Protocol (RTSP)", RFC 2326, Internet Engineering Task Force, Apr. 1998, http://www.ietf.org/rfc/rfc2326.txt.

[21] L. D. Cicco, S. Mascolo, and V. Palmisano, "Skype Video Congestion Control: An Experimental Investigation," *Computer Networks*, vol. 55, no. 3, pp. 558–571, Feb. 2011.

[22] H. Chang, S. Jamin, and W. Wang, "Live Streaming with Receiver-based Peer-division Multiplexing," *IEEE/ACM Trans. on Networking*, vol. 19, no. 1, pp. 55–68, Feb. 2011.

[23] N. M. Markovich and U. R. Krieger, "Statistical Analysis and Modeling of Skype VoIP

Flows," *Computer Communications*, vol. 33, pp. 11–21, Nov. 2010.

[24] S. Guha, N. Daswani, and R. Jain, "An Experimental Study of the Skype Peer-to-Peer VoIP System," in *Proc. of the 5th Int. Workshop on Peer-to-Peer Systems (IPTPS '06)*, Santa Barbara, CA, pp. 1–6, Feb. 2006.

[25] M. Kalman, E. Steinbach, and B. Girod, "Rate-Distortion Optimized Video Streaming with Adaptive Playout," *IEEE Int. Conf. on Image Processing ICIP-2002,* vol. 3, pp. 189–192, Rochester, NY, Sept. 2002.

[26] P. A. Chou and Z. Miao, "Rate-distortion Optimized Sender-driven Streaming over Best-effort Networks," in *Proc. Workshop on Multimedia Signal Processing*, vol. 1, pp. 587–592, Oct. 2001.

[27] C. P. Ho and C. J. Tsai, "Content-adaptive Packetization and Streaming of Wavelet Video over IP Networks," Journal on Image and Video Processing, vol. 2007, Article ID 45201, Jan. 2007.

[28] L. Lamport, "Time, Clocks, and the Ordering of Events in a Distributed System," *Communications of the ACM*, vol. 21, no. 7, pp. 558–565, Jul. 1978.

[29] S. J. Choi and J. W. Woods, "Motion-compensated 3D Subband Coding of Video," *IEEE Trans. on Image Processing*, vol. 8, no. 2, pp. 155–167, 1999.

[30] J. Xu, Z. Xiong, S. Li, and Y. Q. Zhang, "Three-dimensional Embedded Subband Coding with Optimized Truncation (3-D ESCOT)," *Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 290–315, 2001.

[31] B. J. Kim, Z. Xiong, and W. A. Pearlman, "Low Bit-rate Scalable Video Coding with 3-D Set Partitioning in Hierarchical Trees (3-D SPIHT)," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, no. 8, pp. 1374–1387, 2000.

[32] J. W. Woods and G. Lilienfield, "A Resolution and Frame-Rate Scalable Subband/Wavelet Video Coder", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 9, pp. 1035–1044, Sept. 2001.

[33] R. Xiong, J. Xu, F. Wu, S. Li, and Y. Q. Zhang, "Layered Motion Estimation and Coding for Fully Scalable 3D Wavelet Video Coding," in *Proc. Int. Picture Coding Symp. PCS* '04, (San Francisco, CA,USA), Dec. 2004.

[34] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority Encoding Transmission," *IEEE Trans. on Information Theory*, vol. 42, no. 6, pp. 1737–1744, Nov. 1996.

[35] S. Lin and D. J. Costello, "Error Control Coding," Second Edition, Pearson Prentice Hall, 2004.

[36] Y. J. Liang, J. G. Apostolopoulos, and B. Girod, "Analysis of Packet Loss for Compressed Video: Effect of Burst Losses and Correlation Between Error Frames," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 18, no. 7, pp. 861–874, Jul. 2008.

[37] J. Chakareski, J. Apostolopoulos, S. Wee, W. Tan, and B. Girod, "Rate-Distortion Hint Tracks for Adaptive Video Streaming," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 15, no. 10, pp. 1257–1269, Oct. 2005.

[38] J. Chakareski and P. A. Chou, "Rate-distortion Optimized Streaming from the Edge of the Network," *IEEE/ACM Trans. on Networking*, Dec. 2006.

[39] P. Ramanathan, M. Kalman, and B. Girod, "Rate-Distortion Optimized Interactive Light Field Streaming," *IEEE Trans. on Multimedia*, vol. 9, no. 4, pp. 813–825, Jun. 2007.

[40] C. P. Ho, J. Y. Yu, and S. Y. Lee, "Efficient Data Replication for the Delivery of High-quality Video Content over P2P VoD Advertising Networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, Article ID 105, 2011.

[41] J. Kangasharju, K. W. Ross, and D. Turner, "Adaptive Content Management in Structured P2P Communities," *Infoscale*, 2006.

[42] S. Deering, "Host Extensions for IP Multicasting," RFC 1112, Aug. 1989, http://www.ietf.org/rfc/rfc1112.txt.

[43] H. Holbroo, and D. Cheriton, "IP Multicast Channels: EXPRESS Support for Large-scale Single-Source Applications," in *Proc. of SIGCOMM*, 1999.

[44] D. Kosiur, "IP Multicasting: The Complete Guide to Interactive Corporate Networks," *New York: John Wiley & Sons, Inc.*, 1998.

[45] BBC iPlayer, http://www.bbc.co.uk/iplayer/radio.

[46] Babelgum, http://www.babelgum.com.

[47] Jaman, http://www.jaman.com.

[48] GridNetworks, http://globalmediaservices.net.

[49] W. Fenner, "Internet Group Management Protocol," Version 2, RFC 2236, Nov. 1997, http://www.ietf.org/rfc/rfc2236.txt.

[50] B. Cain, S. Deering, I. Kouvelas, B. Fenner, and A. Thyagarajan, "Internet Group Management Protocol," Version 3, RFC 3376, Oct. 2002, http://www.ietf.org/rfc/rfc3376.txt.

[51] M. Ordukaya1 and H. A. Ilgin, "Video Multicasting in Campus Networks," *Int. Journal of Video & Image Processing and Network Security*, vol. 9, no. 10, pp. 55–61, Dec. 2009.

[52] B. Fenner, M. Handley, H. Holbrook, and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)," RFC 4601, Aug. 2006, http://www.ietf.org/rfc/rfc4601.txt.

[53] A. Adams, J. Nicholas, and W. Siadak, "Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)," RFC 3973, Jan. 2005, http://www.ietf.org/rfc/rfc3973.txt.

[54] N. Bonmariage and G. Leduc, "A Survey of Optimal Network Congestion Control for Unicast and Multicast Transmission," *Computer Networks*, vol. 50, no. 3, pp. 448–468, Feb. 2006.

[55] Y. H. Chu, S. G. Rao, S. Seshan, and H. Zhang, "A Case for End System Multicast," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 1456–1471, Oct. 2002.

[56] O. Abboud, K. Pussep, A. Kovacevic, K. Mohr, S. Kaune, and R. Steinmetz, "Enabling Resilient P2P Video Streaming: Survey and Analysis," *Multimedia Systems*, vol. 17, no. 3, pp. 177–197, 2011.

[57] E. Cohen and S. Shenker, "Replication Strategies in Unstructured Peer-to-Peer Networks," in *Proc. ACM SIGCOMM Conf.*, pp. 177–190, 2002.

[58] E. Leontiadis, V. V. Dimakopoulos, and E. Pitoura, "Creating and Maintaining

Replicas in Unstructured Peer-to-Peer Systems," in *Proc. of the 12th Int'l Euro-Par Conf. on Parallel Processing* (*EURO-PAR*). pp. 1015-1025, 2006.

[59]B. Cohen, "Incentives Build Robustness in BitTorrent," in *Proc. of the 1st Workshop on Economics of Peer-to-Peer Systems*, Jun. 2003.

[60]S. Ghandeharizadeh, B. Krishnamachari, and S. Song, "Placement of Continuous Media in Wireless Peer-to-Peer Networks," *IEEE Trans. on Multimedia*, vol. 6, no. 2, pp. 335–342, Apr. 2004.

[61]K. H. Wan and C. Loeser, "An Overlay Network for Replica Placement within a P2P VoD Network," *Int. Journal of High Performance Computing and Networking*, vol. 3, no. 5/6, pp. 320–335, Mar. 2005.

[62]W. F. Poon, J.-Y. B. Lee, and D. M. Chiu, "Comparison of Data Replication Strategies for Peer-to-Peer Video Streaming," in *Proc. of the 5th IEEE Int. Conf. on Information, Communications and Signal Processing*, 2005.

[63]C. Ye and D. Chiu, "Peer-to-Peer Replication with Preferences", in *Proc. of the 2nd Int. Conf. on Scalable Information Systems*, no. 4, 2007.

[64]ISO/IEC MPEG Test Group, "Subjective Test Results for the CfP on Scalable Video Coding Technology," *MPEG Documents N6383*, Mar. 2004.

[65]S. Brangoulo, R. Leonardi, M. Mrak, B. Pesquet Popescu, and J. Xu, "Draft Status Report on Wavelet Video Coding Exploration," *MPEG Documents N7571*, Oct. 2005.

[66]P. A. Chou and Z. Miao, "Rate-distortion Optimized Streaming of Packetized Media," *IEEE Trans. on Multimedia*, vol. 8, no. 2, pp. 390–404, 2006.

[67]A. K. Katsaggelos, Y. Eisenberg, F. Zhai, R. Berry, and T. N. Pappas, "Advances in Efficient Resource Allocation for Packet-based Real-time Video Transmission," in *Proc. of the IEEE*, vol. 93, no. 1, pp. 135–146, 2005.

[68]X. Zhu, E. Setton, and B. Girod, "Congestion-distortion Optimized Video Transmission over Ad Hoc Networks," *Signal Processing: Image Communication*, vol. 20, no. 8, pp. 773–783, 2005.

[69]F. Zhai, C. E. Luna, Y. Eisenberg, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Joint Source Coding and Packet Classification for Real-time Video Transmission over

Differentiated Services Networks," *IEEE Trans. on Multimedia*, vol. 7, no. 4, pp. 716–725, 2005.

[70] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1971.

[71] T. Chu and Z. Xiong, "Combined Wavelet Video Coding and Error Control for Internet Streaming and Multicast," *EURASIP Journal on Applied Signal Processing*, vol. 2003, no. 1, pp. 66–80, 2003.

[72] J. Dong and Y. F. Zheng, "Content-based Retransmission for 3-D Wavelet Video Streaming on the Internet," in *Proc. of IEEE Int. Conf. on Information Technology: Coding and Computing* (ITCC '02), pp. 452–457, Las Vegas, Nev, USA, Apr. 2002.

[73] Y. Zhao, S. C. Ahalt, and J. Dong, "Content-based Retransmission for a Video Streaming System with Error Concealment," in *Proc. of SPIE*, vol. 5438, pp. 63–70, Orlando, Fla., USA, Apr. 2004.

[74] W. T. Tan and A. Zakhor, "Real-time Internet Video using Error Resilient Scalable Compression and TCP-friendly Transport Protocol," *IEEE Trans. on Multimedia*, vol. 1, no. 2, pp. 172–186, 1999.

[75] J. C. Bolot and T. Turletti, "Experience with Control Mechanisms for Packet Video in the Internet," *Computer Communication Review*, vol. 28, no. 1, pp. 4–15, 1998.

[76] M. Kalman and B. Girod, "Techniques for Improved Rate-distortion Optimized Video Streaming," *ST Journal of Research*, vol. 2, no. 1, pp. 45–54, 2005.

[77] H. Wang, F. Zhai, Y. Eisenberg, and A. K. Katsaggelos, "Cost-distortion Optimized Unequal Error Protection for Object-based Video Communications," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 15, no. 12, pp. 1505–1516, 2005.

[78] C. L. Chang, S. Han, and B. Girod, "Sender-based Rate-distortion Optimized Streaming of 3-D Wavelet Video with Low Latency," in *Proc. of the 6th IEEE Workshop on Multimedia Signal Processing* (*MMSP'04*), pp. 510–513, 2004.

[79] C. L. Chang, S. Han, and B. Girod, "Rate-distortion Optimized Streaming for 3-D Wavelet Video," in *Proc. of IEEE Int. Conf. on Image Processing* (*ICIP '04*), vol. 5, pp. 3141–3144, Singapore, Oct. 2004.

[80]F. Zhai, Y. Eisenberg, C. E. Luna, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Packetization Schemes for Forward Error Correction in Internet Video Streaming," in *Proc. of the 41st Allerton Conf. Communication, Control and Computing*, Monticello, USA, Oct. 2003.

[81]E. Martinian and C.-E. W. Sundberg, "Decreasing Distortion using Low Delay Codes for Bursty Packet Loss Channels," *IEEE Trans. on Multimedia*, vol. 5, no. 3, pp. 285–292, 2003.

[82]K. Shimizu, N. Togawa, T. Ikenaga, and S. Goto, "Reconfigurable Adaptive FEC System based on Reed-Solomon Code with Interleaving," *IEICE Trans. on Information and Systems*, vol. E88-D, no. 7, pp. 1526–1537, 2005.

[83]V. Stankovic´, R. Hamzaoui, and Z. Xiong, "Efficient Channel Code Rate Selection Algorithms for Forward Error Correction of Packetized Multimedia Bitstreams in Varying Channels," *IEEE Trans. on Multimedia*, vol. 6, no. 2, pp. 240–248, 2004.

[84]M. Gallant and F. Kossentini, "Rate-distortion Optimized Layered Coding with Unequal Error Protection for Robust Internet Video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 357–372, 2001.

[85]J. Goshi, A. E. Mohr, R. E. Ladner, E. A. Riskin, and A. Lippman, "Unequal Loss Protection for H.263 Compressed Video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 15, no. 3, pp. 412–419, 2005.

[86]S. Dumitrescu, X. Wu, and Z. Wang, "Globally Optimal Uneven Error-protected Packetization of Scalable Code Streams," *IEEE Trans. on Multimedia*, vol. 6, no. 2, pp. 230–239, 2004.

[87]M. Zink, J. Schmitt, and R. Steinmetz, "Layer-encoded Video in Scalable Adaptive Streaming," *IEEE Trans. on Multimedia*, vol. 7, no. 1, pp. 75–84, 2005.

[88]ISO/IEC MPEG Video Group, "Wavelet Codec Reference Document and Software Manual v1.0," *MPEG Document N7573*, Jul. 2005.

[89]R. Fang, D. Schonfeld, R. Ansari, and J. Leigh, "Forward Error Correction for Multimedia and Teleimmersion Data Streams," *Tech. Rep.*, Electronic Visualization Laboratory, University of Illinois at Chicago, Chicago, USA, 2000.

[90]E. W. Biersack, “Performance Evaluation of Forward Error Correction in an ATM Environment,” *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 4, pp. 631–640, 1993.

[91]Y. H. Yu, C. P. Ho, and C. J. Tsai, “Multiple Adaptations and Content-Adaptive FEC Using Parameterized R-D Model for Embedded Wavelet Video," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 70914, 13 pages, 2007.

[92]J. Y. Lee and S. K. Park, “Optimum UDP Packet Sizes in Ad Hoc Networks,” *IEICE Transactions on Communications*, vol. E88-B, no. 2, pp. 815–820, 2005.

[93]B. Birney, “Reducing Broadcast Delay,” Microsoft Technical Report, Microsoft Corporation, Jun. 2006, http://www.microsoft.com/windows/windowsmedia/howto/articles/Broadcast-Delay.aspx#MinimizingDelay.

[94]ISO/IEC JTC 1/SC 29/WG11, *ISO/IEC TR21000-12: MPEG-21 Test Bed for Resource Delivery*, Jan. 2005.

[95]M. Carson and D.Santay, “NIST net: a Linux-based Network Emulation Tool,” *Computer Communication Review*, vol. 33, no. 3, pp.111–126, 2003.

[96]R. Xiong, X. Ji, J. Xu, and F. Wu, “MSRA Scheme for SVC CE1,” *MPEG Input Document M11320*, Palma de Mallorca, ES, Oct. 2004.

[97]J. M. Boyce and R. D. Gaglianello, “Packet Loss Effects on MPEG Video Sent over the Public Internet,” in *Proc. of the 6th ACM Int. Conf. on Multimedia* (*ACM Multimedia ’98*), pp. 181–190, Bristol, UK, Sept. 1998.

[98]K. Lai, M. Roussopoulos, D. Tang, X. Zhao, and M. Baker, “Experiences with a Mobile Testbed,” in *Proc. of the 2nd Int. Conf. on Worldwide Computing and its Applications* (*WWCA* ’98), vol. 1368 of *Lecture Notes in Computer Science*, pp.222–237, Tsukuba, Japan, Mar. 1998.

[99]R. Risueño, P. Cuenca, F. Delicado, L. Orozco-Barbosa, and A. Garrido, “On the Traffic Disruption Time and Packet Lost Rate during the Handover Mechanisms in Wireless Networks,” in *Proc. of the 18th Int. Conf. on Advanced Information Networking and Application* (*AINA’04*), vol. 2, pp. 351–354, Mar. 2004.

[100]W. Li, “Overview of Fine Granularity Scalability in MPEG-4 Video Standard,” *IEEE Trans. on Circuits and Systems for Video Technology,* vol. 11, no. 3, pp. 301–317, Mar. 2001.

[101]W. Li, F. Ling, and X. Chen, “Fine Granularity Scalability in MPEG–4 for Streaming Video,” in *Proc. of the Int. Symp. on Circuits and Systems* (*ISCAS*), pp. 299–302, May 2000.

[102]J. Zhou, H. Shao, C. Shen and M. T. Sun, "Multi-path Transport of FGS Video," *Packet Video Workshop*, 2003.

[103]Y. Charfi and R. Hamzaoui, "Packet Loss Protection of Scalable Video Bitstreams using Forward Error Correction and Feedback,” *Int. Symp. on Image and Signal Processing and Analysis*, Rome, Sept. 2003.

[104]K. W. Stuhlmueller, M. Link, B. Girod, and U. Horn, "Scalable Internet Video Streaming with Unequal Error Protection", *Packet Video Workshop*, New York, 26/27 Apr. 1999.

[105]L. Zhao, J. W. Kim, and C. C. Jay Kuo, "Constant Quality Rate Control for Streaming MPEG-4 FGS Video," *IEEE Int. Symp. on Circuits and Systems*, vol. 4, pp. 544–547, May 2002.

[106]X. M. Zhang, A. Vetro, Y. Q. Shi, and H. F. Sun, "Constant Quality Constrained Rate Allocation for FGS Video Coded Bitstreams, " *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 2, pp. 121–130, Feb. 2003.

[107]D. S. Evans, “The Online Advertising Industry: Economics, Evolution, and Privacy,” *The Journal of Economic Perspectives*, vol. 23, no. 3, pp. 37–60, 2009.

[108]M. Brettel and A. Spilker-Attig, “Online Advertising Effectiveness: a Cross-cultural Comparison,” *Journal of Research in Interactive Marketing*, vol. 4, no. 3, pp. 176–196, 2010.

[109]Blinkx BBTV, http://www.blinkx.com/.

[110]Joost, http://www.joost.com/.

[111]Livestation, http://www.livestation.com/.

[112]J. F. Buford and H. Yu, “Peer-to-Peer Networking and Applications: Synopsis and

Research Directions," *Handbook of Peer-to-Peer Networking*, Springer US, pp. 3–45, 2010.

[113]A. Bikfalvi, J. Garcia-Reinoso, I. Vidal, F. Valera, and A. Azcorra , "P2P vs. IP Multicast: Comparing Approaches to IPTV Streaming based on TV Channel Popularity," *Computer Networks*, vol. 55, no. 6, pp. 1310–1325, 2011.

[114]D. Gomes, R. L. Aguiar, and S. Sargento, "A Cross-System Approach for Multimedia Services with IP Multicast in 4G Networks", *Wireless Personal Communications*, vol. 52, no. 3, pp. 651–668, 2010.

[115]B. Cheng, L. Stein, H. Jin, Z. Zhang, and X. Liao, "GridCast: Improving Peer Sharing for P2P VoD," *ACM Trans. on Multimedia Computing, Communications, and Applications*, vol. 4, no. 4, Oct. 2008.

[116]K. Mokhtarian and M. Hefeeda, "Analysis of Peer-assisted Video-on-Demand Systems with Scalable Video Streams, " in *Proc. of ACM Multimedia Systems* (MMSys'10), pp. 133–143, Scottsdale, AZ, Feb. 2010.

[117]Y. He and Y. Liu, "VOVO: VCR-Oriented Video-on-demand in Large-Scale Peer-to-peer Networks," *IEEE Trans. on Parallel and Distributed Systems*, vol. 20, no. 4, pp. 528–539, Apr. 2009.

[118]U. Abbasi and T. Ahmed, "Architecture for Cooperative Prefetching in P2P Video-on-demand System," *Int. Journal of Computer Networks & Communications*, vol. 2, no. 3, pp. 126–138, 2010.

[119]G. Raczkowski, "Mobile Commerce: Focusing on the Future, a Special to Dash 30," White Paper, 2002.

[120]ZapShares, http://www.zapshares.com/.

[121]MediaDefender, http://www.mediadefender.com/.

[122]P2Pads, http://www.p2pads.com/.

[123]P2Pwords, http://www.brandassetdigital.com.

[124]M. Bender, S. Michel, G. Weikum, and C. Zimmer, "The MINERVA Project: Database Selection in the Context of P2P Search," in *Proc. BTW*, pp. 125–144, 2005.

[125]Mininova, http://www.mininova.org/.

[126]I. J. Taylor, "From P2P to Web Services and Grids: Peers in a Client/Server World," Springer, pp. 23–41, 2004.

[127]M. Yang and Y. Yang, "An Efficient Hybrid Peer-to-Peer System for Distributed Data Sharing," *IEEE Trans. on Computers*, vol. 59, no. 9, pp. 1158–1171, Sept. 2010.

[128]A. Goldfarb and C. Tucker, "Online Display Advertising: Targeting and Obtrusiveness," *Marketing Science*, published online before print, Feb. 9, 2011, DOI: doi:10.1287/mksc.1100.0583.

[129]N. Salamanos, E. Alexogianni, and M. Vazirgiannis, "Ad-Share: An Advertising Method in P2P Systems based on Reputation Management," *8th Hellenic-European Conf. on Computer Mathematics and its Applications* (HERCMA), 2007.

[130]S. Saroiu, P. K. Gummadi, and S. D. Gribble, "Measuring and Analyzing the Characteristics of Napster and Gnutella Hosts," *Multimedia Systems*, vol. 9, no. 2, pp. 170–184, 2003.

[131]D. Hughes, G. Coulson, and J. Walkerdine, "Free Riding on Gnutella Revisited: the Bell Tolls?," *IEEE Distributed Systems Online*, vol. 6, no. 6, 2005.

[132]Z. Chen, H. T. Shen, Q. Xu, and X. Zhou, "Instant Advertising in Mobile Peer-to-Peer Networks," *IEEE Int. Conf. on Data Engineering*, pp. 736–747, 2009.

[133]J. D. Boever and D. D. Grooff, "Peer-to-Peer Content Distribution and Over-the-top TV: An Analysis of Value Networks," *Handbook of Peer-to-Peer Networking*, Springer, Part 8, pp. 961–983, 2010.

[134]T. M. Chung, S. C. Huang, C. T. King, and C. P. Chang, "Optimising Upload Bandwidth for Quality of VCR Operations in P2P VoD Systems," *Int. Journal Ad Hoc and Ubiquitous Computing*, vol. 5, no. 4, pp. 201–208, May 2010.

[135]C. P. Ho, S. Y. Lee, and J. Y. Yu, "Deploying an Agent-driven Advertising Service on P2P Video-on-demand System," *2010 Int. Conf. on Information and Multimedia Technology* (ICIMT 2010), vol. 1, pp. 382–387, Dec. 2010.

[136]A. P. Couto da Silva, E. Leonardi, M. Mellia, and M. Meo, "Chunk Distribution in Mesh-Based Large-Scale P2P Streaming Systems: A Fluid Approach," *IEEE Trans. on Parallel and Distributed Systems*, vol. 22, no. 3, pp. 451–463, Mar. 2011.

[137]C. Li and C. Chen, “Measurement-based Study on the Relations between Users' Watching behavior and Network Sharing in P2P VoD Systems,” *Computer Networks*, vol. 54, no. 1, pp. 13–27, Jan. 2010.

[138]O. A. Lotfallah, M. Reisslein, and S. Panchanathan, “A Framework for Advanced Video Traces: Evaluating Visual Quality for Video Transmission over Lossy Networks,” *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 42083, 21 pages, 2006.

[139]V. Zlokolica, A. Pizurica, and W. Philips, “Wavelet-Domain Video Denoising based on Reliability Measures,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 16, no. 8, pp. 993–1007, Aug. 2006.

[140]A. Mukherjee, “On the Dynamics and Significance of Low Frequency Components of Internet Load,” *Internetworking: Res. Experience*, vol. 5, pp. 163–205, Dec. 1994.

[141]Z. Li and T. Herfet, “MAC Layer Multicast Error Control for IPTV in Wireless LANs,” *IEEE. Trans. on Broadcasting*, vol. 55, no. 2, pp. 353–362, Jun. 2009.

[142]H. Schwarz, D. Marpe, and T. Wiegand, “Overview of the Scalable Video Coding Extension of the H.264/AVC Standard,” *IEEE Trans. on Circuit and System Video Technology*, vol. 17, no. 9, pp. 1103–1120, Sept. 2007.

[143]V. Ponomaryov and E. Ramos, “Real Time Generation of 3D Video Sequences via multi Wavelet Multilevel Technique,” *2nd Int. Conf. on Mathematics and Information Science*, 2011.

[144]J. Garbas, B. Pesquet-Popescu, and A. Kaup, “Methods and Tools for Wavelet-Based Scalable Multiview Video Coding,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 21, no. 2, pp. 113–126, Feb. 2011.

[145]N. Adami, A. Signoroni, and R. Leonardi, “State-of-the-art and Trends in Scalable Video Compression with Wavelet-Based Approaches,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1238–1255, Sept. 2007.

[146]C. Y. Tsai and H. M. Hang, “A Rate-Distortion Analysis on Motion Prediction Efficiency and Mode Decision for Scalable Wavelet Video Coding,” *Journal of Visual Communication and Image Representation*, vol. 21, no. 8, Nov. 2010.

[147]C. E. Shannon, “A Mathematical Theory of Communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, Jul. and Oct. 1948.

[148]J. Chakareski and P. A. Chou, “RaDiO Edge: Rate-Distortion Optimized Proxy-Driven Streaming from the Network Edge,” *IEEE Trans. on Networking*, vol. 14, no. 6, pp. 1302–1312, Dec. 2006.

[149]J. Chakareski and P. Frossard, “Distributed Collaboration for Enhanced Sender-Driven Video Streaming,” *IEEE Trans. on Multimedia*, vol.10, no. 5, pp. 858–870, 2008.

[150]R. E. Bruner and J. Singh, “Video Ad Benchmarks: Average Campaign Performance Metrics,” *a DoubleClick research report*, Feb. 2007.

[151]R. Xiong, X. Ji, D. Zhang, and J. Xu, “Vidwav Wavelet Video Coding Specifications,” ISO/IEC JTC1/SC29/WG11 MPEG, M12339, 2005.

[152]C. Huang, J. Li, and K.W. Ross, “Can Internet Video-on-Demand be Profitable? ,” in *Proc. ACM SIGCOMM*, Koyto, Japan, pp. 133–144, Aug. 2007.

[153]Z. Liu, Y. Shen, K. W. Ross, S. S. Panwar, and Y. Wang, “Substream Trading: Towards an Open P2P Live Streaming System,” in *Proc. IEEE Int. Conf. on Network Protocols*, pp. 94–103, Oct. 2008.

[154]“University of Oregon Route Views Project,” http://www.routeviews.org/.

[155]G. P. Jesi, A. Montresor, and O. Babaoglu, “Proximity-Aware Superpeer Overlay Topologies,” *IEEE Trans. on Network and Service Management*, vol. 4, no. 2, pp. 74–83, Sept. 2007.

[156]H. L. Yu, D. D. Zheng, B. Y. Zhao, and W. Zheng, “Understanding User Behavior in Large-Scale Video-on-demand Systems,” in *Proc. EuroSys*, vol. 40, no. 4, pp. 333–344, 2006.

Publication List

Journal Papers

1. Chien-Peng Ho, Jen-Yu Yu, and Suh-Yin Lee, "Efficient Data Replication for the Delivery of High-quality Video Content over P2P VoD Advertising Networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 105, 2011. [SCI/EI, Impact Factor: 1.012]
2. Ya-Huei Yu, Chien-Peng Ho, and Chun-Jen Tsai, "Multiple Adaptations and Content-Adaptive FEC Using Parameterized R-D Model for Embedded Wavelet Video," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 70914, 13 pages, 2007. [SCI/EI, Impact Factor: 1.012]
3. Chien-Peng Ho and Chun-Jen Tsai, "Content-Adaptive Packetization and Streaming of Wavelet Video over IP Networks," *EURASIP Journal on Image and Video Processing*, vol. 2007, Article ID 45201, 12 pages, 2007. [SCI/EI, Impact Factor: 0.619]
4. Chien-Peng Ho, Chun-Jen Tsai, and Yuh-Feng Hsu, "MPEG-21 Digital Item Adaptation Architecture for Fully Scalable Video Streaming," *CCL Technical Journal*, Vol. 109, pp. 55- 64, 2004

Conference Papers

1. Li-Wu Tsai, Yee-Choy Chean, Chien-Peng Ho, Hui-Zhen Gu, and Suh-Yin Lee, "Multi-Lane Detection and Road Traffic Congestion Classification for Intelligent Transportation System," 2011 International Conference on Machine Learning and Computing (ICMLC 2011), Vol. 4, pp. 477-481, Singapore, February 26-28, 2011. [EI, INSPEC, and Thomson ISI (ISTP)]
2. Chien-Peng Ho, Suh-Yin Lee, and Jen-Yu Yu, "Deploying an Agent-driven Advertising Service on P2P Video-on-Demand System," *2010 International Conference on Information and Multimedia Technology* (ICIMT 2010), Hong Kong,

China, December 28-30, 2010. [EI, INSPEC, and Thomson ISI (ISTP)]

3. Chien-Peng Ho, Suh-Yin Lee, and Jen-Yu Yu, "Cluster-based Replication for P2P-based Video-on-Demand Service," *2010 International Conference on Electronics and Information Engineering* (ICEIE 2010), Vol. 2, pp. 49-53, Kyoto, Japan, August 1-3, 2010. [EI and Thomson ISI]

4. Chia-Ling Hsieh and Chien-Peng Ho, "The Web as a Corpus of Pragmatic Studies," *2009 International Conference on Multi Development and Application of Language and Linguistics* (MDALL-2), pp. 1-17, Taiwan: National Cheng Kung University, May 15-16, 2009.

5. Chia-Ling Hsieh and Chien-Peng Ho, "Corpus Application in Chinese Pragmatic Analysis," *The Proceedings of the 2009 International Conference on Applied Linguistics and Language Teaching* (ALLT-2009), pp. 693-702, Taiwan, April 16-18, 2009.

6. Chien-Peng Ho and Chun-Jen Tsai, "Content-Adaptive Packetization and 3-D Wavelet Video Streaming over IP Networks," *Workshop on Consumer Electronics and Signal Processing* (WCEsp), Taiwan, November 16, 2006.

7. Chien-Peng Ho and Chun-Jen Tsai, "Rate-Distortion Optimized Video Streaming System with Adaptive Interleaved Forward Error Correction," *Workshop on Consumer Electronics and Signal Processing* (WCEsp), Taiwan, November 17-18, 2005.

8. Chien-Peng Ho, Wen-Chieh Chang, Kuo-Cheng Lee, and Chun-Jen Tsai, "Rate-Distortion Optimized Video Streaming with Smooth Quality Constraint," *IEEE International Symposium on Circuits and Systems* (ISCAS 2005), Vol. 4, pp. 3271-3274, Kobe, Japan, May 23-26, 2005.

9. Chien-Peng Ho and Chao-Min Su, "A Chinese Named Entity Identification System for Web Retrieval," *Proceedings of Mechatronic and Industry Interact Cross Strait Conference*, Taiwan, November 23, 2005.

10. Chien-Peng Ho, "Next Steps for 300mm Transportation Manager System," TSMC e-Operations Symposium 2002, Taiwan Semiconductor Manufacturing Company Limited, HsinChu, Taiwan, June 2002.

11. Chien-Peng Ho, "An Effective Software Quality Control for 300mm Material Control System," TSMC e-Operations Symposium 2001,Taiwan Semiconductor Manufacturing Company Limited, HsinChu, Taiwan, November 2001.

MPEG Standard Contributions

1. F.-C. Chang, C.-P. Ho, C.-Y. Tsai, C.-H. Li, J.-D. Cheng, W.-C. Chang, Y.-T. Shih, C.-L. Lin, C.-H. Lu, C.-C. Cheng, C.-Y. Liu, J.-C. Ma, K.-C. Lee, and C.-J. Tsai, "ISO/IEC JTC1/SC29/WG11 M12373: Update to the FGS-Based Multimedia Resource Delivery Test Bed Software," July 2005 (73rd, Poznan, Poland).
2. C.-N. Wang, C.-H. Li, C.-W. Fan, Y.-T. Shih, C.-P. Ho, C.-L. Lin, F.-C. Chang, G.-C. Li, C.-N. Chiu, C.-Y. Tsai, H.-C. Chuang, J.-H. Chen, J.-C. Ma, Chi-Yu Liu, C.-C. Chen, C.-J. Tsai, Tihao Chiang, S.-Y. Lee, and H.-M. Hang, "ISO/IEC JTC1/SC29/WG11 M11117: Scalable Multimedia Streaming Test Bed for Media Coding and Testing in Streaming Environments," July 2004 (69th, Redmond, Washington, USA).

Patent Activities

1. Yao-Hsiung Chang, Chien-Peng Ho, and Jau-Huang Chen, "System and Method for Automated Dispatch and Transportation of Work-In-Process," U.S. Patent, Pub. No.: US 2004/0267641 A1, Dec. 30, 2004.
2. 張耀雄, 何健鵬, 陳昭宏, "在製品自動分派運輸系統及方法以及電腦可讀取儲存媒體," I246108，中華民國專利，2005/12/21 (專利權: 20051221-20240321)。