MDC ADAPTIVE VIDEO STREAMING

By

Zi Ling

B.Sc., Wuhan University, China, 1997

A thesis
presented to Ryerson University
in partial fulfillment of the
requirement for the degree of
Master of Applied Science
in the Program of
Electrical and Computer Engineering

Toronto, Ontario, Canada, 2007

Author's Declaration

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Author's signature:_______ ____________________

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Author's signature:____ ________________________

Borrower's Page

Ryerson University requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Name	Signature	Address	Date

Abstract

MDC ADAPTIVE VIDEO STREAMING

© Zi Ling 2007
Master of Applied Science
Department of Electronic and Computer Engineering,
Ryerson University

Multiple Description Coding (MDC) is designed for multiple path video streaming with channel diversities. In this thesis, we investigate the performance of multi-path video streaming using the MDC technique. The MDC frame loss rate is one of the indicators of the real time video quality. A classification based framework for making mode decisions to minimize the MDC video frame transmission cost that may be defined in terms of the six parameters, number of sub-streams, number of transmission channels, GOP length, the I-frame positions, probability of network transmission states and probability of transmission changes.

This thesis surveys the current status of horizontal decomposition into distributed computation, and vertical decomposition into functional modules such as congestion control, routing, scheduling, random access, and video coding. The focus of this thesis is on the video adaptive coding process to improve performance in terms of one or more of these factors. How to deliver a real-time MDC video from an end user over multi-channels is studied. The traffic is used to probe the network on determining the network conditions and optimizing the coding algorithms appropriately. An efficient transmission statistical model Auto Regression (AR) to capture the properties of the region of interest is also introduced. Both the mode decisions and the error concealment require feedback from the network regarding the available bandwidth, loss probability, video coding methods and coding time spatial manners. The proposed algorithm works in a fully distributed environment, making it suitable for wireless ad hoc networks or other IP networks.

Acknowledgment

I would like to start by thanking my advisor Prof. Ivan Lee who has been a constant source of inspiration through the course. I am very grateful to him for his continuous support, encouragement, and great enthusiasm. His insights have been very valuable in enabling my exploration of new research directions as well helping me overcome the many problems that we encountered. I have also benefited vastly from his emphasis on improving both academic and industry knowledge. He has been a great teacher, motivator and research mentor. I have learnt many valuable lessons through our interaction and hope to carry forward his high standards through my career and life. I am fortunate to have had him as my advisor.

My lab-mates Yifeng He, Xiaoming Fan, Yun Tie, William Shaw have been of great help during the entire course of this Master. They have not only contributed to my research through many suggestions and discussions, but have also enriched my interests through their personal interaction.

I am grateful for the friendship of faculties and schoolmates of Ryerson Multimedia Research Lab and Ryerson Electrical and Computer Engineering Department. I would also appreciate the valuable comments.

Much appreciation goes to my family, my friends, especially my parents for their kindness and support through this and all phases in my life. I sincerely appreciate everything they have done for me and many thanks for believing in and encouraging me to reach my goals.

Publication List

CONFERENCE PAPERS

1. Zi Ling and I. Lee, "Adaptive Multi-Path Video Streaming", *Eighth IEEE International Symposium on Multimedia (ISM'06)*, CA, U.S., pp. 399 - 406, Dec 2006.

Table of Contents

Figure 4.13 Video average PSNR (dB) of the different GOP length and sub-stream63

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Aims and Motivation

An Internet based electronic efficient and effective distribution of a high quality video offers many advantages to the content publisher and customers. The digital delivery reduces the operation cost, such as transportation, storage and selling. It creates more potential customers to the distribution companies and provides more choices for the end users at the meantime. Each user can purchase his desired media and play it back instantly. The explosive growth of the Internet backbone on the computer infrastructure and architectures offers the opportunities for improving the productive solution of this service delivery to fulfill the user requirements.

Recent applications such as the traffic, health and industrial environments monitoring, live streaming, GIS information exchanging and distributed video conferencing, have stimulated research for technologies in the areas of multimedia communications. The multimedia content in general demands a much higher transmission bandwidth than hypertext content. With the advances in consumer high-tech electronic products, the affordable prices of the multimedia gadgets make the data exchange inadequate due to the client-server transmission scalability. The decentralized approach, such as Peer-to-Peer (P2P) network technique [1] has widely been applied to extent the client-server transmission bandwidth with all the active resources in the peer networks. As well, the distributed networks as a hybrid with centrally managed manners provide the available virtual transmission bandwidth to the multimedia deliveries.

The increasing multimedia data bit-rate requires better transmission, resource allocation schemes, and compression to ensure the data delivery, especially for the real-time applications such as video streaming and video conferencing. A high quality video real-time delivery via the complex Internet architectures may be the desirable for the most end users. The video quality would be a key factor to evaluate the overall delivery results.

The all aspects of the delivery channel resources will affect the qualities of those received videos at real time. The changing Internet environment requires the robust transmission and data compression manner to achieve a sound delivered multimedia content at any single point of transmission period.

We investigated the challenges of video streaming and content delivery qualities. Researches of applying the adaptive resource allocation robust into the distributed network infrastructure, and analysis of the improved performance on video delivery are the main objectives of this thesis.

1.2 Video Streaming Over The Distributed Network

Video streaming is the method of video delivery to the end users who are playing back the content at real-time over the networks, such as Internet, intranet or wireless network. The video contents can either be pre-coded or living streamed. The major challenge posed in video streaming is how to fulfill the real-time requirement. The information delivery task has the strictly time efficient and effective manners. The overall performance of streaming is associated with the successful delivery within the required period of time. Either the network packets failed passing through, or those transferring does not meet the required time, the streaming transmission tasks are regarded as failure. This is significantly different from the play-after-download approach where transmission delay is not a factor of error, and the video is playing back with the well prepared and all packets received contents. In today's Internet environment, there are more challenges to achieve the best-effort nature for playing back the streaming video. The physical network channels are consisted of various network types and protocols depending on the real-time participating channels and are hard to control due to the loose end-user management. The overall transmission bandwidths keep changing at each single point of streaming time and the packet delivery is not guaranteed in general. Numerous techniques in transmission feedback control, adaptive source encoding algorithm, efficient packet sizing, optimized resource allocation, and error control coding have been proposed to improve the quality of the video communication [2] [3].

The client-server network is the fundamental architecture for video streaming, because historically clients do not provide sufficient bandwidth or computational power to forward requests from neighborhood peers. Conventional video streaming uses the client-server infrastructure with uni-casting transmission. To reduce the amount of duplicated data transmission with uni-cast, multicast technique was introduced to broadcast the data destination only to the subscribers. Multicast routing and multicast delivery have evolved from laboratory research [4] to practical applications such as MBONE [5]. Multicast requires protocol support on the routers, the cost of deploying the multicast application over the Internet backbone represents the major drawback.

In a client-server environment, the clients' primary applications and files are stored in a common location. Video contents are often set up so that each user on the network has access to his or her "own" subscribed contents once the access token is retrieved, along with a range of "public" information where applications are stored. If the two clients above want to communicate with each other, they must go through the central server to do it. A message from one client to another is first sent to the file server, where it is then routed to its destination. With tens or hundreds of clients, the central server is the only way to manage the often complex and simultaneous operations that large networks require. The client-server architecture suffers from the several bottlenecks:

- Inefficient network utilization: With the explosive growth of computer computation power, each video streaming peer has the high computational power and high speed network connections. The client-server framework has the central processing manner, and will under-use the overall resources.

- Scalability: Despite of the various techniques of broadcasting with multicasting to use the available network channel bandwidth, once the streaming video is consist of the non-synchronized data, the central managed server need to use different manner to transferring the encoded contents, especially, in the Internet environment, there are massive number of available interactive media to end-users.

- Vulnerability: The centralized streaming servers are the only means for data distribution, no matter they are cluster server to help improving the reliability or multiple standalone servers function with the culminated bandwidth, the centralized model is vulnerable to large scale Denial of Services (DoS) [6] or Distributed DoS attacks [7].

The peer-to-peer computer network exploits diverse connectivity between participants in a network and the cumulative bandwidth of network participants rather than conventional centralized resources where a relatively low number of servers provide the core value to a service or application. A peer-to-peer network does not have the notion of clients or servers, but only equal peer nodes that simultaneously function as both "clients" and "servers" to the other nodes on the network. This model of network arrangement differs from the client-server model where communication is usually to and from a central server. The peer to peer networks are mainly characterized by there behaviors: self-organization, symmetric communication and distributed control [8]. P2P technologies have been successfully applied in the recent years for applications such as instant messaging service (IMS) [9], file sharing [10], voice and video conferencing [11] [12].

The P2P content distribution system presents a feasible alternative to the client-server architecture, for the feature of offloading the traffic from the bottleneck link to the under-utilized channels. The peers in the P2P network function as both client and server, thus, each peer can work independently and the need of a centralized server can be eliminated. Although the P2P architecture has the potential to overcome some limitations of the client-server architecture, the P2P system possesses other constraints by themselves characters.

- Heterogeneous: Different peer has physically different bandwidth and processing capacity. The peer's transmission condition is independent and is various upon the transferring rate over the peer itself and keeps changing at each slot of service time.

- Unpredictable: Each peer is free to join and leave the service infrastructure. There is no reinforcement to guarantee the service level from the peers.

The performance of video streaming and the service level will be affected by these constraints, special designs and considerations are required to function as a relatively measured product. The distributed video streaming proposed in this thesis addresses some of these constraints while improving the streaming quality and promoted service level.

1.3 MDC Adaptive Coding With Multi-Channels

There are many challenges in order to play back real-time video over today's Internet. Reconstructed video content at the client will be regarded lost due to network congestion or an unacceptable response time that exceed the playback threshold. Multiple Description Coding (MDC) has been proposed for use in packet audio and video transmission systems to combat both packet losses and component failures in a variety of application scenarios [13]. A popular approach for video streaming over IP network is the simultaneous streaming from multiple senders. This approach yields a higher throughput and a better tolerance to loss and delay due to network congestions [14]. Others suggest that multiple descriptions are stripped across multiple packets, and transmitted to a number of clients using IP multicast, thereby easing the loss of packets due to congestions [15]. The MDC descriptions serve the property that a baseline quality video can be decoded upon receiving any one description, and an improved quality video can be decoded upon receiving more descriptions. MDC, therefore, becomes a popular technique for real time application as it provides acceptable video quality without the need for a retransmission when degradations occur [16]. Some studies have also investigated frame dropout rate of simultaneous multiple streamed video over multiple independent paths [17], and over a hybrid streaming infrastructure with a centralized server and P2P networks [18]. These works assume that multi-path networks possess identical network property for loss pattern and delay characteristics.

Due to the heterogeneous nature of today's Internet, data transmission over a massive number of channels is difficult to control by a centrally managed solution at a low cost. Multimedia content in general has a highly time varying bandwidth requirement since media data are variable bit rates (VBR) in nature using modern coding techniques [19] [20]. In addition, streaming applications demands guaranteed delivery in order to meet specified temporal and spatial constrains. The services which provide guaranteed delivery in Network Multimedia System (NMS) was improved in the control management level of the host and the underlying network architectures [21]. Physically copying data is expensive for the guaranteed services for network transmission. The integrated processing loops for performing manipulation functions over a single common unit instead of performing them serially with the concept of Application Level Framing [22]. Media synchronization is also need to guarantee jitter-free playback requirements [23] [24]. The high bandwidth requirement and a real time delivery constraint are the two major challenges for streaming video. Driven by the goal of improving long-term system performance, dynamic resource allocation schemes with application specific adaptation capabilities are integrated in the solution. Previous work enables complex adaptations by use of general models of target systems [25]. With large variations in bandwidth requirements of multimedia content, rate adaptation [26] adjusts the bandwidth used by the transmission channel according to the existing network conditions.

A single level quality streaming using multiple transmission channels can be extended to use layered media stream according to the bandwidth heterogeneity for different receivers in real-time video transmission [27] [28]. On the Internet, the bandwidth of each participated peers keep changing in real-time, and the transmission is the mixed up manner within all the transmission channels. The pre-organized and processed the video contents need to be recognized at the receiver end to figure out both the sequence and layer information. Alternatively, a layered video codec can combine with MDC associated with the different service quality requirements.

1.4 MDC Streaming Adaptive to Network Traffic

Multimedia transmission is new and rapidly growing field which is concerned with all aspects of processing and manipulating multimedia data for transmission and storage. Fundamental issues in this area include data compression and coding, preprocessing (such as pre-filtering), interaction with physical transmission storage elements, and post-processing such as voice or video restoration. There are many applications which range from video-on-demand consumer services to ad hoc network design to data storage, transmission, and access [29]. Optimization technique has been widely used in the multi-path routing. People uses genetic algorithm (GA) to compute two optimal paths by minimizing the expected distortion [30]. The video is encoded into multiple descriptions, each transporting over an independent path. However, GA is computationally intensive and essentially centralized. Distributed algorithm is desired to compute the routing in wireless ad hoc networks. However, the link quality (e.g. failure probability due to fading) has not yet been considered in [27]. Another issue is the scalability of the source rate. Most of the exist works [31] [32] [33] allocate the rate for each path in advance, a form of resource reservation, which may not be optimal when the path conditions change.

For video streaming, we examined the transmission status to reflect the video streaming quality for the real-time requirements. The network traffic has the history free character, but it will affect the streaming quality instantly. The satisfactions of the end users need the pro-active streaming adjustment according to those changing transmission conditions. We proposed an optimized multi-path routing scheme for video streaming over the IP networks. This algorithm jointly optimizes the source rate and the routing scheme. It works in a fully decentralized manner, which is extremely suitable for wireless ad hoc networks where the power and computation capacity of each node is limited. We use a scalable source coding called forward error correction (FEC) based Multiple Description coding (MDC) proposed in [34]. In multi-path transport, there are risks that duplicated packets may be delivered to the same destination node. In this project, we assume such delivery redundancy problem is solved by applying a certain coding scheme, e.g., network coding [35], and there is no need to reconcile the content difference during transport. The linear program can be solved using some general algorithms, such as the simplex, ellipsoid and interior point methods [36].

The growing popularity of the wireless technology enriches the multimedia streaming applications. Wireless communication suffers from unreliable and time-varying channels, caused by fading and interferences. These severely impact the real-time video streaming due to the channel loss or instability. Many approaches have been investigated previously for the error resilience [37]and error concealment [38]. Various channel-adaptive video streaming schemes have been investigated in [39] [40] , for energy efficient wireless video streaming in [41], channel adaptive scalable video streaming was investigated over the 3G network [42]. Splitting the multimedia content and applying unequal error protection (UEP) [43] is another popular approach to improve the quality of video streaming. UEP is applied to multiple scalable image codec to achieve the graceful quality degradation [44], and the end-to-end streaming architecture with UEP and Automated Repeat Request (ARQ) was proposed in [45].

Many techniques have been applied over the TCP friendly congestion control for adaptive streaming. The roles of their algorithms are to control adding and removing quality layers, where the control decisions are based on a rate-driven feedback. The designs of their controls are based on analysis of additive-increase multiplicative-decrease (AIMD) congestion control and an assumption of a priori knowledge of video rate requirements [46]. Feamster et al extends this work to more general congestion control mechanisms [47]. Recently, Dai et al. [48] have presented an approach that proposes to integrate scaleable compression (MPEG-4 FGS) with adaptive streaming, using alternative congestion controls based on Kelly controllers. In contrast to these systems that explicitly attempt to match rates, Feng et al describes an adaptive streaming algorithm that uses a sliding window over video frames, sending data from low to high quality, in best effort fashion [49], this algorithm has the advantage of working without direct assumptions about the design of the underlying congestion control. Kang et al. [50] propose a priority-driven adaptation, but assuming fixed bandwidth channels.

The MDC characters has the all means to fill the above adaptive adjustment requirements. The number of sub-streams can be a good object related to the network transmission condition for the channel availability and the real-time transmission states can be the indicator of the prioritized transferring of different video content quality level (proposed in layered coding). The encoded video consists with GOP, and the length of MDC GOP is a keep factor associated with the probability of successful transmission and the picture quality. Those combination of the adjustments applied in MDC, will further smooth the real-time video streaming. At the mean time, we are motivated by the benefit optimization over the computation and routing techniques for the maximum network utilization and full utilizing those under utilized resources distributed in the multi-channel communications without extra hardware cost. However, these algorithms are inherently centralized and costly. We use the Auto Regression (AR) to perceive the transmission status changing automatically and to trigger the coding adaptation depending on the prediction of the upcoming transmission status. This may reduce the computation and is a well distributed manner. The streaming frame dropout rate is used to evaluate the service quality with rate explicitly.

1.5 Key Contributions

In this thesis I investigated the multimedia applications over the distributed networks with the main focus on the video streaming. Streaming techniques over different infrastructures are introduced. The main contributions of this thesis are:

- Introducing the MDC frame dropout rate statistic models that are flexible enough to capture the characteristics of the video traffic. We describe models with H.264 baseline standard MDC I and P frames that are doubly Markov in nature, to account for trace having different activity levels as well as different types of frames.

- Adjusting the MDC codec parameters to best fit into the network transmission channels and transferring rate to full use the under utilized network resources. Dynamic resource allocation is being applied according to the transmission capacity and better video quality was retrieved.

- Introducing an efficient MDC algorithm with spatial diversities, using cubic-spline interpolation for loss description recovery.

- Using autoregressive (AR) processes to capture the temporal correlations between successive frames based on network transmission status for the adaptive MDC coding adjustment associated with the real network transferring conditions.

We also studied the frame dropout rate and PSNR video quality over the MDC constructed video contents for the acceptable streaming performance and adaptive transmission during the transmission time.

1.6 Organization of this Thesis

The main objective of this thesis is to propose an adaptive video streaming system depending on the statistic modeling for distributed network video MDC coding optimization and error concealment. Following the introduction, the thesis is composed of five chapters and the remained four chapters are organized as follows:

Chapter 2 introduces the background knowledge of video coding and streaming. The fundamentals of video coding algorithms are described. These algorithms are used across many of the prevalent video coding standards. We then build a block diagram of the encoding and decoding process.

Chapter 3 examines the key MDC coding six parameters and streaming qualities. The experiments are implemented under the considerations of both transmission and visual qualities.

Chapter 4 proposes a decision model for MDC streaming adjustment, and compared the adjustment priorities of number sub-stream, length of GOP and GOP over multiple channel transmission diversity re-arrangement, and network transmission status.

Chapter 5 provides conclusions and directions for the future works. A summary of this thesis is presented as well.

CHAPTER 2

VIDEO CODING AND STREAMING CONCEPTS

Video coding is the process of compressing and decompressing a digital video signal, thus those contents can fit into the network bandwidth for streaming. This chapter introduces the frameworks in which this thesis may be included. We introduce the fundamentals of video coding, then the video encoder and decoder. We also briefly introduce the prevalent video coding standards, such as H264, etc. We compared client-server, centralized Peer-to-Peer and decentralized Peer-to-Peer architectures which advance the streaming techniques are the assumed the network infrastructures in this thesis.

2.1 Video Sequence Redundancy

Video compression refers to compact or condense a digital video sequence into small packets which can fit into the network bandwidth for streaming or limited precious storage resource for archiving. Compression has a complementary pair of systems, an encoder and a decoder. Video content is usually encoded by different video processing models to remove the redundancies, such as spatial redundancy and temporal redundancy. Then, the data is further converted to the bit stream for transferring or storage. When the end user received the encoded video, after inversely running the compression process, he can play back the video with an acceptable quality. Figure 2.1 shows the video data flow, and we proposed "transmission or coding decision model" was highlighted to show our video streaming adaptive adjustment objects.

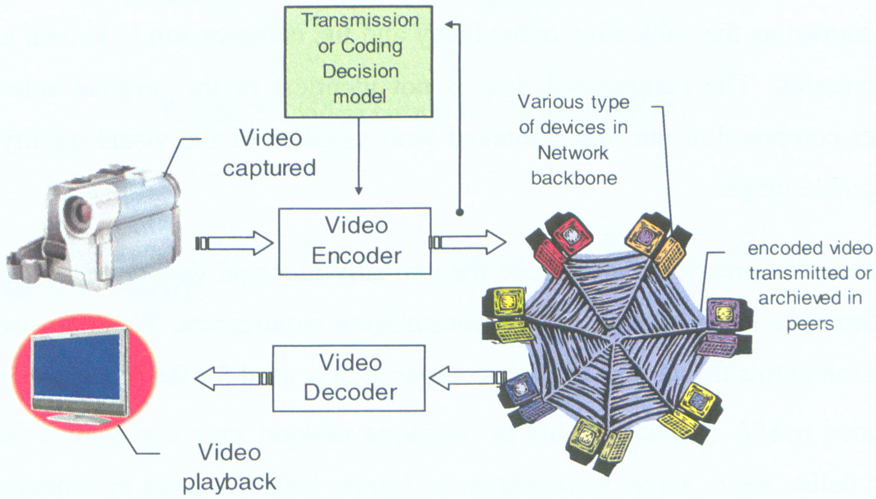

Figure 2.1 video encoder/decoder concept introduction

A video sequence is consisted of multiple image frames which are organized on a regular periodic sampling interval. Video coding translates video sequences into an efficient bit-stream suitable for network transmission and storage. This translation is to compress the raw video data by removing the redundant information in the video sequence. Video sequences contain three kinds of redundancies, spatial, temporal and psycho-visual [51].

- Spatial redundancy refers to the correlation present between different parts of a single frame. Removal of spatial redundancy thereby involves processing within a single frame and is known as the Intra Frame Coding.

- Temporal redundancy refers to correlation between the previous and subsequent pictures within a video. The successive frames in the video have the high probabilities that are very similar to the previous one. A lot of duplicated information present in a frame is also present in the frame that preceded it. Hence, removal of such temporal redundancy involves looking between frames and is called Inter Frame Coding.

- Psycho-visual redundancy refers to the non-sensitive information to Human Vision System (HVS). For example, certain elements of the image or video sequence can be removed without significantly affecting the viewer's perception of visual quality. It is also known as the subjective redundancy and the compression is known as the lossy compression. The compressed data is not identical to the original source, thus, a higher compression rate was achieved with sacrifice of the visual quality within an acceptable range.

The lossless compression can achieve theoretically identical decompressed video at the present time due to the signal noise or transmission requirement. We don't use this lossy compression in this thesis. The lossless compression is used for the full scope research.

2.2 Video Frames

Video usually is formed by Group of Pictures (GOP), and GOP contains three kinds of frames, Intra (I), Predictive (P) and Bi-directionally-predictive (B). Figure 2.2 shows the relationship within the coded video GOP. The intra frame (I-frame) is coded in isolation from other frames using transform coding, quantization and entropy coding. The P frame is coded using motion compensation followed by transform coding and entropy coding. The B frame is predicted bi-directionally, which means that the prediction is formed using both its previous frame as well as the successive frame. An I frame is often the first frame in a GOP, and used to efficiently code frames corresponding to scene changes. Frames within a scene are similar to preceding frames and hence may be coded as P or B for increased efficiency. Frames between two successive I frames, including the leading I frame, are collectively called a group of pictures (GOP). Usually, multiple B frames are inserted between consecutive P or between I and P frames.

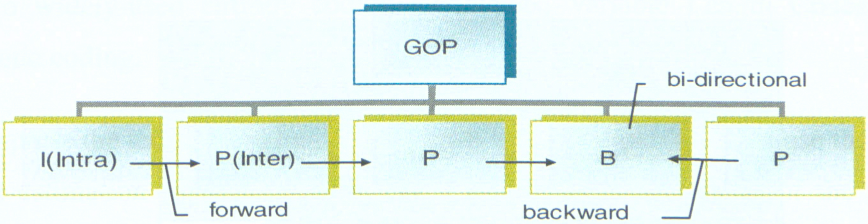

Figure 2.2 video Group Of Pictures (GOP)

2.3 Video Coding

Block based coding is a very popular approach to video encoding. Video compression typically operates on square-shaped groups of neighboring pixels, often called a macro-block. These pixel groups or blocks of pixels are compared from one block to the next and the video compression codec (both the encoder and the decoder) sends only the differences within those blocks. This works well if the video has no motion. For example, A still picture can be repeated with very little transmitted data. In areas of video with more motion, more pixels change from one frame to the next. When more pixels change, the video compression scheme must send more data to keep up with the larger number of pixels that are changing. If the video content includes an explosion, flames, or any other image with a great deal of high frequency detail, the quality will decrease, or the bit rate must be increased to render this added information with the same level of detail. These blocks are processed in the scan order as shown in Figure 2.3.

Those blocks are processed left to right and top to bottom. Spatial redundancy is removed through the use of Transform Coding techniques. The Discrete Cosine Transform (DCT) is a popular technique which operates on the blocks. This manner reduces the memory requirement compared with the image based transforms, such as the Discrete Wavelet

Transform (DWT).

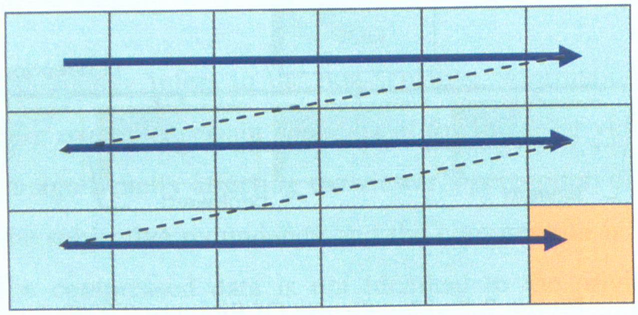

Figure 2.3 video compression scan order of blocks in frame

The DCT transform image data or residual values after prediction into coefficients. The action of DCT and its inverse the IDCT can be described in terms of a matrix manipulation. This can be given by:

$$Y = C X C^T \tag{2.1}$$

where X is a matrix of original image, Y is the resulting DCT coefficients, C is a N x N transform matrix. The elements of C are:

$$C_{ij} = k_i \cos(\frac{(2j+1)i\pi}{2N}), \text{ where } k_i = \sqrt{\frac{1}{N}} \ (i=0),\ k_i = \sqrt{\frac{2}{N}} \ (i>0) \tag{2.2}$$

Equations 2.1 and 2.2 can be written the bellow summation form:

$$Y_{xy} = C_x C_y \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} X_{ij} \cos(\frac{(2i+1)x\pi}{2N}) \cos(\frac{(2j+1)y\pi}{2N}) \tag{2.3}$$

$$X_{ij} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C_x C_y Y_{ij} \cos(\frac{(2i+1)x\pi}{2N}) \cos(\frac{(2j+1)y\pi}{2N}) \tag{2.4}$$

Quantization involves loss of information, and is the operation most responsible for the compression. The zigzag scanning converts the quantized DCT coefficients into one dimensional (1D) array for entropy coding. Entropy coding takes these quantized coefficients and converts them to bit stream suitable for transmission or storage. There are two widely-used entropy coding techniques, Variable Length Codes (VLC) and arithmetic coding.

To compress the temporal redundancies, most video coding standards use the block-based motion estimation and compensation techniques. As an extension of the conventional block based DCT techniques, 3D-DCT is proposed as one method to simplify this motion compensation technique. The 3D-DCT is known as XYZ coding, where X and Y represent the horizontal and vertical dimensions of the video pictures, and Z represents the temporal dimension. With the *8* x *8* video block sizes, similar to 2D, the 3D-DCT partition the video into *8* x *8* x *8* Group Of Blocks (GOB). The summation forms are:

$$F(a,b,c) = C_a C_b C_c \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\sum_{z=0}^{N-1} f(x,y,z)\cos(\frac{(2x+1)a\pi}{2N})\cos(\frac{(2y+1)b\pi}{2N})\cos(\frac{(2z+1)c\pi}{2N}) \quad (2.5)$$

$$f(x,y,z) = \sum_{a=0}^{N-1}\sum_{b=0}^{N-1}\sum_{c=0}^{N-1} C_a C_b C_c F(a,b,c)\cos(\frac{(2x+1)a\pi}{2N})\cos(\frac{(2y+1)b\pi}{2N})\cos(\frac{(2z+1)c\pi}{2N}) \quad (2.6)$$

where $C_i = \begin{cases} 0 & for \quad i=0 \\ 1/\sqrt{2} & for \quad i>0 \end{cases}, \quad i \in \{a,b,c\}$.

The 3D zigzag is used for the DCT coefficients transform as the regrouping of a collection of eight 2D zigzag arrays.

2.4 The Video Coding Standards

Video coding techniques were developed from the academic research area into a highly commercial business for almost twenty years. International standardization requires collaboration between regions and countries with different infrastructures, different technical backgrounds, and different commercial interests. There are two sets of video coding standards. One is the "H26x" series standards developed by the International Telecommunication Union (ITU-T), another is the MPEG series standards developed by the International Organization for Standardization (ISO). The ITU-T standards are defined by volunteers in open committees and are agreed upon based on the consensus of all the committee members.

2.4.1 H.264 Comparison With Major Video Standards

H.264 is one of the latest standards to provide similar functionality to earlier standards, such as H.263+ and MPEG-4 Visual (Simple Profile). It has significant better compression performance and improved support for reliable transmission. It is suitable for the video communication applications, such as video conferencing or video telephony, coding for high quality video streaming over the packet networks.

The other major organization involved in the development of standards is the ISO. The ISO standards are the MPEG-1, 2 and 4. The ISO is currently working on the next standard MPEG-7. Both these organizations have defined different standards for video coding. These different standards are summarized in Table 2.1. The major differences between these standards lie in the operating bit rates and the applications they are targeted for. Each standard allows for operating at a wide range of bit-rates, hence each can be used for a range of applications.

Video Coding Standard	Typical Range of Bit Rate	Typical Applications
H.261	px64 kbits/s, p=1...30	ISDN Video Phone
MPEG-1 Video	1.2 Mbits/s	CD-ROM
MPEG-2 Video	4-80 Mbits/s	SDTV, HDTV
H.263	200~12000 Kbps	PSTN Video Phone
MPEG-4 Video	24-1024 kbits/s	wide range of applications
H.264	wide range	wide range of applications

Table 2.1 Main Stream Video Coding Standards

2.4.2 H.264 Profiles and Levels

H.264 defines three sets of Profiles, each supporting the particular coding functions. The Baseline Profile supports intra (I-frame) and inter-coding (P-frame) and entropy coding with Context-Adaptive Variable Length Codes (CAVLC). Baseline Profile is a very good solution for network communication applications, such as video conferencing, video telephony and wireless communication for its high ratio compression. The Main Profile supports the interlaced video, inter-coding with B-frames, inter coding by weighted prediction and entropy coding by Context-Based Arithmetic Coding (CABAC). The Extended Profile adds modes to enable efficient switching between coded bit streams (SP and SI frames) and improved error resilience by data partitioning. Main Profile is mainly used for television broadcasting and video storage. Extended Profile is particularly useful for media streaming. Each profile has the sufficient flexibility to support a wide range of applications. Besides, the High Profile is developed for broadcast and disc storage applications, especially for high-definition television applications and Blu-ray Disc. The table 2.2 compared H.264 major profiles and we can see the high-end profile has the certain features can best fit the various and changing network application requirements.

<table>
<tr><th rowspan="2"></th><th colspan="2">GOP frames</th><th>Entropy Coding</th><th colspan="3">Chroma Format</th></tr>
<tr><th>SI, SP</th><th>B</th><th>CABAC</th><th>4:0:0</th><th>4:2:2</th><th>4:4:4</th></tr>
<tr><td>Baseline</td><td>n/a</td><td>n/a</td><td rowspan="2">n/a</td><td rowspan="3">n/a</td><td rowspan="3">n/a</td><td rowspan="3">n/a</td></tr>
<tr><td>Extended</td><td>have</td><td rowspan="3">have</td></tr>
<tr><td>Main</td><td rowspan="2">n/a</td><td rowspan="2">have</td></tr>
<tr><td>High</td><td>Hi10P</td><td>Hi422P</td><td>Hi444PP</td></tr>
</table>

notes:

1. GOP frame: all H.264 profiles have the I, P-frames
2. Entropy Coding: all H.264 profiles have Context-Adaptive Variable Length Codes (CAVLC)
3. Chroma Format: all H.264 profies have 4:2:0 type

Table 2.2 H.264 Profile Comparison

2.5 Peer-to-Peer Streaming Over Network Infrastructure Diversity

Peer-to-Peer streaming protocol (P2PSP) addresses the bottleneck of the conventional client-server streaming techniques, to deliver higher perceived quality of services. Centralized Peer-to-Peer (P2P) streaming is designed as an application operating on the Internet backbone, thus representing a cost advantage over the techniques such as DiffServ [52]. Centralized P2PSP offloads video traffic from the bottleneck link and decentralizes the processing power from the streaming server, while maintaining the high integrity with centralized Authentication, Authorization, and Accounting (AAA) [53]. Simulation results indicate that centralized P2PSP consistently delivers a high video quality at the receiver, while client/server video streaming starts experiencing packets loss and degrades the video quality under the same network condition due to the multiple available transmission paths.

We use the MDC video coding techniques with spatial diversity to further offload the video traffic. Video contents are sub-sampled into multiple sub-streams, those sub-videos are compressed with H.264 baseline encoder. The available streaming sources in real-time packet network thus get the increasing data sources, especially in Peer-to-Peer network. When the streaming availability increases, the end user will have higher possibility to receive the requested video contents.

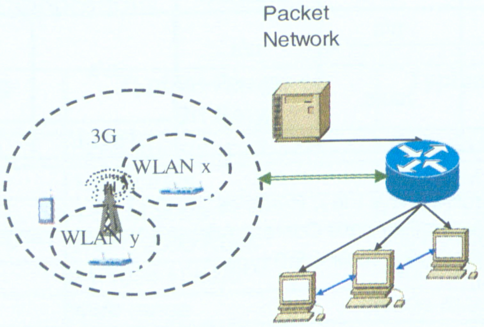

Figure 2.4 MDC Infrastructure with packet networks

The most challenge for video streaming comes from the wireless channels due to the issues of its latency, limited bandwidth, and variability of the wireless link quality. Video streaming demand the real-time transmission to satisfy the instantly playing-back characters while transmission rather than the traditional playing-back after downloading services, such as the commonly used HTTP applications or Multimedia Messaging Services (MMS). Figure 2.4 illustrates the various network nodes in the today's Internet and wireless network environments. Besides the common the computers (PC and Servers) and network devices (routers and switches), the wireless network has two main sources: the conventional cellular network and the data network. The cellular network is mostly dealing with the voice communication and it advances to 3G network for broader bandwidth and to carry non-voice data. The data network involves with the WLAN hotspot techniques with the mobility extension. In practice, the 3G network such as CDMA-1X has the throughput of around 60-100 kbps while the data network WLAN system such as 802.11b/g/n have throughput up to 2 Gbps. 3G networks cater for applications with real-time constraints while WLAN provides the higher throughput.

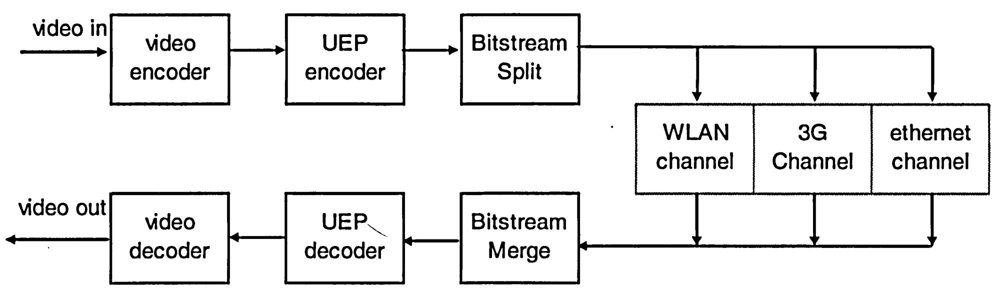

Figure 2.5 Block Diagram for Peer-to-Peer video streaming

Video streaming over those various network infrastructures with the fully Peer-to-Peer manner will benefit from the server coverage of 3G network and high throughput of the WLAN and common Ethernet devices. Advanced video coding techniques adopt the Unequal Error Protection (UEP) [54] scheme to integrate the unbalanced bandwidth and form the consolidated broader virtual transmission channel to video streaming applications. The block diagram of the Peer-to-Peer streaming system is summarized in

Figure 2.5. Multiple coding techniques can split video into different sub-stream, such as MDC, Layered coding, etc... In this thesis, we apply the MDC technique and follow the H.264 baseline profile to encode video stream into multiple sub-streams, then feed into the multiple transmission channel. According to the channel diversity, different transmission rate was studied in the following chapters.

2.6 Video Quality Measurement

In order to specify, evaluate and compare video communication system it is necessary to determine the quality of the video images displayed to the end viewer. Visual quality is inherently subjective and is influenced by many factors that make it difficult to obtain the completely accurate measures for the human behavior is deeply involved. Measuring visual quality using objective criteria gives accurate, repeatable results is still the challenge of today's major schemes.

The Peak Signal to Noise Ratio (PSNR) is commonly used to present quality of the reconstruction in image compression. It presents the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. The PSNR is calculated on a logarithmic scale and depends on the mean squared error (MSE) of between an original and an impaired image or video frame, relative to the square of the highest-possible signal value in the image (equation 2.8). High PSNR usually indicates high quality and low PSNR usually indicates low quality.

$$MSE = \frac{1}{MN} \sum_{i-0}^{M-1} \sum_{j=0}^{N-1} [f'(i,j) - f(i,j)]^2 \tag{2.7}$$

$$PSNR = 10 \log_{10} [\frac{255^2}{MSE}] \tag{2.8}$$

$$SAE = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} | f'(i,j) - f(i,j) | \qquad (2.9)$$

SAE (equation 2.9) is also the most widely-used measure of residual energy for reason of computational simplicity. The H.264 reference model software [55] uses SA(T)D, the sum of absolute differences of the transformed residual data, as its prediction energy measure for both Intra and Inter prediction.

In this thesis, we use the PSNR value to evaluate the broadcasting video quality. Unless the certain declaration, the PSNR with the unit dB will be mostly used to evaluate the picture pixel luminance we focused.

As the recent studies state, the perception of a visual scene is formed by a complex interaction between the components of the Human Visual System (HVS). And the HVS is more sensitive to luminance errors compared with high frequency chrominance errors. We use the luminance PSNR with the unit dB in the thesis to measure the video quality.

CHAPTER 3

VIDEO STREAMING WITH MDC ERROR RESILIENCE

The Multiple Description Coding (MDC) is a fault tolerance coding technique which fragments a media stream into multiple independent sub-streams which consist of multiple complementary descriptions. The encoded packets of each description are routed over multiple, independent paths. MDC provides an interoperable and flexible distributed environment with the primary goal of solving interoperability problems in heterogeneous, networked environments. We use the spatial diversity to divide the video into multiple sub-streams, and the multi-path video transmission is simulated using the two-state Markovian Gilbert Model. The analysis applied to MDC coding with frame dropout rate evaluation is covered with the MDC encoder parameters: number of sub-streams (n) and MDC GOP length (g). Meanwhile, the probability of network transmission status at each single point of time, the probability of transmission status changing and the collaborations of the transmission channels, such as the transmission with aligned or un-aligned GOP I-frame positions in all the channels are also included in the study.

3.1 MDC Video Streaming

MDC techniques are designed for path diversity. It is one attempt to resolve the drawback caused by single path transmission with Single Description coding (SDC). MDC encodes media contents into multiple independent descriptions, and these descriptions are also known as sub-streams. Once a sub-stream is received at the client end, the granules within the stream can be decoded. The overall quality of recovered content is depending on the number of successfully delivered sub-streams. The more sub-streams are received, the higher the reconstructed video quality can be achieved. MDC provides loss tolerance, and it is therefore beneficial for delay-sensitive, real-time streaming applications where data losses are highly disruptive. The concept of the MDC encoder is illustrated in Figure 3.1. Two perceived available channels are used for the MDC encoded video streaming. Once the packet received at the other end of the transmission channel, it is decoded by both site decoder and central decoder, thus, the picture can be played back.

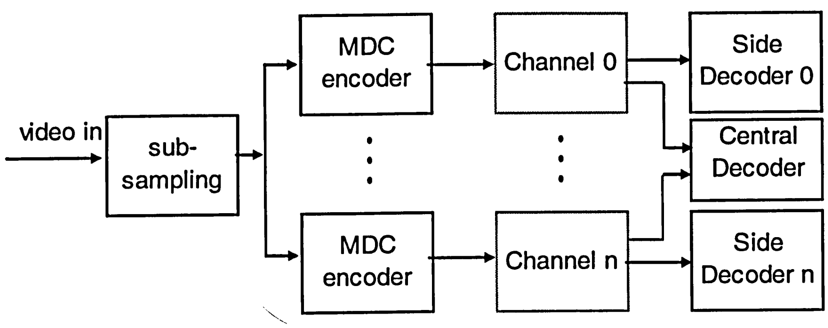

Figure 3.1 block Diagram Multiple Description Coding

3.1.1 MDC with Spatial Diversity

MDC is typically applied in scenarios where multiple nodes independently forward video content to the client node over physically connected networks. Prior research in [56] has proposed to apply MDC for a decentralized peer-to-peer streaming system, where the MDC sub-streams (*M1+M2+M3+...+Mn*) received at server nodes are forwarded to the end user who request those sub-streams. Therefore, the more forwarding servers exist, the more paths can be used for transmission. The combination of the under utilized network bandwidth of multiple forwarding servers may give an overall broader bandwidth for MDC streaming, hence yielding a better reconstructed video quality at the receiver.

Each intra frame (I-frame) in the MDC encoded video contains the macro blocks in which each *16* x *16 or 4* x *4* luminance region and each *8* x *8* chrominance region is predicted from the previous sample in the same frame. After prediction, the residual data for each MB and DCT transforms with *4* x *4* integers are quantized. The coefficients and syntax elements are entropy coded with context-adaptive variable length coding scheme (CAVLC). According to the available transmission channels, the MDC can adaptively encoder the video content with different size of Macro block, thus, different sub-streams are generated. Figure 3.2 shows the video bit rate of the original one and the MDC decoded video from received sub-stream. As the sub-steam video has less luminance information after the compression, the video bit rate decreased to the desired level for network transmission.

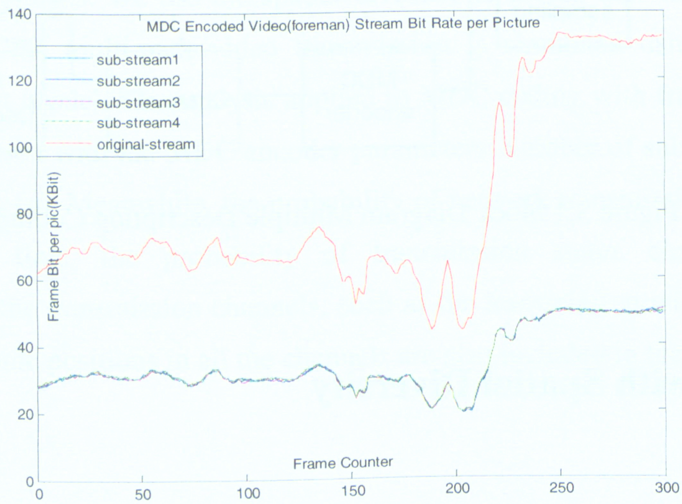

Figure 3.2 MDC foreman 4 sub-streams bit rate (bit/picture)

3.1.2 MDC with Temporal Redundancy Compression

MDC also has the feasibility to compress video streams into different sizes of GOP, thus the temporal redundancy between two or multiple consecutive frames can be removed efficiently. Figure 3.3 shows the experimental result of video bit rate of MDC encoded video foreman with 1-15 different GOP length. When the GOP length increases, the Intra frame bit decreases and Inter frame increases as more information is included into the Predictive frames. When we evaluated the total video bit rate, it decreased along with the GOP length increase. This indicates the MDC encoder efficiently removed the temporal redundancy with the predictive frame compression manner.

The MDC system is capable of reconstruct video with different quality levels depending on the number of MDC sub-streams received at the destination. The more description the end user received, the enhanced the picture quality will be achieved.

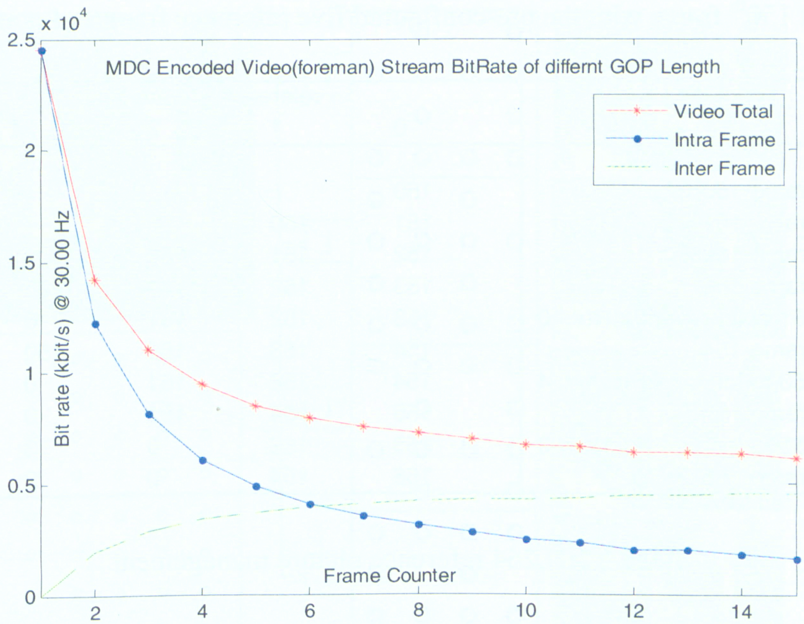

Figure 3.3 MDC video bit rate with different GOP length

During the encoding process, the reference picture which is the previously processed is stored in a reference buffer with different states, and the buffer is managed according to the marked states. The encoder manages the marks of reference picture as bellow:

- Short term picture: it is ordered with decreasing order by picture number. When the buffer is full, the oldest picture will be dropped from the buffer.
- Long term picture: it is ordered with increasing order by picture number
- Output to display or unused for reference: it is not used for further prediction; frame swapped out of the buffer after content processed.

The encoder uses the reference pictures from list 0 for encoding each macro block partition in an inter-coded macro block. The choice of reference picture is signaled by an index number, where index 0 corresponds to the first frame in the short term section and the indices of the long term frames start after the last short term frame. The table 3.1 shows the procedures of reference picture selection in compression procedure, the sample time is set at 150^{th} frame with the pre-configured five reference frames of compression.

operation	reference picture list				
	0	1	2	3	4
initial					
encode frame 150	150				
encode frame 151	151	150			
encode frame 152	152	151	150		
encode frame 153	153	152	151	150	
assign 151 to longtermPicNum 0	153	152	151	0	
encode frame 154	154	153	152	150	0
assign 153 to longtermPicNum 4	154	152	151	0	4
encode frame 155	155	154	152	0	4
assign 155 to longtermPicNum 3	155	152	0	3	4
encode frame 156	156	154	0	3	4

Table 3.1 H.264 reference picture management

In this thesis, H.264 baseline profile encoder was applied and the number of reference frames was set as five as showed in table 3.1.

3.1.3 MDC H.264 Baseline Profile Encoder and Decoder With Error Resilience

MDC encoder with H.264 baseline profile sub-samples each video frames into multiple sub-frames over the spatial domain. Each sub-frame is encoded by using H.264 baseline video encoder. As figure 3.4 showed, the picture was sub-sampled into 4 QCIF, (*4* x *4*), then compressed with H.264 encoder. The encoded sub-videos are streaming via multiple independent channels to the end users. At the receiver end, the corresponding MDC decoder is shown in Figure 3.5, and the cubic-spline interpolation is applied to reconstruct the missing sub-frames. The more valid descriptions from all the sub-streams, the information can be detective and invested into the picture re-displaying.

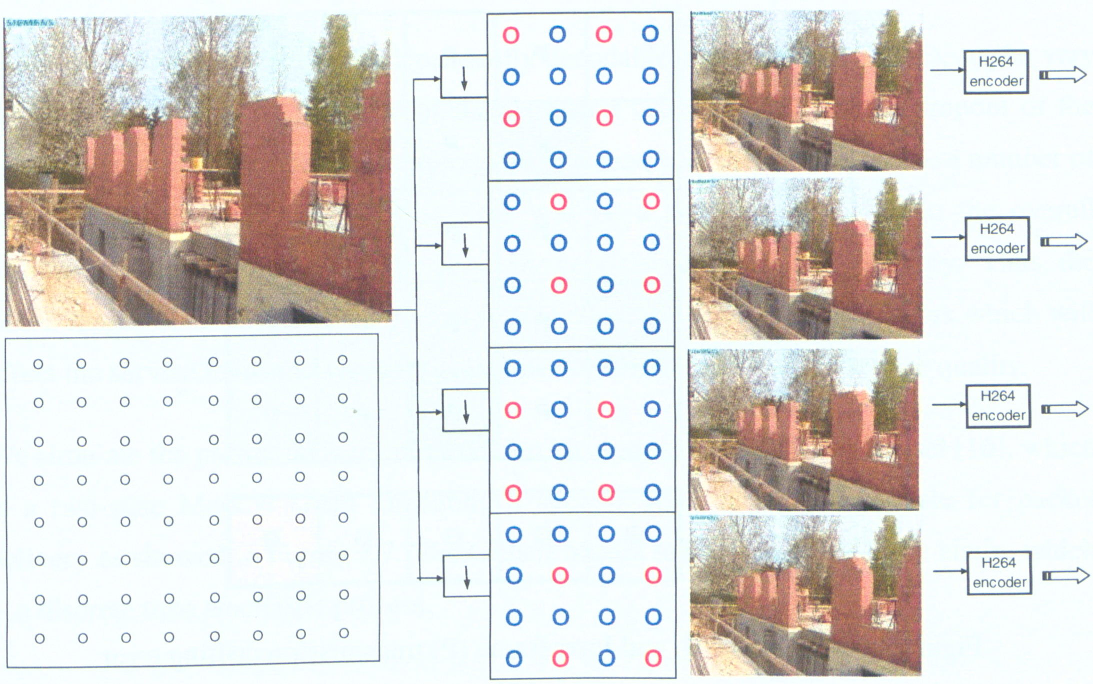

Figure 3.4 MDC with spatial diversity encoding

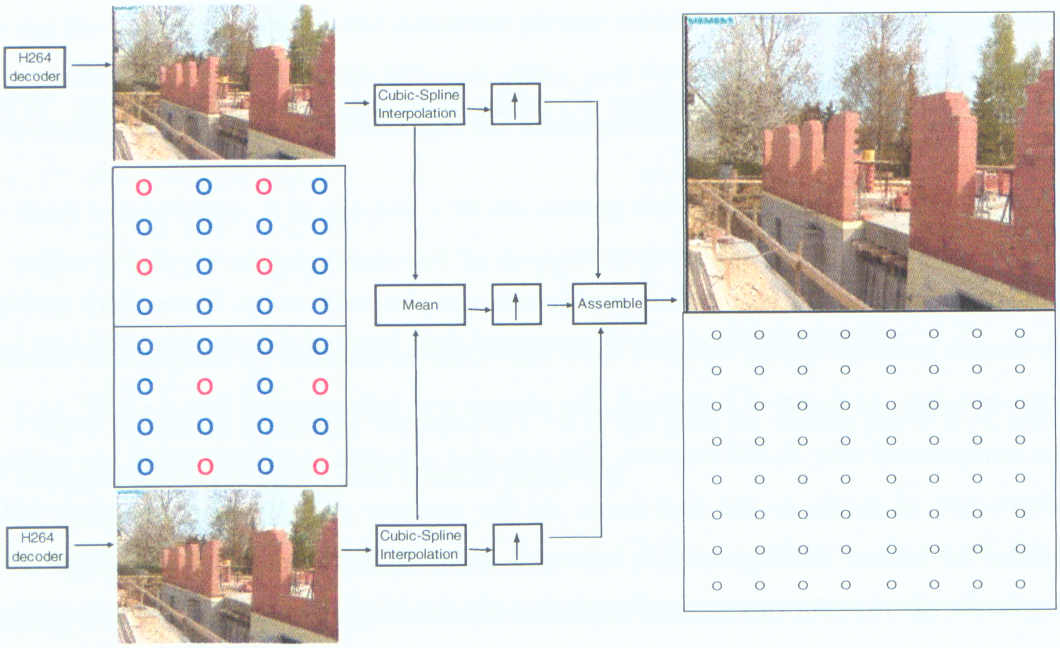

Figure 3.5 MDC with spatial diversity decoding

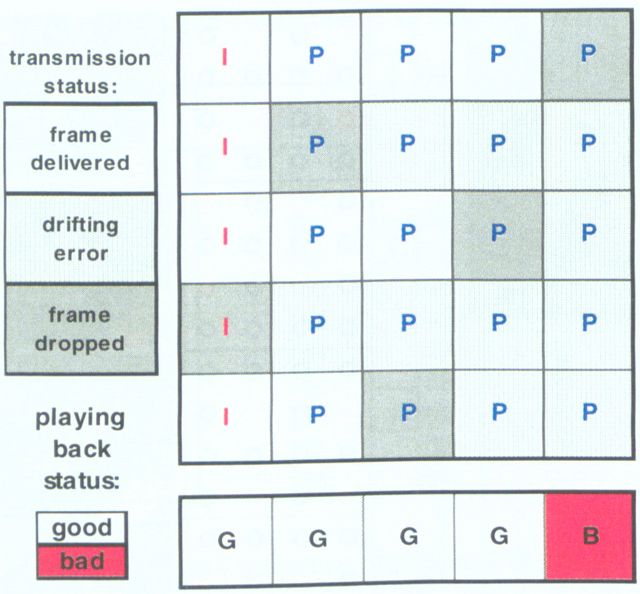

Figure 3.6 MDC Intra (I) and Inter frame (P) transmission drifting error

3.2 MDC Gilbert Model Video Frame Dropout Rate

The video stream is consisted of group of pictures. We apply H.264 baseline profile which uses two types of frames: Inter (I)-frame and Predictive (P)-frame as showed in Figure 3.6. I-frames are encoded independent of prior frames. P-frames are encoded with respect to a prior I-frame or P-frame, where motion compensation techniques are applied for improving the compression efficiency. A group of frames that starts with an I-frame, followed by a set of P-frames and ends before the next I-frame is called a Group of Pictures (GOP). Losing an I-frame or a P-frame will result in distortion of the following frames within the same GOP. This is known as the drifting error. Because of the drifting error(s), a lost frame will affect the subsequent frame(s) until the end of GOP. The end user will have to wait until to the next GOP begin because of the streaming video can only be reconstruct back with a valid I-frame when the previous GOP content is collapsed.

GOP usually started from the independently decodable Intra frame (I-frame), so the very first I-frame is also the indicator of the very first GOP arrival after the dropout of the previous frame. In the open distribution environment, considering the massive number of end users, the summarized waiting time will be a significant volume to the overall resource consuming when consider about the services of information delivery. Thus, the MDC length of GOP is also a key factor of the waiting time of the end users which will affect the service level and would be a key performance factor of the service quality.

We simulate the packet delivery of the video bit streams using Gilbert Model [16], which is a two-state Markov chain indicating a success state and a failure state for packet delivery as showed in Figure 3.7. The Gilbert Model is based o the Markov chain, which is a discrete time stochastic process.

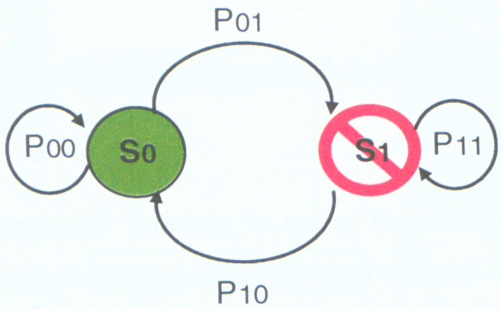

Figure 3.7 Gilbert Model of MDC transmission states

To simplify the analysis, we assume that each packet contains one encoded video sub-frame, the network conditions are the same in all the participated streaming sub-channels. Let S_0 (good) denote the good state when the IP network packets are received correctly and timely, and S_1 (bad) denotes bad state when the packets are lost. The probabilities of the network transition from S_0 to S_1 and from S_1 to S_0 are denoted as P_{01} and P_{10} respectively. The probabilities of staying in the same state are denoted as P_{00} for state S_0 and P_{11} for state S_1. Steady state analysis shows that the overall probabilities of good state P_{S0} and bad state P_{S1} are:

The probabilities of states are: $\begin{bmatrix} P_{S0} \\ P_{S1} \end{bmatrix}$ and the probabilities of states changes are: $\begin{bmatrix} P_{00} & P_{10} \\ P_{01} & P_{11} \end{bmatrix}$

the probabilities of steady states of being in Good (Bad) state are:

$$\begin{bmatrix} P_{S0} \\ P_{S1} \end{bmatrix} = \begin{bmatrix} P_{00} & P_{10} \\ P_{01} & P_{11} \end{bmatrix} \times \begin{bmatrix} P_{S0} \\ P_{S1} \end{bmatrix} \tag{3.1}$$

we can transfer equation (3.1) as:

$$\begin{bmatrix} P_{00} & 1-P_{11} \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} P_{S0} \\ P_{S1} \end{bmatrix} = \begin{bmatrix} P_{S0} \\ 1 \end{bmatrix} \quad (3.2)$$

Solving equation (3.2), the probability of successful delivery at each MDC transmission status is:

$$P_{S0} = \frac{P_{11} - 1}{P_{00} + P_{11} - 2} \quad (3.3)$$

and the probability of unsuccessful frame delivery is:

$$P_{S1} = \frac{P_{00} - 1}{P_{00} + P_{11} - 2} \quad (3.4)$$

Once the peer nodes in the video streaming network build up the communications between each other, each connection is regarded as an independent transmission channel. Let D denote the possibility of frame dropout of the MDC streams and the transmission state is bad, g denote the length of GOP, n denote the number of MDC sub-streams, at each transmission time i, the probability of good frame transmission is:

$$P_{S0} P_{00}^{i-1} \text{ where } i = 1, 2, \cdots g \quad (3.5)$$

And its complement $1 - P_{S0} P_{00}^{i-1}$ indicates the probability of all the situations that have a frame dropout before or at frame i. When all the sub-streams have the aligned Intra frame (I-frame), resulting in the same GOP sequence pattern, the mean frame dropout rate of all n sub-streams transmission is:

$$D = \frac{1}{g} \sum_{i=1}^{g} (1 - P_{S0} P_{00}^{\ i-1})^n \quad (3.6)$$

The start-up time is also a key component of the frame dropout rate over the MDC streaming periods. It can be various according to the sub-stream transferring conditions. There are always the un-aligned I-frames with the same sub-stream contents in the

Internet environments, which have different transmission start-up times or the un-controlled transferring lags. The MDC frame position i within a GOP with length (g) during the transmission time t is:

$$i = \begin{cases} \mathrm{mod}(t, g) & \mathrm{mod}(t, g) \neq 0 \\ g & \mathrm{mod}(t, g) = 0 \end{cases} \tag{3.7}$$

For each decoded MDC sub-stream, especially in peer-to-peer environment, there are high possibilities that the available transmission channels have the duplicated video sub-stream contents. Let m denote the number of channels with the duplicate sub-stream content, thus, the m channels with duplicated streaming contents can be regarded as a consolidated single sub-stream with the cumulated bandwidth. Assuming the transmission is homogeneously distributed and the probability of transmission over m consolidated channel in good state $\bar{S}_0$ is:

$$\bar{P}_{S0} = 1 - P_{S1}^{m} \tag{3.8}$$

and the probability of the output of m channels staying in the same state good state $\bar{S}_0$ is:

$$\bar{P}_{00} = 1 - P_{11}^{m} \tag{3.9}$$

thus, the overall frame dropout rate $D_{m \times n}$ of m channels is:

$$D_{m \times n} = \frac{1}{g} \sum_{i=1}^{g} (1 - \bar{P}_{S0} \bar{P}_{00}^{(i-1)})^{n} \tag{3.10}$$

$$D_{m \times n} = \frac{1}{g} \sum_{i=1}^{g} (1 - (1 - P_{S1}{}^{m})(1 - P_{11}{}^{m})^{(i-1)})^{n} \tag{3.11}$$

3.3 MDC Frame Dropout Rate At Number Of Connections

In the Peer-to-Peer packet network, the transmission channels have few limitations during the streaming time. The end user can join the video broadcasting or receiving at any their desired time. Then the number of streaming connections is theoretically infinity. A well managed transmission over the massive number of connection is desired asset of the overall service delivery.

Let's firstly start from the fundamental analysis of the frame dropout rate on different number of MDC sub-streams. It has the decreasing function character. For simplify the interpretation, let any two different integers n_1 and n_2 denote the numbers of MDC sub-streams respectively, where $1 \le n_1 \le n_2$, and from the probability definition, $(1 - P_{S0}P_{00}^{(i-1)}) \in [0,1]$, then, $0 \le (1 - P_{S0}P_{00}^{(i-1)})^{n_2} \le (1 - P_{S0}P_{00}^{(i-1)})^{n_1} \le 1$, thus, the frame dropout rate is a decreasing function of number of MDC sub-streams. The more MDC sub-streams exist, the less frames will drop, and the better quality for video playing back. In real world, it can be interpreted as the video streaming quality will be improved.

Figure 3.8 compares the MDC frame dropout rate over number of sub-streams at fixed GOP lengths *15*. The network transmission status are simulated as $P_{11}=0.2\sim0.6$ and P_{S1} changed accordingly at the same time with P_{11} in this scenario. The frame dropout rate of multiple streams transmission is lower, thus the better quality of service compared with the single or less number of sub-streams.

Figure 3.9 compares the MDC frame dropout rate over number of transmission connections at fixed GOP lengths *15*. The network transmission status are simulated as $P_{11}=0.2\sim0.6$ and P_{S1} changed accordingly at the same time with P_{11} in this scenario. The frame dropout rate of multiple connections decreases as the under-used network bandwidth is consolidated to form a broader one, thus the better transmission quality is achieved.

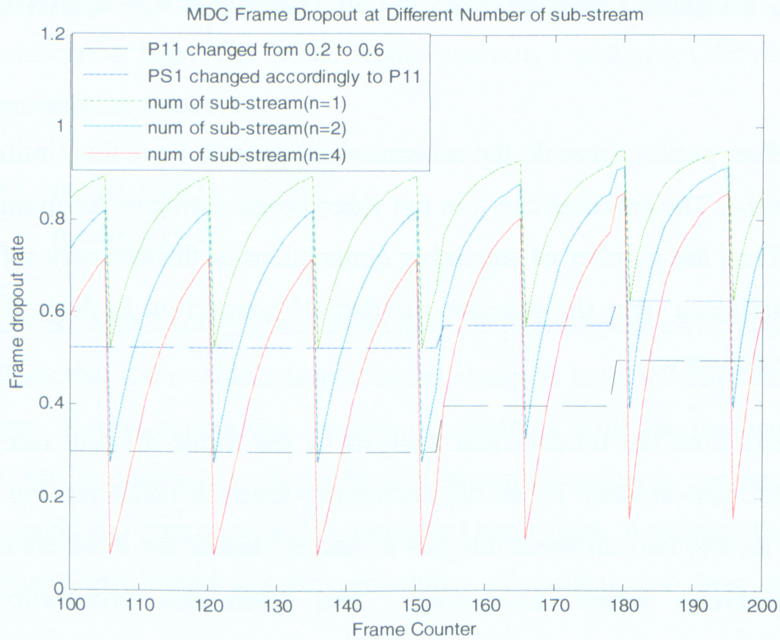

Figure 3.8 MDC frame dropout rate at different number of sub-streams *(n=1,2,4 & g=15)*

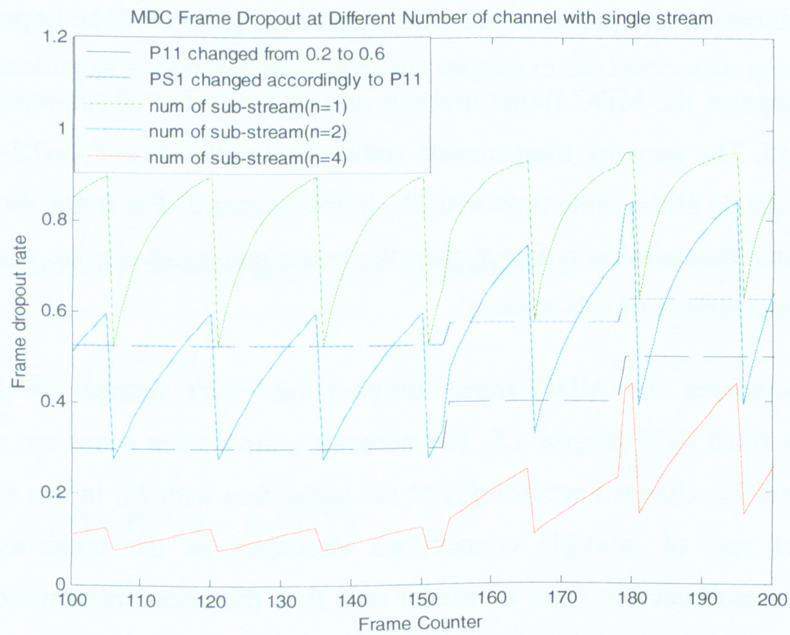

Figure 3.9 frame dropout rate at different number *(m)* of channel with sub-stream *(n=1)*

Figure 3.10 exams the MDC frame dropout rate of multiple sub-video transferring with different number of channels, each channel has the duplicated video contents, sub-video is encoded at the fixed GOP lengths *15*. The network transmission status are simulated as $P_{11}=0.2\sim0.6$ and P_{S1} changed accordingly at the same time with P_{11} in this scenario. The overall quality is improved by both of the broader bandwidth of those consolidated m channels and the more sub-stream video content.

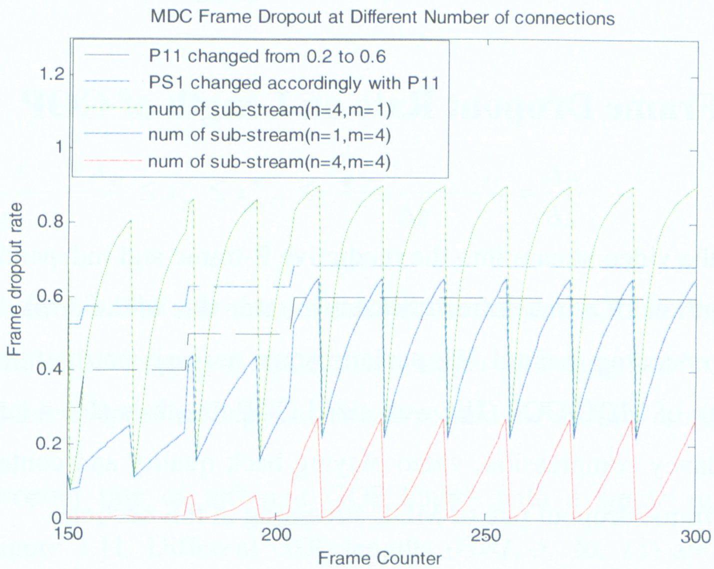

Figure 3.10 MDC sub-videos *(n)* streaming at different number of connection *(m)*

We exams the differences between the frame dropout rate of streaming with multiple sub-streams with a single connection or single sub-stream with multiple connections. Though the transmissions of single sub-stream with multi-connection achieve better results than multiple sub-streams with single connection, the delivered video contents, the MDC encoded descriptions are not at the same level. As the MDC decoder manner, the more descriptions received and be decoded, the better quality of video playing back. Once the object was pre-set for the video quality, the transmission can be set to send as many as possible different sub-streams with the total available network channels. When the end user has the tolerance of video quality, the multiple channels transmission can set the lower frame dropout rate as a goal. MDC provide the feasibility to fit into the peer-to-peer network environment to achieve the desired video qualities or better transmission depending on the end user's subjectively choices. No matter which decision the end user make, the MDC streaming have the linear relationship with the transmission connection, the more network connections are available, the better transmission qualities. So the capacity of all the connections, which is determined by the end user's own transmission equipments or applications, is the key of the streaming.

3.4 MDC Frame Dropout Rate on Length of GOP

MDC compress the video stream into the predictive P-frame and independent Infra frame I-frame. The length of GOP may affect the transmission due to the error drifting with the interlaced frame encoding method. We perform some analysis on the frame dropout rate of different length of MDC GOP. The well sized GOP length is also a key factor affects the video redundancy compression, video playing back quality and content transferring. We focus on the frame dropout rate of MDC streaming in this section.

Let two integers g_1 and g_2 denote any two different length of GOP respectively, where $1 \le g_1 < g_2$, and let $x_i = 1 - P_{S0}P_{00}^{i-1}$, where $i = 1,2,\cdots g$.

then the frame dropout rates of those two MDC stream at length of GOP g_1, g_2 are:

$$D_1 = \frac{1}{g_1}\sum_{i=1}^{g_1}[(1 - P_{S0}P_{00}{}^{i-1})]^n = \frac{x_1{}^n + x_2{}^n + \cdots + x_{g_1}^n}{g_1}$$

$$D_2 = \frac{1}{g_2}\sum_{i=1}^{g_2}[(1 - P_{S0}P_{00}{}^{i-1})]^n = \frac{x_1{}^n + x_2{}^n + \cdots + x_{g_1}^n + x_{g_1+1}^n + \cdots + x_{g_2}^n}{g_1 + (g_2 - g_1)}$$

let $w = x_1{}^n + x_2{}^n + \cdots + x_{g_1}^n$, $\Delta w = x_{g_1+1}^n + x_{g_1+2}^n + \cdots + x_{g_2}^n$, $\Delta g = g_2 - g_1$

As $g_2 - g_1 > 0$ and $x_i \in [0,1]$, so $x_1^n \le x_2^n \le \cdots \le x_g^n$

then the difference between the two frame dropout rates $D_2 - D_1$ can be represented as:

$$D_2 - D_1 = \frac{w + \Delta w}{g_1 + \Delta g} - \frac{w}{g_1} = \frac{\dfrac{\Delta w}{\Delta g} - \dfrac{w}{g_1}}{\dfrac{(g_1 + \Delta g)g_1}{g_1 \Delta g}}$$

as

$$\frac{w}{g_1} = \frac{x_1^n + x_2^n + \cdots + x_{g_1}^n}{g_1} \le x_{g_1}^n \le x_{g_1+1}^n \le \frac{x_{g_1+1}^n + \cdots + x_{g_2}^n}{\Delta g} = \frac{\Delta w}{\Delta g}$$

Therefore, $D_2 - D_1 \ge 0$. We can draw the conclusion that the frame dropout rate is the increasing function of length of GOP. The longer the length of GOP, the higher probability of frame dropout will appear. This is so said the drifting errors.

MDC frame dropout rate of different GOP length with a single sub-stream was is compared in Figure 3.11. Different GOP lengths (*G=1, 5, 10, 15*) are simulated. The network transmission status are simulated as $P_{11}=0.2\sim0.6$ and P_{S1} changed at the same time with P_{11} in this scenario. The frame dropout rate of all the selected GOP length has the transmission error drifting characters, and the when GOP length of MDC encoded video increases, the streaming frame dropout rate increases accordingly to the frame positions. When GOP length is set as 1, which means the video is encoded only with I-frame, the error drifting character is removed. The frame dropout rate showed with the green line in Figure 3.11 and the frame dropout rate matches the network transmission condition, the same rate as the status in bad (P_{S1}).

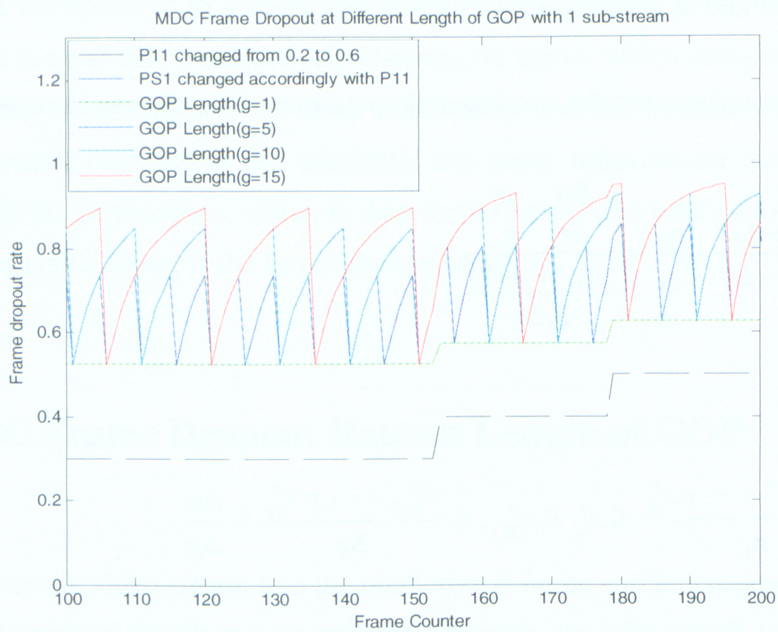

Figure 3.11 MDC single stream frame dropout rate at different GOP length

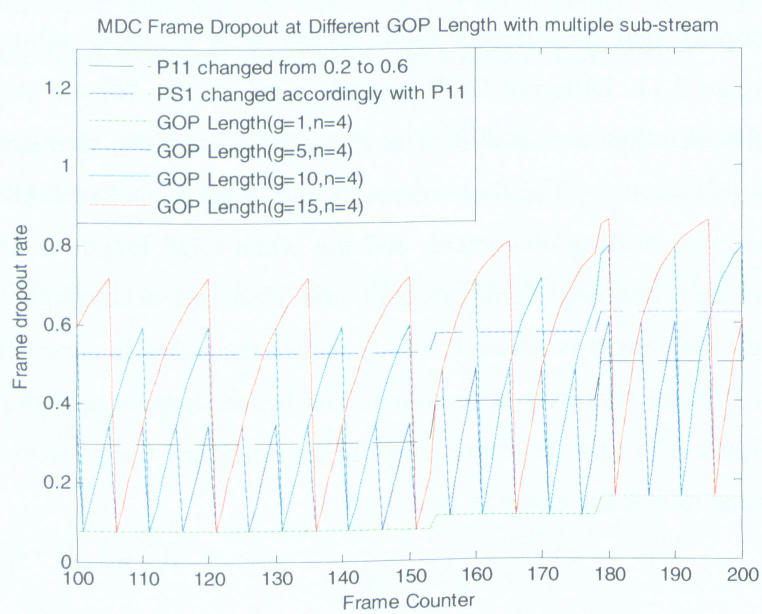

Figure 3.12 MDC sub-stream frame dropout rate at different GOP length

Figure 3.12 compares the MDC frame dropout rate over number multiple GOP lengths at 1, 5, 10, *15* with four sub-streams. The network transmission status are simulated as P_{11}=0.2~0.6 and P_{S1} changed accordingly at the same time with P_{11} in this scenario. The frame dropout rate was improved from the single sub-stream.

Different MDC GOP lengths provide probabilities to compress the video stream into different compression level for transmission and the streaming will the drift error character depending on the GOP length. The GOP length also has the linear relationship with the transmission. When GOP length increase, the average frame dropout rate increases.

3.5 MDC Video Transmission Start-up Time

Video streaming has the time constraints not because of the time is not sufficient, but due to the real-time transferred video contents need to be playing back. To fulfill the streaming requirement, many techniques are applied to relief the strike limitations of the transmission time, such as playing back buffer. That will increase the resource requirement, and may be not suitable for the mobile devices. Some network devices like the cellular phones may have very limited storage rather than the increasing computing powers. MDC have the transmission time diversity as the multiple encoded video descriptions can route to end user as long as the transmission time meet the video re-displaying requirements. With number *n* independent sub-streams, output video can be re-constructed from any received sub-streams. Only if all sub-streams are lost, the video frame cannot be re-constructed. When transmitting MDC sub-streams over different paths, each path will handle less traffic; hence will reduce the packet drop rate caused by traffic congestion.

3.5.1 MDC Multi-path Frame Dropout Rate On Different Start-up Time

We study the MDC sub-stream transmission without identical start-up time. The multiple servers transmit MDC sub-streams at the unaligned I-frame start-up time with the same or various GOP lengths via the multiple independent paths. The Intra (I-frame) and Inter (P-frame) pictures within a GOP will have the same sequence pattern, which can be decoded once they are successfully received at the user end.

As I-frame is not always the first frame during the processing period in certain stream, we need to determine the position of the I-frame in each channel. Let t denote the processing period, which is equal to the length of one GOP, let n denote the number of streams and s_n denote the offset in n^{th} path which equal to the period before the I-frame arrival. From the Markovian Gilbert Model, and the good states of transmission only if the start-up state has to be good, otherwise, the transmission will be regarded as frame dropout. So the probability of good frame transmission is: $P_{S0}P_{00}^{(t_n-1)} \in [0,1], t_n = 1,2,\cdots,t$, and its complement indicates that at any time t with all the situations that have a frame dropout as the MDC sub-stream transmission paths are independent from each other. Thus, the frame dropout rate is $1-P_{S0}P_{00}^{(t_n-1)}$, the overall frame dropout rate of N paths at anytime t_n is: $\prod_{n=1}^{N}(1-P_{S0}P_{00}^{(t_n-1)})$. For easy simulate, we introduce the s_n as the distance from the streaming time at the n$^{\text{th}}$ path as the time t and MDC GOP length is independent to the time distance. Thus, the MDC frame position t_n is presented as $t_n = (t - s_n) \bmod g$, $t_n = 1,2,\cdots,t-1$, $s_n = 1,2,\cdots,t-1$.

when offset $s_n = 0$, which means the aligned start-up time, the mean frame dropout rate for simultaneous transmission is the same as equation (3.6). At each point of time i_n , the frame dropout rate of each path is $1 - P_{S0} P_{00}^{i_n}$.

3.5.2 MDC multi-path frame dropout rate on homogeneous and aligned I-frame

MDC streaming use the overall network bandwidth and computing time with the encapsulated MDC frame by frame delivery method. During the streaming time, all the available network resources can be fully utilized with the diversified forms. It is not constrained by the infrastructures as the streaming video contents have been spited by the MDC encoder into the picture frame based cells. The data delivery is located at the logically resources. As along as computer resources located in the peer-to-peer network are available for each single cell, the streaming can keep going smoothly.

As illustrated in Figure 3.13, in the right side, the MDC frames are arranged with different I-frame positions, and the left side illustrates the aligned I-frame which is commonly used. Both MDC streams have the network congestions at the same time. And the overall output at the receiver end give out different frame dropping result. The pre-arranged I-frame start-up can also be used to improve the overall transmission, hence to get the better video quality.

Figure 3.14 shows the simulation result of the homogeneously distributed I-frame within 4 sub-streams in red, and the blue line shows the aligned I-frame with the same network transmission condition. The output of the I-frame distribution with the managed manner outperforms the commonly used aligned MDC I-frame.

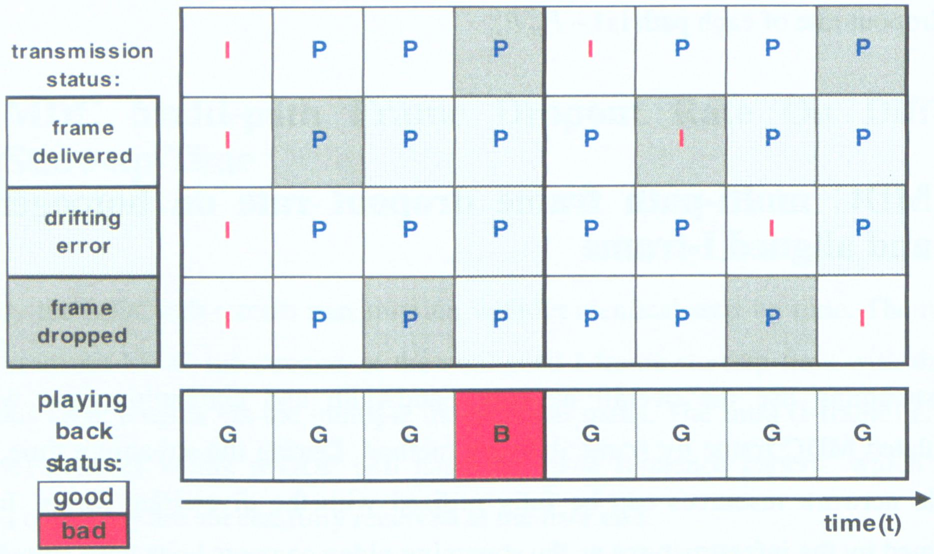

Figure 3.13 MDC playing back: aligned I-frame V.S. un-aligned I-frames

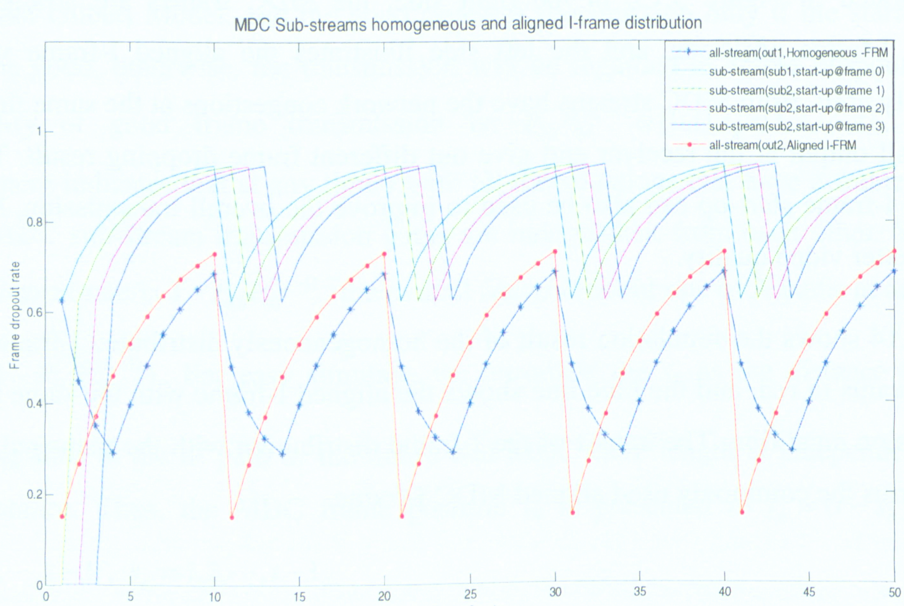

Figure 3.14 MDC frame dropout rate with aligned I-frame V.S. unaligned

In this section, the special scenario of homogeneous start-up offset was studied. As the number of sub-streams is theoretically infinite large, and GOP length has a finite size, usually we have $N \geq g$. When the number of sub-streams N equal to the GOP length g $(N=g)$, and offsets $s_n = n$, we can get a square matrix where the start-up time is homogeneously distributed in the each path and each transmission moment. Then at any time i_n, we can always find a path has offset $n=i$, the frame dropout rate is:

$$D_{homo} = \frac{N}{g}(1-P_{S0}P_{00}^{0})(1-P_{S0}P_{00}^{1})\cdots(1-P_{S0}P_{00}^{g-1}) \text{ as N=g,}$$

$$D_{homo} = (1-P_{S0}P_{00}^{0})(1-P_{S0}P_{00}^{1})\cdots(1-P_{S0}P_{00}^{N-1}) \tag{3.13}$$

let $\alpha_i = (1-P_{S0}P_{00}^{i-1})^N$ then equation (3.6) and (3.13) can be represented respectively as:

$$D_{align} = \frac{1}{N}(\alpha_1 + \alpha_2 + \cdots + \alpha_N) \tag{3.14}$$

$$D_{homo} = \sqrt[N]{\alpha_1}\sqrt[N]{\alpha_2}\cdots\sqrt[N]{\alpha_N} = (\alpha_1\alpha_2\cdots\alpha_N)^{\frac{1}{N}} \tag{3.15}$$

As $\alpha_1, \alpha_2, \cdots, \alpha_N$ is a set of positive real numbers, N is a positive integer, then the arithmetic mean of set $\alpha_1, \alpha_2, \cdots, \alpha_N$ is D_{align}, geometric mean is D_{homo}, then $D_{homo} \leq D_{align}$.

Thus, the homogeneously distribute the start-up time in multiple sub-streams can improve the overall transmission.

3.6 Summary

The overall performance of MDC streaming is driven by the promptly received content as well as the network transmission. In the Internet environment, we can theoretically get infinity transmission channels. Meanwhile, the content that can be decoded depends on the active frames received from each channel at any point of time. Under the same transmission condition, when MDC sub-stream transmission start-up time is different, the output of active frames is different. To get better real-time performance, we can adjust the start-up time in each channel. And the overall output is determined by the consolidation result of network transmission rate, number of transmission path, GOP length and the streaming start-up time. With streaming, the client browser can start displaying the data before the entire file has been transmitted. For streaming to work, the client side receiving the data must be able to collect the data and send it as a steady stream to the application that is processing the data and converting it to pictures. This means that if the streaming client receives the desired information more quickly than required, it needs to save the excess raw data in a buffer before the information converting processes. If the data doesn't come quickly enough, however, the presentation of the data will not be smooth.

This chapter introduces the video streaming in distributed environment with the multi-path streaming infrastructure and the MDC technique. We evaluate the performance of different real-time streaming systems in terms of the frame dropout rate of the reconstructed videos, which reflect the probably of a frozen playback video frame at the receiver devices. We studied different scenarios by those four factors: (1) number of MDC sub-streams, (2) number of streaming paths with the duplicated contents, (3) length of group of pictures, (4) transmissions with the re-organized position of I-frames in multi-paths. The number of sub-stream and length of GOP are the parameters of MDC encoder, usually it is interpreted as computation power and the number of channel with duplicated video content is associated with the overall network transmission capacity. The last but not the least, they are the transmission collaboration with those sub-streaming channels. This can be addressed as to well plan the I-frames start-up time of all the transmission channels.

CHAPTER 4

ADAPTIVE MDC VIDEO STREAMING OF MULTI-CHANNEL BROADCASTING

In today's Internet heterogeneous environment, the resources available to multimedia applications are various at all aspects of the real time network conditions. The physical hardware, logical software and the time of the resource availability are different and keep changing. To adapt to the changes in network conditions, both networking techniques and application layer techniques have been proposed. In this section, we focus on the application techniques, including methods based on MDC compression algorithm features to make the parameter adjustment. And the adaptive error control based on network transmission smoothing is included. We discuss how feedback from lower networking layers can be used by these application-level adaptation schemes to deliver the highest quality content. We describe a prototype implementation of the adaptive video streaming service and present the results of a performance evaluation of this prototype system. This chapter investigates multimedia applications over distributed networks, with the main focus on the MDC H.264 baseline video streaming. We integrate the video encoder and distributed network infrastructures to improve the perceived the video quality. The MDC parameters of video encoding, such as number of sub-streams, length of GOPs are used for the best effort services. And the collaboration with multiple sub-streams by adjusting the I-frame positions are investigated.

4.1 Adaptive Streaming Architecture

Multimedia architectures have undergone significant changes from centralized to distributed, self-organizing structures. Consequently, multimedia content processing has employed new paradigms that include distributed, collaborative, and resource-constrained models, whereas multimedia communications have addressed the new application challenges by focusing on cross-layer design. From the ISO/OSI Network Model lower to top, multiple techniques have be widely adopted. At the data link layer, error control and adaptive reservation techniques can be used to relief the errors in bit rate variance. Dynamic routing mechanisms are used to avoid congestion in the mobile environment at the network layer. At transportation layer, dynamic re-negotiation of connection parameters can be used for adaptation. Applications can use protocols such as Real Time Streaming Protocol (RTSP) [57] and Real Time Protocol (RTP) [58]. At application layer, the application can be adaptive to network condition by using multiple techniques, such as hierarchical coding, efficient compression, bandwidth smoothing, error control and adaptive synchronization.

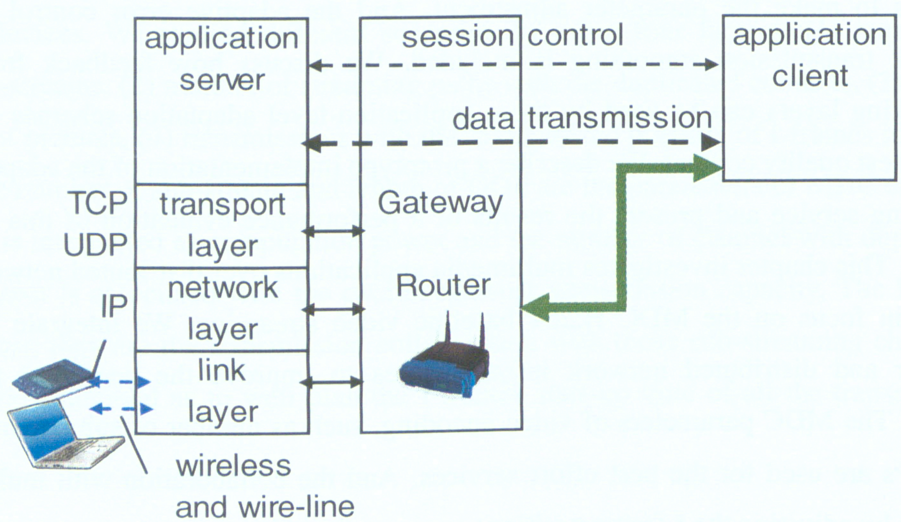

Figure 4.1 streaming architecture

The end-to-end architecture in Figure 4.1 preserves the Internet paradigm of stateless routing with connection oriented services implemented in the terminals. We split the application layer into application server and application client to highlight our main focus on the adaptations of video streaming. Channel-adaptive streaming methods, as discussed above, would be implemented in the client and the server only. In order to solve the problem of sharing bandwidth fairly both in the wire-line and the wireless links, reliable loss differentiation algorithms (LDA) [59] are required that can distinguish loss due to congestion and a deteriorating wireless channel. Wireless media protocol could even be a circuit-switched multimedia protocol stack, such as H.324M [60].Channel-adaptive streaming techniques would be used between the gateway and the streaming media server, while packet-oriented streaming media techniques, such as dynamic packet scheduling, might not be applicable to the wireless link. With H.324M, error resilience of the video stream is important, as is rate scalability or rate control to accommodate variable effective throughput even on a nominally fixed-rate link. The 3G-PP consortium has evolved the ITU-T recommendation H.324M into 3G-324M [61], which also supports MPEG-4 video, in addition to H.263v2, for conversational services.

4.2 Video Coding and Service of Quality

A good adaptive video streaming system usually needs the following two goals on the service improvement. Firstly the efficient content processing, secondly the effective real time network transmission and eventually the overall picture quality will reflect the many aspects in those end to end user environment.

One of the main issues that span each of these areas is the granularity of adaptation. Ideally, a quality-adaptive streaming system will select video quality to match the average available network bandwidth. In practice, adaptation tends to be limited to discrete steps, and consequently the rate match is only approximate. A system that supports steps with finer-granularity generally results in a better match, which manifests itself in higher quality and better reliability of streaming. Frame dropout rate is a well known work-around, and is probably the most popular video adaptation mechanism, having been used since the first quality-adaptive Internet streaming systems appeared [62]. In addition to frame dropping, there are other important ways to adapt video, which will be discussed in the following sections. When multiple adaptations are combined, it can help to increase the granularity of the space of adaptations.

With few exceptions, such as joint source channel coding and adaptive streaming, processing of multimedia content has been approached as a network-independent problem by the research and academic communities. In turn, the works that addressed the content delivery aspects have typically not included the characteristics of the source content and have primarily studied interactions among lower layers of the protocol stack. However, the processing and delivery of multimedia content are not independent and their interaction has a major impact on the Quality-of-Service (QoS) [63] aspects. To address this interaction, network-aware multimedia content processing and delivery methods are needed. Network-aware delivery aspects include solutions with emphasis on a stronger integration among higher layers of the protocol stack including the application layer. Those reconstructed video content will lost due to network congestion or an unacceptable response time that exceed the playback threshold. The network-aware methods make use of the network characteristics in their design, to intelligently take advantage of infrastructure and content characteristics. Network-aware processing aspects include in-network solutions such as event detection, tracking, data aggregation, spatial-temporal query and summarization.

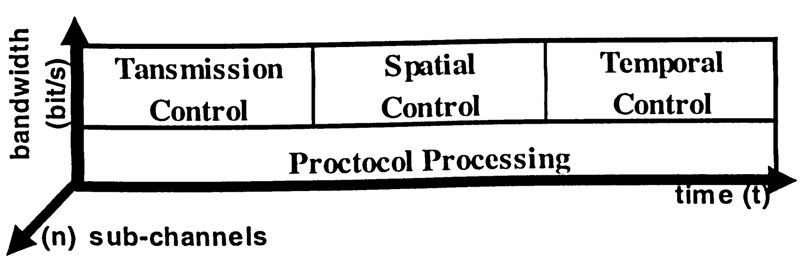

Figure 4.2 MDC Video Streaming Controls

We separated the video streaming congestion control into three components as shown in Figure 4.2. The three components are functionally independent so that they can be designed separately and upgraded asynchronously. During the streaming time *(t)*, the transmission conditions of the sub-channels *(n)* are independent from each other with the separate bandwidth *(bit/s)*. With the MDC channel diversity feature, all the sub-channel transmission can collaborate with each other. The overall quality will be improved with a managed effort of encoding and transferring. The transmission control component determines which packets to transmit, spatial control determines how many sub-pictures to transmit, and temporal control determines how many video predictive frames are suitable. These decisions are made based on information provided by the estimation component. The estimation component computes the feedback information for each sent encoded video frame. The transmission control selects the next video frame to send from two pools of candidates: different spatial compressed frame, and different temporal compressed frame. Transmission control regulates packet transmission at the timescale to smooth the fluctuations in service delivery.

The adaptive video streaming service exploits the inherent IP network environment to perform controlled and graceful adjustments to the perceptual quality of the displayed video in response to fluctuations in the Quality of Service (QoS) delivered by the primary components in the end-to-end path of a video stream. The design supports multiple video adaptation techniques that can be applied individually or in combination to adapt the transmitted video stream in response to fluctuations in the QoS provided by the underlying system. A novel aspect of the design is a video adaptation algorithm that selects the adaptation delivering the best perceptual quality for the video playback for a given QoS delivered by the underlying infrastructure. In networking circles, the QoS term is often used to denote support for various forms of resource reservation. Good support for video was an often touted payoff for the transition to QoS support. A core design is to use of small fixed size packets to fit into the current infrastructure, thus, the predicable service quality can be achieved and the streaming quality can be guaranteed at the certain level. With the MDC adaptive compression, the video streaming can fit into large range of network bandwidth to meet different user's requirements.

4.3 MDC Encoder Adjustments Over Frame Dropout Rate

The video sequence is composed of multiple image frames which are captured on a regular periodic sampling interval, and video coding is the technique to compress the raw video data to eliminate the redundancies presented in the original video sequence. Video start streaming firstly from broadcast video stream with an Intra frame, the real time end user will have a nature transmission wait, such as the routing, computing lags as showed in Figure 4.3. In most cases, those kinds of lags are not perceived and won't affect the end user's satisfaction with the streaming services. Once the end users start receiving packet from the overall transmission channels, they may be benefit from the MDC feasibility to consolidate all the available sub-streaming bandwidth. But either the available contents or the network connections may change at any time (illustrated as multiple blue arrows in figure 4.3), as long as the end user can receive any valid Intra or Inter frame, video can be re-construct back for display. In certain time, all the received frames are not valid, due to the network congestion or due to the fame cannot be used for decoded as the previous Intra frame loss or Inter frame missing, the end user need wait until the next GOP circle for the valid I-frame which can be decoded independently.

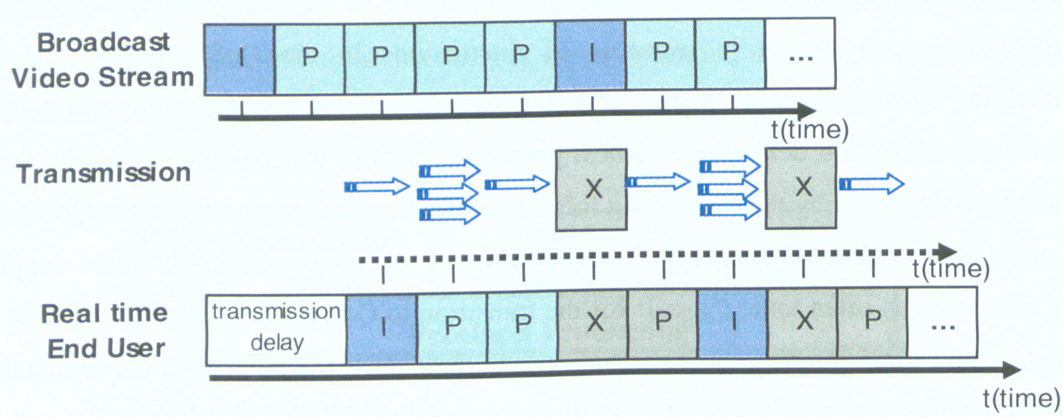

Figure 4.3 GOP of H264 baseline video streaming

The number of P frame is various depending on the real time MDC encoder processing. Meanwhile, those encoded I (P) pictures were transferred over the available IP network channels, which have the independent bandwidths in each transmission channel at each slot of transmission time. Once the encoded video pictures were received at user end, MDC decoder processes those pictures to reconstruct back the video stream. For an alliance based on joint, collaborative innovation, perhaps the best indicator of success is the video stream frame dropout rate.

We proposed an infrastructure which allows the receiver to feedback the streaming condition to the servers, and adaptively adjust the encoded bit streams to improve the performance of the streaming session. Figure 4.4 shows the key parameters of the MDC streaming which are involved in our adaptive frame dropout rate consideration. The original video is encoded by MDC, then, is transferred via multiple paths to end users. Until the video playing back, there are distortions at each step (encoding, transferring, decoding, playing back). As MDC use the theoretically lossless compression, the major challenges in the video quality would be the network transmission. And the MDC has the frame drop drifting errors, so the overall frame dropout rate will be a good subject to be adjusted for better video re-displaying quality.

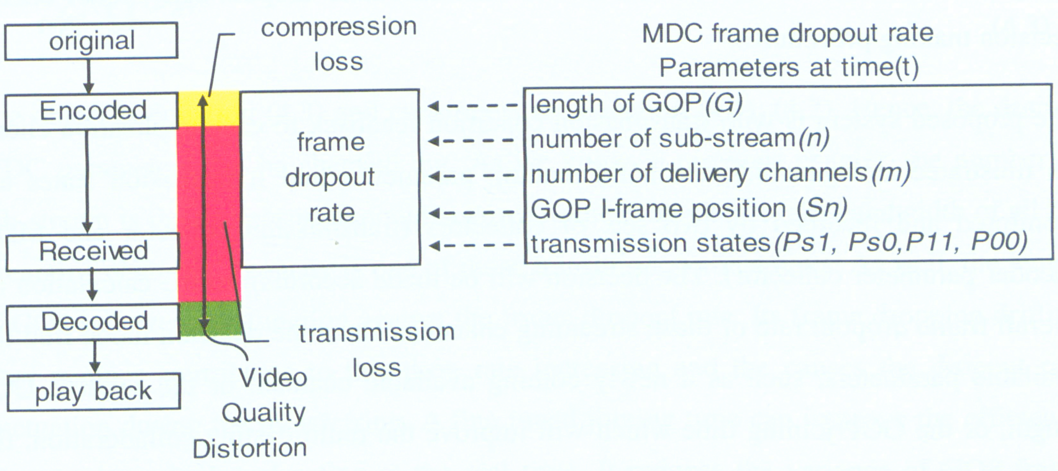

Figure 4.4 frame dropout rate of MDC adaptive streaming

Once the bit rate of encoded video streaming exceed network available bandwidth, for best service quality guarantee, the MDC adaptation needs the prompt relieving from the transmission congestion and the real-time playing back requirements, all the parameters have the time tags to reflect the frame dropout rate instantly as we state in chapter 3.

4.4 Adaptive MDC Sub-Streams with Real Time Feedback

More than a pooling of resources, the alliance provided the framework for truly collaborative innovation that played to the complementary strengths for the benefit of their mutual peers within the video broadcasting network. Under the alliance, each video source is aware to the sub-stream content, and the transmission is managed to achieve the high picture quality or efficient transmission rate.

MDC video streaming has multiple parameters can be adjusted to achieve an optimized broadcasting quality with the same computing and transmission resources consuming. People usually adjust the most significant parameter to get the real-time quick response and positive results. We use the MDC transmission frame dropout rate model on the decision making procedures.

The proposed system is with transmission condition feedback to each application client. As illustrated in figure 4.5, the MDC coding parameters and transmission states are monitored and recorded by two sets of collectors (transmission collector and MDC encoder parameter collector). The decision will be made according to the calculation on overall frame dropout rate of those streaming channels. Once the decision mode find the available parameters, such as a newly coming available channel, or the suitable GOP length, or the GOP joining time which will improve the multi-casting collaboration, the streaming over number of sub-streams *(n)* and the I-frame distribution will be re-organized, the GOP length (g) or the number of picture sub-samples (n) can be changed accordingly to generate the video bit rate fit the bandwidth capacity of multi-channel.

The streaming frame dropout rate is sensitive to the number of sub-streams and the sub-stream joining time. Along with the length of GOP at each sub-stream (gn), the sub-stream joining time is a factor affecting the collaboration with the multiple sub-streams. To decode a video, the GOP is the smallest granulated function unit as it is firstly to be recognized then the picture decompression start. We introduce the GOP joining time (Sn), which also indicates the I-frame position within each GOP, thus the position of each frame (I or P-frame) during the transmission is:

$i_n = \mathrm{mod}[(t - S_n), g_{nt}]$, where gnt is the length of GOP in nth sub-stream at time t.(4.1)

Let Dn denote the frame dropout rate at the nth channel out of the total number of sub-streams (N), at time t, the frame dropout of the nth channel is:

$$D_n = 1 - P_{S0} P_{00}^{i_n - 1} \qquad (4.2)$$

and the frame dropout of all sub-stream (N) is: $\prod_{n=1}^{N}(D_n)$, the average frame dropout rate of N streams over the time t is:

$$D = \frac{1}{t}\sum_{i=1}^{t}\left(\prod_{n=1}^{N}(D_n)\right) \qquad (4.3)$$

The object is equation (4.3) and constraints are equation (4.1), (4.2). Hence, the desired MDC parameters can be checked out. As we know in previous chapter, the number of sub-stream is the decreasing function against frame dropout rate, the bandwidth of all the streaming channels' capacity is one of those transmission constraints. As well, the length of GOP is increasing function against the frame dropout rate. Its frame dropping drifting error manner contributes to this drop rate increasing and the causes the dropout rate fluctuation during the transferring. A fine tuned joining time can improve the efficiency of multiple paths broadcasting at the real time. It reduces the variance of GOP frame dropout drifting by efficiently collaborating with those multiple channels bandwidth. In the real world, once the quality of service has been set with the frame dropping threshold, the streaming parameters adaptation is the service guarantee.

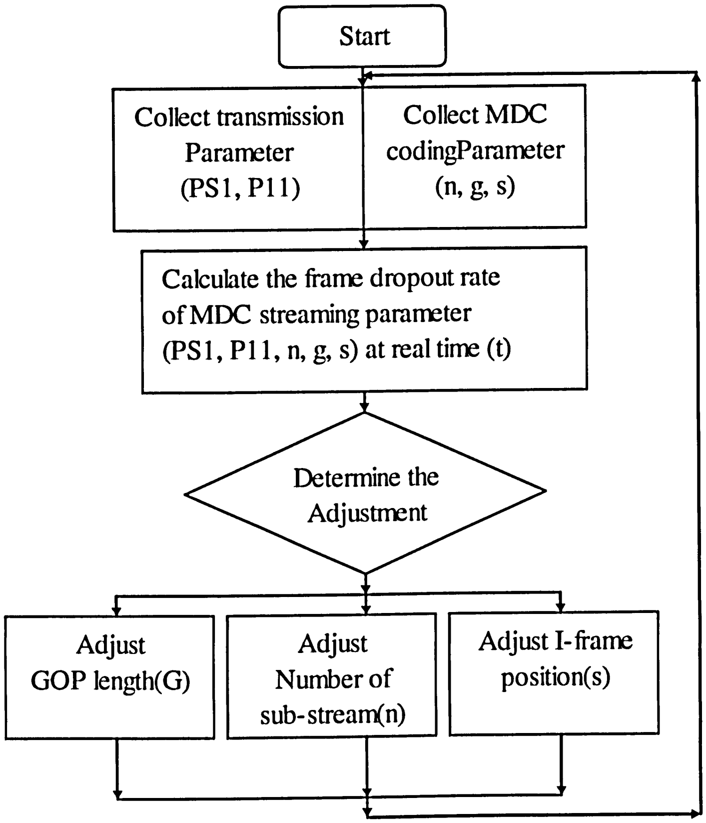

Figure 4.5 Adaptive MDC Streaming Decision Model Work-flow

Simulation is setup as the video sub-streams keep joining the broadcasting network from sub-stream *1* to *4*, and the encoded GOP lengths are initialized at *15*. The probabilities of transmission states remain at bad (P_{11}) increase from *0.2* to *0.6*, the probability of transmission to be state bad (P_{S1}) changed according to P_{11} randomly in each independent sub-channel. The expected Quality of Service threshold is pre-set as frame dropout lower than 70%.

Figure 4.6 illustrates the sub-stream joining scheme at random time. After 2 or more sub-streams join, the frame dropout rate decreases under the pre-set service threshold *0.7*, the overall dropout rate meets the quality of service requirement. While sub-stream joins, the fluctuation of overall frame dropout reduces from the initialized GOP length 15.

Figure 4.7 compares the combination of the adaptations over number of sub-stream *(n)*, GOP Length *(g)* and their unaligned I-frame distribution. Adapted to the P_{SI} and P_{II} increasing, the 4 independent sub-streams decrease the GOP length for less frame dropout rate, we can see the overall output decrease accordingly. Once the all the *4* sub-stream reach a stable network condition, they are with the aligned I-frame, we can see from the frame after *150*, the overall output fluctuation is enlarged by the consolidated *4* sub-streams. There is still chance the service quality threshold to be exceed.

Figure 4.8 illustrates the *4* sub-streams adaptively GOP length change according to a gradually deteriorated channel condition (P_{SI}) increase. After the GOP length changes to the pre-configured smallest *5*, and the GOP join time is arranged to be homogeneous, the frame dropout rate and fluctuation range reach a stable level to provide the streaming services.

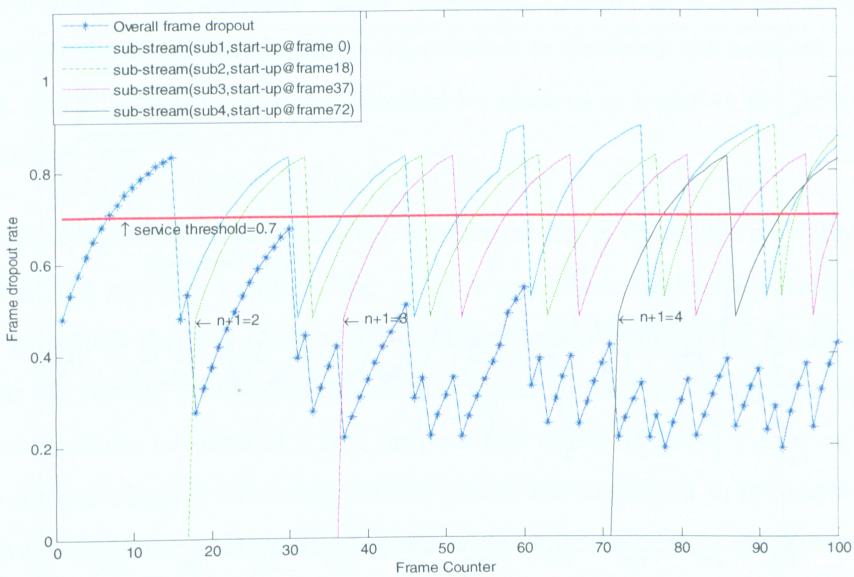

Figure 4.6 MDC adaptive number of channels with random start-up time

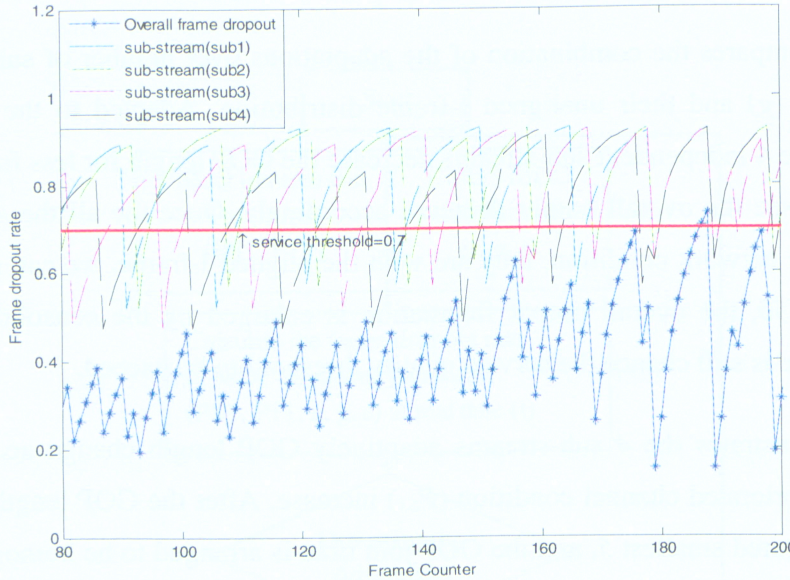

Figure 4.7 MDC adaptive GOP lengths with multiple sub-streams

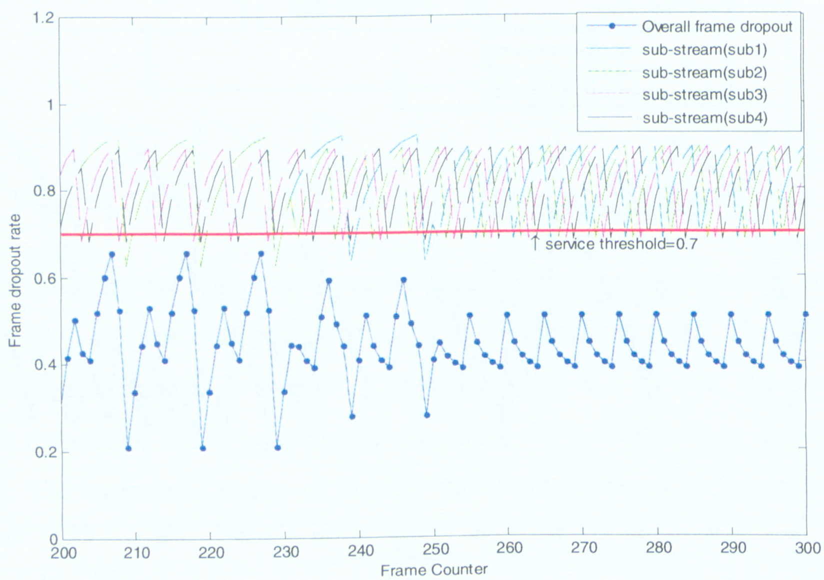

Figure 4.8 MDC adaptive GOP Length with channel collaboration

We observe the frame dropout rate of the multi-channel adaptive streaming can improve the overall result to meet the service quality requirements by adjusting the number of sub-stream and length of GOP according to the transmission conditions. Furthermore, the rate fluctuation on overall output can be tuned to make the streaming quality consistent. The object equation (4.3) provides the scalability of the adjustment choose. In Internet environment, the massive number of content provider has the required potential GOP in all manners, once the adaptation size is determined, the end user can start seeking for the GOP certain size to fit into his own playing back.

4.5 MDC Transmission Status Awareness And Prediction With AR Model

We consider the nature of today's Internet where the delay and throughput values are varied dynamically. An adaptive video streaming system is therefore required. We proposed an infrastructure which allows the receiver to feedback the network condition to the servers, and adaptively adjust the encoded bit streams to improve the performance of the streaming session.

The proposed adaptive streaming decision maker model is illustrated in Figure 4.9. The work flow of the receiver monitors the network condition, and using Auto Regression projection to predict the loss pattern of subsequent video frames. The predicted value is delivered to the servers as a feedback to adjust the GOP lengths. In-depth analysis of the performance measure in terms of frame dropout rate under different scenarios, as well as the details for the adaptive streaming infrastructure, are explained in the subsequent sub-sections. Autoregressive (AR) processes are wide sense stationary stochastic processes that model certain temporal relationships between consecutive samples. There is some dependence of every sample on the previous sample for the network condition changing. The AR model can get the fundamental prediction of the upcoming transmission status. Thus, the MDC can product a different compression to pass through the network congestions or during the thinking time until the next decision making circle, the MDC can find the desired video frames with the suitable size in the broadcasting network, especially in the Internet environment, massive number of channels and resources are available. Compared with the direct network monitoring, it reduces the requirement from the computation consuming encoder by offloading the video compression if there are upcoming desired video content with proper GOP length in the network, the AR model is based on a set of time sequence samples. Until the transmission condition exceeds the threshold, the adjustment is triggered. It can works as a quality guarantee to limit the transmission fluctuations to the acceptable level.

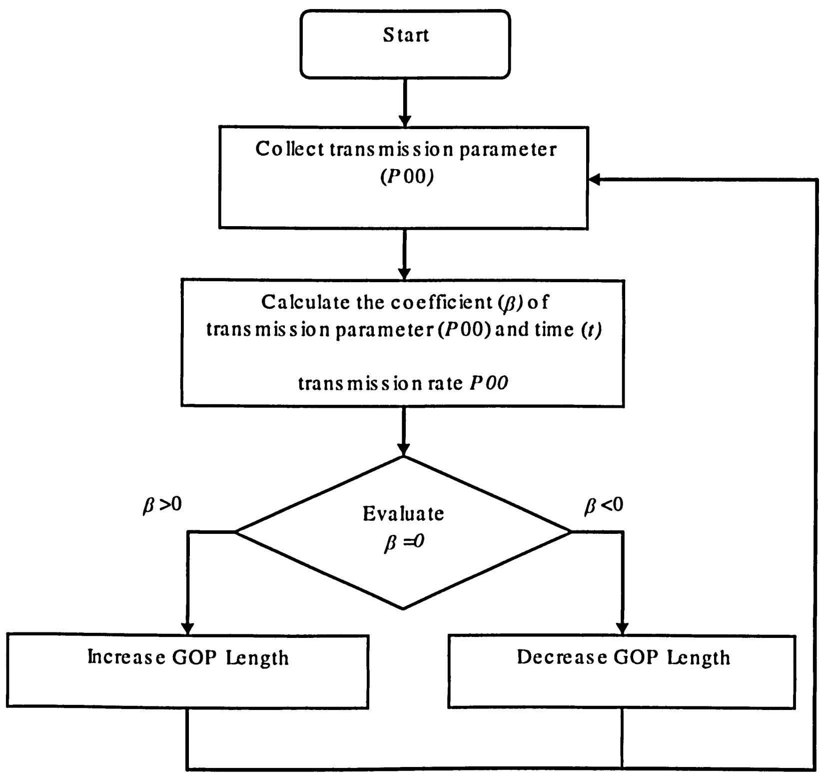

Figure 4.9 network transmission aware with AR prediction

Figure 4.10 compares experimental frame dropout rates from adaptive and fixed streaming systems. The simulation is based Gilbert Model with a fixed good state probability P_{S0} at *0.3*. The transition state probability P_{00} decreases from *0.9* to *0.7*. The initial GOP length equals *15*, and at each transition the GOP length may be adjusted by a length of *5*.

For the fixed approach, we observed that the peak frame dropout rates increase with P_{00} decreases. For the adaptive approach, GOP lengths are adjusted after changes of P_{00} are observed. We noticed that between each P_{00} transitions, multiple I-frames are observed. This is because the linear projection model is evaluated based on ten historical frames in this experiment. The proposed method resets GOP with different lengths multiple times before stabilization.

Figure 4.11 compares different adjustment step sizes. We observed that the frame dropout rate is independent of adjustment step sizes. Also, we observe that for adjustment step size equals 2, frame dropout rate increases when P_{00} decreased from *0.8* to *0.7*, which is because the GOP lengths are increased. Nevertheless, the performance is still better than the fixed solution shown in Figure 4.10.

Multi-path scenarios were studied in Figure 4.12. Two transmission channels are used to transmit duplicated copies of video streams, and both channels are consisting of identical transition state probabilities. We observed that the adaptive approach again yields superior performance over the fixed approach.

In the experiments conducted for Figure 4.10 to 4.12, the system can adaptively vary the GOP lengths at any time. We studied another scenario where the GOP length may only be altered at certain time spots. Figure 4.11 shows the experimental results when the GOP lengths may only be modified at the end of a complete GOP. The GOP length was initialized to 15, and the step size for the GOP length modification is restricted to 5. Therefore, as indicated in Figure 4.12, the I-frames can only occur at frame counts equal to multiples of 5. This approach successfully eliminated multiple I-frames occurring subsequently, hence results in an improved coding efficiency.

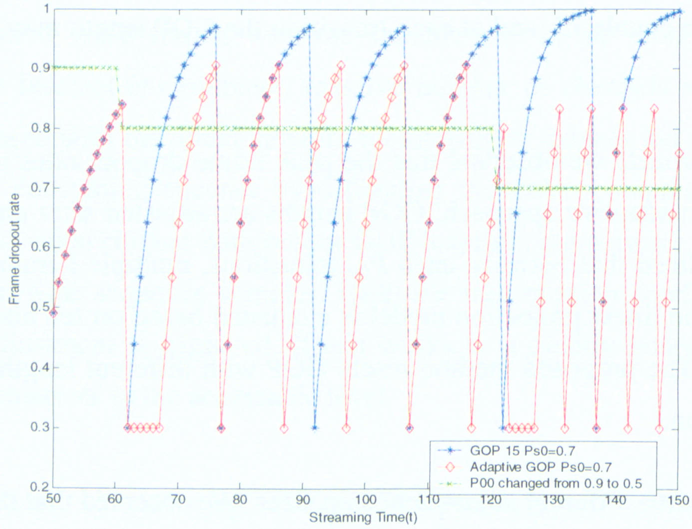

Figure 4.10 Frame dropout rate of adaptive GOP length versus fixed GOP length

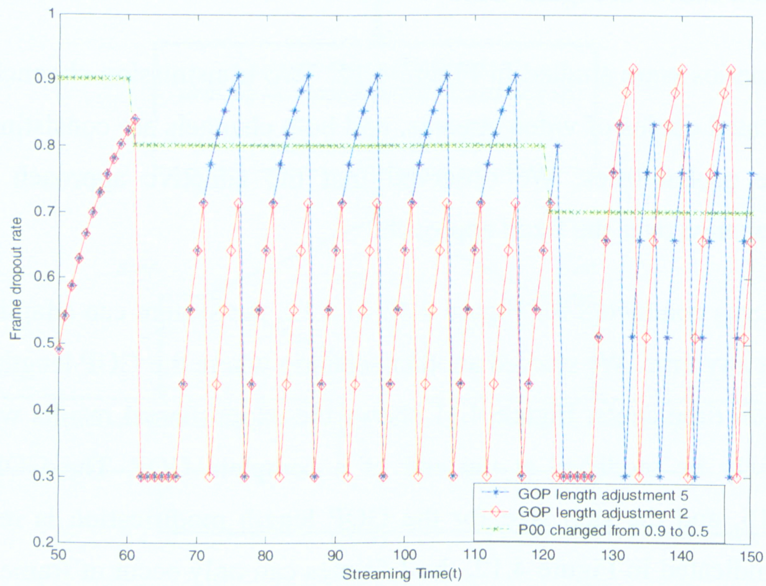

Figure 4.11 Adaptive GOP length with different adjustment step sizes

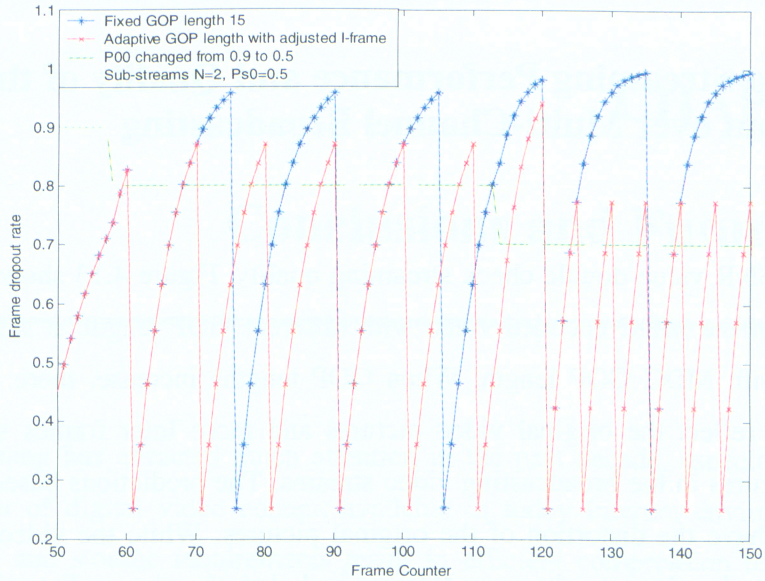

Figure 4.12 multi-path frame dropout rate of GOP adaptive length versus fixed length

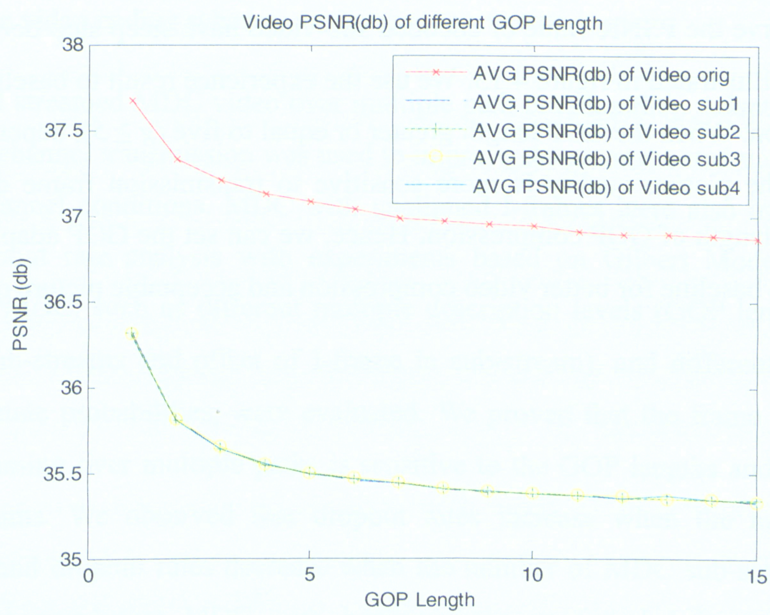

Figure 4.13 Video average PSNR (dB) of the different GOP length and sub-stream

4.6 Video Streaming Performance and Quality of the Overall Output over Multi-Channel Broadcasting

We use the PSNR value double check streaming quality. Figure 4.13 shows the average PSNR value of the MDC encoded video with different GOP length. It shows the close correlations with MDC GOP length. When GOP lengths increase, there are less Intra frames which reflect the original video pictures and more Inter frames which are the calculated pictures in the broadcasting video streams. The predictions based on previous frames (I, P) have the distortion of the original pictures. While the video contents are compressed into less bit rate, the more P frame included, the picture distortion is more in the temporal compression.

The sub-stream PSNR values are included as the broadcasting is based on sub-stream video. We observe the PSNR value of encoded sub-video have steep slop before the fifth frame which is illustrated in figure 4.13. We use the experience result to baseline the GOP length adaptation. When the GOP length greater or equal to five $(g \geq 5)$, depending on the PSNR value, the picture quality is more sensitive to transmission frame dropout rate rather than the length of GOP compression. Hence, we can set the GOP adaptation range greater than the baseline for better video compression and acceptable picture quality.

CHAPTER 5

Conclusions and Future Work

Video streaming has attracted much attention in the past decade, especially due to the large amount of digital video content available in today Internet environment. Video transmission and storage requirements result in efficient compression techniques with many different evolving compression standards, such as H.264 and MPEG-4. Besides efficient compression, video coding techniques have to ensure good video quality while involving real-time processing. Complexity, quality and bit rate are factors that measure success of a video coding scheme.

We studied streamed MDC video over multiple path broadcasting. Adaptive GOP length in the sub-channel transmission was used to improve the reconstructed video quality over varying channel conditions. MDC with unaligned I-frames were also investigated. The frame dropout rate analysis with experiments based on Gilbert Model is examined. Factors of MDC, such as different multiple description levels (GOP length, number of multiple sub-streams and offset of I-frame in sub-stream), and different Gilbert model transition state probabilities, were evaluated. We proved that the frame dropout rate of MDC streaming over multiple paths is sensitive to the GOP lengths and the number of MDC streams. We observed that dropout rates increase when the number of GOP increases; and dropout rates decrease when the number of MDC sub-streams increases. We also observed that MDC with homogeneous unaligned I-frame distribution in transmissions performed better than the aligned cases in our analysis. We investigated adaptive GOP lengths and their impacts on the frame dropout rate. Scenarios under observation include deteriorated network conditions, different adjustment step sizes, as well as multi-path transmissions. We observed that the proposed adaptive solution outperforms the fixed solution in terms of lowering the frame dropout rates.

5.1 Conclusions

All the work in this thesis is directed toward using stochastic modeling techniques to improve the performance of video coding. We target the adaptation of video coding in terms of the complexity-quality-bit rate tradeoff and to achieve this goal we focus both on encoder optimizations as well as decoder post-processing. More specifically, we target the mode decisions in the encoding process and error concealment as part of the decoding process. Simultaneously we also realize that these adaptations require information regarding the network condition. In order to get such information from the network we build accurate models for the video traffic so that these may be used to probe the network.

The first contribution of this thesis is in building a classification based framework for making mode decisions to adaptive the video encoding with transmission quality aware. We use features that may be easily computed from the video data to provide an indication of the cost of making a mode decision. It has the distributed computing manner, can best fit into the under utilized peer-to-peer network environment. This framework is independent of the level at which the mode decision is to be made or the cost that we need to minimize. We use this framework to improve the performance of the Intra-Inter MDC compression decision.

Secondly, we evaluate the MDC frame dropout rate of the multi-channel video streaming. With the channel transmission variances, the adaptation is applied to dynamically allocate the network resources for better streaming quality or to fit the multi-channel diversities. The GOP forms the fundamental function unit rather than a single picture, which affect the overall video playing back due to the picture dropping error drifting, transmission channel collaboration, and multi-cast bandwidth consolidation. The frame dropout rate explicitly reflect the streaming quality and thus the adaptation with MDC video streaming parameters give out direct efforts of the picture playing back.

We propose flexible models to capture the different frame types, activity levels and varying scene lengths present in real video traffic data. We use doubly Markov models with autoregressive, AR processes to capture all these data variations.. These models may be used by the network adaptation, transmission rate aware applications. They may provide certain performance guarantees with the threshold setting. We examine the performance of the models in terms of the stochastic properties of the trace as well as using network simulations.

5.2 Suggestions for Future Research

Due to the massive transmission methods and increasing infrastructure resources, the real-time streaming has more and more flexibilities and solution diversities. The parallel computing, server virtualization and high speed IO/network subsystems free up the application server to end peer user with more functions. An effectively and efficiently resource organizing and dynamically allocation will greatly improve quality of services without extra cost. Streaming application has the character demanding resources at real-time, thus the pre-configured system standard or information delivery method may short the transaction time in general. And the distributed manner requires the logically central managed standard to fine tune the performance at each unit and at each single time.

5.2.1 Video on-demanding

The existing P2P VoD systems can be classified into two categories: buffer-forwarding architecture [64] [65] [66] and storage-forwarding architecture [67] [68] [69]. In buffer-forwarding architectures, each peer buffers the recently received content, and forwards it to the following peers. The participating peers can be organized into a tree-structure for quick index seeking. In storage-forwarding architectures, the video content is distributed over the storage of peers. When a peer wants to watch a video, it first finds the serving peers who are storing the content, and then requests the content from them in parallel. Typically an individual peer may not be able to, or willing to store the whole video. The MDC encoded sub-streams are distributed in the peers, and the receiver connects to multiple serving peers to retrieve distinct sub-streams simultaneously.

5.2.2 Mobile Ad-hoc Network infrastructure Transmission

Mobile Ad-hoc network is a collaborative communication technique within a collection of mobile nodes (MN), without the aid of centralized access point or existing infrastructures [70] [71] The participating mobile nodes can move freely, the network topology varies consistently, and sophisticated routing protocols are mandatory to recover the end-to-end connectivity from path-losses. The optimized routing schema and both LC and MDC is a possible future extension for research works.

5.2.3 Network Coding

The concept of the network coding is first stated in the [72], which proves that with network coding, the achievable throughput of a multicast session is the minimum of the maximum flow from the sender to any receiver. A brief introduction of network coding can be found in [73]. Network coding has been introduced in the application of video streaming. In [74], the authors use the Linear Information Flow (LIF) algorithm to find the network code at each node to maximize the total rate to the heterogeneous receivers in the layered video multicast. However, this method requires centralized computation. [75] introduced a sequence of decoding at the destination, In [76], the authors propose a distributed scheme for practical network coding that obviates the need for centralized knowledge of the network topology. They mention a priority encoding transmission technique that is especially useful in both video and audio broadcast applications. With this concept, further investigation of the combination of the MDC video coding and network coding to deliver the quality video in the lossy transmission

.

5.2.4 Integrated CPU and IO Management

The integrated quality of service management system to manage CPU, network and I/O resources is proposed in [77] to enable the cooperation of multimedia application and OS. This is called Adaptive Quality of service Architecture(AQUA) [78]. In AQUA, the application specifies the desired the QoS when it starts. Once the resource changes, the application re-negotiate with OS and adapt to provide the predictable QoS with the current resource constraints. The AQUA framework can use the network feedback to detect the available resource and predict the QoS. This enriches the flexibility of multimedia application and QoS is manageable.

Bibliography

[1] J. Risson and T. Moors, "Survey of research towards robust peer-to-peer networks: search methods", *tech. rep.,* University of New South Wales, 2004.

[2] D. Wu, Y. T. Hou, W. Zhu, H.-J. Lee, T. Chiang, Y.-Q. Zhang, and H. J. Chao, "On end-to-end architecture for transporting MPEG-4 video over the Internet", IEEE *Transactions Circuits and System for Video Technology*, vol. 18, pp. 977-995, 2000.

[3] D. Wu, Y. T. Hou, W. Zhu, Y.-Q. Zhang, and J. M. Peha, "Streaming video over the Internet: Approaches and directions", *IEEE Transactions Circuits and System for Video Technology*, vol. 11, pp. 282-300, 2001.

[4] S. E. Deering, "Multicast routing in Internet networks and extended LANs", *in Proc. of ACM SIGCOMM 88*, pp. 55-64, 1998.

[5] H. Eriksson, "MBONE: The multicast backbone", *ACM Communication*, vol. 37, pp. 54-60, 1994.

[6] Wood, A.D. Stankovic, J.A., "Denial of service in sensor networks", *Computer*, vol. 35, pp. 54-62, Oct, 2002.

[7] Yau, D. K., Lui, J. C., Liang, F., and Yam, Y., "Defending against distributed denial-of-service attacks with max-min fair server-centric router throttles", *IEEE/ACM Trans. Netw.* 13-1, pp. 29-42, Feb. 2005

[8] M. Roussopoulos, M. Baker, D. S. H. Rosenthal, T. J. Giuli, P. Maniatis, and J. Mogul, "2 p2p or not 2 p2p?", in Proceedings of *International Workshop on Peer-to-Peer Systems*, pp. 33-43, 2004.

[9] "MSN messenger homepage", in http://messenger.msn.com/

[10] "Napster home page", in http://www.napster.com

[11] "Vbuzzer home page", in http://www.vbuzzer.com/

[12] "Skype homepage", in http://www.skype.com

[13] V. K. Goyal, "Multiple description coding: Compression meets the network", *IEEE Signal Processing Magazine*, pp.74-93, September 2001.

[14] T. Nguyen and A. Zakhor, "Distributed video streaming over the Internet", *Proc. of SPIE Conference on Multimedia Computing and Networking*, pp. 186-195, Jan. 2002.

[15] P. A. Chou and K. Ramchandran, "Clustering source/channel rate allocations for receiver-driven multicast with error control under a limited number of streams", *Proc. of IEEE International Conference on Multimedia and Expo*, July 2000.

[16] V. K. Goyal, J. Kovacevic, and M. Vetterli, "Multiple description transform coding: Robustness to erasures using tight frame expansions", *Proc. of IEEE International Symposium on Information Theory*, August 1998.

[17] I. Lee and L. Guan, "Reliable video communication with multi-path streaming", Proc. *of IEEE International Conference on Multimedia and Expo*, July 2005.

[18] I. Lee, Y. He, and L. Guan, "Centralized P2P Streaming with MDC", *Proc. of IEEE Multimedia Signal Processing Workshop*, Shanghai, China, pp. 189-192, Oct 2005.

[19] W. Effelsberg, O. Spaniol, A. Danthine and D. Ferrari, editors, "High-speed networking for Multimedia Applications", *Kluwer Academic Publishers*, 1996.

[20] H. Kanakia, P. Mishra, and A. Reibman, "Adaptive Congestion Control Scheme for Real-time Packet Video Transport", *SIGCOMM Symposium on Communications Architecture and Protocols*, San Francisco, California, pp. 20-31, ACM/IEEE, Sept. 1993.

[21] Klara Nahrstedt and Ralf Steinmetz, "Resource Management in Networked Multimedia Systems", *IEEE Computer*, volume 28, pages 52-63, 1995.

[22] D.D. Clark, D. L. Tennenhouse, "Architectural Considerations for a New Generation of Proctocals", *SIGCOMM'90*, Philadelphia, PA, pp. 200-208, Sep 1990.

[23] J. Roberts, U. Mocci, and J. Virtamo, "Broadband Network Teletraffic Final Report of Action Cost 242", *Springer*, 1996.

[24] R. Steinmetz, "Human Perception of Jitter and Media Synchronization", *IEEE Journal on Selected Areas in Communications*, Vol. 14, pp.61-72, January 1996.

[25] Daniela Rosu, "Dynamic Resource Allocation for Adaptive Real-Time Applications", *thesis fulfillment of the requirements for degree doctor of philosophy in computer science*, Georgia Institute of Technology, Sep. 1999.

[26] X. Wang and H. Schulzrinne, "Comparison of Adaptive Internet Multimedia Applications", *IEICE Trans. Commun.*, Vol. E82-B, No. 6, pp. 806-818, June 1999.

[27] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard", vol. 11, pp. 301-317, 2001.

[28] F. Wu, S. Li, and Y. –Q. Zhang, " A framework for efficient progressive fine granularity scalable video coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, pp. 332-344, 2001.

[29] X. Xu, Y. Wang, S. P. Panwar, and K. W. Ross, "A Peer-to-Peer Video-on-Demand System using Multiple Description Coding and Server Diversity," *in Proc. of IEEE ICIP*, vol. 3, pp. 1759–1762, Oct. 2004.

[30] S. Mao, Y. T. Hou, X. Cheng, H. D. Sherali, and S. F. Midkiff, "Multi-path routing for multiple description video over wireless ad hoc networks", *in Proc. of IEEE INFOCOM*, vol. 1, pp.740-750, Mar.2005.

[31] X. Zhu and B. Girod, "A distributed algorithm for congestionminimized multi-path routing over ad hoc networks", *in Proc. of IEEE ICME*, pp. 1484-1487, Jul. 2005.

[32] S. Mao, S. Lin, S. S. Panwar, Y. Wang, and E. Celebi, "Video transport over ad hoc networks: Multistream coding with multipath transport", *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1721-1737, Dec. 2003.

[33] W. Wei and A. Zakhor, "Robust multipath source routing protocol (RMPSR) for video communication over wireless ad hoc networks", *in Proc. of IEEE ICME*, vol. 2, pp. 1379-1382, Jun. 2004.

[34] R. Puri, K. Ramchandran, "Multiple description source coding using forward error correction codes", *in Proc. of IEEE 33rd silomar Conference on Signals, Systems, and Computers*, vol. 1, pp. 342-346, Oct. 1999.

[35] R. Khalili and K. Salamatian, "On the capacity of multiple input erasure relay channels", *in Proc. WINMEE, RAWNET and NETCOD 2005 Workshops*, Apr. 2005.

[36] R. J. Vanderbei, "Linear programming: foundations and extensions, 2nd Edition", *Springer Press*, 2001.

[37] Y. Wang, G. Wen, S. Wenger, and A. K. Katsaggelos, "Error resilient techniques for video communications", *IEEE Signal Processing Magazine*, vol. 17, pp. 61-82, 2000.

[38] M. C. Hong, L. Kondi, H. Scwab, and A. K. Katsaggelos, "Video error concealment techniques", *Signal processing: Image Communications*, vol. 18, pp. 437-492, 1999.

[39] B. Girod, M. Kalman, N. J. Liang, and R. Zhang, "Advances in channel adaptive video streaming", *in Proc. of IEEE International Conference on Image Processing*, vol. 1, pp. 9-12, 2002.

[40] S. Aramvith, I. M. Pao, and M. T. Sun, "A rate control scheme for video transport over wireless channels", *IEEE Transactions on Circuit System and Video Technologies*, vol. 11, pp. 569-580, 2001.

[41] C. Luna, Y. Eisenberg, R. Berry, T. N. Pappas, and A. K. Katsaggelos, "Joint source coding and data rate adaptation for energy efficient wireless video streaming", *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 1710-1720, 2003.

[42] Q. Zhang, W. Zhu, and Y. -Q. Zhang, "Network adaptive scalable video streaming over 3G wireless network", *in Proc. of IEEE International Conference on Image Processing*, pp. 579-582, 2001.

[43] A. Alatan, M. Zhao, and A. Aliansu, "Unequal error protection of spatial encoded image bit streams", *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 814-818, 2000.

[44] A. Mohr, E. Riskin, and R. Ladner, "Unequal error protection: Grateful degradation of image quality over packet erasure channels through forward error correction", *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 819-828, 2000.

[45] D. Mukherjee and S. Mitra, "A vector set partitioning noisy channel image coder with unequal error protection", *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 829-840, 2000.

[46] REJAIE, R., HANDLEY, M., AND ESTRIN, D. "Quality Adaptation for Congestion Controlled Video Playback over the Internet", *In Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM),* pp. 189 - 200, Cambridge, MA, Oct 1999.

[47] FEAMSTER, N., BANSAL, D., AND BALAKRISHNAN, H. "On the Interactions Between Layered Quality Adaptation and Congestion Control for Streaming Video", *In Proceedings of the International Packet Video Workshop*, Kyongiu, Korea, pp. 128–139, Apr 2001.

[48] DAI, M., AND LOGUINOV, D. "Analysis of Rate-Distortion Functions and Congestion Control in Scalable Internet Video Streaming", *in Proceedings of the International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV)*, pp. 60–71.Jun 2003.

[49] FENG, W., LIU, M., KRISHNASWAMI, B., AND PRABHUDEV, A. "A Priority-Based Technique for the Best-Effort Delivery of Stored Video", *In Proceedings of the SPIE Multimedia Computing and Networking Conference*, San Jose, California, pp.129-139, January 1999.

[50] KANG, S. H., AND ZAKHOR, A. "Packet Scheduling Algorithm for Wireless Video Streaming", *In Proceedings of the International Packet Video Workshop* pp. 121–133, Pittsburgh, April 2002.

[51] Benk, S., "Psycho-visual model of the human optical system as base forlossy-lossless compression of visual information", *Telecommunications in Modern Satellite, Cable and Broadcasting Service 5th International Conference TELSIKS 2001*, vol.22, Issue 2001, pp. 430 - 433.

[52] Li, Z. and Mohapatra, P. "QoS-aware multicasting in DiffServ domains", *SIGCOMM Comput. Commun.* Rev. 34, 5, pp. 47-57, Oct. 2004

[53] López, G., Cánovas, O., Gómez, A. F., Jiménez, J. D., and Marín, R., "A network access control approach based on the AAA architecture and authorization attributes", *Netw. Comput. Appl.*30, 3 pp. 900-919, Aug. 2007.

[54] Wang, Y. and Srinath, M. D., "Error Resilient Packet Video with Unequal Error Protection", *In Proceedings of the Data Compression Conference (DCC '01), IEEE Computer Society*, 523, Washington, DC, March 2001.

[55] H.264 reference model software version JM6.1b, http://bs.hhi.de/~suehring/tml/, March 2003.

[56] H. Schulzrinne, A. Rao, and R. Lanphier, "Real Time Streaming Protocol (RTSP)", *RFC 2326*, Apr 1998.

[57] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol fro real-time applications", *audio-video transport working group. RFC* 1889, Sep 1987.

[58] S. Cen, P. C. Cosman, and G. M. Voelker, "End-to-end differentiation of congestion and wireless losses", *SPIE Multimedia Computing and Networking (MMCN2002)*, San Jose, CA, Jan 18-25, 2002.

[59] N. Farber, B. Girod, and J. Villasenor, "Extensions of the ITU-T Recommendation H.324 for error resilient video transmission", *IEEE Communications Magazine*, vol. 36, no.6, pp. 120-128, June 1998.

[60] J. Wu. And P. Wang, "The performance analysis of sipt signaling system in carrier class voip network", *in Proc. IEEE AINA '03*, vol. 2, mar 2003, pp. 39–44.

[61] CEN, S., PU, C., STAEHLI, R., COWAN, C., AND WALPOLE, J. "A Distributed Real-Time MPEG Video Audio Player", *in Proc. Network and Operating System Support for Digital Audio and Video*, pp. 142–153, 1995.

[62] Li, B., Kalter, W., and Nahrstedt, K. 2000. "A hierarchical Quality of Service control architecture for configurable multimedia applications", *High Speed Netw.* 9, 3,4 (Dec. 2000), 153-174.

[63] T. T. Do, K. A. Hua, and M. A. Tantaoui, "P2VoD: Providing Fault Tolerant Video-on-Demand Streaming in Peer-to-Peer Environment", *in Proc. of IEEE ICC*, vol. 3, pp. 1467–1472, Jun. 2004.

[64] Y. Cui, B. Li, and K. Nahrstedt, "oStream: asynchronous streaming multicast", *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 1, pp. 91–106, Jan. 2004.

[65] Z. Li and A. Mahanti, "A Progressive Flow Auction Approach for Low-Cost On-Demand P2P Media Streaming", *in Proc. of QShine*, Aug. 2006.

[66] C. Huang, J. Li, and K. W. Ross, "Peer-Assisted VoD: Making Internet Video Distribution Cheap", *in Proc. of IPTPS*, Feb. 2007.

[67] Y. Shen, Z. Liu, S. S. Panwar, K. W. Ross, and Y. Wang, "Streaming Layered Encoded Video Using Peers", *in Proc. of IEEE ICME*, Jul. 2005.

[68] Lechner, R. J. Sequentially encoded data structures that support bidirectional scanning", *In Proceedings of the 2nd Annual Symposium on Computer Architecture ISCA '75. ACM*, New York, NY, 188-194.

[69] W. P. Yiu, X. Jin and S. H. Chan, "Distributed storage to support user interactivity in peer-to-peer video streaming", *in Proc. of IEEE ICC*, vol. 1, pp. 55–60, Jun. 2006.

[70] Galluccio, L., Morabito, G., and Palazzo, S. "Spontaneous group management in mobile Ad Hoc networks", *Wirel. Netw.* 10, 4 , pp. 423-438, Jul. 2004.

[71] Wu, K. and Harms, J., "Performance study of proactive flow handoff for mobile ad hoc networks", *Wirel. Netw.* 12, 1, pp. 119-135, Feb. 2006.

[72] R. Ahlswede, N. Cai, S.-Y. R. Li and R. W. Yeung, "Network information flow", *IEEE Trans. on Information Theory,* vol. 46, pp. 1204–1216, Jul. 2000.

[73] C. Fragouli, J. Boudec, J. Widmer, "Network Coding: An Instant Primer", *in Proc. of ACM SIGCOMM*, vol. 36, no. 1, pp. 63–68, Jan. 2006.

[74] N. Sundaram, P. Ramanathan, and S. Banerjee, "Multirate Media Streaming Using Network Coding", *in Proc. of 43rd Allerton Conference on Communication, Control, and Computing*, Sep. 2005.

[75] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding", *in Proc. of 41st Annual Allerton Conference on Communication, Control, and Computing,* Oct. 2003.

[76] Woldegebreal, D. H., Valentin, S., and Karl, H. 2007. "Outage probability analysis of cooperative transmission protocols without and with network coding: inter-user channels based comparison", *In Proceedings of the 10th ACM Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems, MSWiM '07.* 36-44., ACM, New York, NY, 2007.

[77] K. Lakshman, R. Yavatkar, R. Finkel, "Integrated CPU and network –I/O QoS management in an end system", *Computer communications,* 21:325-333, 1998.

[78] Lu, S., Lee, K., and Bharghavan, V. "Adaptive Quality of Service Support for Packet-Switched Wireless Cellular Networks", *Multimedia Tools Appl.* 17, 2-3 (Jul. 2002), 157-179.