INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A

Performance Evaluation and De-jittering for the Transport of Multimedia Traffic over ATM Networks

By

Khaled Shuaib

A dissertation submitted to the Graduate Faculty in Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1999

UMI Number: 9924852

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

This manuscript has been read and accepted for the Graduate Faculty in Engineering in satisfaction of dissertation requirements for the degree of Doctor of Philosophy.

4/19/1999
Date

Chairman of Examination Committee:

Dr. Tarek N. Saadawi

Professor of Electrical Engineering, The City College of The City University of New York.

4-19-1999
Date

Mumtaz K. Kassir
Executive Officer:

Dr. Mumtaz K. Kassir

Dean of the Graduate Studies at the Engineering School, The City College of The City University of New York.

Dr. Myung J. Lee	Department of Electrical Engineering, The City College of The City University of New York.
Dr. Ibrahim W. Habib	Department of Electrical Engineering, The City College of The City University of New York.
Dr. M. Umit Uyar	Department of Electrical Engineering, The City College of The City University of New York.
Dr. Bert Basch	GTE Laboratories, Waltham, Massachusetts.
Dr. Thomas Mills	Army Research Laboratories, Maryland

The City University of New York

Abstract

Performance Evaluation and De-jittering for the Transport of Multimedia Traffic over ATM Networks

By

Khaled Shuaib

Advisor: Professor Tarek N. Saadawi

In this work we evaluate the performance of AAL5 in transporting MPEG-2 streams. In the case of AAL5, the consequences of using the CRC and Length fields of the CPCS PDU Trailer for error detection on the perceptually received video quality are observed. A new cell loss padding technique is proposed, which takes advantage of the Length field in the case of AAL5 to conceal the impact of cell losses. The proposed technique is based on passing the incomplete PDUs and the Length field to the decoder TS layer or to an AAL-SSCS where parsing and padding is performed based on the available information and the kind of incomplete TS packet. The sensitivity of MPEG-2 Transport Stream packet types to communication network impairments specifically bit errors and data loss, is observed and evaluated. Through an intensive experimental study, the characteristics of CDV and CTD in an ATM network transporting CBR traffic under diverse operating conditions are evaluated. We examine the effect of changing network parameters, such as the number of nodes and background load conditions, on the performance parameters of both CDV and CTD. Our results revealed the distributions of CTD and CDV could not be fit into any of the standard distributions (such as normal, gamma, ... etc.). The results also present the auto-correlation behavior of CTD as a function of both load and number of traveled hops. Also presented a verification study of the Square Root Law (SRL) the ATM FORUM proposed to estimate the standard deviation of the CTD for a multi node network. In addition, this work includes a de-jittering scheme, which can be used in the transport of MPEG-4, and MPEG-2 video to absorb any introduced network jitter is proposed. The essence of the de-jittering scheme is based on the statistical approximation of delay variation in the arrival times of video packets carrying encoded clock reference values and a filtering and re-stamping mechanism.

Acknowledgments

This is the right place and time to acknowledge all the persons who directly or indirectly contributed in the successful completion of this work.

First, I would like to thank my supervisor Prof. Tarek Saadawi for his endless support and encouragement, for providing me with great ideas and thoughts, for the time he spent in supervising and directing me to the way of success, and for his patience with me when things got out of hand. I also would like to thank Dr. Mynug Lee for supporting my work and for his invaluable directions. Great thanks to Steve Gringeri and Bert Basch from GTE Laboratories for helping us in choosing the research topic and ideas, and who without them the search for MPEG-2 video streams would have been endless. I would also like to thank the rest of my examination committee, Prof. Uyar, Prof. Habib, and Dr. Mills for taking the time to review my work. Last but not least, many thanks to those responsible for the Advanced Telecommunication Information Distribution Research Program (ATIRP) Consortium sponsored by the U.S. Army Research Laboratory for partially sponsoring my work.

Most of all, I would like to thank my family for their constant support especially, my parents, my wife, my daughter and my son. I probably would have graduated without them, but I wouldn't have nearly as much fun!

Contents

List of Tables

List of Figures

1 Introduction

Asynchronous Transfer Mode (ATM) is considered the core technology for B-ISDN networks, and it is reaching a certain level of maturity that allows for its deployment in local as well as wide area networks. The ATM standards [8,36,38,39] define a fast packet switched network in which data is fragmented into fixed size, 53 byte cells. It defines the manner in which cells are switched and routed through network packet switches and links. The ATM standards are currently servicing and will continue to serve as the transport technology for a wide spectrum of applications and traffic types with different performance requirements ranging from real time video to best-effort data. A key benefit of ATM technology is its ability to provide Quality of Service (QoS) guarantees specific to the application type. Data, voice, and video have different QoS requirements. Real time multimedia applications, such as videoconferencing and video on demand, impose requirements on an inter-network because the traffic they produce must be delivered on a certain schedule and with upper bounds on delay, jitter, bit errors and cell losses to be useful. However, in traditional best effort data services, such as File Transfer Protocol (FTP), or Simple Mail Transfer Protocol (SMTP) in which variation in delay or latency often go unnoticed, and therefore retransmission of corrupted data is always possible.

The ATM Forum Traffic Management group has completed its Traffic Management 4.0 specification [36] which defines the following five service categories: constant bit-rate (CBR), real-time variable bit-rate (rt-VBR), non-real-time variable bit-rate (nrt-VBR), unspecified bit-rate (UBR), and available bit-rate (ABR). The CBR service category is intended to provide circuit emulation for compatibility with existing constant bit-rate synchronous circuits such as T1 or T3 circuits. It is also well suited to carry real-time streams such as audio and video. The CBR traffic class has the highest priority in the network and, therefore, will provide bounded cell delay, cell delay variation, and cell loss characteristics. The VBR service category is divided into two subclasses, one for real-time and the other for non-real-time services. The rt-VBR service category is intended for services that have variable bit-rates combined with stringent real-time requirements. Rt-VBR traffic provides strict bounds on cell delay, cell delay variation, and cell loss. The nrt-VBR service category can be used when timely delivery of information is important, but some amount of jitter can be tolerated. Nrt-VBR does not strictly bound cell delay and cell delay variation like real-time rt-VBR does.

The UBR service category provides no bandwidth guarantees. If congestion occurs in the network, cells will be discarded with no feedback to the sender. UBR is the lowest priority traffic class and provides no quality of service guarantees. The ABR service category is designed for bursty traffic whose bandwidth requirements are roughly known. With ABR service the network will provide a guaranteed minimum bandwidth, but a higher maximum bandwidth is available if the network is not congested. Table 1 provides a summary of the service characteristics for the various service categories.

The ATM Forum and the ITU-T [7,38] have defined QoS parameters which quantify end-to-end network performance at the ATM layer. The QoS parameters include a traffic descriptor which specifies the rate and characteristics of users traffic, negotiable network performance parameters, and fixed non-negotiable characteristics of the network. A traffic descriptor may contain two or more of the following parameters: peak cell rate (PCR), sustained cell rate (SCR), minimum cell rate (MCR), cell delay variation tolerance (CDVT), and maximum burst size (MBS). The traffic descriptor describes the rate (both peak and average), the burstiness, and jitter tolerance for a connection. The PCR defines the maximum instantaneous rate at which a user can send data while the SCR is the long term average rate over an interval. For CBR services the peak and sustained cell rates will be the same while for other service classes the sustained rate will be lower than the peak rate. For each ATM connection, one or more ATM layer QoS objectives are negotiated between the network and the end-system. The network agrees to meet or exceed the negotiated QoS as long as the end-system maintains the negotiated traffic contract. Since QoS commitments are probabilistic in nature and the actual QoS can vary over the duration of the connection, short term performance observations can be worse or better than the agreed upon QoS commitment. QoS commitments can, therefore, only be evaluated over the long term and over multiple connections with similar QoS commitments. ATM uses usage parameter control (UPC) to limit and police traffic at the ingress to the network so that quality of service can be maintained. Traffic is policed in an ATM network using the Generic Cell Rate Algorithm (GCRA). Both CBR and VBR traffic is policed based on a peak cell rate and a cell delay variation tolerance while VBR traffic is also policed for sustained rate and maximum burst size.

Table 1 : ATM Layer Service Categories

Feature	ATM Layer Service Categories				
	CBR	VBR-RT	VBR-NRT	UBR	ABR
Guaranteed bandwidth	Yes	Yes	Yes	No	Minimum
Suitable for bursty traffic	No	No	Yes	Yes	Yes
Feedback based rate control	No	No	No	No	Yes (Using RM cells)
Traffic Descriptor Parameters Policed by the Network	Peak Cell Rate, Cell Delay Variation Tolerance	Peak and Sustained Cell Rates, Maximum Burst Size, Cell Delay Variation Tolerance	Peak and Sustained Cell Rates, Maximum Burst Size, Cell Delay Variation Tolerance	None	Minimum and Peak Cell Rate Cell Delay Variation Tolerance
Network QoS Parameters	Cell Transfer Delay, Cell Delay Variation, Cell Loss Ratio	Cell Transfer Delay, Cell Delay Variation, Cell Loss Ratio	Cell Loss Ratio	None	Cell Loss Ratio

The ATM Forum [38] recommends the use of the CBR traffic class and AAL-5 for the transmission of MPEG-2 video. Other service types such as rt-VBR can also be used as a cost-effective delivery mean of MPEG video. If variable rate MPEG-2 is to be carried using rt-VBR service, it must be shaped to conform to an ATM service contract. When traffic shaping is performed, there is a tradeoff between the burstiness of the traffic, the quality of video, and the buffering requirements. Our discussion in this work is limited to CBR type ATM services for constant rate video transmission. The new emerging video/multimedia applications such as Video On Demand (VOD), video teleconferencing, distance learning, and HDTV are based upon the efficient and reliable transmission of video, audio, graphics, and data over computer networks. In multiplexing, compressing, and transporting such voluminous information over even a reliable network like ATM, problems such as network congestion, and information loss become unavoidable. End-to-end performance and Quality of Presentation (QoP) of the transported traffic over an ATM network is greatly influenced by the choice of an ATM adaptation layer (AAL) [2,8], which enhances the services provided by the ATM layer to support the functions required by the next higher layer. While AAL1 can be used to transport Constant Bit Rate applications, AAL5 was chosen by the ATM Forum to carry both constant and Variable Bit Rate real and non real time services and applications.

This work deals with the transport of real-time MPEG multimedia traffic over an ATM network. MPEG is the most recently standardized and adopted coding system for multimedia applications. MPEG compression and encoding technique is based on exploiting both spatial and temporal redundancies; it achieves compression ratios up to 50:1 and can encode a video or audio source with several levels of qualities. Figure 1 presents a high level over view of how MPEG video is transported over a computer network.

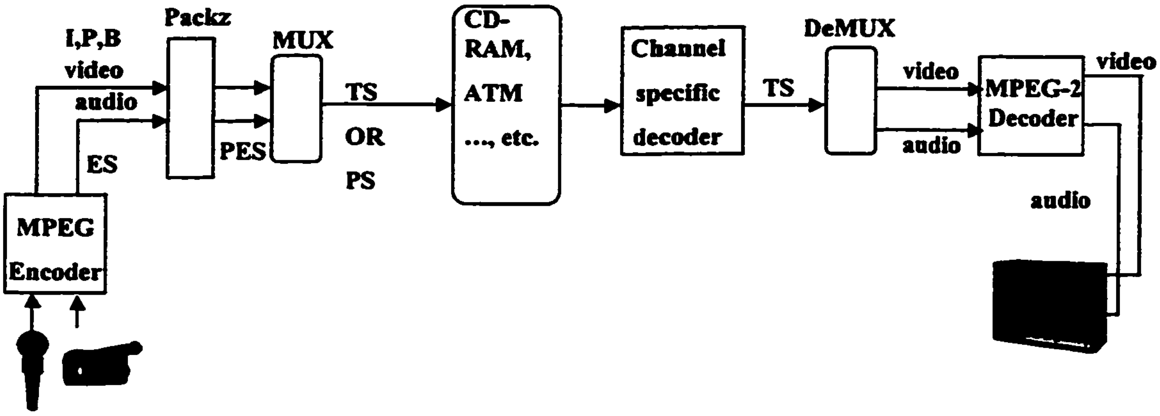


Figure 1: MPEG Transporting Overview

In transporting MPEG over ATM several issues arise that must be addressed in order to achieve the desired quality of service on an end-to-end basis. These include the choice of an ATM adaptation layer, encapsulation methods of MPEG into AAL protocol data units (PDU), the choice of jitter and delay filtering techniques on the cell level, error concealment schemes for data loss, and investigation of the sensitivity of MPEG-2 packets to bit errors and data loss for possible application of selective error correction methods.

In this work we first, address the sensitivity of MPEG-2 Transport Stream packet types to bit errors and cell losses by means of emulating ATM network impairments, and observing their effects on the perceptually received video quality when decoded and presented on a TV monitor.

Second, we evaluate the performance of AAL5 in transporting, real time, MPEG-2 Transport Streams in the presence of bit errors, cell losses and jitter. Third, we propose a cell loss concealment technique which could be used as part of the decoder or the used AAL to minimize the effect of cell losses on the subjectively received and evaluated quality of MPEG-2 Transport Streams.

Fourth, we conduct a performance evaluation of the end-to-end ATM layer in supporting CBR MPEG-2 traffic with the focus on Cell Delay Variation (CDV) and Cell Transfer Delay (CTD) by studying the impact of changing different network parameters (such as background load, and assigned bandwidth) on the CDV and CTD for MPEG-2 CBR connections and consequently on the observed multimedia Quality of Presentation (QoP). We envision that statistical modeling of CDV and CTD based on real time measurements and under various network conditions, is a key issue in designing and minimizing the cost of CDV filtering techniques while preserving the desired multimedia QoP. Fifth, we propose a de-jittering scheme that can be used for both MPEG-2 and MPEG-4 and provide an overview of the MPEG-4 system. Finally, we conclude the thesis along with our future work.

2 Motivation

As the communication networks are forced to support an ever-increasing number of bandwidth-hungry applications, there is an urgent need to efficiently consolidate the existing transmission methods throughout the public and private networks in order to accommodate the new telecommunications economic and marketing imperatives. B-ISDN technology offers such a solution by providing a uniform method of transporting a variety of services including voice, video, data and images using the Asynchronous Transfer Mode (ATM). An important benefit of ATM technology is its ability to provide Quality of Service (QoS) guarantees for any application it supports regardless of how stringent characteristics it may have. Today, ATM is considered the core technology for B-ISDN networks, and is currently commercially deployed in local as well as wide area networks. However, as any other data communication system, ATM networks have their share of common transmission impairments, such as bit errors, cell losses, the end-to-end cell delay, and Cell Delay Variation (CDV). The impact of the cell-specific impairments on the ATM performance needs to be quantified in order to use this technology as efficient as possible.

End-to-end performance and Quality of Presentation (QoP) of the transported MPEG-2 video over an ATM network is greatly influenced by the choice of an ATM adaptation layer (AAL), which enhances the services provided by the ATM layer to support the functions required by the next higher layer. While AAL1 can be used to transport Constant Bit Rate (CBR) video, AAL5 was chosen by the ATM Forum to carry both constant and Variable Bit Rate (VBR) video. The ATM adaptation layer type 5 originally was designed for data application then was chosen by the ATM FORUM for the transport of multimedia traffic for its design and implementation simplicity, without any real evaluation of its performance in doing so. In data application retransmission is always possible since data is not delay sensitive; However, live multimedia application such as video and audio are transmitted under a delay constraint condition where retransmission is usually not possible or very costly. Moreover, all ATM based multimedia devices such as MPEG-2 codecs were designed to support AAL5; an empirical investigation of the performance of ALL5 for multimedia application is a must. The novel aspects of this thesis are as follows:

- Since MPEG-2 was chosen and accepted as the international standardized compression technique for multimedia applications, researchers have worked hard on evaluating its video layer (compression layer); however, little work has been done in the evaluation of the MPEG system layer. In a real audio-visual application, the system layer is used to multiplex different streams such as video and audio. In our work the system layer has been investigated and shown to have a significant impact on the quality of the decoded video and audio.

- A new technique developed by us to improve the received transported video quality under ATM/AAL5 network impairment.

- An Empirical evaluation of the ATM layer in transporting CBR traffic under diverse operating conditions is essential for accurate analytical modeling and characterization of its performance parameters such as end-to-end delay, cell delay variation and cell losses, such evaluation was conducted and results are presented in this thesis.

- A new de-jittering scheme is proposed which could be used to absorb any introduced network jitter and preserve the video quality. The MPEG standards assume a constant delay network and since this is not usually the case de-jittering is a must to preserve the temporal relationship between video objects

3 MPEG Overview

3.1 History and Applications.

Video compression methods use mathematical algorithms to reduce (or compress) video data by eliminating and/or grouping and/or averaging similar data found in the video signal. Numerous attempts have been made to find the best method of capturing, storing and playing back digital video from the desktop. While each attempt brings us closer to the ideal, these advances have come at the cost of incompatibility and conflicting standards. The International Standards Organization (ISO) has worked very hard during this evolutionary period to overcome these conflicts and has been successful in creating internationally accepted standards for digital video.

While the number of competing compression methods has diminished greatly over the last few years, there is still some confusion as to which method is the right one. First rule of thumb is stick to the standards. Standards don't guarantee that the solution is the best, but they are there for a reason. Years ago two large companies fought over Beta versus VHS video tape formats. Beta clearly had better quality, but millions of dollars were lost when VHS was adopted because of other factors and became the defacto standard.

MPEG (Motion Picture Expert Group) is not a defacto standard, it is an internationally accepted ISO standard. Some of the most brilliant minds in the video and computer industries spent several years looking at every possible full motion compression solution and are responsible for the MPEG specification and standard. MPEG is also not an attempt by one company to push a proprietary format on the computer and video industry. This is an open format available to all. Unlike the Beta/VHS example, MPEG not only is the standard, but it is clearly the winner in terms of quality and performance.

There are a wide variety of MPEG resolutions and the ISO MPEG committee has defined several MPEG formats (MPEG-1, MPEG-2, MPEG-4). The MPEG compression algorithm is intended for compression of full-motion video. The compression method uses two forms of energy compaction: prediction is used to exploit temporal redundancy between frames through storing only the differences between them, and discrete cosine transform [1,5,6] is applied spatially within each frame. In MPEG there are three types of video frames: the intraframe coded frames (I), predictive coded frames (P), and bi-directional coded frames (B). The MPEG approach is optimized for motion-intensive video applications [1,5,6], and its specification also include an algorithm for the compression of audio data at the ratios ranging from 5:1 to 10:1.

The MPEG first-phase standard (MPEG-1) [1] is targeted for compression of full-motion video at rates of 1 to 1.5 Mbps in applications, such as interactive multimedia, CD storage, and broadcast television. MPEG-2 standard [5,6] is an extension of MPEG-1 and intended for higher resolutions at bit rates from 1.5 to 15 Mbps in applications, such as video on demand, telemedicine, HDTV, and cable over satellite and broadband network broadcasting. The MPEG-2 standards also address scalable video coding for a variety of applications with different image resolution, such as video communications over ISDN networks using ATM. The MPEG-4 [30,31,32] standard is not yet finalized and although its encoding rage spans a large spectrum (6 Kbits- 35Mbits) it is mostly intended for low bit rate applications from 6-64 Kbps such as video telephony, interactive application over the Internet, and multimedia applications over error prone wireless network.

3.2 MPEG Video Compression and Coding.

3.2.1 Compression Method and Coding Principles.

The goal of video compression is to massively reduce the amount of data required to store or transport the digital video file or sequence of images by means of spatial and temporal redundancy reduction, while retaining the quality of the original video. With this in mind, there are several factors which need to be taken into account when discussing digital video compression.

- Real-time Versus Non-Real-time
- Compression Ratios
- Lossless versus Lossy
- Inter frame Versus Intra frame

- Bit Rate Control

Real-time Versus Non-Real-time - The term "real-time" has been badly abused. In the compression world it means the ability to process 30 frames per second sustained. Some compression systems can capture, compress & store to disk, and decompress & play back video (30 frames per second) all in real time. Other systems are only capable of capturing some of the 30 frames per second and/or are only capable of playing back some of the frames. Insufficient frame rate is one of the most noticeable video deficiencies. Without a minimum of 24 frames per second, the video will be noticeably "jerky". In addition, the missing frames will contain extremely important lip synchronization data. In other words, if the movement of a person's lips is missing due to dropped frames during capture or playback, it is impossible to match the audio with the correct video.

Compression Ratios - A second issue is often referred to when working with compressed video. The compression ratio describes how much smaller than the original video the compressed video is. If the compression ratio is higher, then more compression has been applied to the video. Generally, the higher the compression ratio, the potentially poorer the video quality. With MPEG, compression ratios of 50:1 are common, with good image quality. Motion JPEG (Joint Photographic Experts Group) provides ratios ranging from 5:1 to 20:1, although 10:1 is about the maximum for maintaining a good quality image.

Lossless Versus Lossy - The “loss” factor determines whether there is a loss of data (and therefore quality) between the original image and the image after it has been compressed and played back (decompressed). The more compression there is, the more likely that quality will be affected. Virtually all compression methods lose some data when compressed. Even if the quality difference is visually lossless, these are considered “lossy" compression methods. At this point in time, the only "lossless" algorithms are for still image compression. Using current methods, lossless compression can usually only compress a photorealistic image by a factor of 2:1.

Inter frame Versus Intra frame - This is probably the most widely discussed and debated compression issue, and constitutes the basic difference between MPEG and Motion JPEG. Motion JPEG uses the “intraframe” method, which compresses and stores each video frame as a discrete picture. MPEG, in contrast, uses a combination of the intra frame and “inter frame” method. Inter frame compression is based on the idea that although action is happening, the background in most video scenes remain stable, a great deal of the scene is redundant. Under MPEG, compression is started by creating a reference frame, called an “I” or intra frame and it acts as a random access point. Each subsequent frame of the video is compared to the previous frame and the next frame, and only the differences between the frames is stored. Since only the differences are stored, the amount of data is substantially reduced. This process is repeated for a user definable number of frames (usually 9 to 120 depending on the application), then another "I" frame is created, and the whole process is repeated. This combination of frames is referred to as a "GOP" or Group of Pictures.

Bit Rate Control - The final factor to be aware of with video compression is bit-rate control, which is especially important if your system has a limited bandwidth. A good compression system should allow the user to control the total or average data rate.

Compressed digital video such as MPEG is inherently variable rate since its complexity and motion content affect the encoding bitrate required to maintain picture quality. ATM networks have the flexibility to support both constant rate and variable rate services and can provide statistical gains when using variable rate MPEG. However, uncontrolled burstiness will lead to inefficient use of network resources by occasionally requiring excessively high processing, storage (buffering), and transmission capacity from the network. These requirements motivated the early implementers and providers of digital video services to use a constant bitrate channel for delivering real-time video to customers. Digital video encoder manufacturers, accordingly, implemented additional rate-control and buffering in the encoder to generate a constant bitrate stream for transmission and distribution applications. The fullness of the rate-control buffer dynamically controls the quantization resolution so that the number of bits generated per picture satisfies the bitrate constraints of the video stream. These rate adaptation methods not only lead to variable video quality but also may result in poor utilization of bandwidth since the rate must be selected to accommodate the most complex scenes, and the encoder may need to use stuffing bits to maintain the constant bitrate.

Variable bitrate encoding can achieve improved coding efficiency by better matching the encoding rate to the video complexity. Variable rate encoding is currently used in storage applications such as Digital Versatile Discs (DVDs) to achieve significant storage savings. Using variable bitrate MPEG, it should be possible to achieve substantial savings in networking resources while maintaining video quality, if the burstiness of the video can be controlled using some pre-specified constraints. In addition, it is possible to offer a more constant video quality than can be provided using constant rate encoding. Savings in storage resources, such as server capacity, allow more movies or video sequences to be stored using the same capacity. Similarly, savings in networking resources imply that the same amount of switching, buffering, and transmission capacity can be used to deliver more video content to the customer while maintaining the application level Quality of Service (QoS) requirements. Either source coding or output shaping or a combination of the two can be used for adapting MPEG video streams for transmission over variable bitrate ATM channels. To achieve such savings, it is imperative that we control the burstiness and bitrate variability of a single stream or a set of multiplexed streams in order to make the stream(s) adhere to a traffic contract while maintaining the desired video quality. The bandwidth saving for transmission applications may be smaller than for storage applications due to additional restrictions that the network places on the video's burstiness [17,43].

3.2.2 Basic Coding Units in MPEG.

The MPEG algorithm uses the following basic units as in Table 2:

Block: A block is the smallest coding unit in the MPEG intraframe DCT coded frames. It is made up of 8 x 8 pixels and can be one of three types: luminance (Y), red chrominance (Cr), and blue chrominance (Cb).

Macroblocks: A macroblock is defined as the minimum coded unit and consists of four 8 x 8 blocks of Y, one 8 x 8 block Cr and one 8 x 8 Cb block. The maximum dimension of a macroblock is 16 pixels, and each video picture is divided into a series of macroblocks, from left to right and from top to bottom.

Table 2: Basic Units of MPEG-2 Video Encoding

Basic Units of MPEG-2 Video Encoding	
Sequence Layer	Random Access Unit: Context
Group of Pictures Layer	Random Access Unit: Video Coding
Slice Layer	Primary Coding Unit
Macroblock	Motion Compensation Unit
Block	DCT Unit

Slice: A slice is the basic resynchronization unit since coding of each slice is done independently from its adjacent one. A slice forms the horizontal strip of a video picture and it is the main processing unit in MPEG.

Picture: A single frame of a video sequence in MPEG is called a picture.

Group-of-Pictures: A Group-of-Pictures (GOP) is a sequence of coded pictures arranged in a specific pattern for example IBBPBBPBBI. It provides random access and its characterized by tow parameters: M, the distance between an I frame and the next P frame, typical values (0,1,2); and N the distance between two consecutive I frames, typical values (0 to 15).

Sequence: A sequence is a series of GOPs.

3.2.3 Picture Types in MPEG-2.

MPEG is quite flexible and offer trade off between coding efficiency and random access based on application-specific parameters. MPEG-2 defines three different types of pictures:

Intra-picture s (I-pictures): I-Pictures are compressed using intra frame coding (spatial redundancy reduction only); that is, they do not reference any other pictures in the coded bit stream. They provide for random access, but offer only moderate compression.

Predicted pictures (P-pictures): P-pictures are coded using motion-compensated prediction from past I-pictures or P-pictures. The compression for P-pictures is higher than for I-pictures and both spatial and temporal redundancy reductions are achieved. P-pictures can be used as reference point for additional motion compensation.

Bi-directionally predicted pictures (B-pictures): B-pictures provide the highest degree of compression. They are coded using motion-compensated prediction from either past and/or future I or P pictures. Since B-pictures are not used in the prediction of other B or P pictures, such pictures can accommodate more distortion and hence yield more compression than I or P pictures. Moreover, because B-pictures reference both past and future pictures the coder has to reorder the pictures that are involved in the codec process so that each B-picture is produced after all the referenced pictures. This introduces a reordering delay which depends on the interval between consecutive B-frames. Figure 2 shows a typical MPEG-2 video sequence and the relationship among picture types. The I-picture is coded first then next the P-picture and then the interpolated B-picture between the two.

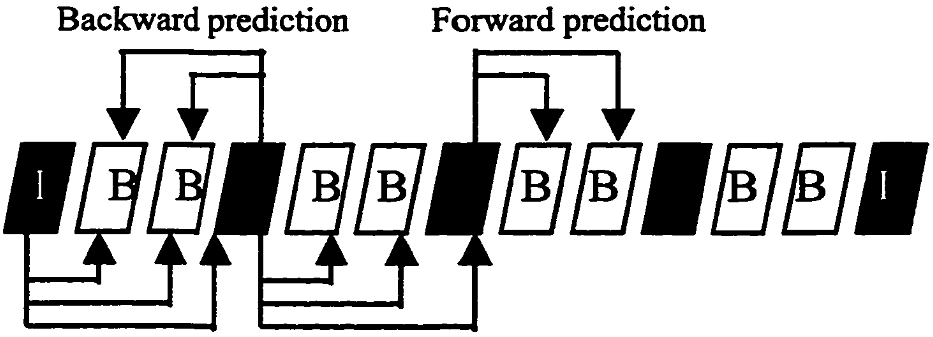


Figure 2: Frames Types in MPEG

3.2.4 Statistical Characteristics of MPEG Video.

Traffic from video sources is envisaged to be a substantial portion of the traffic in broadband networks [17,18,21]. Statistical sources models of different types of video traffic are needed to design networks that achieve acceptable picture quality at a minimum cost, and to control and shape the output rate. For instance, source models are needed to do admission control and bandwidth allocation [24,25,37,43]. Since a significant fraction of video traffic is expected to be generated by MPEG sources, it is important to understand the statistics of traffic from MPEG sources. It is not our intend in this thesis to model sources of MPEG video, but to show statistics of some MPEG video clips and their bursty nature. For example, Figure 3 and Figure 4 show how the bit distribution and burstiness of MPEG-2 clips depending on the video content. The MPEG-2 video clips were: an action movie encoded as a variable bit rate with a peak of 6 Mbits and an average rate of 3 Mbits and a talking head clip encoded as a variable bit rate with a peak of 3 Mbits and an average rate of 2 Mbits. Both clips were 15 minute long encoded using the same GOP structure IPBBPBBPBBPBBPBBI using a commercial MPEG-2 encoder. Figure shows the bit distribution of an MPEG-4 video clip encoded at 30 fps, with a GOP structure IPBBPBBPBBPBBI. The MPEG-4 video clip was a 15 minute action movie clip encoded as a variable bit rate with a peak of 268 kbps and an average rate of 75 kbps using a software MPEG-4 encoder. As seen from Figure 3 and Figure 4, the MPEG-2 action movie exhibit much more variation and burstiness in its video frame sizes than the talking head clip which is expected due to the fact that action video has more frequent scene changes and fast motion scenes than talking head. Similar characteristics of burstiness and wide frame bit distribution as that of the MPEG-2 action movie clip can be seen in Figure 5 for the low bit rate MPEG-4 action movie clip. For more details on the statistical modeling of MPEG video sources the reader is referred to [24,25,27].

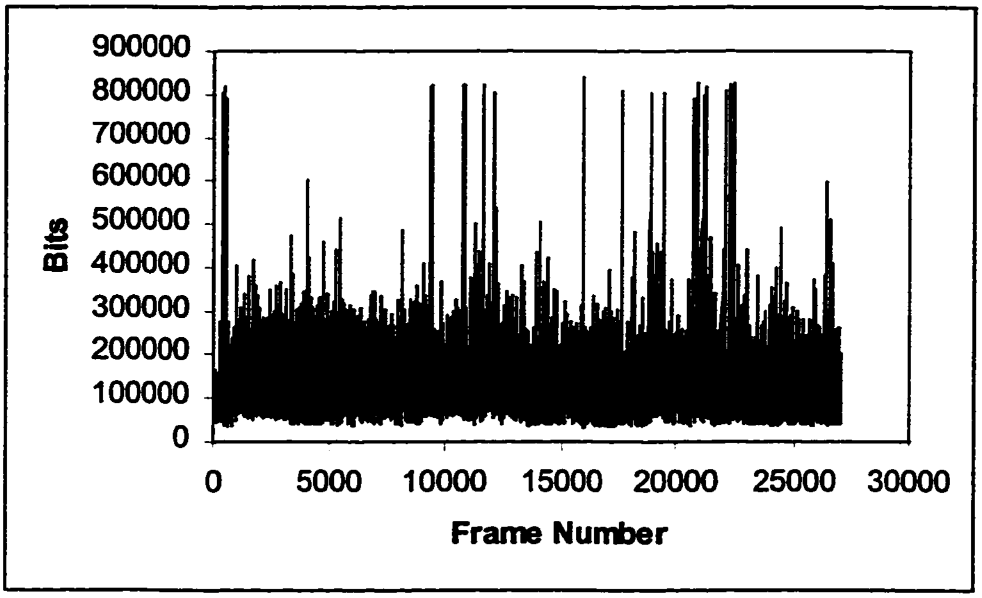


Figure 3: Bit Distribution of the MPEG-2 Action Movie Clip

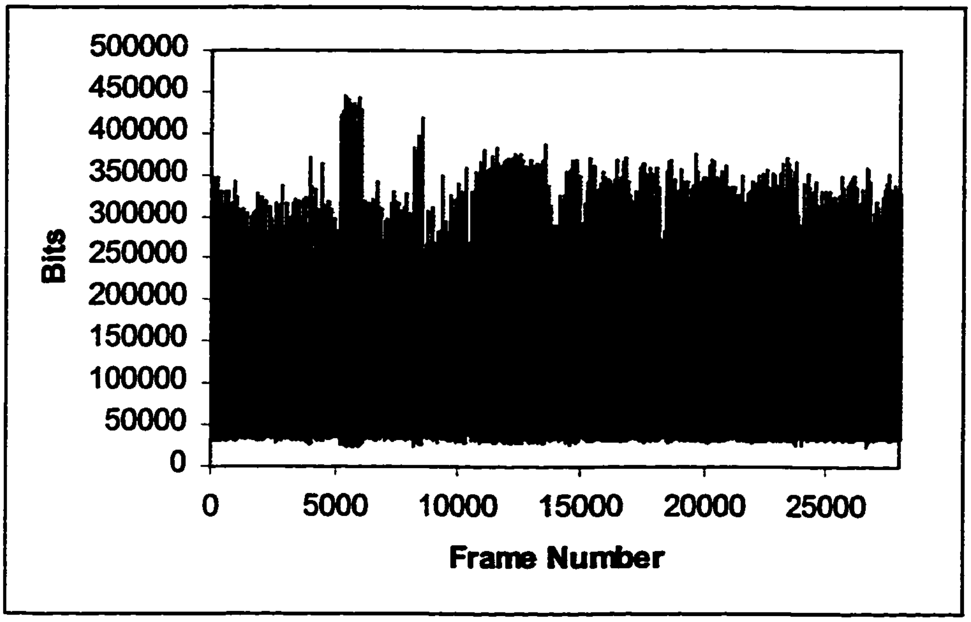


Figure 4: Bit Distribution of the MPEG-2 Talking Head Video Clip

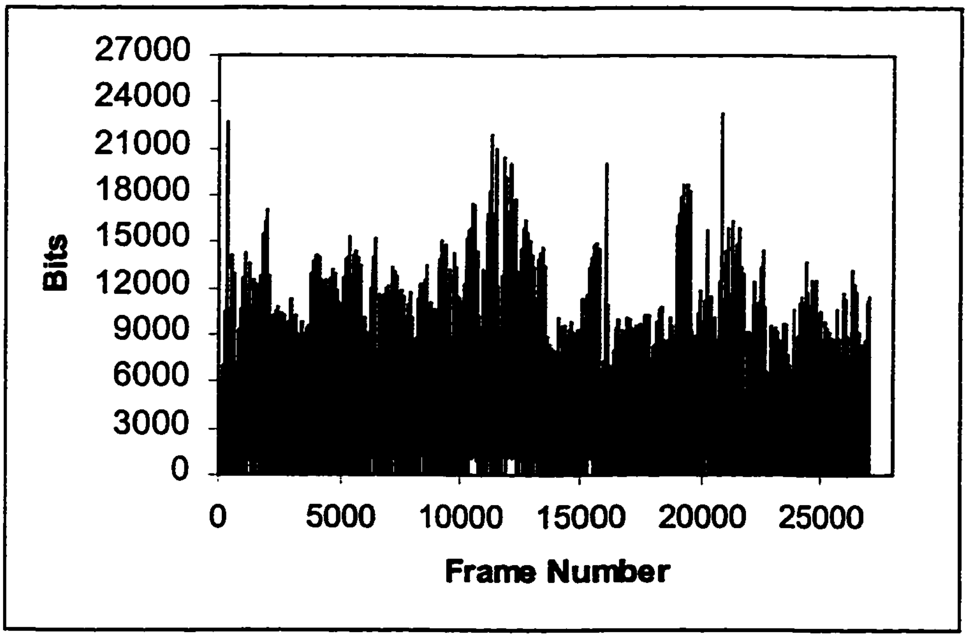


Figure 5: Bit Distribution of the MPEG-4 Action Movie Video Clip

3.2.5 Compression Stages in MPEG.

Compression of digital data is based on various computational algorithms, which can be implemented either in software or in hardware. To achieve the desired compression ratio both spatial and temporal redundancy reduction techniques are needed. Here we discuss the MPEG encoder structure and its compression stages. Figure 6 describes a block diagram of an MPEG encoder with its various stages. The idea behind the buffer at the output of the encoder is to maintain a constant output data rate. When it begins to get empty, it sends an underflow message to the quantizer to increase the bit rate, and when it begins to get full an overflow message is sent to the quantizer to reduce its bit rate.

It is important to understand that the MPEG standards do not define an encoding process, but they specify the syntax of coded bit stream.

DCT coding:

The Discrete Cosine Transform (DCT) is based on and related to the Discrete Fourier Transform (DFT). In principle, the DCT introduces no loss to the source image samples; it merely transforms them to a domain in which they can be more efficiently encoded. To transform and concentrate the video energy of an image into few coefficients, the actual image is divided into blocks of 8 x 8 and then transformed using a two dimensional (2D) DCT which can be considered as a one dimensional DCT on the columns and a one dimensional DCT on the rows. Each 8 x 8 block of source image samples is effectively a 64-point discrete signal which is a function of two spatial dimensions.

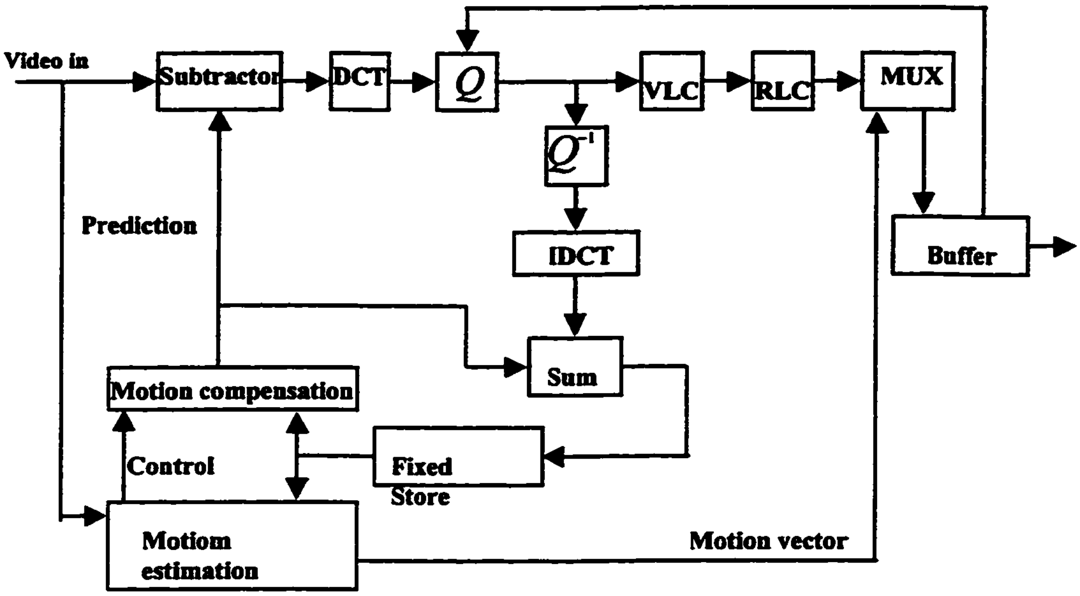


Figure 6: Typical MPEG Encoder

The DCT takes such a signal as its input and decomposes it into 64 orthogonal basis signals. Each contains one of the 64 unique two dimensional spatial frequencies which comprise the input signal spectrum. The output of the DCT is the set of 64 basis signal amplitudes or "DCT coefficients" whose values are uniquely determined by the particular 64-point input signal.

The DCT coefficient values can thus be regarded as the relative amounts of the 2D spatial frequencies contained in the 64-point input signal. The coefficient with zero frequency in both dimensions is called the "DC coefficient" and the remaining 63 coefficients are called the "AC coefficients". Because sample values typically vary slowly from point to point across an image, the DCT processing step lays the foundation for achieving data compression by concentrating most of the signal in the lower spatial frequencies. For a typical 8 x 8 sample block from a typical source image, most of the spatial frequencies have zero or near-zero amplitude and need not be encoded

Quantization:

The next stage of the encoding process after the DCT is Quantization. Quantization is defined as division of each DCT coefficient by its corresponding quantizer step size using a 64-element quantization table specified by the application, followed by rounding to the nearest integer. Quantization is a many-to-one mapping, and therefore is fundamentally lossy. It is the principal source of lossiness in DCT-based encoders. The quantization reduces the amplitudes of the coefficients which contribute little or nothing to the quality of the image, with the purpose of increasing the number of zero coefficients. Quantization also discards information which is not visually significant.

When the aim is to compress the image as much as possible without visible artifacts, each step size ideally should be chosen as the perceptual threshold for the visual contribution of its corresponding cosine basis function. These thresholds are also functions of the source image characteristics, and display application.

In MPEG-2 and when both iner- and intra coded pictures are present, quantization is done based on the picture type. The intra coded blocks contains energy in all frequencies and are very likely to produce "blocking effects" if too coarsely quantized; on the other hand, prediction error type blocks (inter coded) contain predominantly high frequencies and can be subjected to much coarser quantization

Entropy Coding:

The final DCT-based encoder processing step is entropy coding. This step achieves additional compression by encoding the quantized DCT coefficients more compactly based on their statistical characteristics. Compression is done by serializing the DCT coefficients and arranging them in a zig-zag manner as shown in Figure 7. The scanning starts from the DC coefficient and follows the zig-zag pattern till it reaches the last AC coefficient

In MPEG-2 an alternative scanning method might be used as shown in Figure 7. The sequence of coefficient is then entropy coded using a VLC. The output of the VLC consists of a string of zeroes which can be increased by the data scanning method. The string of zeroes is then replaced with a code length representing the number of zeroes in the string.

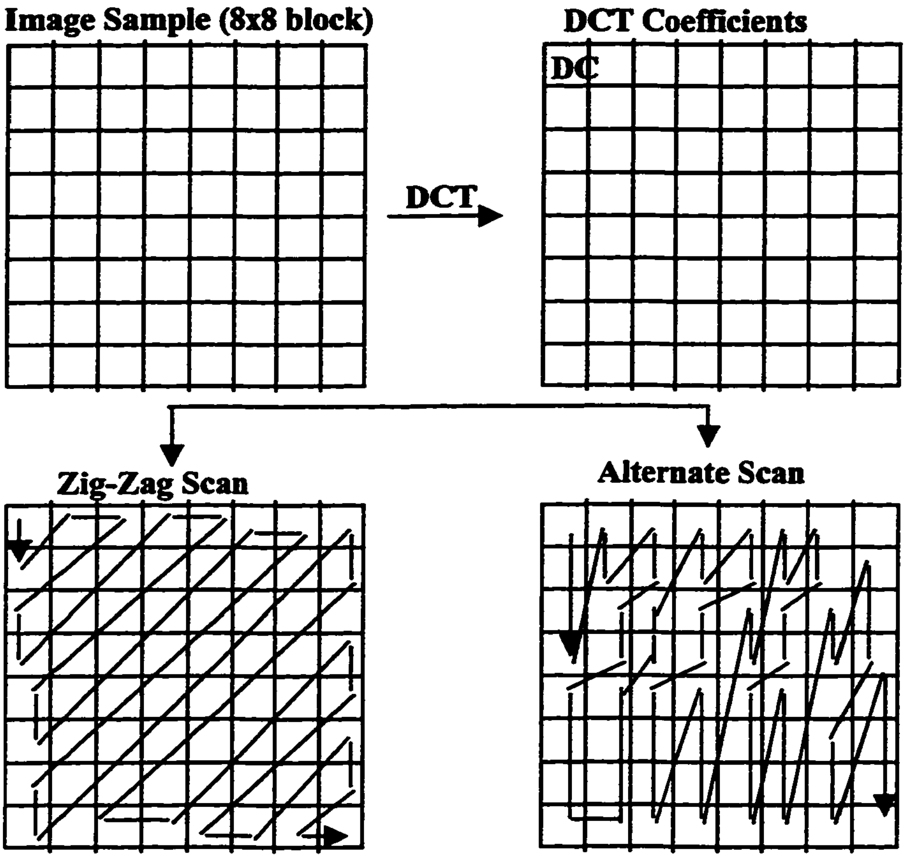


Figure 7: Scanning in MPEG

Motion Compensation and Estimation:

For temporal redundancy reduction of a video signal, MPEG uses the motion compensation prediction (P) and interpolation or bi-directional (B) techniques (inter-picture coding). The basic motion compensation units used are called macroblocks (16 x 16 blocks). Prediction motion compensation is based on the assumption that the current picture can be modeled as a translation of the picture at some previous time. Interpolation motion compensation is based on the reconstruction of the video signal by adding a correction term to combination of past and future reference. Although interpolative prediction introduces coding delay and complexity, it has great advantages such as: noise reduction because of the averaging process, complete coverage of uncovered scenes, and bit rate reduction.

The coding process for P and B frames includes the motion estimator, which finds the best matching macroblock in the available reference frames as shown in Figure 8. Motion estimation is used to extract the information from video sequence. For every 16 x 16 block (macroblock) of P and B frames, one or two motion vectors are calculated. One motion vector is calculated for P frames, while two motion vectors are calculated for interpolated B frames.

The MPEG standards do not specify the motion estimation technique, however block matching techniques are usually used. In block matching techniques, the goal is to estimate the motion of a block of size (n x m) in the present frame in relation to the pixels of the previous or future frames. The block is compared with a corresponding block within a search area of size (m + 2p x n + 2p) in the previous or the future frame, as seen in Figure 8. In a typical MPEG system, a macroblock or a match block is 16 x 16 pixels i.e. n = m = 16, and the parameter p = 6 (Figure 8).

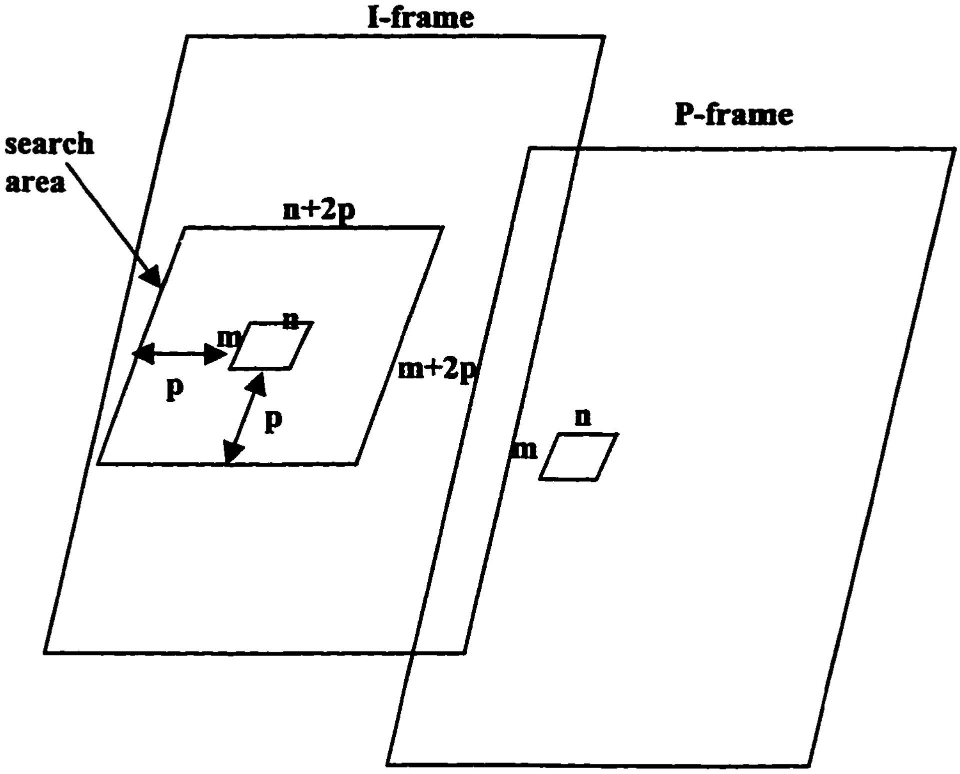


Figure 8: Motion Compensation and Prediction in MPEG

MPEG-2 Scalability Features:

One of the most attractive features about MPEG-2 is its scalable syntax. Scalable syntax [6,20,21,26,27] allows reconstruction of useful video with a user defined acceptable video quality from the total bit stream. Different scalability schemes are used for modeling MPEG-2 video traffic as two layers. This is achieved by structuring the total bit stream as layers, starting with the base layer and adding a number of enhancement layers as desired by the application type. In order to achieve a scalable syntax codecs must support Layered Source Coding (LSC) techniques. With LSC it is possible to generate and control a high priority (HP) layer (or base layer) which contains the most important video information and whose transport over a communication network should be guaranteed to ensure a user defined acceptable video quality.

The difference between the HP layer and the original bit stream is to be encoded as a low priority (LP) enhancement layer and to be transmitter on demand or as best effort traffic since loss of video at this level is not so critical. Many forms of scalability exist. For example, temporal, spatial, and SNR scalability provide variations in temporal, spatial, and amplitude resolution of the video.

3.3 MPEG-2 System Layer.

The MPEG-2 (ISO 13818-2) is known as the compression specification for audio and video, where the MPEG-2 system layer specification is known as (ISO 13818-1). The system layer provides a transmission medium independent coding and encapsulation technique to build bit streams containing one or more MPEG programs. This is made possible through a formal grammar and a set of semantic rules for the construction of bit streams that include provisions for the following functions:

1) Synchronization of multiple compressed streams on decoding.
2) Interleaving and multiplexing of multiple compressed streams into a single stream.
3) Initialization and management of buffering at the decoder.
4) Clock recovery and time identification
5) Error detection.
6) Encryption, and data privatization.

Each program in MPEG-2 is composed of one or several Elementary Streams (ES) with a common time base. An Elementary Stream is a generic term for any single encoded bit stream (audio, video, or data) with a common time base. An Elementary Stream consists of successive access units (video pictures or audio frames). Each ES is packetized into a Packetized Elementary Stream (PES), the multiplexed with other PES to form either a Program Stream (PS) or a Transport Stream (TS). The process describing the different stages of the MPEG-2 system layer is described in Figure 9.

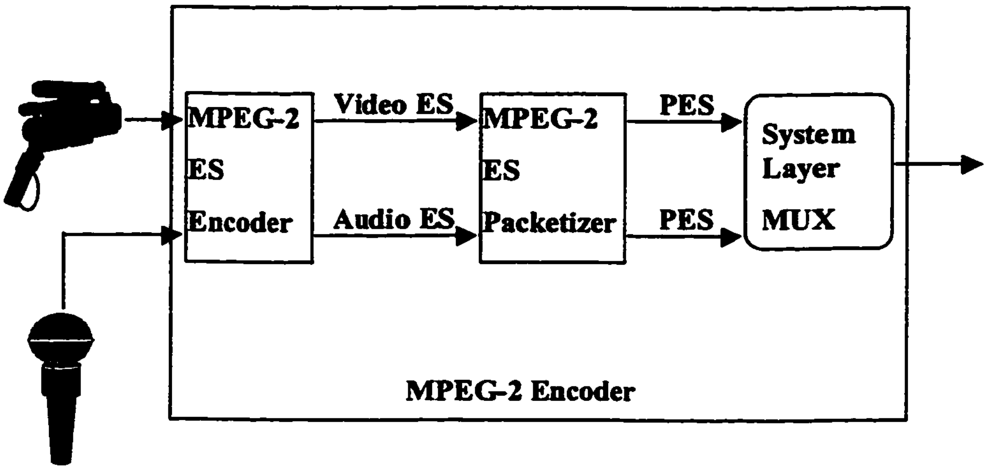


Figure 9: MPEG-2 System Layer

3.3.1 Packetized Elementary Stream (PES).

The MPEG-2 Elementary Streams are packetized individually to form individual Packetized Elementary Streams. The structure of a PES provides a method of packaging subparts of a longer video, audio, or data ES where timing and header indicators are added to facilitate re-assembly and decoding at the decoder side. PES packets are of either constant or variable length, and the length was not specified by the MPEG-2 standards. A PES packet consists of a header and a payload. While the payload is nothing but information from an ES, the header contains vital and crucial information and indicators. For example, an optional Cyclic Redundancy Check (CRC) might be included for error detection and protection mechanisms. A PES header starts with a 24 bit field to identify the start of it, the Stream_ID 8 bit field specifies the type and number of the ES in the PES packet, and the PES_Packet_Length 16 bit field is to indicate the number of bytes in the PES packet following its last byte. Among the many optional PES header fields are the following:

- Scrambling _Control field (2 bits) for encryption.
- Priority_Indicator field (1 bit) to indicate priority among PES packets.
- PTS_DTS field (2bits) and takes one of the following values:
    1. "00" : No Presentation Time Stamp (PTS) or Decoding Time Stamp (DTS) is present in the current PES packet.
    2. "10" : Only a PTS present in the current PES packet.
    3. "11" : Both PTS and DTS are present in the current PES packet.
    4. "01" : Forbidden, reserved for future applications.

Figure 10 describes the structure of a PES, and shows some of the header fields. For a complete PES structure format the reader is referred to [6].

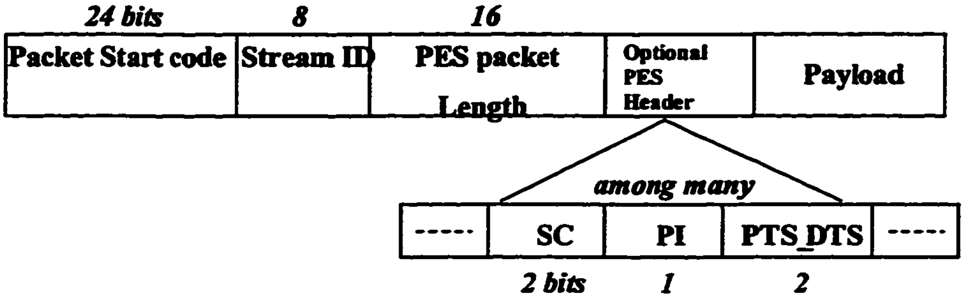


Figure 10: MPEG-2 PES Packet Format

3.3.2 Program Stream (PS).

A program stream is defined as a collection of one or more Elementary Streams with a common time base. A PS is similar to the MPEG-1 syntax since it utilizes variable length packets called packs and each pack consist of a number of constant or variable size multiplexed PES packets. Because of the long size packs used, a loss of one pack may lead to a loss of a video frame, therefore; a PS is only suitable for an error free environment. The variability in the pack size makes the MPEG decoder dependent on a pack-length field encoded within the pack header to distinguish the boundaries (start and finish) of a pack, therefore; a PS is intended for synchronized applications. Relevant PS applications would include program storage and playback from a local digital storage media such as a CD ROM or a local video on demand server.

3.3.3 Transport Stream.

The MPEG-2 Transport Stream (TS) is a multiplex of a one or more digital programs which may consist of an ES, a PS or a TS. Each program may have its own individual time base or it may share a time base with another program. In the TS syntax, PES packets are segmented and encapsulated into fixed size TS packets of 188 bytes each. The TS is designed for transmission in errored conditions and includes features for error resiliency and packet loss detection. Figure 11 describes the general structure format of a TS.

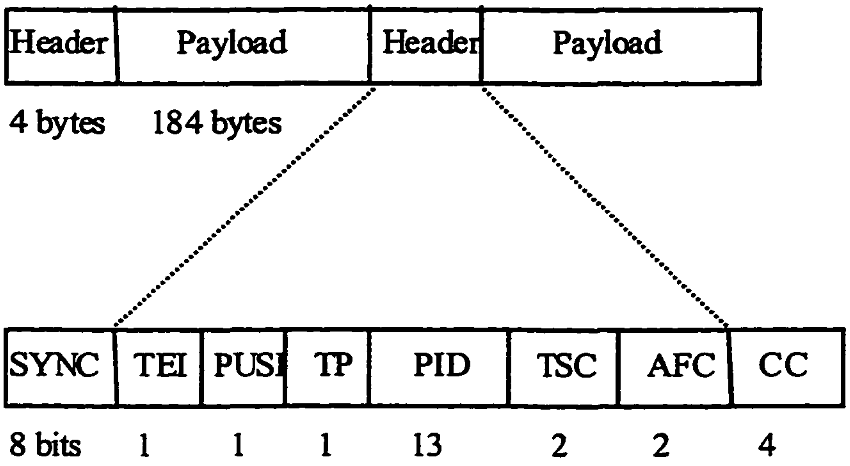


Figure 11: MPEG-2 TS Format

Transport Stream packets begin with a 4 byte prefix containing the following fields:

Sync: the Sync_byte is a fixed 8 bit field whose value is '0100 0111' (0x47), it is used by the decoder to identify the start of a TS packet every 188 bytes.

TEI: the transport error indicator is a 1 bit flag. When set to '1' it indicates the presence of at least one corrupted bit in the TS packet.

PUSI: the payload unit start indicator is a 1 bit flag. When set to '1', it indicates that the payload of the TS packet should start with the first byte of a PES packet or a Program Specific Information (PSI) section.

TP: the transport priority bit is used to prioritize TS packets within the same stream.

PID: the packet ID is a 13 bit field used to distinguish the kind of data stored in the TS packet payload and the reconstituting of ES's from the multiplexed TS.

TSC: the transport scrambling control is a 2 bit field which specifies the encryption mode of the TS packet payload.

AFC: the adaptation field control is a 2 bit field which indicates the contents type of the payload.

CC: the continuity counter is a 4 bit field which gets incremented every time a non duplicate TS packet with the same PID is sent.

The payload of a TS may consist of the following:

1) Information from PES packets which consist of more headers, timing information like picture decoding and presentation time stamps, and media information (video, audio).

2) Adaptation Field (AF), which may contain MPEG related information like the Program Clock Reference (PCR) field used by the decoder for clock recovery, and private information like the encryption key management.

3) PSI tables which provide the decoder with the necessary and sufficient information for de-multiplexing, decrypting and decoding a TS. The PSI tables carried in the TS are of four types:

PAT: the program association table provides the correspondence between a program number and the PID values of the TS packets which carry the program map table (PMT).

PMT: provides the mapping between program numbers and ES's that comprise them, and it is the complete collection of all program definitions for a TS.

NIT: the network information table provides the physical network parameters such as FDM frequencies, and it is transmitted as a private section with a reserved program number (0x0000).

CAT: the conditional access table provides the association between one or more conditional access systems, their entitlement management messages (EMM) streams, and any special parameters associated with them, such as entitlement control messages (ECM). In the broadcasting applications the CAT is used to control the subscriber accessible channels and to enable the user to block certain programs.

3.4 MPEG-4.

After completing the MPEG-1 and MPEG-2 standards, MPEG is currently completing the work on developing a new coding standard for very low bit-rates. This activity, known as MPEG-4 [30,31,32,33], is targeted at video bit rates from 6 kbps to 4 Mbps for applications such as:

- Videotelephony and videoconferencing with low delay requirements
- Symmetric and asymmetric remote expertise and classrooms (tele-medicine, tele-mechanic, and distant learning).
- Interactive information based retrieval.
- Very low bit-rate scalable multimedia distribution over error prone networks for portable and mobile receivers.

MPEG-4 is an object based compression technique where natural and synthetic audiovisual objects (AVO) can be arranged into an audiovisual scene using scene description. The MPEG-4 specifications describe three layers, the compression layer, the system or sync layer and the delivery layer.

At the compression layer elementary streams (ES) are formed, where an ES may contain one of the following:

- Object Descriptor Stream: Identify describe and associate ESs together.
- Scene Description Streams: A scene descriptor is encoded as a Binary Format for Scene (BIFS) and it describes the spatial and temporal positioning of audiovisual objects in a scene.
- Media Streams: Coded audiovisual information.
- Object Content Information Streams: Carries information such as content classification descriptors, keywords, language...etc.
- Timing information: Clock reference values for synchronization and decoding AVOs.

The system or sync layer (SL) is where each ES is divided into one or more access unit (AU). Each AU is then divided into SL-packets to form a SL-packetized stream (SPS) where timing, synchronization and random access information as will as padding bytes are inserted to facilitate proper encapsulation, decoding, composition and playback of media objects. However, a sync layer packet does not contain information on its length or on ES identification, this information is to be provided by the lower layer protocols by which the SL packetized stream is encapsulated for delivery. The interface between the compression layer and the sync layer is called the elementary stream interface (ESI).

An optional and application dependent layer called the FlexMux layer exists between the sync layer and delivery layer. The FlexMux layer is used to interleave one or more elementary streams into a FlexMux stream. The FlexMux layer also can be utilized for the proper framing of SL-packets to fit the profile of the underlying transport protocol.

The Delivery layer in MPEG-4 consists of the Delivery Multimedia Integration Framework (DMIF) [32]. This layer is media unaware but delivery technology aware. The delivery layer ensures transparent access to MPEG-4 content irrespective of the transport network technology used (e.g. IP or ATM). The boundary between the sync layer and the delivery layer is called DMIF application interface (DAI). The DAI is a generic API dependent on the operation system and the programming language used. It is used to define the functions of the DMIF layer and to allow the DMIF user to specify the quality of service (QoS) requirements of the desired stream to be transported using any permissible transport protocol stacks (TransMux layer). The TransMux layer is not specified by MPEG-4 and could consist of any stack of protocols with properties consistent with the MPEG-4 requirements.

4 MPEG-2 over ATM Networks

4.1 Introduction.

Asynchronous Transfer Mode (ATM) is considered the core technology for B-ISDN networks, while MPEG-2 is the most recently standardized and adopted coding system for multimedia applications. The new emerging video/multimedia applications such as Video On Demand (VOD), video teleconferencing, distance learning, and HDTV are based upon the efficient and reliable transmission of video, audio, graphics, and data over computer networks. In multiplexing, compressing, and transporting such voluminous information over even a reliable network like ATM, problems such as network congestion, and information loss become unavoidable. End-to-end performance and Quality of Presentation (QoP) of the transported MPEG-2 video over an ATM network is greatly influenced by the choice of an ATM adaptation layer (AAL) [14,22], which enhances the services provided by the ATM layer to support the functions required by the next higher layer. While AAL1 can be used to transport Constant Bit Rate (CBR) video, AAL5 was chosen by the ATM Forum to carry both constant and Variable Bit Rate (VBR) video.

Through a combination of spatial and temporal compression techniques, MPEG-2 video can be compressed up to approximately 50:1 while preserving image quality. The spatial compression techniques are same as those used by JPEG and include Discrete Cosine Transform (DCT), quantization and entropy coding. The temporal compression techniques used rely on block-based motion compensation to reduce the temporal redundancy. In MPEG three main picture types are defined. Intra-coded pictures (I-pictures) are coded without reference to other coded pictures. They provide access points to the coded sequence where decoding can begin, but are coded with only moderate compression. Predictive coded pictures (P-pictures) are coded more efficiently using motion compensated prediction from a past intra or predictive coded picture and are generally used as a reference for further prediction. Bidirectionally-predicted coded pictures (B-pictures) provide the highest degree of compression but require both past and future reference pictures for motion compensation.

An MPEG-2 program is a single bit stream that may consist of multiplexed video, audio, and data. The individual program components called elementary streams are packetized into either constant or variable size packets to make the individual packetized elementary streams (PES). The MPEG-2 system layer then multiplexes the different PES's into either a Program Stream (PS) intended to be transported over a lossless error free network, or into a Transport Stream (TS) designed for transmission over noisy and lossy networks. A transport stream consists of fixed size packets of 188 bytes, the first four bytes are the header fields, while the remaining 184 bytes form the payload.

Figure 12 shows the format of TS packets. The payload may consist of one or more of the following:

a) PES packets which besides the element stream data, contain timing information like picture decoding and presentation time stamps.

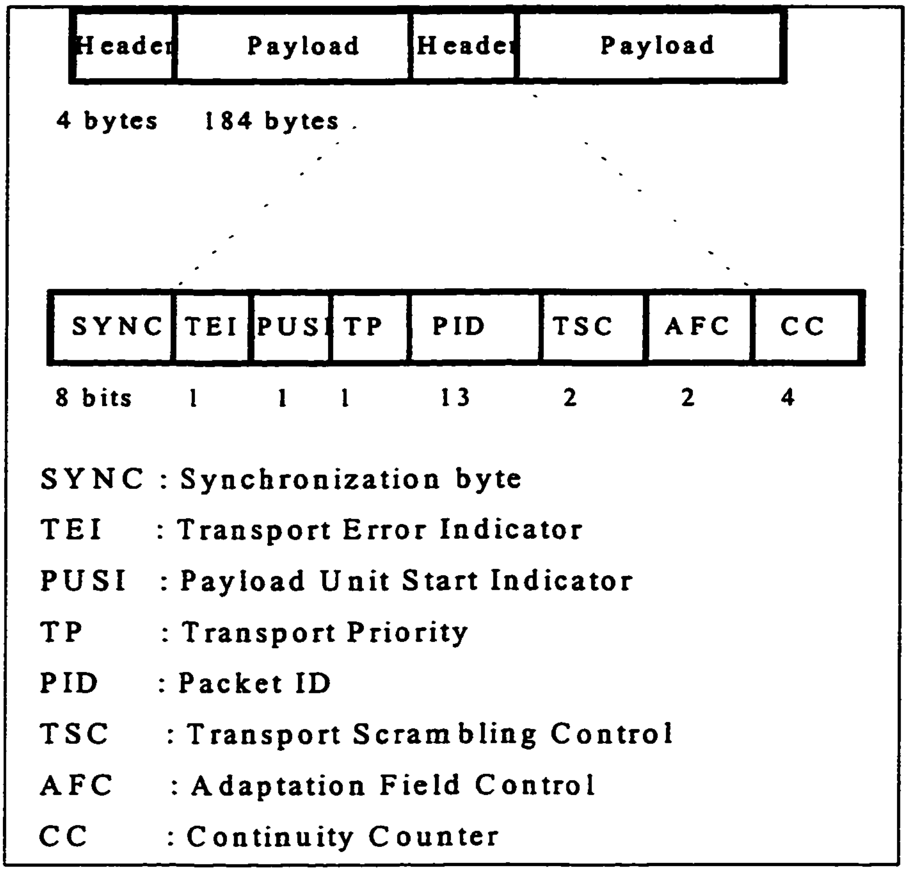
b) Adaptation Field (AF), which may contain one or more of the following: MPEG related information like the Program Clock Reference (PCR) field used by the decoder for clock recovery, private information like the encryption key management or simply padding bits for alignments of TS packets.

c) Program Specific Information (PSI) [3,6] tables to provide the decoder with the necessary information for demultiplexing, decrypting and decoding a TS.

4.2 Sensitivity of MPEG-2 Transport Stream to Network Impairments.

The sensitivity of various types of the MPEG-2 TS packets to errors is valuable information when selectively applying Forward Error Correction (FEC) codes to minimize their overhead while preserving the video quality. In investigating the sensitivity of the MPEG-2 TS for bit errors and cell losses we used two groups of TS video clips. Group 1 has the encoding sequence IBBPBBPBBPBBPBB encoded at 5.77 Mbps with automatic I frame insertion on scene changes. Group 2 consists of MPEG-2 ESs which we downloaded from [4], and we used a software multiplexer to form TSs. Group 2 has the encoding sequence IBBPBBPBBPBB encoded at 1.5 Mbps.

Figure 13 and Figure 14 show the bit distribution of two encoded sequences, Nature at 5.77 Mbps and Suzi at 1.5 Mbps.



SYNC : Synchronization byte
TEI : Transport Error Indicator
PUSI : Payload Unit Start Indicator
TP : Transport Priority
PID : Packet ID
TSC : Transport Scrambling Control
AFC : Adaptation Field Control
CC : Continuity Counter

Figure 12: MPEG-2 TS Packet Format

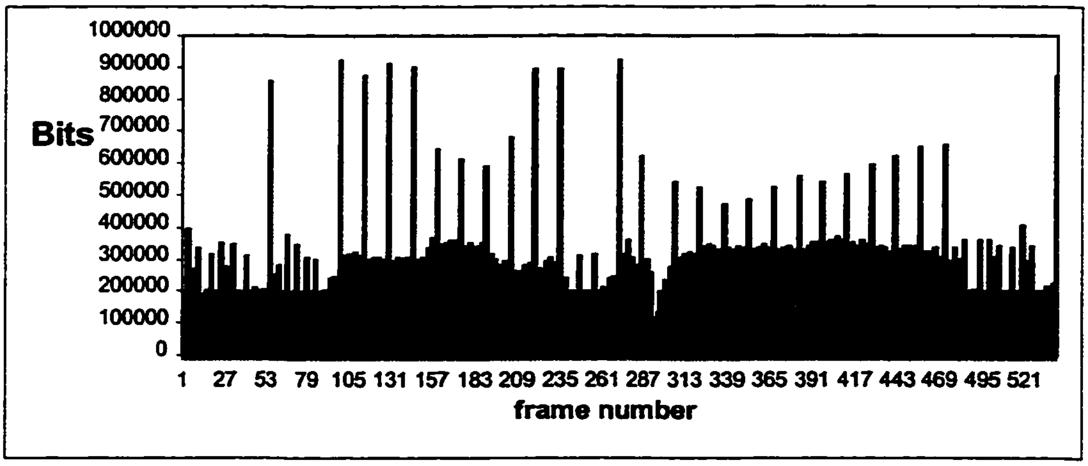


Figure 13: Bit Distribution of the MPEG-2 Video Clip Nature

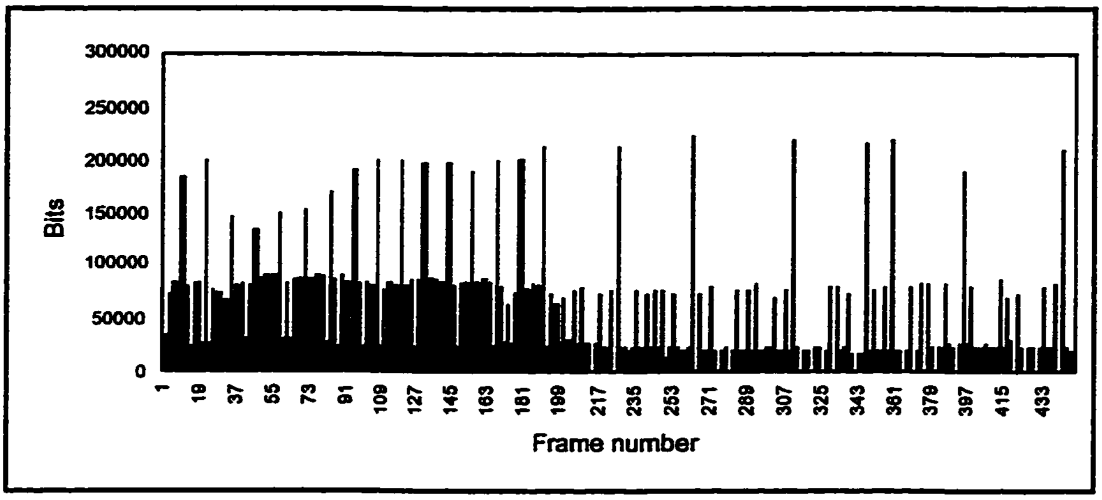


Figure 14: Bit Distribution of the MPEG-2 Clip Susi.

For decoding and playing back the TSs we used an Optivision hardware decoder at 30 frames per second.

Errored (with bit errors) and incomplete MPEG-2 TS video clips were generated in two ways:

1) Bit Error Rates (BER) ranging from 10^{-3} to 10^{-8}, and Cell Loss Ratio (CLR) ranging from 10^{-2} to 10^{-6} with normal or uniform distributions.

The resulting errored and incomplete streams were decoded in two ways:

a) With AAL5 PDUs dropped for each cell lost or for a single bit error.

b) With no AAL5 segmentation, the resulting streams were decoded as they were.

2) Bit errors or cell losses in specific fields and packets were as follow:

a) Bit errors in each field of the 4 byte header were inserted with a certain frequency, and the resulting transport stream was decoded to observe the impact.

b) Bit errors were generated within the payload of transport stream packets of the same type according to a header field (PUSI =1, PID = 0x000, etc.), and the errored transport streams were decoded.

d) Transport stream packets of each desired type, as in part b, were segmented using AAL5, and ATM cells were dropped, one at time. The resulting corrupt AAL5 PDU was dropped, and the resulting data stream was decoded.

To investigate the sensitivity of TS packets based on frame types, we prioritized packets from I frames using the TP bit and forced bit errors or cell losses within TS packets of P, and B frames. The resulting transport streams were decoded and compared with unprioritized errored and incomplete transport streams. As expected, protecting TS packets of I frames from bit errors or cell loss greatly improved the received video quality even with much higher error rates.

The video quality was evaluated based on the frequency of picture freezing, picture tiling, color changing, and irregular motion of still pictures, or "blurring". Table 3 and Table 4 summarize the observed artifacts of bit errors in the header fields of a TS packet and cell losses in different TS packet types respectively.

When an error occurs in the Sync. or PID field, the effect is equivalent to a TS packet loss. A bit error in the AFC will cause the decoder to interpret the bytes that follow the header incorrectly. For example, if the AFC value is changed from (01), which indicates to the decoder that the payload only consists of video information, to (11), the decoder will assume the payload comprises both video and adaptation field information, and the presented picture is affected.

A PUSI field of one will indicate whether the payload of a TS packet starts with the first byte of a PES packet or with a one byte pointer to a section of a PSI table within the TS packet. The TS video clips we used were formed using PES packets of variable lengths, each containing one video frame; therefore, a bit error in the PUSI field or bit errors and cell losses in the payload of TS packets with PUSI=1 caused screen blanking and picture mixing which greatly degraded the video quality.

Table 3: Sensitivity of TS Packet Header to Bit Errors

Field	Artifacts
Sync	Tiling, Blurring, Picture Freezing, Color change
TEI	No artifacts (decoder dependent)
PUSI	No playback, Screen Blanking
TP	No artifacts
PID	Tiling, Blurring, Picture Freezing, Color change
TSC	No artifacts
AFC	Tiling, Blurring, Picture Freezing, Color change
CC	No artifacts

Table 4: Sensitivity of TS Packet Types to Bit Errors and Data Loss

TS Packet with	Artifacts: Bit Errors	Artifacts: Cell Losses
PID 0x0000	No Playback	No Playback
PID 0x1FFF	No Artifacts	Tiling, Blurring
RAI=1	Screen Blanking, Picture Freezing	Screen Blanking, Picture Freezing
PUSI=1	Screen Blanking, Picture Freezing	Screen Blanking, Picture Freezing
Packets from I frames	Tiling, Blurring, Color Change	Picture Freezing, Tiling, Blurring, Color Change
Packets from P,B frames	Tiling, Blurring, Color Change	Tiling, Blurring, Color Change

The Random Access Indicator (RAI) is a one-bit field of the AF. The payload of TS packet with a RAI=1 is greatly sensitive to errors since it contains the first byte of a video sequence header, and acts as a random access point.

TS packets with PID=0x0000 make up the Program Association Table (PAT) [6] which is used to enable the decoder to map the different PIDs of TS packets to their associated programs. Errors in the PAT may block the decoder from starting a program. Cell losses in null packets (PID = 0x1FFF) causes a temporary loss of synchronization among packets, and tiling and blurring artifacts are observed.

4.3 Performance Evaluation of AAL5 in Transporting MPEG-2 Transport Stream.

When an MPEG-2 TS is encapsulated using AAL5, as in Figure 15, two fields of the 8 byte AAL5 trailer are used for error detection. The Cyclic Redundancy Check (CRC), 4 bytes, is used for bit error detection over the entire PDU and the Trailer, while the Length Field (LF), 2 bytes, is used for cell loss detection. When there is a CRC error, a complete Protocol Data Unit (PDU), in this case 8 ATM cells, is discarded even for a single bit error. A discrepancy in the PDU length (caused by cell loss) will also result in the loss of two transport stream packets or a total of 376 bytes. Dropping the last cell in a datagram, which contains the end of datagram flag, could cause four packets to be dropped. In our experiments, we distinguished between two basic network impairments, bit errors and cell losses. PDUs with errored bits are referred to as errored PDUs, while those with lost cells are referred to as incomplete PDUs. In the case of errored PDUs, the experiments we conducted prove (Figure 16) that passing the errored PDU to the higher layer greatly improves the received video quality. This encourages us to believe that enabling the CRC should be optional. However, passing incomplete PDUs to the higher layer yielded no improvement over discarding the PDU when no corrective measures are taken.

Recently, deliberation at the ATM Forum has concluded that passing incomplete PDUs along with the LF is highly desirable. Next, we present how the length field when passed along with the AAL5 CPCS PDU to the upper layer can be used to improve the video quality.

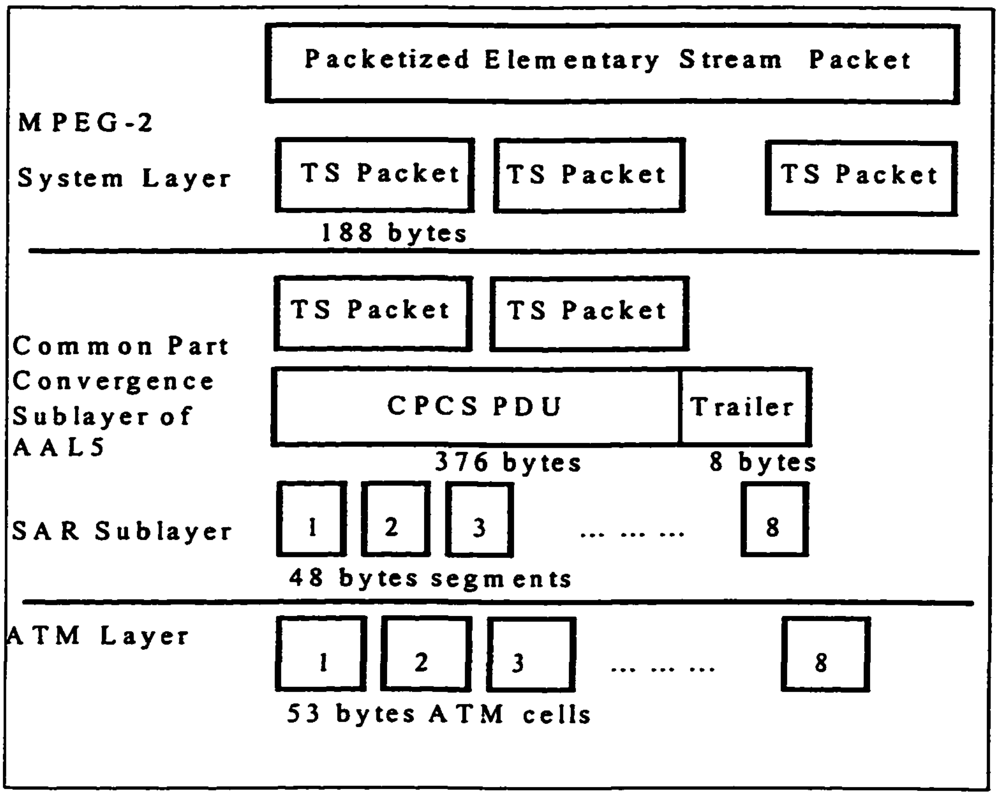
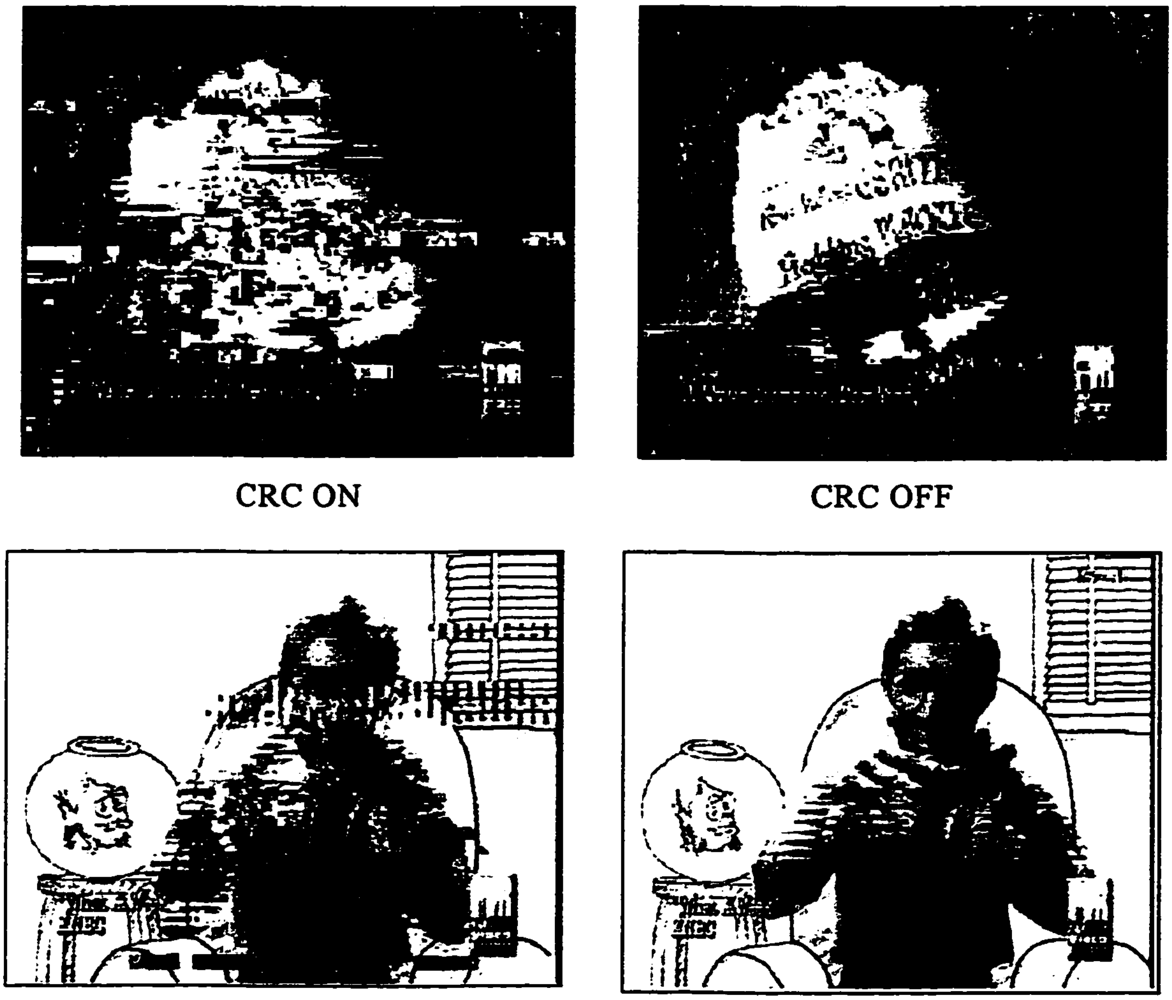


Figure 15: Encapsulation of MPEG-2 TS Packets Using AAL5

4.4 A cell Loss Padding Technique for the Transport of MPEG-2 Transport Stream over AAL5.

In dealing with information loss when MPEG video is transported over a computer network where retransmission is not possible, there usually are two approaches, Forward

Error Correction (FEC) which is best suited for bit errors but involves transmission overhead, and error concealment at the MPEG decoder. The use of error concealment techniques is based on utilization of the available information to estimate the lost portions of the video signal and usually increases computational overhead and complexity. We present a simple, yet proven, effective cell loss padding technique which is suitable to be used with AAL1 or AAL5.



CRC ON

CRC OFF

Figure 16: Effect of CRC on Video Quality in the Presence of Bit Errors

Why is padding necessary? Padding preserves synchronization among packets and PCR values, and maintains the integrity of CBR TSs. In the case of AAL5, where two TS packets are encapsulated as a PDU, the essence of the proposed technique is based on passing the incomplete PDUs and the LF to a SSCS of the AAL5 or to the decoder TS layer. If the actual length matches the encoded length, no cell loss occurred and no action is taken. However, when the length field is incorrect, padding is performed based on the following set of parameters {LF, AFC, PID, Adaptation Field Length (AFL)} to conceal artifacts resulting from data loss.

At the encoder side, padding with stuffing bytes of ones within the AF of TS packets is performed when necessary, and null TS packets are also generated to form CBR TSs. Such facts become significant when considering padding at the decoder side.

4.5 Cell Loss Padding Strategies.

Among several padding strategies conducted, the one that yielded the best results in concealing data loss is described as follows:

1) Actual length is greater than the encoded length (parameters are parsed for second TS packet):

   a) {LF > encoded one, AFC = '01'}, 40 bytes of zeroes were inserted at the end of second TS packet.

b) {LF > encoded one, AFC = '11', AFL > 148}, at the end of second TS packet, bytes of ones were inserted to complete the AF and packet is completed with bytes of zeroes for video data.

c) {LF > encoded one, AFC = '10'}, 40 bytes of ones were inserted at the end of second TS packet.

d) {LF > encoded one, PID = 0x1FFF}, 40 bytes of ones or zeroes were inserted at the end of second TS packet.

2) Actual Length is less than the encoded length:

a) {LF < encoded one}, and header of first TS packet is lost, padding was done with the cell format as in Figure 17 a at the beginning of first TS packet, or the incomplete first TS packet was replaced by a zero energy packet as in Figure 17 b.

b) {LF < encoded one}, and header of first TS packet is present, parsing for the second TS packet Sync. and PID fields was done at byte number 189 and 191-193 within the PDU, if present, padding of 48 bytes of zeroes, ones or a combination of both is inserted at the end of second TS packet based on its {AFC, AFL, PID} as in 1. Otherwise, the lost cell was padded for at byte 144 as in Figure 18 a or the rest of the second TS packet was removed, and the PDU was completed as in Figure 18 b.

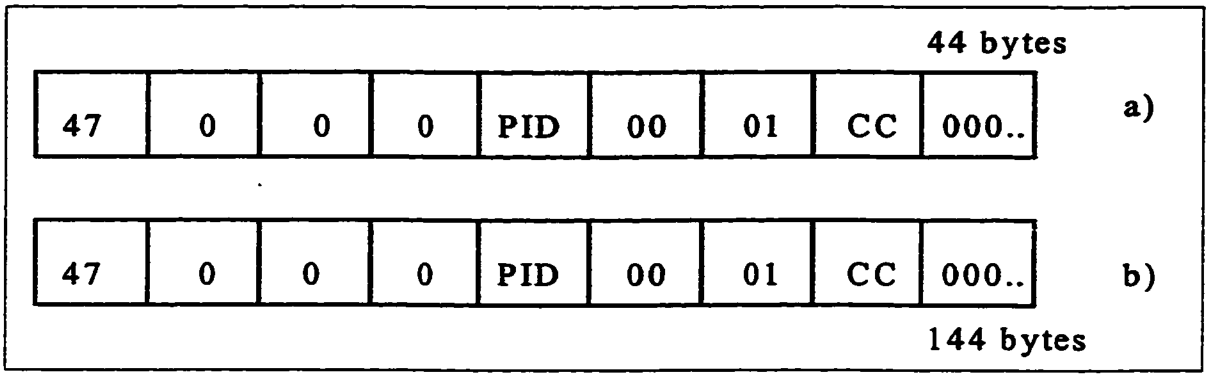


Figure 17: Cell Loss Padding 1

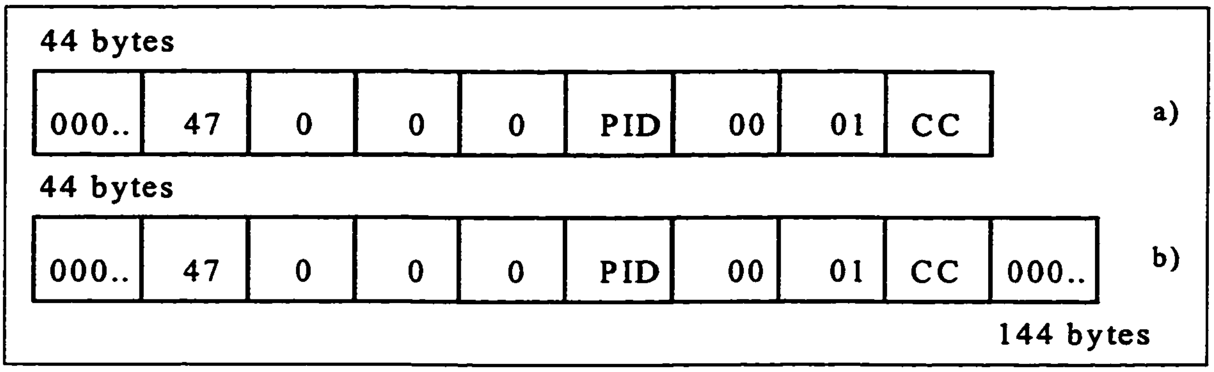
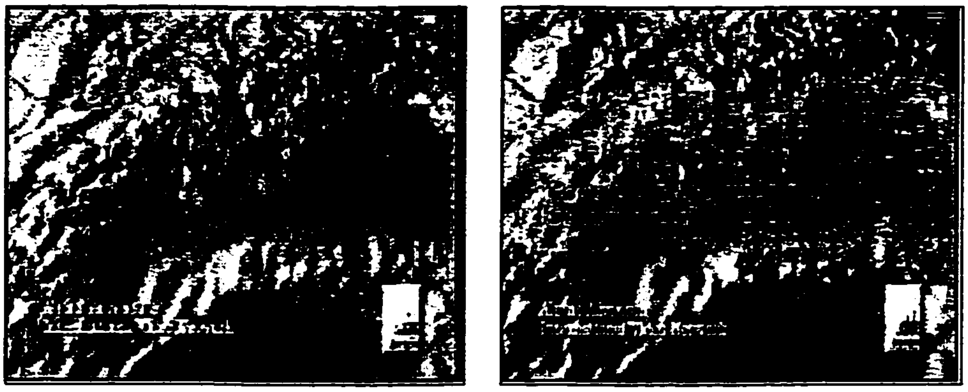


Figure 18: Cell Loss Padding 2

Figure 18 summarizes the overall performance and effectiveness of the proposed cell loss padding technique. Figure 19 shows a snapshot of the video sequence Nature in the presence of cell losses to show the effect of padding on the video quality.

Table 5: Performance of Padding Technique

Scene Motion	Without Padding	With Padding
Intense	Picture freezing, Blurring, Tiling, and Color Change (Continuos)	Blurring ,and Tiling (Continuos))
Moderate	Picture freezing, Blurring, Tiling, and Color Change (Occasional)	Blurring, and Tiling (Occasional)
Low	Picture freezing, Tiling. (Occasional)	Blurring and Tiling (Very Infrequent)
Still	Tiling, Color Change (Occasional)	Tiling (Very Infrequent)



With Padding Without Padding

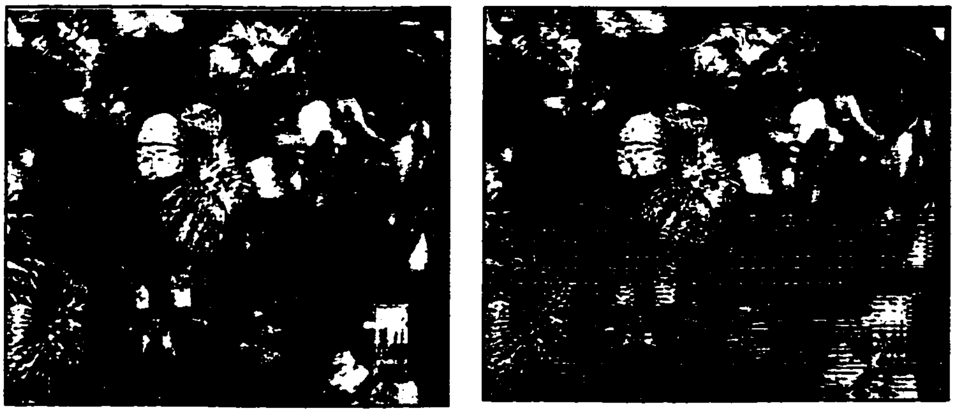


Figure 19: Effect of Cell Loss Padding

When AAL1 is used, each TS packet of 188 bytes is segmented into 4 segments of 47 bytes each, and a 1 byte header which contains a 4 bits Sequence Number (SN) is attached to each segment to form the 48 bytes necessary for the ATM cell payload. The sequence count value of the SN makes it possible to detect the loss or misinsertion of cells. Using the SNs, padding is done based on the position of the lost cell in a similar fashion as in the case of AAL5

4.6 Advance Coding Techniques for the Transport of MPEG-2 Video over Impaired Computer Networks.

4.6.1 Error Concealment Techniques.

The MPEG-2 standard specifies several techniques to improve video quality when operating over impaired computer network networks. These include scalability and error concealment. An MPEG-2 decoder can utilize available spatial and temporal information to regenerate lost or errored video information. The network has the capability to pass valuable information about the placement of bit errors to the decoder using the AAL-5 cyclic redundancy check and length fields. In most cases these fields provide enough information to isolate the bit error to an area of a picture (a particular macroblock or slice). In addition, the MPEG-2 standard makes the frequency of resynchronization points in the stream (the occurrence of a new slice) flexible so that errors may be localized to a small portion of the picture. Once an errored macroblock is identified, it may be concealed by using surrounding macroblocks from the same picture or from previous pictures.

In spatial concealment, the decoder can estimate a lost macroblock from surrounding macroblocks in the same frame by averaging the DC coefficients of its neighboring blocks. This method will estimate the low frequency characteristics of the lost macroblock and should be sufficient for concealing monochrome macroblocks or macroblocks with a low level of detail.

Temporal concealment utilizes the redundant information in previous or future frames to estimate a lost macroblock. The simplest implementation is to replace the lost macroblock with the macroblock in the same position of the previous frame which works well for low-motion video sequences. The effectiveness of temporal concealment techniques can be improved by averaging the motion vectors of surrounding macroblocks to estimate the lost macroblock's movement from the previous to the current frame.

4.6.2 Hierarchical Coding and Scalable Syntax.

Hierarchical coding allows reconstruction of useful video from pieces of the total bit-stream. The MPEG standard specifies scalable syntax to support this process. Scalability is achieved by structuring the total bit-stream in two or more layers starting with a stand-alone base layer and adding a number of enhancement layers. When video streams are transmitted over a network, the QoS parameters can be set differently for each layer. The base layer should be transmitted with a higher priority to ensure low cell loss while the enhancement layers can be transmitted with lower priority. This allows video quality to be controllable by the user or to vary with the available network bandwidth.

Layered or scalable coding is one approach to provide robust transmission of video based multimedia traffic in impaired transmission environments. Scalable coding provides the ability to achieve video of more than one resolution or quality simultaneously. Scalable video also supports the concept of prioritized video streams to improve transmission robustness, error resiliency, and compliance with network bandwidth constraints. The following paragraphs discuss the various forms of scalability including their potential applications.

In data partitioning the video bit stream is split at the source into two layers, the base layer and the enhancement layer. The base layer contains all of the key header fields, motion vectors and low-frequency DCT coefficients. The enhancement layer contains less critical information such as high-frequency DCT coefficients. The video stream is partitioned using a Priority Break Point (PBP) that specifies the contents of each layer or partition. The choice of a priority break point can be based on the MPEG-2 video frame type, the characteristics of transmission link, the available bandwidth, or any other application specific requirements. For example, a priority break point could be set so that most of the Intra-coded (I) frame information is present in the base layer, while only minimum information of Predictive-coded (P) or Bi-directional-coded (B) frames is contained in the base layer. When partitioned video is transmitted over an ATM network, the base layer is transmitted using a channel with guaranteed quality of service requirements to preserve its integrity, and the enhancement layer is transmitted using a less reliable channel. At the receiver side, the two layers are combined to produce the original video stream whenever possible. If errors occur in the enhancement layer, the video will be reconstructed using only the base layer.

SNR scalability provides a way of transmitting two layers of the video stream at the same spatial and temporal resolutions but with different quality levels. The base layer encoding process is identical to that for non-layered encoding. The quantized DCT coefficients from the base layer (after being inverse quantized) are subtracted from the input DCT coefficients of a block. The resulting quantization error from each block is next re-quantized with the same or different quantization parameters and encoded to form the enhancement layer bit stream. In SNR scalability the base layer is identical to a non-layered bit stream and could be decoded as a stand-alone layer using a non-layered or SNR decoder. However, since the enhancement layer is encoded using the DCT coefficients of base layer, it is useless in the decoding process without the base layer. When both base and enhancement layers are added and decoded, they regenerate a higher quality reproduction of the input video. In a SNR scalable decoder, the dequantized base and enhancement layer DCT coefficients are obtained independently from their respective bit streams before they are combined to form the refined higher quality bit stream. Potential applications of SNR scalability include providing a high degree of error resilience to transmission errors since the more important data (base layer) can be sent over a channel with better error performance. In addition video services with multiple qualities can be provided over a heterogeneous network.

There are three more forms of scalability: spatial, temporal, and hybrid. These forms of scalability are considerably more complex than data partitioning and SNR scalability. Spatial scalability generates two video layers with different spatial resolutions from the same bit stream. The lower layer may consist of a different video format, providing interoperability between the lower layer and the upper layer (e.g. lower layer H.261, MPEG-1 and higher layer MPEG-2). Temporal scalability partitions frames from the input video stream into two or more layers, each with either the same or different temporal resolutions. When layers are combined, they provide the full temporal resolution as in the input video stream. Hybrid scalability provides a way of combining two or more types of scalability such as SNR, temporal, and spatial to generate three or more layers of video. However, the degree of complexity in hybrid scalability is proportional to the number of desired layers.

5 Analysis of Cell Transfer Delay and Cell Delay Variation in the Transport of CBR Traffic over ATM

5.1 Introduction.

As the communication networks are forced to support an ever-increasing number of bandwidth-hungry applications, there is an urgent need to efficiently consolidate the existing transmission methods throughout the public and private networks in order to accommodate the new telecommunications economic and marketing imperatives. B-ISDN technology offers such a solution by providing a uniform method of transporting a variety of services including voice, video, data and images using the Asynchronous Transfer Mode (ATM). An important benefit of ATM technology is its ability to provide Quality of Service (QoS) guarantees for any application it supports regardless of how stringent characteristics it may have. Today, ATM is considered the core technology for B-ISDN networks, and is currently commercially deployed in local as well as wide area networks. However, as any other data communication system, ATM networks have their share of problems caused not only by the common transmission impairments such as bit errors, but are also related to the variations in end-to-end cell delay (cell delay variation), and the cell transfer delay. Those factors are functions of the network loads, number of nodes/switches involved, etc. The impact of the cell-specific impairments on the ATM performance needs to be quantified in order to use this technology as efficient as possible.

In this work we examine the two end-to-end ATM QoS parameters: cell delay and cell delay variation or cell jitter. The end-to-end delay an ATM cell experiences travelling the ATM network is mainly due to the cell buffering performed by the ATM switches that are built to support the asynchronous statistical multiplexing and to switch dynamic traffic load. Constant Bit Rate (CBR) communication traffic services such as real time voice, and video applications are very sensitive to cell delay and cell delay variation, requiring tight conformance from the network to meet their specified Quality of Service (QoS) contracts.

In our multimedia laboratory, experiments have been conducted to characterize CTD and CDV in ATM network. Several simulation studies [9,10,11] have revealed that the amount of traffic load carried by the ATM switch greatly influences the characteristics of both CTD and CDV. The uniqueness of our work is due to its empirical and experimental nature under a controlled environment, which more accurately reflects the real behavior and true characteristics of CTD and CDV in ATM networks under diverse operating conditions. This chapter starts with a description of our system and test-bed configuration. Experimental results, analytical modeling, and characterization of CDV and CTD under diverse network condition then discussed.

5.2 System Description.

5.2.1 Test-bed Architecture.

Our laboratory measurements were performed using two ATM switches and a single B-ISDN ATM test set (used as both a source and sink). Figure 20 illustrates this architecture.

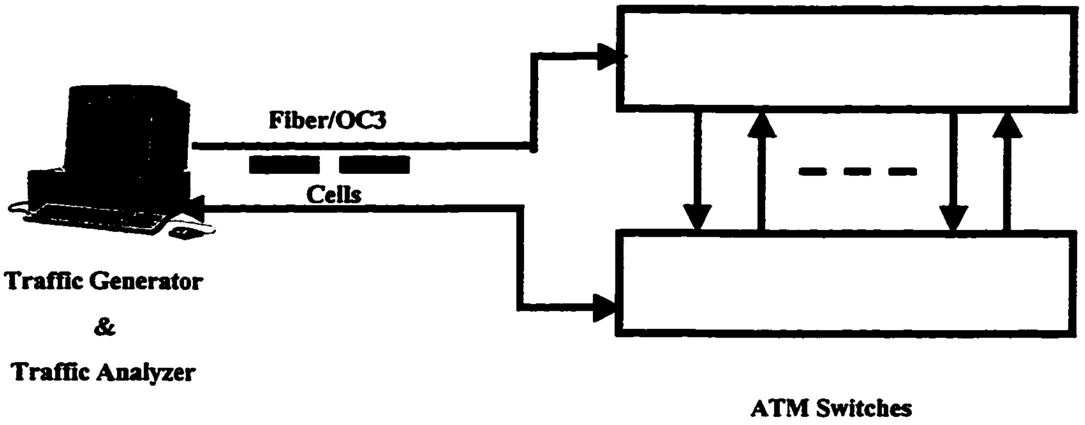


Figure 20: Test-Bed Architecture for Delay and Jitter Measurements

This system provided a controlled environment for our tests and allowed us to emulate networks of different size (with up to 16 ATM nodes/ switches), and to exam them under almost any set of conditions that we considered important for our study. It was always carefully configured and synchronized, to preserve all the major characteristics of a real network except perhaps for the delay associated with signals transmission over large WAN (a small factor of negligible impact especially on jitter analysis). Thus, our end-to-end measurements and the performance analysis presented in this chapter should quite adequately depict behavior of a typical WAN.

5.3 ATM Test Set.

The ATM test set that we used, designed for high-speed protocol testing (at constant and variable bit rates), provided complete physical, ATM and AAL layer generation capability. It also analyzed data received from the switches, determined the network's end-to-end Quality of Service (QoS) performance, and monitored its overall operation. For example, data for the two QoS parameters associated with the network, the Cell Delay Variation CDV and the Cell Transfer Delay (CTD) was captured and then analyzed off line.

The test equipment comprised of several modular components allowing custom configuration for various test applications. For our study we used two separate OC3 ATM Generator/Analyzer modules, one for the foreground traffic (duplex operation), and one for the background traffic (duplex or simplex transmission). ATM cells from the two created streams were independently generated then received by the first ATM switch on two distinct ports. The background traffic source was subsequently routed internally through the switch to all desired output ports, then taken with a physical OC-3 connection to the second switch where it was also internally routed in a similar fashion. The foreground source was routed physically with an OC-3 connection from one emulated node to the next as shown in Figure 20. On the output of each node (used port) both traffic sources (background and foreground) meet to compete and share the buffer as they exit from the ATM switch. All generated cells are time stamped as they exit a generator module, and time recorded as the analyzer module receives them

Each Generator/Analyzer module was capable of providing one physical port and up to eight programmable traffic sources with distinct bandwidths, and type of services (e.g., CBR, VBR, UBR, ABR).

5.3.1 Cell Generator.

Each Generator/Analyzer module was capable of providing one physical port and up to eight programmable traffic sources with distinct bandwidths, and type of services (e.g., CBR, VBR, or UBR). Their cells accepted any user-supplied values including the virtual path identifier (VPI), the virtual channel identifier (VCI), the cell loss priority (CLP), and the generic flow control (GFC). In addition, the Generator/Analyzer modules could insert a time-stamp to mark each test-cell departure instant from the generator. The test cells contain 6-byte timestamp in their payload, stored low byte first, starting at payload byte 39. The timestamp in the payload is protected by a 10-bit CRC and is ignored when an error is detected. During our tests, the cell generator was setup to provide CBR service (only) for both the foreground and background signals. The permanent virtual circuits (PVC) were used exclusively to route them across our test-bed.

5.3.2 Cell Analyzer.

Each Analyzer port included 16 programmable cell filters for separating incoming cell streams into 16 sub-streams that could be simultaneously analyzed in real time to provide cell traffic statistics for any or all of them.

To obtain statistics such as mean, variance and standard deviation of the CTD, CIT and subsequently CDV parameters, we used data from histograms provided by the test-set. Each histogram contained large quantity of data correlating measured values of a given parameter (displayed on the horizontal or x-axis) with the corresponding number of occurrences (indicated on the vertical or y-axis). These important relationships, captured by the Adtech were further processed using standard spreadsheet techniques. At the end, each tested sub-stream had its own set of statistics that when combined revealed vital aspects of the ATM cell behavior under diverse conditions created in our test-bed.

5.3.3 ATM Switch.

The heart of our test-bed were two ForeRunner ASX/200 WG self-contained ATM switches. Both of them had a non-blocking switching capacity of 2.5 Gbps and were capable to support up to 16 users or networking devices, each running at speeds up to 155.52 Mb/s via dedicated multimode fiber optical links. They had a single switch fabrics, used UNI 3.0 dual leaky bucket support for traffic policing. Their major hardware components consisted of a switch board, a switch controller, and network modules. The switch-board contained the VPI/VCI lookup tables, and routing circuits to ensure a cell received from an input port was correctly switched to one or more output ports. The ASX-200 WG switch-board could accept up to four network modules, which themselves contained up to four OC-3 I/O ports each.

The switch controller provided the distributed connection set-up engine for a network of ATM switches. The switch controller primarily provided management access through SNMP (Simple Network Management Protocol) and was responsible for storing and updating all SNMP management information. The network modules acted as the I/O ports with four OC-3 physical interfaces (for maximum of 622 Mb/s aggregated traffic).

5.3.4 Timing Arrangements.

Since the objective of our experiments was to qualify the variation in end-to-end cell delay, cell delay jitter, in a WAN environment, our test-bed timing/ synchronization arrangements reflected this assumption by ensuring that the transmit data streams used the same clock as the received data streams. To accomplish it both switches used the network clock as their timing source, i.e., they used the recovered clock from the incoming signals as the transmit clock.

5.4 Experimental Results and Analysis.

Cell Delay Variation or Jitter is considered to be one of the major problems of packet switched networks. It is often claimed that the application introduces much more jitter than the network so that CDV due to the network becomes negligible. However, in applications with tight delay requirements and constraints such as video and voice it becomes important to study and control jitter even at the ATM layer level.

Several studies [9,12] have been done to provide modeling and estimation of CDV and CTD characteristics in ATM networks. In order to evaluate the high-speed and broadband characteristics of ATM technology, it needs absolute accuracy for dependability and objectivity of evaluation results. In [9], the obtained results were based on simulation, where in [12] the vBNS wide area network was used under limited control over the network conditions. Our study is based on a fully controlled LAN network were more variations of network parameters are possible for the emulation of WAN type networks.

For the measurements and analysis of CDV and CTD of CBR traffic, we used a 4001 Kbps CBR (periodic test cells) as our reference connection. A single background traffic source (0-149,760 Kbps, periodic CBR) was used for loading purposes. The background load was incremented on a 10% of the OC-3 rate bases. Networks of 4, 6, 8 and 10 nodes were emulated under all possible load conditions by symmetrically switching between modules on the two used ATM switches as shown in Figure 21.

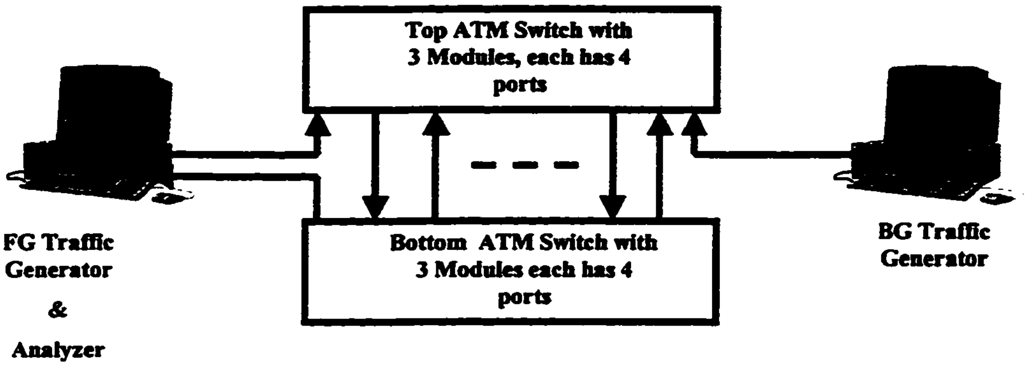


Figure 21: Test-Bed Configuration for a Multi-node ATM Network

5.5 Measurement Results of CDV/CTD.

CDV is an essential measure of ATM performance evaluation since ATM networks support various types of simultaneous applications with unique characteristics and QoS requirements. In an ATM network it is often the case where several input streams are routed within the switch to the same output port sharing a common buffer. CDV is defined as the variation in the end-to-end delay of cells within the same media stream. The value of CDV and its accumulation within an ATM network is dependent on several parameters such as, the number of nodes, traffic load, switch buffering capacity and mechanism, and the number of internal switch operations [13,15,16,18,19].

The data for CTD and CDV analysis was collected using a special feature in the ATM test set called Save Histogram Results. This feature allows the collection of data (contents of histogram bars) from runs for any time period to be saved and then imported to a spreadsheet program for off line processing and report generation. All data were collected in µsec then converted to ATM cell time slots (2.735 µsec, transmission time of ATM cell).

All runs were made for a period of 3 minutes several times then analyzed and compared for accuracy with runs made for 1 hour period, the three minute run was most of the times more than 98% accurate as the data for the 1 hour run. Our accuracy criterion was based on error differences in the mean, and variance.

In performing the experiments and for computation and analysis purposes CDV was defined as the difference between the inter-departure and inter-arrival of two consecutive cells. If D_I, D_{I-1}, A_I and A_{I-1} represent the departure and arrival times of two successive cells then CDV (J_I) is given by

$$J_I = (A_I - A_{I-1}) - (D_I - D_{I-1})$$

where J_I could be positive indicating cell spreading ,or negative indicating cell clustering. For CBR traffics, as in our case, (D_I -D_{I-1}) was replaced by a constant T, the reciprocal of the used reference traffic rate.

CTD of a cell (L_I) was obtained automatically by the ATM tester used, using the difference between the departure and arrival times, that is

$$L_I = A_I - D_I$$

5.5.1 Analysis of Mean and Standard Deviation.

For the analysis of CTD and CDV three network performance parameters were evaluated to study the impact of various network conditions on the referenced 4001 Kbps CBR traffic connection, the mean, variance and Cell Loss Ratio (CLR). All captured data with cell losses were excluded from analysis of CTD and CDV. Cell losses were only observed for runs over 8 and 10 hops with load conditions between 93% and 97% of the OC-3 rate, Table 6 summarizes the CLRs observed for those runs.

Table 6: Cell Loss Ratios Observed for 93% and 97% of Load Conditions

8 hops (load 93 to 97 % of OC-3 rate)	10 hops (load 93 to 97%of OC-3 rate)
CLR= 1.5×10^{-6} to 5.5×10^{-3}	CLR= 2.5×10^{-6} to 6.2×10^{-3}

Figure 22 and Figure 23, respectively, show how the Standard Deviation of CTD and CDV of the baseline traffic source tends to grow as the traffic load increases. The growing is slow under low and moderate network load conditions (0%-80% of OC-3 rate) and increases with a much sharper slope when the background load exceeds 80%, a similar result was also reported in [9]. The analysis of the mean as a function of the traffic load revealed that there was no change in the mean of CDV which contradict the results of simulation as presented in [9]. Figure 24 shows the change in the mean of CTD for the 2, 6, and 10 node cases as the traffic load increases. Although, the CTD mean tend to grow with an increasing traffic load, its maximum change is much less than with the results presented in [9,12]. The change in mean of CTD as a function of number of hops is shown in Figure 25 for background load cases of no load, 50% load and 90% load. From Figure 25 it is apparent that the change in the mean of CTD is approximately linear, with its slope being dependent on the load conditions.

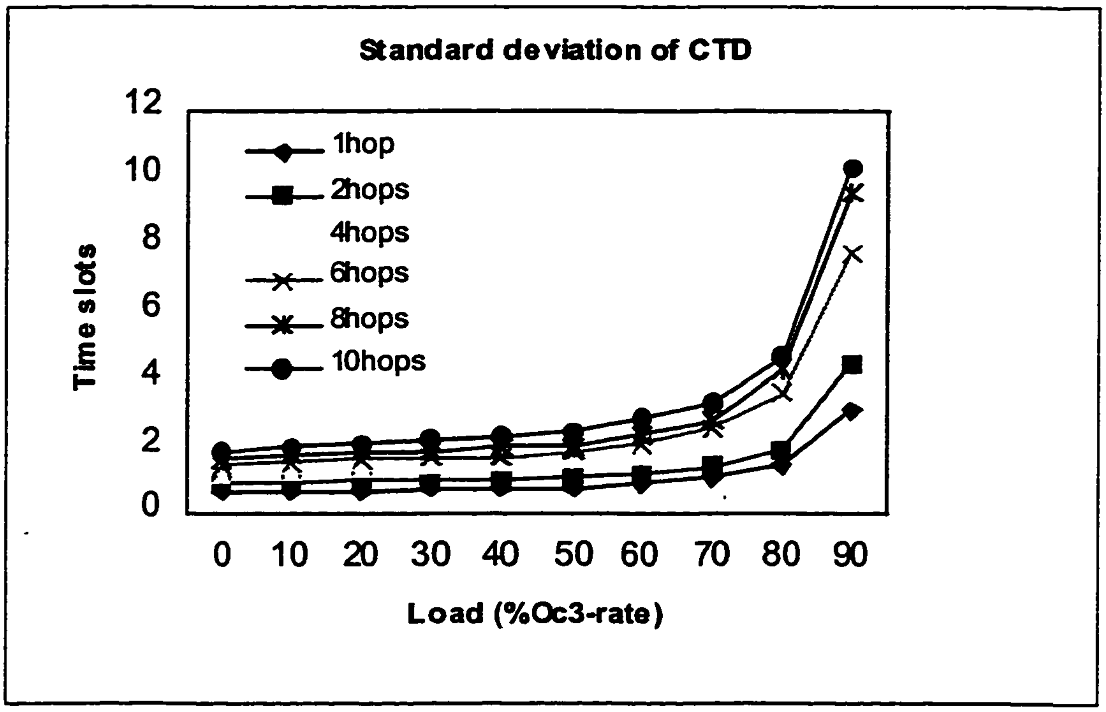


Figure 22: Change in the Standard Deviation of CTD as a Function of Background Load Conditions

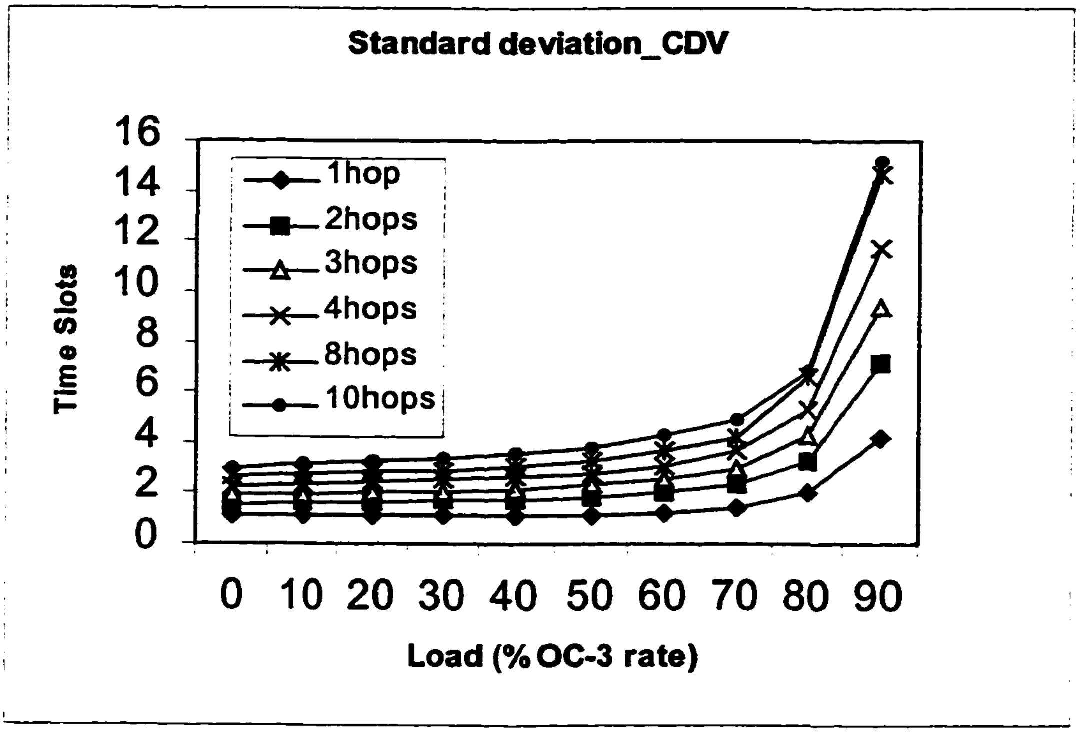


Figure 23: Change in the Standard Deviation of CDV as a Function of Background Load Conditions

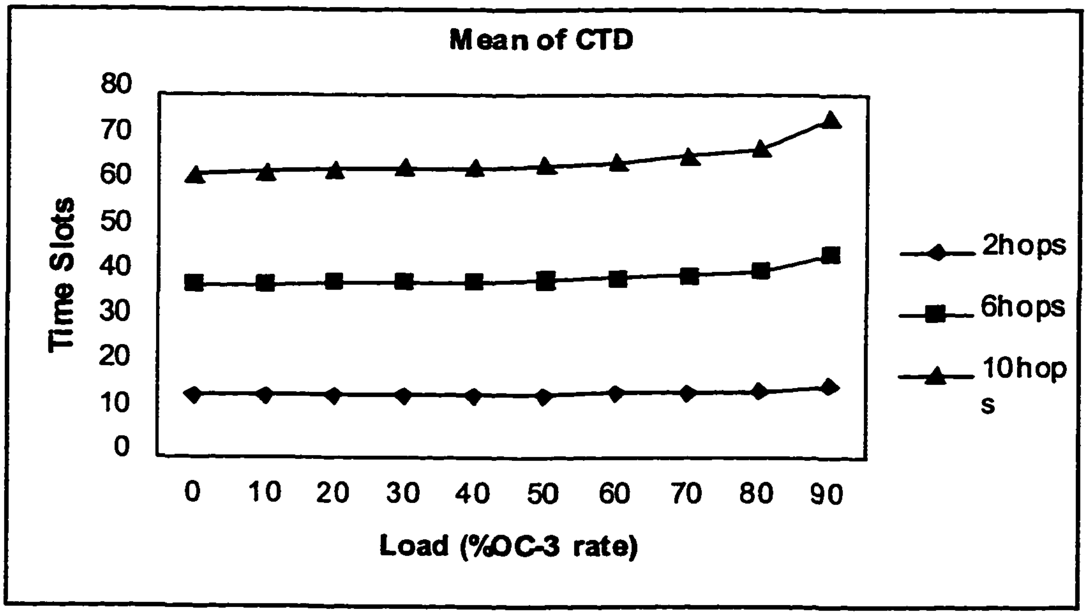


Figure 24: Change in the Mean of CTD as the Network Load Changes

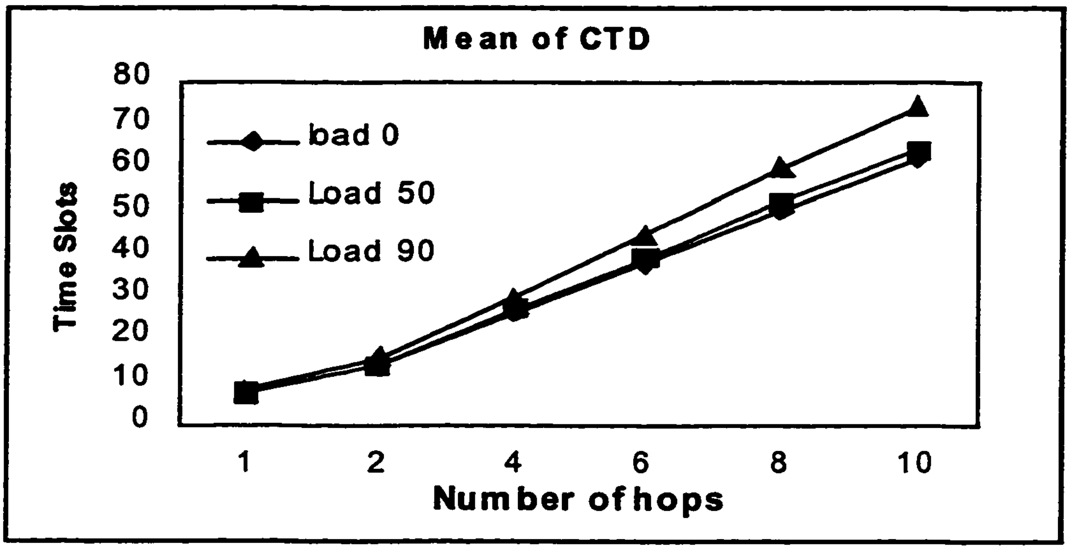


Figure 25: Change in the Mean of CTD as the Number of Hops Changes

In the ATM Forum B-ICI specification [8,37], it is suggested that the accumulation of CDV across multiple switching nodes can be approximated by the Square Root Law (SRL). The SRL state that the end-to-end CDV a cross a network of N nodes is approximately equal to the CDV for an individual node times the square root of N (e.g. $\sigma_N = \sigma_I \sqrt{N}$). The SRL was based on the assumption of normal distribution for CTD across a node regardless of traffic load conditions, and also assumed that the CDV across each node is independent from the next node. Assuming that $\sigma_I = \sigma_{I+1} = \sigma$ (the standard deviations of the CTD across a node), then the SRL is used to estimate the standard deviation of end-to-end delay of node N as follows.

Our preliminary results verify that, the SRL can fairly approximate the change in the standard deviation of the CTD only when the traffic load distribution is uniform across all nodes, that is, both the referenced estimation node and any other node on the network have the same or close applied background load conditions. For our verification of the SRL we used the standard deviation of CTD of the first hop, to estimate the standard deviation of CTD at any other node.

The SRL tend to over estimate the accumulation of end-to-end CDV, with an increasing percentage of error as number of switching nodes increases. The error percentage in the SRL estimation is also increasing as the difference in load conditions between the referenced and desired node increases. However, the estimation accuracy of the SRL tends to grow with the intensity of equally applied load conditions. The reason for that being that the auto-correlation of CTD tend to die as the load increase as will be shown in the next section. Table 7 summarizes the over all performance of the SRL under uniform load conditions.

Table 7: Performance of the SRL under Uniform Load Conditions

Load		Hop Number				
		2	4	6	8	10
0	σ_{act}	0.900	1.190	1.452	1.645	1.829
	σ_{est}	0.944	1.335	1.636	1.889	2.111
	%error	4.9	12.2	12.7	14.8	15.5
30	σ_{act}	1.005	1.432	1.711	1.863	2.240
	σ_{est}	1.044	1.477	1.809	2.089	2.335
	%error	3.91	3.13	5.73	12.09	4.23
60	σ_{act}	1.208	1.721	2.070	2.411	2.856
	σ_{est}	1.258	1.779	2.179	2.516	2.813
	%error	4.20	3.39	5.29	4.39	-1.50
90	σ_{act}	4.438	6.021	7.739	9.154	10.28
	σ_{est}	4.356	6.160	7.545	8.712	9.741
	%error	1.85	-2.32	2.50	4.85	5.24

5.5.2 Auto-correlation of CTD/CDV.

The auto-correlation function is a measure of dependency between events. To investigate the dependency between the cell transfer delays *k* cells a part, the normalized auto-correlation function was used. The normalized auto-correlation function for cell *i* and cell *i+k* was defined as follow

$$a(k) = \frac{1}{(n-k)\sigma^2} \sum_{i=1}^{n-k} (CTD_i - \mu)(CTD_{i+k} - \mu)$$

In the case of CDV, the CTD term in the above equation was replaced by the CDV, where *n* is the total number of referenced cells, CTD_i is cell transfer delay of cell *i*, μ is the mean of CTD, and σ^2 is the variance of CTD. This measure of interest is useful in studying the effect of varying network parameters on the CTD behavior of a cell relative to another. In calculating correlation between the CTD/CDV of cells we used end-to-end delays of 8192 cells captured and stored by the ATM test set, Figure 26 and Figure 27 are the normalized auto-correlation curves obtained. The correlation curves revealed the following observations. First of all, by increasing the background traffic load while keeping the number of traveled hops the same, the correlation tends to decay. This phenomena is due to the fact that, greater traffic load means more randomness in the cell spacing and cells delay. Second, from Figure 27, it is shown that the cells are becoming less positively correlated with more nodes added to the traveled path, which contradict the founding in the studies done by [9,12].

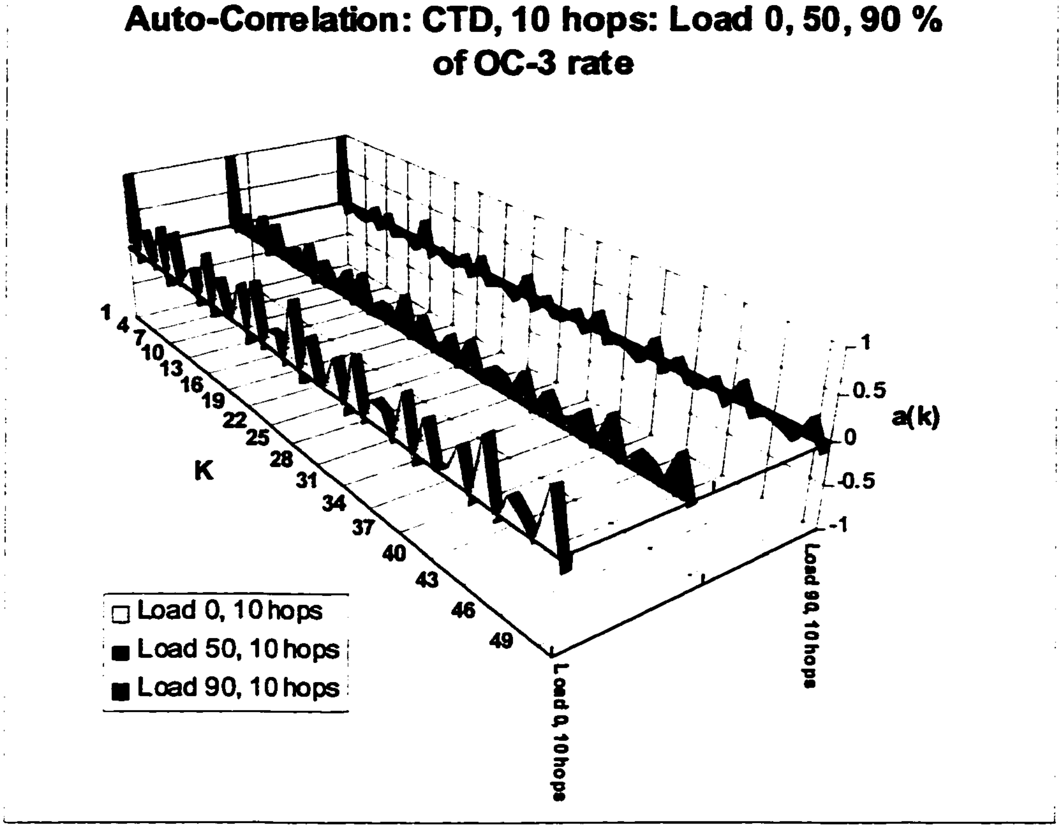


Figure 26: Normalized Auto-correlation of CTD as a Function of Background Load

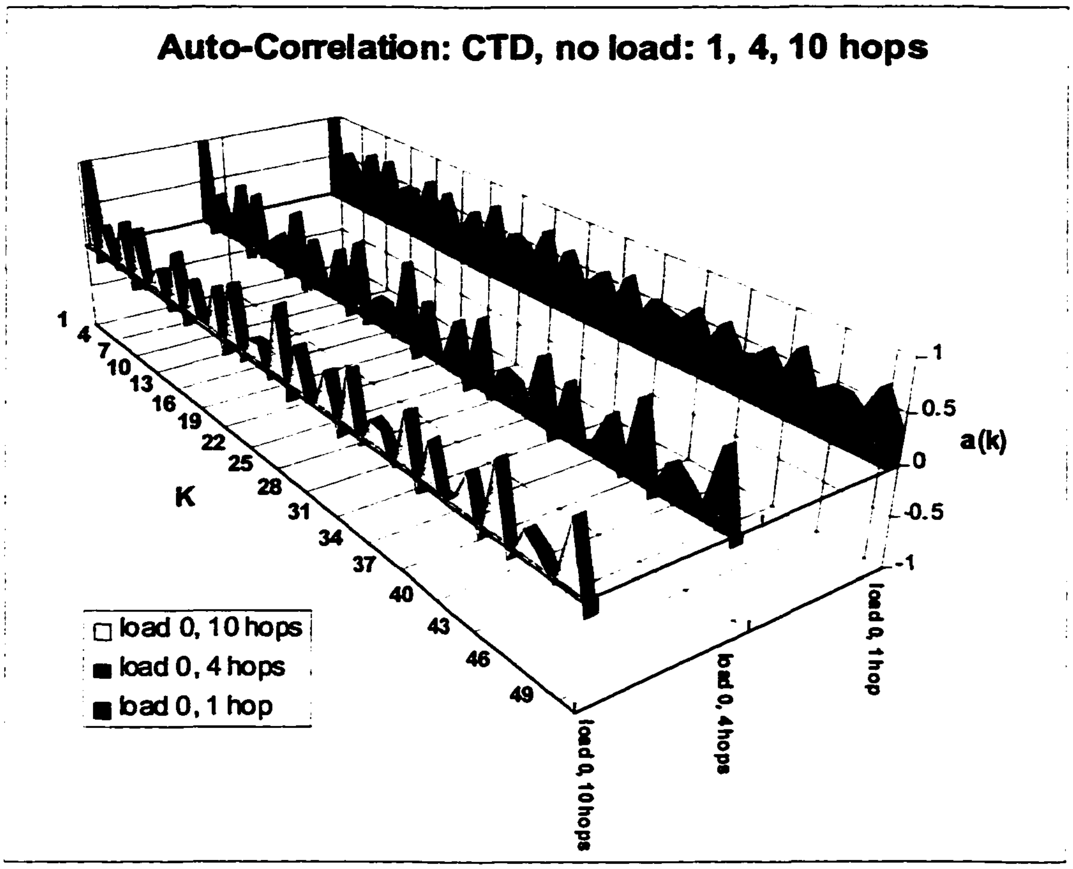


Figure 27: Normalized Auto-correlation of CTD as a Function of Number of Hops

Third, the addition in number of nodes to the path a cell is travelling implies less deterministic experienced queuing behavior among cells through the new longer path. Fourth, although in our study, both foreground and background traffic sources used were periodic, however our results revealed that no periodicity in the local maxima of the auto-correlation function, as reported in the simulation study [9]. Fifth, although it is not obvious in the presented figures because of the window size taken, the correlation decays as the cell number increases, because as the distance between the two referenced cells increases the associated delay correlation between them decreases.

Finally, by studying the normalized auto-correlation behavior of CDV, as in Figure 28, no specific pattern can be concluded as a function of number of hops for the no load condition. However, correlation between CDV of cell *i* and cell *i+k* is decaying as the load increases which is consistent with the CTD case, Figure 29. It is also worth mentioning how the load becomes the dominant factor in shaping up the correlation behavior as seen in Figure 30. By comparing Figure 28 and Figure 30 its clear how by increasing the load from 0% to 90% the correlation as a function of number of hops changed. In Figure 28 there were more correlation in the 4 hop case than both the 1 and 10 hop node situations; however, when the load increased to 90% the most correlation is noticed to be in the 1 hop data set and similar for data sets from the 4 and 10 hops.

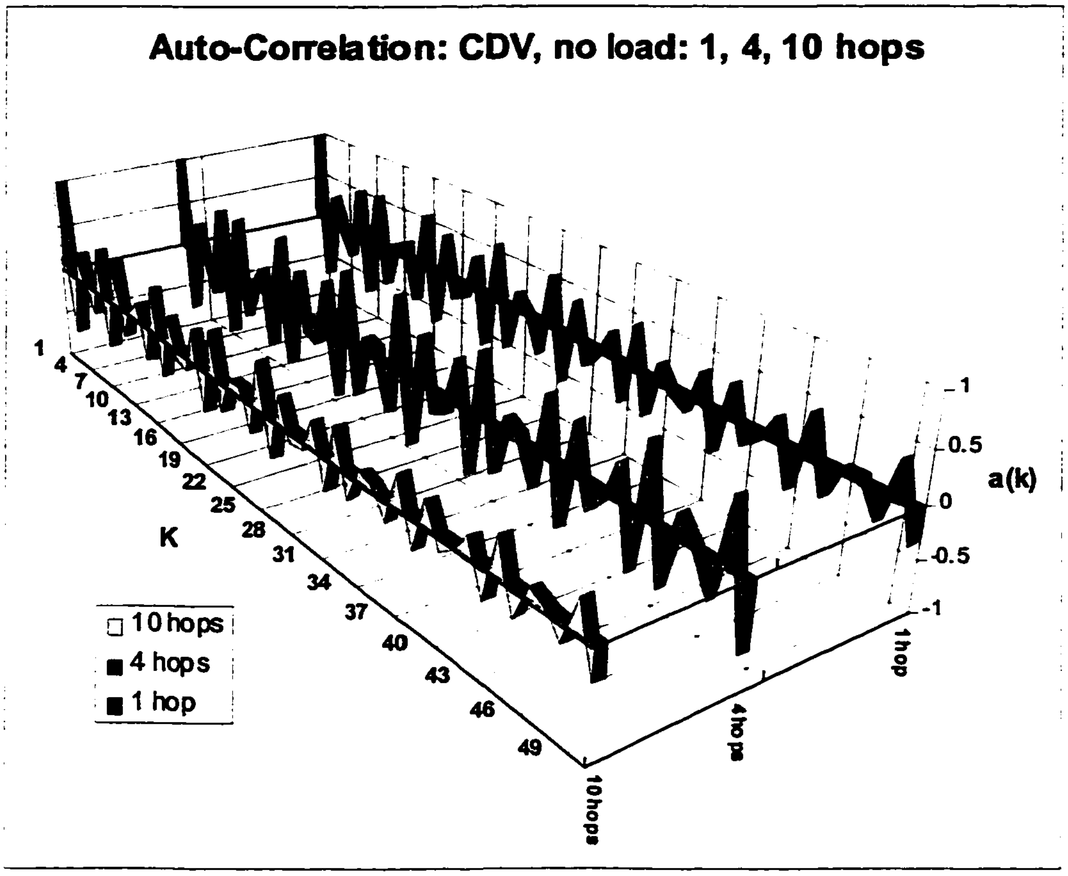


Figure 28: Auto-correlation of CDV as a Function of Number of Hops

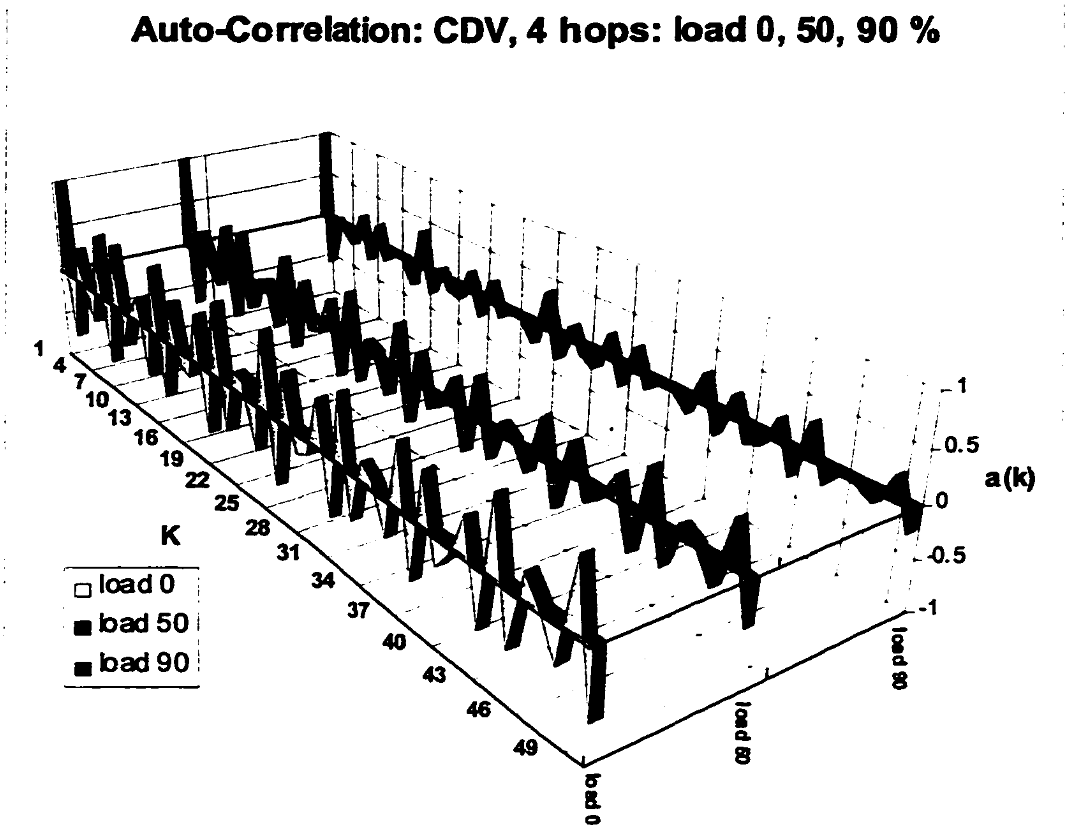


Figure 29: Auto-correlation of CDV as a Function of Background Load

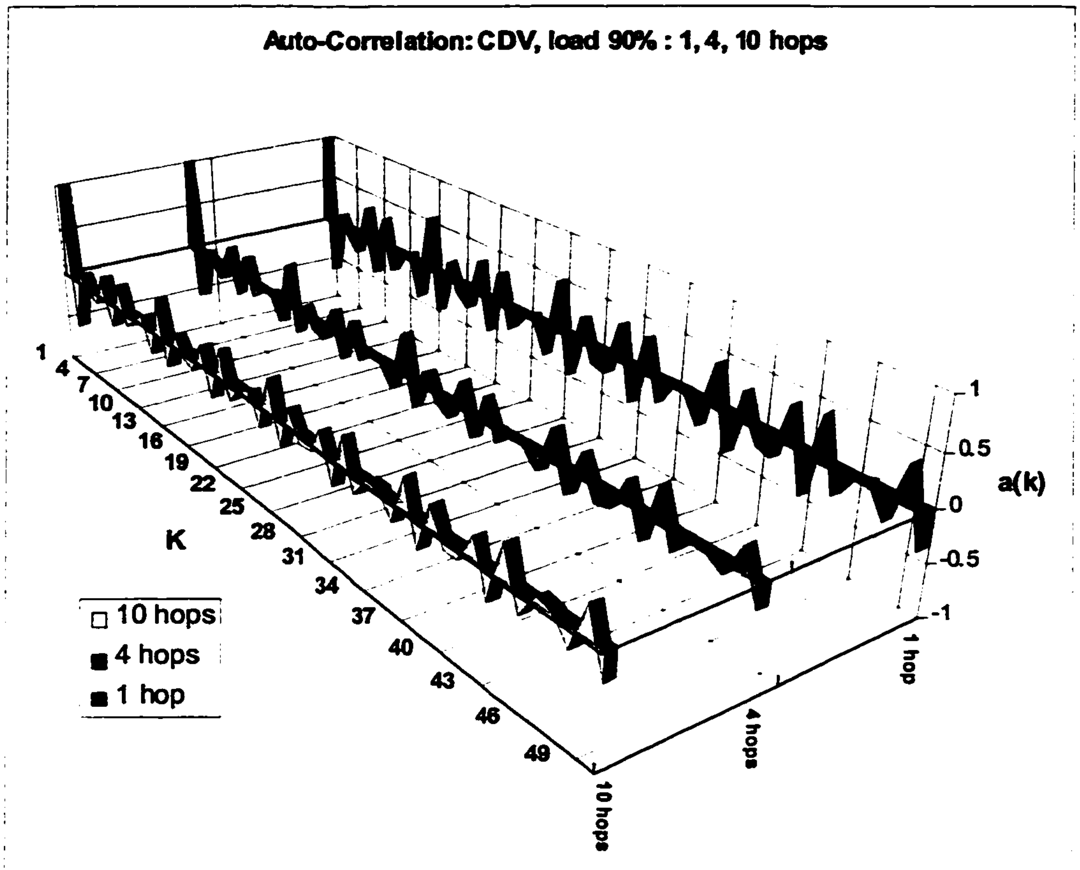


Figure 30: Effect of Background Load on the Auto-correlation of CDV for Multiple Hops

5.5.3 CDV/CTD Distribution and Curve Fitting.

In determining the type of distribution for the collected CTD and CDV data of a run, we first used a smoothing filter provided by the ATM test set. The filter averages all histogram bars with different heights between each cell interval, then the Mean Square Error (MSE) criteria was used to optimize the parameters of each distribution type to best fit the distribution of the on hand data. Based on the MSE obtained for several standard distributions, attempts were made to fit the CTD data to both the Normal and Raleigh distribution and the CDV to the normal distribution. As a goodness of fit measure the Chi Square test was performed. Table 8 summarizes the results of the attempts made to fit actual data for the runs of {(0% load, 10 hops, CDV), (50% load, 2 hops, CDV)} to the normal distribution and the runs of {(0 load, 10 hops, CTD), (50% load, 2 hops, CTD)} to the Normal and Rayleigh distributions.

$$f(x) = \frac{x}{\alpha^{2}} e^{\frac{-x}{2\alpha^{2}}} \quad \text{..........} \quad \text{..........} \quad \text{...} \ \textit{Rayleigh}$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^{2}}{2\sigma^{2}}} \quad \text{..........} \quad \text{..} \ \textit{Normal}$$

The *p.d.f* of the Normal and Rayleigh distributions are given by the two following equations respectively

$$\sigma^{2} = (2 - \frac{\pi}{2})\alpha^{2}$$

$$E\{x\} = \alpha\sqrt{\pi/2}$$

Where μ and E{x} are the means, and σ is the standard deviation.

The Chi-Square test as shown from the results in Table 8 reveal the failure of fitting attempts. Figure 31, Figure 32, Figure 33 and Figure 34 show the histograms of the rums for {(CDV, CTD: Load 0, 10 hops), (CDV, CTD: Load 50, 2 hops)} and the attempts to fit them to a normal or Rayleigh distribution. The Chi-Square test for a population on N samples is described by the following equation

$$\chi^2 = \frac{\sum_i^N \left(x_i - x_i^{est}\right)^2}{x_i^{est}}$$

Where X_i is the measured value and X_i^{est} is the estimate value

Table 8: Chi-Square Fitting Results

		Test case	χ^2	95% χ^2
CTD	Normal	Load 0, 10 hops	1163.4	14.7
		Load 50, 2 hops	3355.2	12.0
	Rayleigh	Load 0, 10 hops	17620.1	14.7
		Load 50, 2 hops	8388.3	12.0
CDV	Normal	Load 0, 10 hops	935.2	25.0
		Load 50, 2 hops	5733.2	23.7

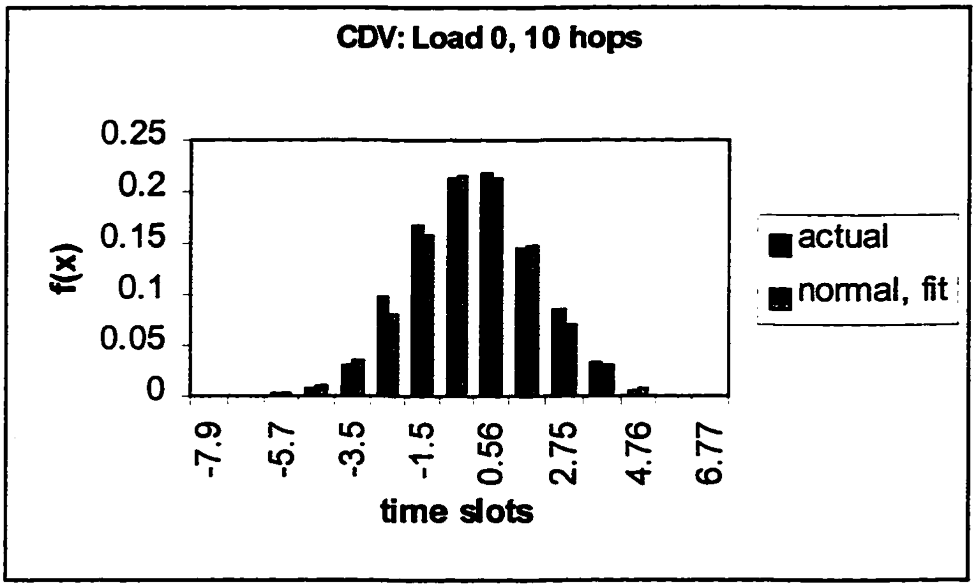


Figure 31: Data Fitting of CDV, 0% Load and 10 Hops Case

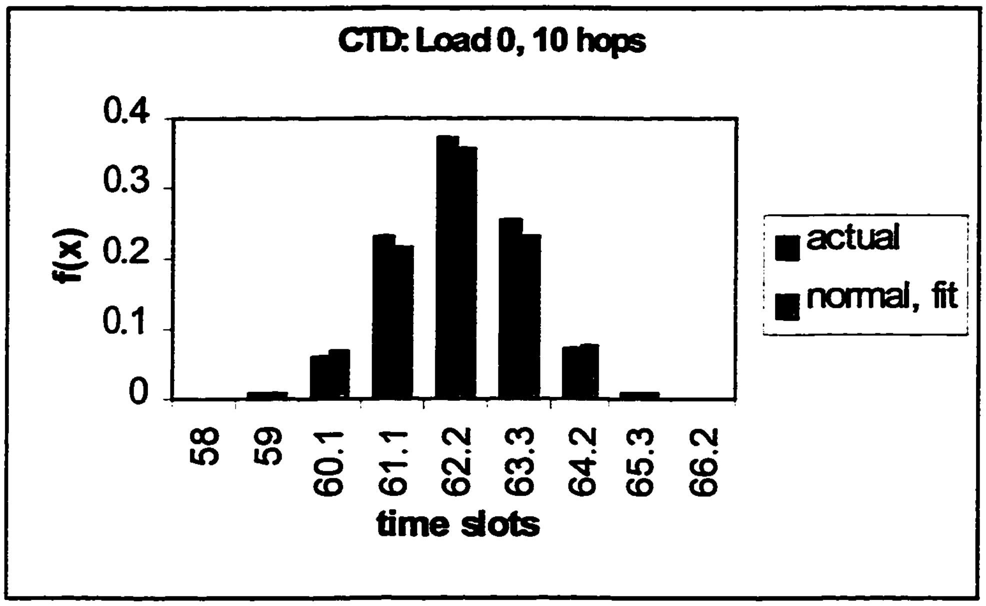


Figure 32: Data Fitting for CTD, 0% Load and 10 Hops

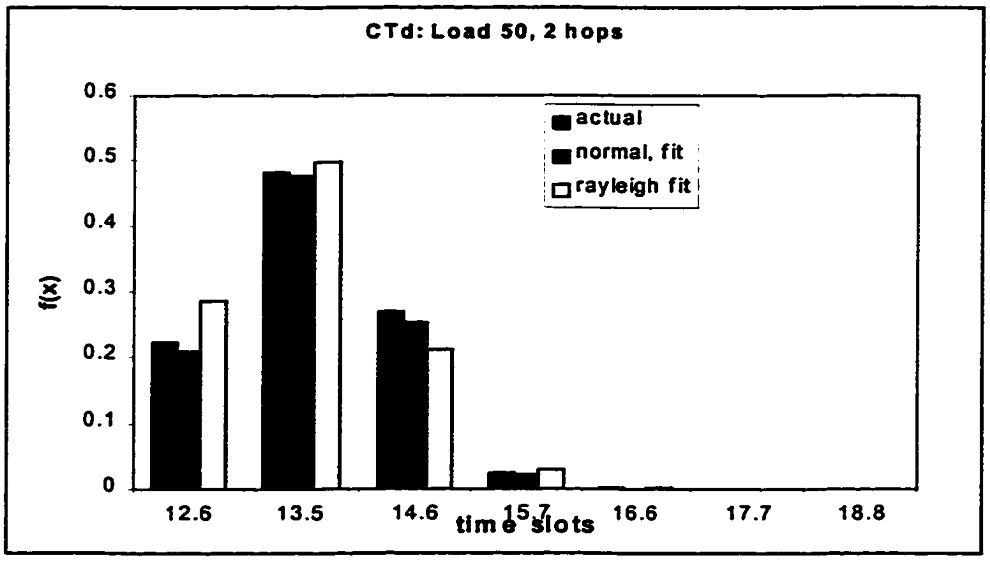


Figure 33: Data Fitting for CTD, 60% Load and 2 Hops

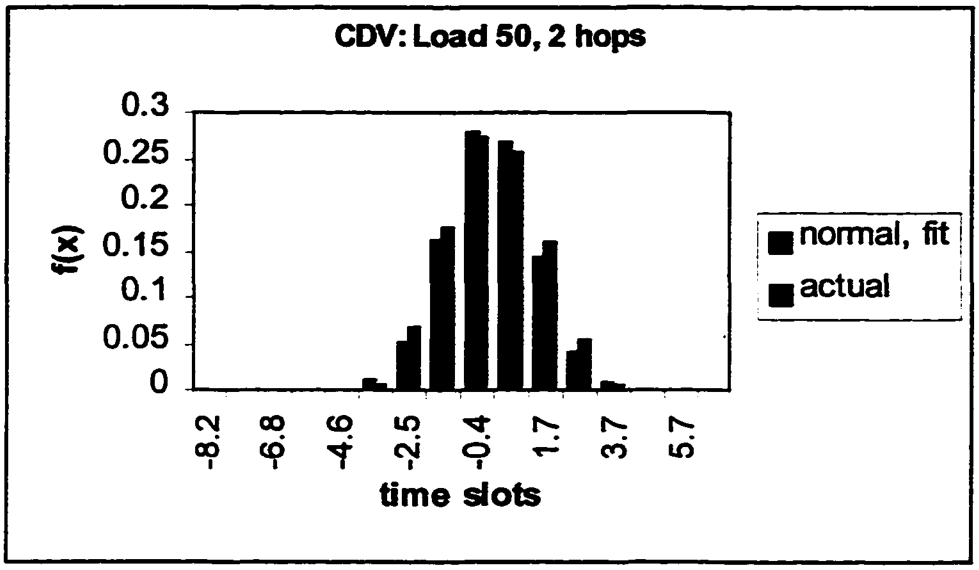


Figure 34: Data Fitting for CDV, 50% Load and 2 Hops

6 MPEG-4 System: Timing Model and Multiplexing

6.1 Introduction.

The Motion Pictures Expert Group (MPEG) was established to develop a common format for coding and storing digital video and associated audio information. After completing the MPEG-1 and MPEG-2 standards, MPEG is currently completing the work on developing a new coding standard for very low bit-rates. This activity, known as MPEG-4 [30,31], is targeted at video bit rates from 6 kbps to 4 Mbps for applications such as:

- Videotelephony and videoconferencing with low delay requirements
- Symmetric and asymmetric remote expertise and classrooms (tele-medicine, tele-mechanic, and distant learning).
- Interactive information based retrieval.
- Very low bit-rate scalable multimedia distribution over error prone networks for portable and mobile receivers.

6.2 MPEG-4 System Overview.

MPEG-4 is an object based compression technique where natural and synthetic audiovisual objects (AVO) can be arranged into an audiovisual scene using scene description. The MPEG-4 specifications describe three layers, the compression layer, the system or sync layer and the delivery layer.

At the compression layer elementary streams (ES) are formed, where an ES may contain one of the following:

- Object Descriptor Stream: Identify describe and associate ESs together.
- Scene Description Streams: A scene descriptor is encoded as a Binary Format for Scene (BIFS) and it describes the spatial and temporal positioning of audiovisual objects in a scene.
- Media Streams: Coded audiovisual information.
- Object Content Information Streams: Carries information such as content classification descriptors, keywords, language...etc.
- Timing information: Clock reference values for synchronization and decoding AVOs.

The system or sync layer (SL) is where each ES is divided into one or more access unit (AU). Each AU is then divided into SL-packets to form a SL-packetized stream (SPS) where timing, synchronization and random access information as well as padding bytes are inserted to facilitate proper encapsulation, decoding, composition and playback of media objects. However, a sync layer packet does not contain information on its length or on an ES identification, this information is to be provided by the lower layer protocols by which the SL packetized stream is encapsulated for delivery. The interface between the compression layer and the sync layer is called the elementary stream interface (ESI).

An optional and application dependent layer called the FlexMux layer exists between the sync layer and delivery layer. The FlexMux layer is used to interleave one or more elementary streams into a FlexMux stream. The FlexMux layer can also be utilized for the proper framing of SL-packets to fit the profile of the underlying transport protocol.

The Delivery layer in MPEG-4 consists of the Delivery Multimedia Integration Framework (DMIF). This layer is media unaware but delivery technology aware. The delivery layer ensures transparent access to MPEG-4 content irrespective of the transport network technology used (e.g. IP or ATM). The boundary between the sync layer and the delivery layer is called DMIF application interface (DAI). The DAI is a generic API dependent on the operation system and the programming language used. It is used to define the functions of the DMIF layer and to allow the DMIF user to specify the quality of service (QoS) requirements of the desired stream to be transported using any permissible transport protocol stacks (TransMux layer). The TransMux layer is not specified by MPEG-4 and could consist of any stack of protocols with properties consistent with the MPEG-4 requirements.

6.3 Multiplexing and Transporting MPEG-4 Video.

The generic term TransMux layer implies the use of any potential protocol stacks that supply a transport multiplex for content compliant with the MPEG-4 standards. Examples of such protocol stacks are the RTP/UDP/IP, H.223/PSTN or MPEG-2

TS/AAL/ATM. As an optional intermediate step and where it applies the MPEG-4 specified FlexMux tool can be used. Figure 35 shows the MPEG-4 packetization and multiplexing process. As could be seen from Figure 35 the FlexMux layer is used to interleave one or more elementary streams into one FlexMux stream, which is, then mapped into a TransMux stream. Each SL-packetized stream is identified by a unique FlexMux channel. The FlexMux layer could be bypassed if only one elementary stream is to be mapped into the TransMux stream.

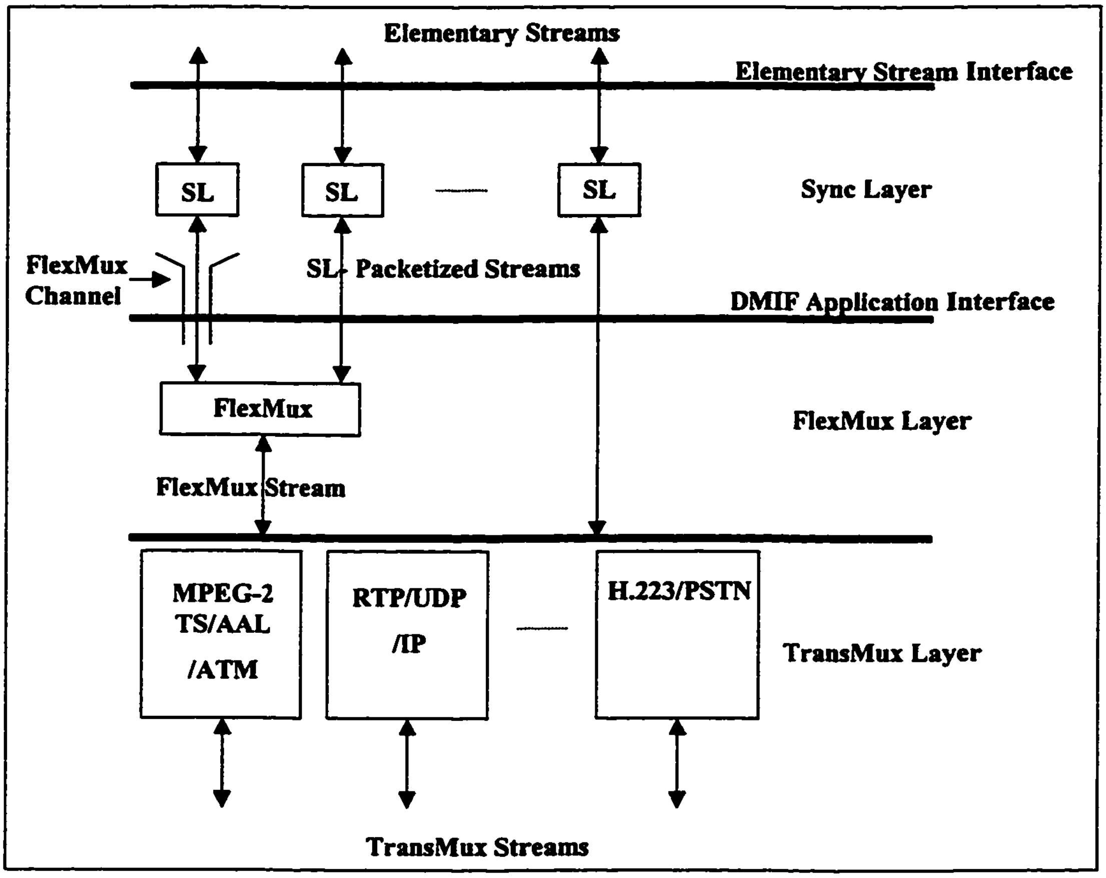


Figure 35: MPEG-4 Packetization and Multiplexing Process

The structure of the FlexMux layer is designed to provide two modes of operation, the simple mode and the MuxCode mode as shown in Figure 36.

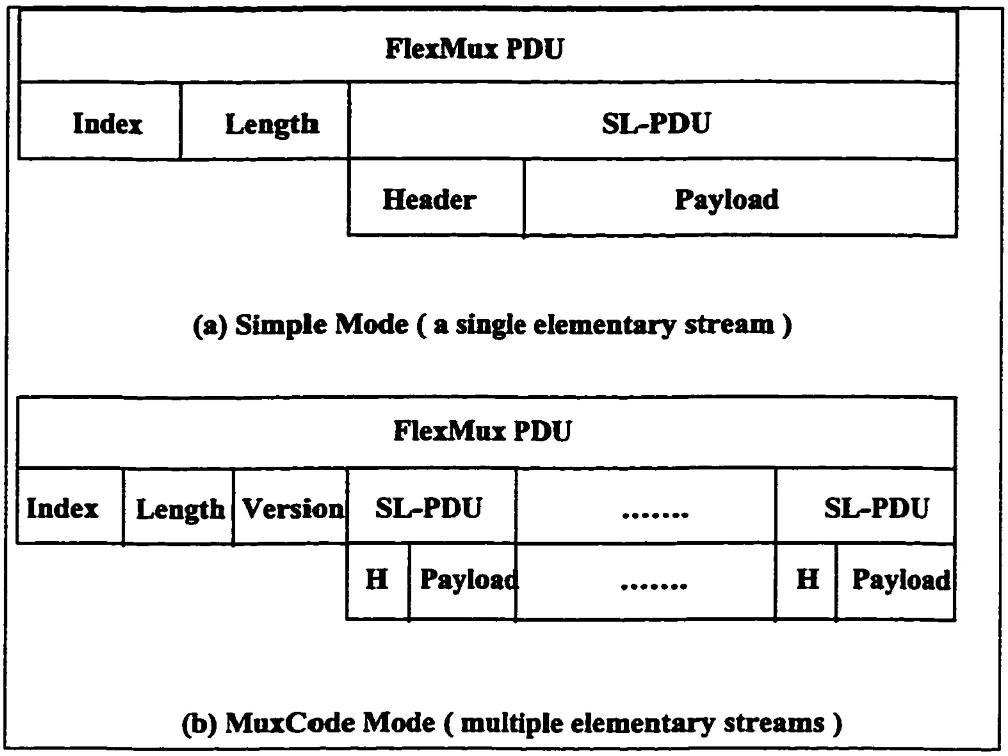


Figure 36: FlexMux Mode Operation

The basic data unit of the FlexMux is the FlexMux packet. The FlexMux packet is variable in length and encapsulates one or more SL-packets originated either from the same or different elementary streams. SL-packets originated from different elementary streams are identified by the number of the FlexMux channel carrying them. Therefore, arbitrary interleaving of SL-packets is possible and the sequence of resulted FlexMux packets is called a FlexMux stream. Although the FlexMux stream can be parsed as a single data stream when retrieved from storage or transmission, it requires proper framing of its packets by the underlying layer for random access, error detection and error recovery.

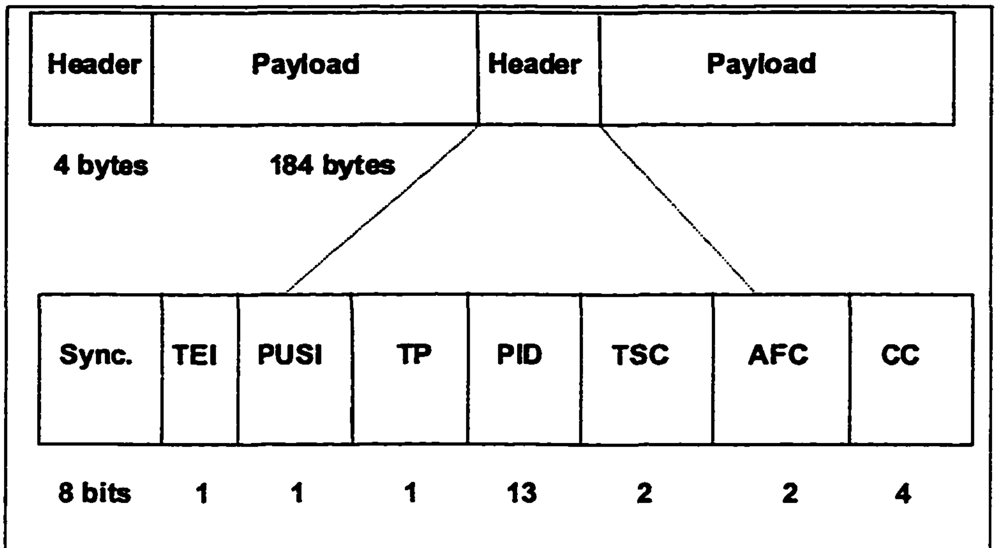
As mentioned above, two different modes of operation are defined for the FlexMux and identified by the value of the index field of the FlexMux packet header, the simple mode (index <=239) and MuxCode mode (240 <=index <=255). In the simple mode one SL packet is encapsulated as a FlexMux packet and tagged by an index equal to the FlexMux channel number. In the MuxCode mode one or more SL packets are encapsulated in one FlexMux packet. In this mode the index field of the FlexMux packet header is used to associate and define the allocation of the FlexMux packet payload to different FlexMux channels.

6.4 Transporting MPEG-4 using the MPEG-2 Transport Stream Syntax.

Due to the desire of making MPEG-4 as independent from the network as possible, the MPEG-4 standards do not specify the TransMux layer to transport the MPEG-4 objects over the network. The MPEG-2 standards defined a transport stream protocol [3,6] to multiplex and transport MPEG-1 and MPEG-2 elementary streams. This transport stream could be easily extended and used for the transport of MPEG-4 data as well. From the set-up box manufacturer point of view, it is desired to make a universal transport protocol so that only a single device is needed at the receiving end to decode and playback an

MPEG coded bit stream. For that reason the MPEG-2 transport stream protocol is considered to be a good candidate for the transport of MPEG-4 data.

A transport stream as defined by the MPEG-2 standards consists of fixed size packets of 188 bytes, the first four bytes are the header fields, while the remaining 184 bytes form the payload and any other additional headers. Figure 37 shows the structure of TS packets. The payload may consist of one or more of the following: a) Encapsulated elementary stream packets (SL-packets, FlexMux packets). b) Adaptation Field (AF), which may contain one or more of the following: MPEG related information like the Program Clock Reference (PCR) field used by the decoder for clock recovery, private information like the encryption key management or simply padding bits for boundary alignments of TS packets. c) Program Specific Information (PSI) tables to provide the decoder with the necessary information for de-multiplexing, decrypting and decoding of TS.



Sync : Synchronization Byte
TEI : Transport Error Indicator
PUSI : Payload Unit Start Indicator
TP : Transport Priority
PID : Packet ID
TSC : Transport Scrambling Control
AFC : Adaptation Field Control
CC : Continuity Counter

Figure 37: MPEG-2 TS Packet Structure

When framing SL packets for encapsulation within an MPEG-2 TS and the FlexMux tool is not used, a TS packet will contain a complete or a part of a SL packet, but if the FlexMux tool is used a TS packet may contain one or more FlexMux packets.

Since neither the SL nor the FlexMux packet header syntax in MPEG-4 provides for a sync word identifying the start of such packet, the SL packet or the FlexMux packet should start at the first byte of the TS packet payload. In this case, the payload unit start indicator flag of the TS packet header should be set to 1. To adjust the TS packet size to 188 bytes, padding bits of the SL can only be applied through the FlexMux packet when the FlexMux tool is used, otherwise, padding bits of the TS adaptation field are used.

When an MPEG-4 elementary stream is carried by an MPEG-2 TS, it is identified by the Elementary_Stream_PID field within the TS program mapping table. The Elementary_Stream_PID field specifies the PID of TS packets that convey the initial object descriptor section specified by MPEG-4. The MPEG-4 initial object descriptor section consists of an initial object descriptor and a stream map table and it is transmitted as other PSI sections. Using the stream map table and the PID field of the TS packet header, TS packets of MPEG-4 are mapped to their specific elementary streams.

6.5 Timing Model Overview.

In MPEG-4 the timing model is very similar to MPEG-2 since it relies on clock references and time stamps to properly decode, compose, and playback media objects [46,47]. The concept of a clock with its clock references is used to synchronize the clock at the receiver to the clock at the video source. Time stamps are used to indicate the precise decoding, composition and playback instances of events (video frames in MPEG-2, video object planes and composition units (CU) in MPEG-4).

The MPEG-4 sync. layer packet consists of a SL packet header and a SL packet payload. The SL packet header consists of many specific fields such as sequence number for continuity check in case of data loss, time stamps such as the object clock reference (OCR), composition time stamps (CTS) and decoding time stamps (DTS). OCR values are encoded as sample values of the source object time base (OTB) and inserted periodically in the SL packet header with a frequency depending on the application.

When a SL packet caries an OCR value, an OCR flag is set indicating that. Each ES when generated by an encoder may have its own OTB or may be slaved to an OTB of another ES. In the latter case, the OCR_ES_ID field in the SL packet header indicates the ES from which the time base for the slaved ES should be derived. The decoder uses the OCR values to reconstruct and adjust the OTB of each stream and map it to the encoders' OTB using well-known PLL techniques. When all media objects in a session are using the same OTB or when one of OTB of an ES is declared as a master clock it is then possible to lock the system time base (STB) of the receiving terminal to that OTB. Otherwise, since the OTBs of all objects may run at different speeds and resolutions than the STB, a mapping method is necessary to map the values of time stamps expressed in terms of any OTB to the STB of the receiving terminal. A detailed explanation of how this can be done is described in [30]. When MPEG-4 is transported using the MPEG-2 TS syntax, the PCR of the TS can be used to reconstruct the STC while all other OTBs can be mapped to it. Figure 38 shows the timing and decoding model in MPEG-4 for two elementary streams.

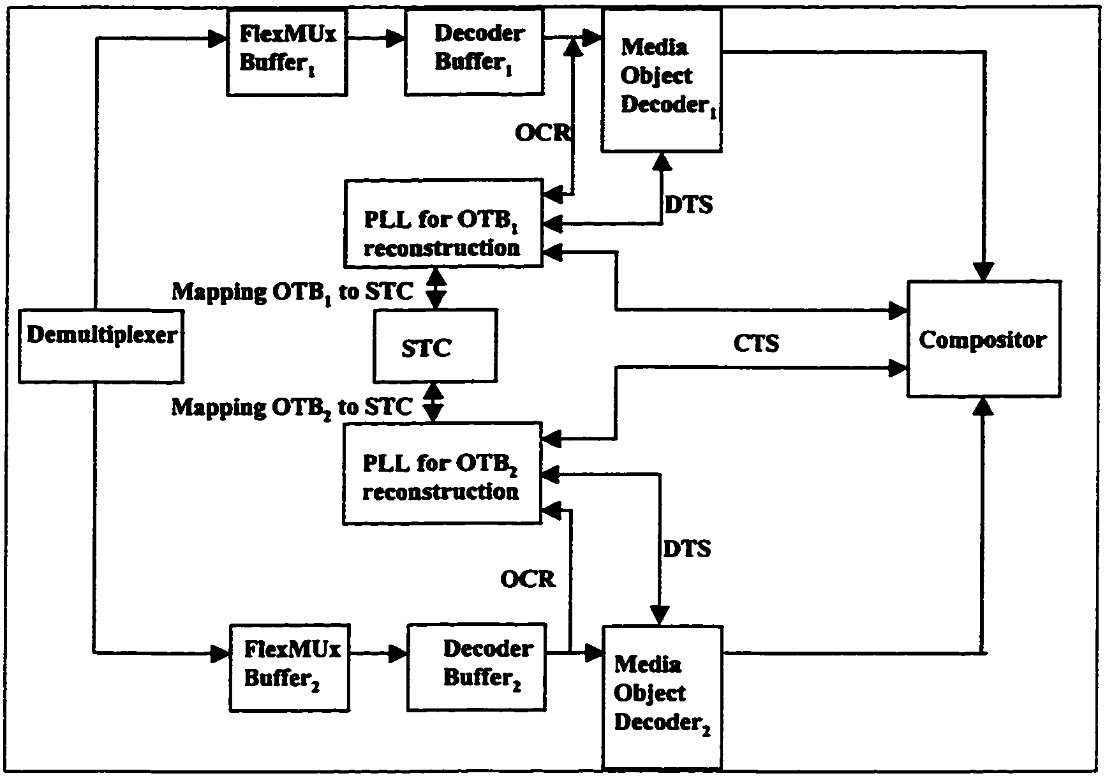


Figure 38: Timing and Decoding Model for Two MPEG-4 Elementary Streams

As shown in Figure 38 data enters the FlexMux buffer from the network-decoder interface demultiplexer. Information is received by the decoder buffer in the form of SL packets. The SL packet headers are parsed and removed at the entrance to the decoder buffer and access units are formed. Access units with OCR values are extracted for OTB re-construction via a PLL. The reconstructed OTB then mapped to STC to account for any resolution or speed differences. The media object decoder extracts access units from the decoder buffer at precisely defined point in time and decode them according to their

DTSs to form composition units which are to be composed and integrated within a scene according to their CTSs and the scene descriptor.

7 De-jittering in MPEG

7.1 Introduction.

When multimedia information traverses over packet switched networks, it experiences impairments that may lead to degradation in the required quality of service (QoS). Multimedia systems store, retrieve, and communicate complex representation of information. Maintaining timing relationships among packets in a single media stream or between packets from different media streams is an essential criterion in multimedia networking. The Motion Pictures Expert Group (MPEG) is the most recent and international accepted multimedia standards. The MPEG standards assume a constant network delay when the multimedia information is transported. However, since it is difficult if not impossible to maintain a constant network delay, it is the function of the synchronization and de-jittering algorithms to re-adjust the timing relationship between multimedia packets from the same or several media streams to assure a synchronized playback of information [10,23,28,29,35,38,40,41,42,43].

Using the assumption of a constant delay network between the source and destination the MPEG standards recommend the use of a Phase-Locked Loop (PLL) to keep the local decoder clock synchronized to the encoder clock. This synchronization is achieved by comparing the clock stamp references in the received MPEG stream with the instant time from a local register, which is driven by the decoder clock. From this comparison, a phase error can be derived to adjust the decoder clock, as shown in the block diagram of Figure 39. When the phase error is positive, the decoder STC is running behind the source clock, and data has to be held in memory for an additional delay before being decoded and therefore more memory is needed to be allocated. When the phase error is negative the STC at the decoder is running ahead of the source clock and data is being decoded early.

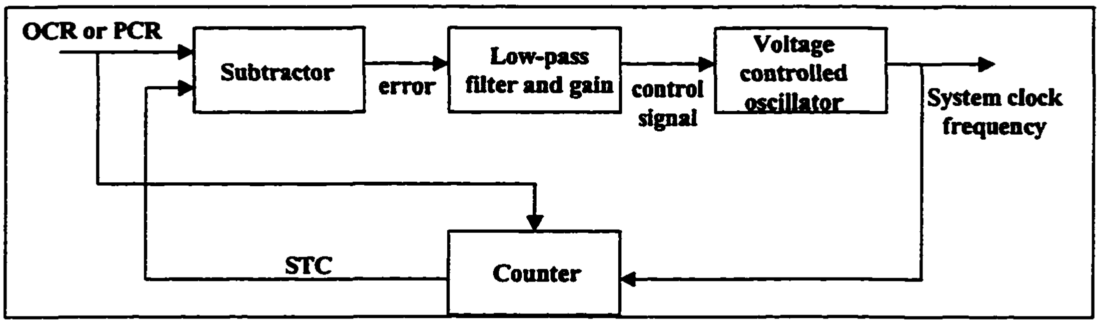


Figure 39: STC Recovery Using PLL

If the network delay varies, jitter will be passed to the system clock reference. As a result, the PLL must perform additional filtering in order to correctly estimate the STC at the decoder. This in turn slows down the responsiveness of the PLL and affects the maximum phase error introduced by the PLL between the encoded clock reference values based time line and the local system time clock reference line.

The error filtered by the PLL is dependent on the delay of the received packet carrying the PCR, which will only reflect the true jitter value if the delay of that packet is equal to the average network delay. To assure an unbiased reconstructed STC, the filtered jitter value by the PLL must be adjusted to account for the difference between the delay of the received packet with a PCR and the average network delay. By choosing not to adjust for this difference, the decoder buffer might overflow or underflow.

To account for any introduced network jitter to assure a stable recovery of the STC, de-jittering must be performed before the encoded clock reference values are passed to the PLL used in recovering the STC. The next section describes our proposed de-jittering scheme and how it can be applied when a TS carrying MPEG-2, or MPEG-4 elementary streams is being used.

7.2 Description of the De-jittering Scheme.

The proposed de-jittering scheme is based on the statistical estimation of jitter added to packets carrying reference clock values from which the system time clock is to be recovered. Such packets are usually sent periodically and can be identified by means of a header flag or a field indicating the presence of a reference clock value. For example, packets carrying program clock reference values in a TS are identified by the PID field of the TS packet header. In MPEG-4, the header of SL packets also carries a flag indicating the presence of an object clock reference. Once a packet with a PCR or an OCR is identified, it is parsed for clock adjustment to account for introduced jitter and therefore, minimize the phase error of the PLL to provide a more stable system time clock. The applied adjustment step is calculated by converting the jitter from seconds to clock tics based on the resolution of the clock reference (i.e. OCR, or PCR).

7.2.1 Basic Assumptions and Steps of De-jittering.

1. The sender periodically transmits packets with a nominal period T.
2. A reference clock value carrying a snap shot of the encoder time base is periodically inserted and sent periodically every T_{ref} second (N packets). The first packet with reference clock value received is used to initiate the decoder clock by setting its system clock to that value, and this is done by default as per the MPEG standards. To perform de-jittering a local counter is also initiated to the reference clock value. This local clock is used for the registration of arrival times and updated along with the STC. In the case of MPEG T_{ref} could be the inter PCR or inter OCR period depending on the specific transport mechanism used.
3. Register the arrival times of all packets (N) received between two successive packets carrying a reference clock value and calculate their jitter values (J_i, $i=1,\ldots N$) as the difference between their theoretical (A_{ti}) and actual (A_{ai}) arrival times. The theoretical arrival times of the N packets are calculated using the arrival time of the packet carrying the reference clock value ($A_{a,refj}$) as a reference point.

$$A_{ti} = A_{a,refj} + i\,T) \qquad (1)$$

$$J_i = A_{ti} - A_{ai} \qquad (2)$$

4. Calculate the sample mean jitter ($\bar{\mu}_k$) as

$$\bar{\mu}_k = \frac{1}{N}\sum_{i=1}^{N} J_i \qquad (3)$$

where the value of $\bar{\mu}_k$ can be positive, negative or zero depending on the delay experienced by the reference packet (D_{rcv}). The sample mean jitter, $\bar{\mu}_k$, will be positive, negative or approach zero depending on D_{rcv} and N. Figure 40 shows a typical timing diagram for the received packets.

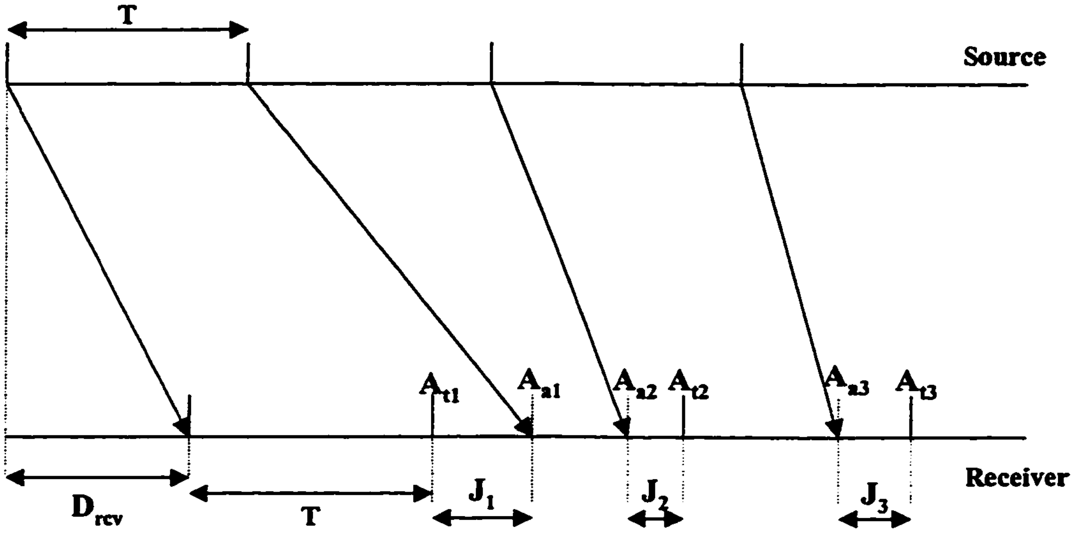


Figure 40: A timing Diagram for Received Packets

5. Estimate the jitter of the TS packet carrying the next reference clock value as the difference between its actual arrival time and its corrected theoretical arrival time ($A_{c.t,refj}$). The corrected theoretical arrival time of the video packet carrying a reference clock value can be found as in equation (4).

$$A_{c.t,refj} = (A_{a,refj-1} + (N+1)T) - \bar{\mu}_k \qquad (4)$$

$$J_{refj} = A_{c.t,refj} - A_{a,refj} \qquad (5)$$

By using the $A_{c,t,refj}$ the value J_{refj} is now estimated based on the average network delay rather than on the delay of a single packet. The corrected theoretical arrival time of the new received packet carrying a reference clock value will now be used as a reference point for the calculation of the sample mean jitter of the next N packets as in step 2 and 3 with $A_{a,refj}$ replaced with $A_{c,t,refj}$. Since the jitter calculation of the next N packets is now based on a reference packet with an estimated average network delay, $\bar{\mu}_k$ should be close to zero and exhibit little variation under the same network conditions.

6. Using the resolution of the STC, translate the jitter value J_{refj} to an adjustment step (Δ) in terms of number of STC tics and re-stamp the reference clock value to account for the adjustment before sending it to the PLL.

$$\Delta = J_{refj}\,(sec) \cdot STC\ resolution\ (tics\ per\ sec)) \qquad (6)$$

This will minimize the phase error of the PLL and provide a more stable STC reconstructed based on the average network delay.

7. Once a new packet with a reference clock value is received, the steps of de-jittering (2 through 6) are repeated till a new video packet with a reference clock value is received.

7.3 Experimental and Simulation Setup.

7.4 Experimental setup.

To test the performance of the proposed de-jittering scheme, experiments were conducted to collect jitter data when MPEG video is transported over an ATM network as shown in Figure 41.

An MPEG-2 video was generated at constant TS rate of 4.096 Mbps from the MPEG encoder and transported over the ATM network over an OC-3 connection using AAL5 with two TS packets per AAL-5 PDU [2, 5]. A program clock reference value was inserted in transport stream packets every 80 ms by the encoder, therefore given the length of a TS packet being 188 bytes, N is approximately 218 TS packets or 109 AAL5 PDUs since two TS packets are encapsulated as an AAL5 PDU. The network test equipment (NTE) traffic generator module was used to generate and load the ATM switch with background traffic to introduce jitter within the MPEG stream. The MPEG stream was passed through the NTE's analyzer module, just before it was being decoded, to register the arrival times of AAL-5 PDUs carrying the video stream and mark those with a reference clock value, in this case a PCR.

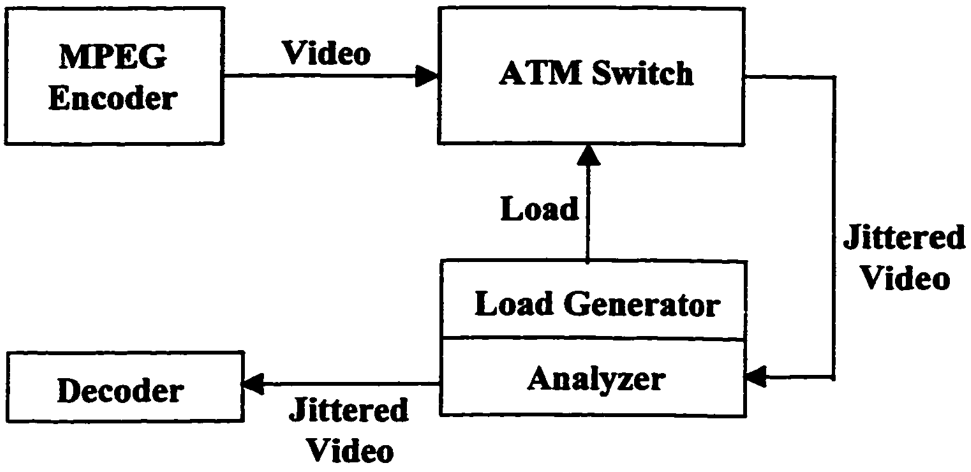


Figure 41: Experimental Setup

7.5 Simulation.

Using both simulated and real time jitter data collected from the experiments, simulations were conducted to test the performance of the de-jittering scheme. Figure 42 shows the flow of the simulation.

Beside the jitter introduced by the network, there is additional jitter introduced as a result of encapsulating two transport stream packets to form an AAL-5 PDU without any regard to the position of a reference clock value within the PDU. This encapsulation scheme introduces packing jitter since when a reference clock value is in the first transport packet it has to wait for the second packet to constitute a PDU [34]. The reference clock value packing jitter occurs when the reference clock value insertion mechanism switches from inserting a reference clock value from the first to the second TS packet of the PDU, we will refer to this phenomenon as a reference clock value switch or a PCR switch.

TS packets carrying the PCR value can easily be identified by parsing for their 13 bit Packet ID (PID) located at the second and third bytes of each TS packet. The parser function used in the de-jittering scheme parses the header fields of TS packets for their PID and continuity counter (CC) (a 4-bit field). When there is data loss, TS packets are lost in pairs in case of bit errors or cell losses due to the AAL-5 Cyclic Redundancy Check (CRC) being applied over the whole PDU. Loss of TS packets is indicated by a discontinuity in the CC.

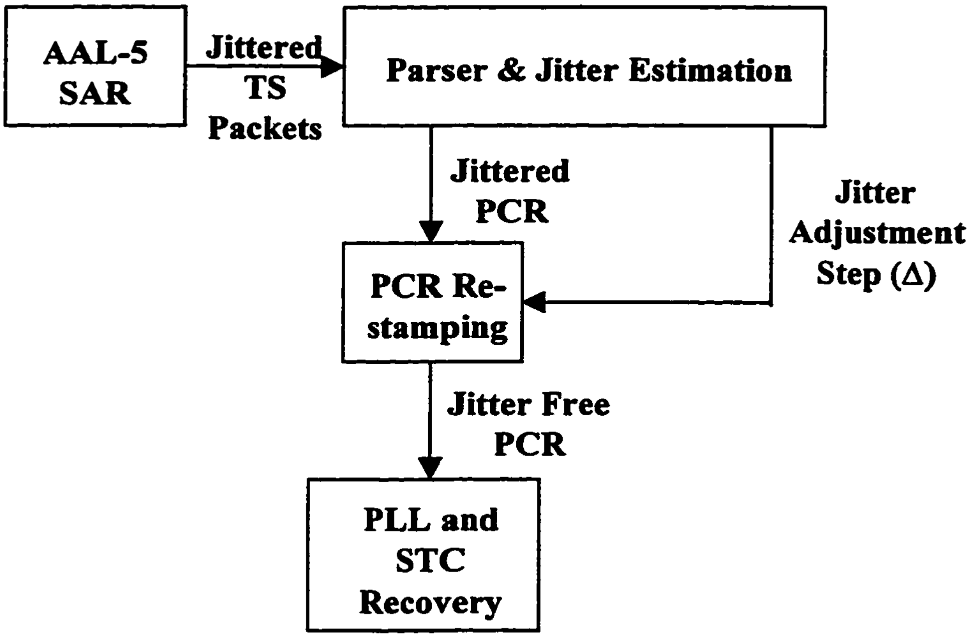


Figure 42: De-jittering Simulation

The de-jittering scheme as described in the previous section was simulated and applied using both simulated and real time jitter data. Certain modifications were applied to the de-jittering scheme due to the use of AAL-5 and MPEG-2 transport stream. For example, when MPEG video is received, two TS packets are buffered in AAL-5 to form a PDU, then transported to the decoder system layer. Network jitter in this scenario is only due to variation in the arrivals of the last ATM cell of a PDU and that implies jitter calculations with a nominal period of $2T$. Moreover, T_{ref} in the de-jittering scheme was replaced with the inter PCR period (T_{pcr}) since the MPEG-2 transport stream was used.

In the case of a PCR switch the jitter value was adjusted to offset the packing jitter as in the following equation

$$J_{pcr\,j} = A_{c,t,pcr\,j} - A_{a,pcr\,j} \pm (188 \cdot TS\ rate) \tag{7}$$

A PCR switch in the simulation was identified by a similar method as in [34].

7.6 Simulation results and analysis.

The performance of the de-jittering scheme can be judged by looking at the inter frame decoding jitter. In MPEG a number of video frames[1] are buffered before the start of the decoding process. A video frame is decoded when its DTS as compared by the STC is reached. If video is encoded at M frames per second, the inter frame decoding period will be $1/M$ second. However, since the STC is reconstructed and adjusted based on the received PCR values, the inter frame decoding period will only be constant if the inter PCR jitter is zero. The inter frame decoding jitter (J_d) can be found as

$$J_d = (C_k - C_{k-1}) - (1/M) \tag{8}$$

where C_k and C_{k-1} are the values of the STC at the decoding times of frame *k* and frame *k-1*.

Before applying the de-jittering scheme, every time a PCR arrives at the decoder, the PLL had to adjust the STC to correct for any introduced inter PCR jitter (J_{pcrj} -J_{pcrj-1}). Once de-jittering is performed and the inter PCR jitter is removed through the applied adjustment step Δ, the PLL will only adjust for (E_k) as in equation (9) to reconstruct the STC based on the average network delay. Figure 43 shows the jitter values the PLL has to correct for before and after de-jittering is applied when an MPEG TS at 30 fps, transported over a single hop as per the encoding conditions described in the experimental setup. The background traffic load used for this setup was Poisson at 90% of the OC-3 line rate.

---

[1] The number of frames buffered is dependent on the buffer size used.

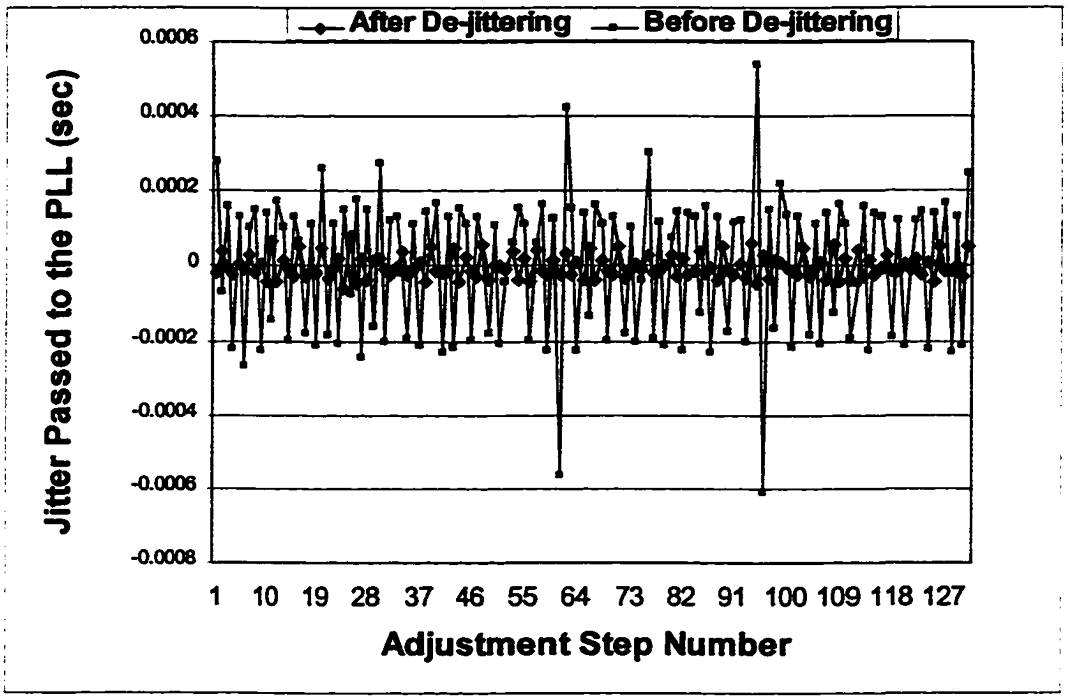


Figure 43: PCR jitter the PLL has to Correct for Before and After De-jittering

$$E_k = (e_{pcr} + \tilde{\mu}_k) - (e_{pcr} + \tilde{\mu}_{k-1}) \tag{9}$$

where e_{pcr} *is* a constant error value in estimating J_{pcrj} due to the fact that the nominal period of PCR insertion (T_{pcr}) is not equal to an integer number of the nominal period T. This error value is dependent on the transmission rate and the T_{pcr} used and can be calculated as follow:

$$e_{pcr} = \left\lceil \frac{T_{pcr}}{T} \right\rceil T - T_{pcr} \tag{10}$$

The value of $\bar{\mu}_k$ should be close to zero and exhibit little variation under the same network conditions as was explained in step 5 of the de-jittering algorithm.

Figure 44 and Figure 45, respectively show the inter frame decoding jitter distribution before and after de-jittering for an MPEG TS at 30 fps, transported over a single hop as per the encoding conditions described in the experimental setup. The background traffic load used for this setup was Poisson at 90% of the OC-3 line rate. Statistical analysis of the inter frame jitter for this case revealed an 84% reduction in the inter frame jitter standard deviation after de-jittering.

Simulated jitter values were also used to test the performance of the de-jittering technique. As an example of simulating a worst case network jitter scenario, a uniform network jitter distributed between ± 10 *msec* was used assuming an MPEG TS rate of 1.536 Mbps with a frame rateat 30 frame per second and a T_{pcr} of 100 *msec* (N is around 102 TS packets, 51 AAL5 PDUs).

Figure 46 and Figure 47 show the inter frame decoding jitter histograms before and after de-jittering for this case.

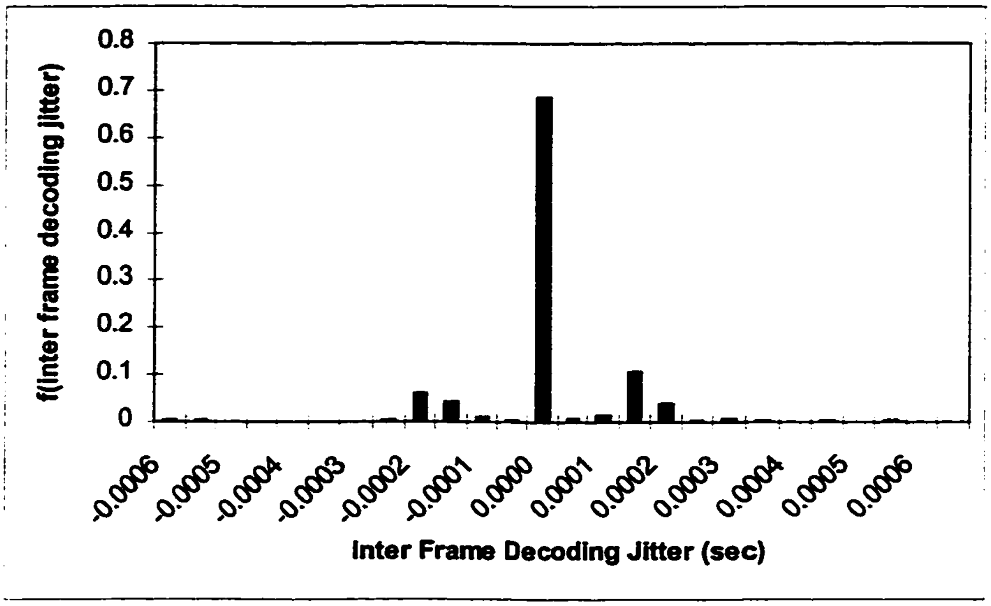


Figure 44: Inter Frame Decoding Jitter Before De-jittering, Experimental Case.

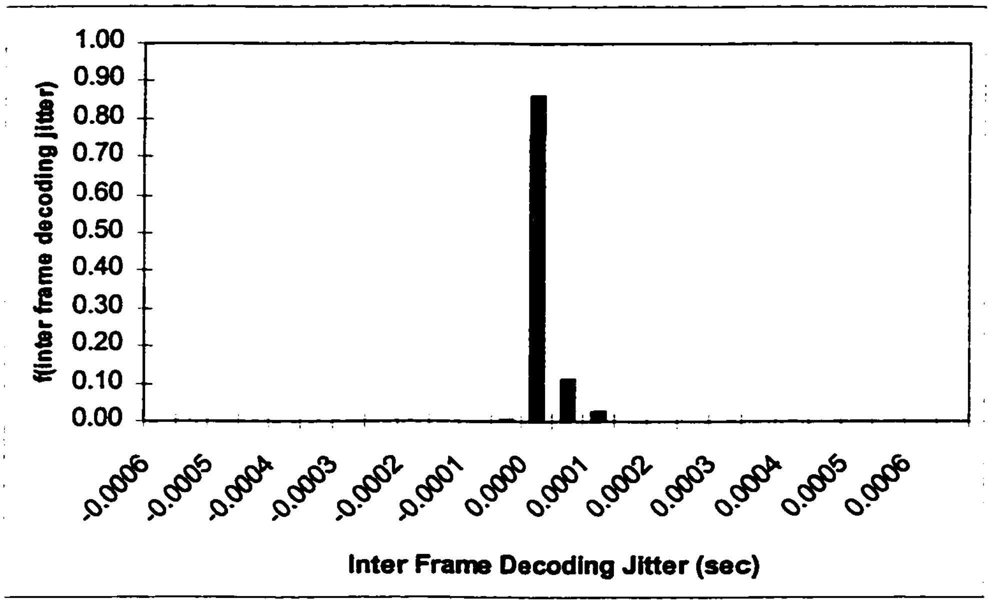


Figure 45: Inter Frame Decoding Jitter After De-jittering, Experimental Case.

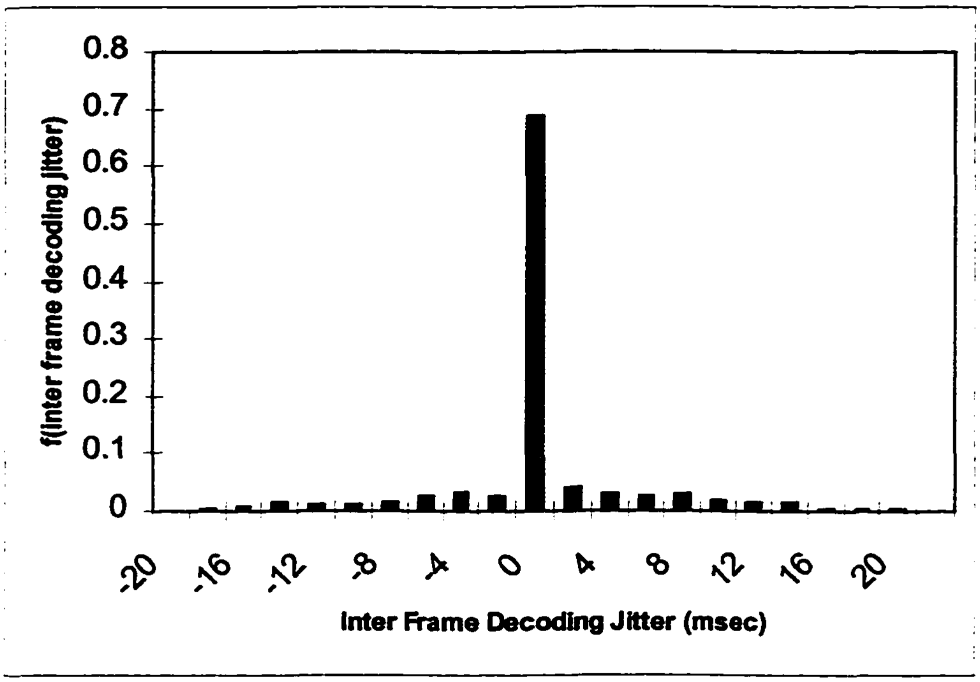


Figure 46: Inter Frame Decoding Jitter Before De-jittering, Simulation Case.

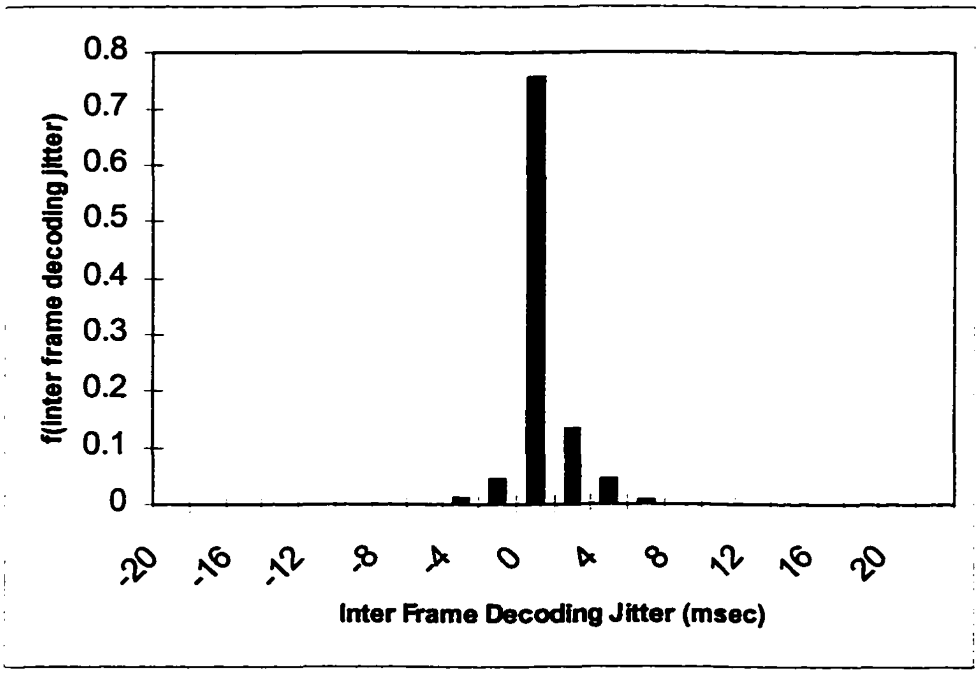


Figure 47: Inter Frame Decoding Jitter After De-jittering, Simulation Case.

7.7 Performance Bounds and Analysis of the De-jittering Scheme.

In this section we examine the performance of the de-jittering scheme by studying the estimation error in the sample mean jitter ($\tilde{\mu}_k$) due to the use of finite number of packets (N). An upper bound on the probability of error in the estimation of $\tilde{\mu}_k$ will be derived. Assuming that the jitter values of received video packets *(J_i)* are independent and identically distributed *(i.i.d)* with standard deviation σ then according to the Central Limit Theorem, the sample mean jitter $\tilde{\mu}_k = \frac{1}{N}\sum_{i=1}^{N} J_i$ has a *p.d.f* close to normal with a standard deviation equal to $\sigma / \sqrt{N}$, providing that N is large enough.

The Tchebycheff Inequality is applied as a loose bound in (11) to find the relationship between the probability of error, *P(e)* in the estimate of μ using $\tilde{\mu}_k$, the number N, and the tolerance interval ε. The tolerance interval is defined as $\varepsilon = \alpha J_{\max}$, where J_{max} is the maximum jitter value allowed for the desired application and $\alpha \in (0,1)$. The tolerance interval as defined then will represent the amount by which the sample mean can differ from the true mean.

$$P\left\{\left|\mu - \tilde{\mu}_k\right| \geq \varepsilon\right\} \leq \frac{\sigma^2}{\varepsilon^2 N} = P(e)$$

$$P\left\{(\tilde{\mu}_k \geq \mu + \varepsilon) or (\tilde{\mu}_k \leq \mu - \varepsilon)\right\} \leq \frac{\sigma^2}{\varepsilon^2 N} = P(e) \text{ or}$$

$$P\left\{(\bar{\mu}_k \geq \mu+\alpha J_{\max})or(\bar{\mu}_k \leq \mu-\alpha J_{\max})\right\} \leq \frac{\sigma^2}{\alpha^2 J^2{}_{\max} N} = P(e) \qquad (11)$$

As was shown in [44,45] the uniform probability density function (*p.d.f*) exhibits the largest entropy of all other *p.d.fs*, that is, it has the largest standard deviation. For a uniform *p.d.f* distributed between $-J_{max}$ and $+J_{max}$, and its standard deviation σ will be $\frac{J_{\max}}{\sqrt{3}}$. Since the Tchebycheff Inequality is only dependent on the standard deviation, substituting $\frac{J_{\max}}{\sqrt{3}}$ for the standard deviation in equation (11) as follow:

$$P\left\{(\bar{\mu}_k \geq \mu+\alpha J_{\max})or(\bar{\mu}_k \leq \mu-\alpha J_{\max})\right\} \leq \frac{J_{\max}{}^2}{3\alpha^2 J^2{}_{\max} N} = P(e) \text{ or}$$

$$P\left\{(\bar{\mu}_k \geq \mu+\alpha J_{\max})or(\bar{\mu}_k \leq \mu-\alpha J_{\max})\right\} \leq \frac{1}{3\alpha^2 N} = P(e) \qquad (12)$$

gives, for the same number N, an upper bound on the probability of error in the estimation of μ using $\bar{\mu}_k$. Therefore, a minimal value of N can be found to satisfy the de-jittering design parameters $(\varepsilon, J_{\max}, andP(e))$. In the next section we show the relationship and trade-off between $(\mathrm{N}, \varepsilon, J_{\max}, andP(e))$ using equation (11) for the experimental jitter distribution and equation (12) for a uniform simulated jitter distribution.

7.7.1 Performance Examples.

The de-jittering scheme proposed in this paper is dependent on the number of video packet received (N) between two consecutive clock reference values, used to reconstruct the clock at the receiver. Numerical examples are introduced to show how the different parameters of the proposed de-jittering scheme are related. The video packet jitter values of the experimental example described in the previous section will be used. In that example J_{max} was 481 *usec* and the standard deviation σ was 120 *usec*. If a tolerance interval $\varepsilon = \alpha J_{\max} = 0.1\ J_{max}$ was desired with a probability of error equal to 0.1, equation (11) gives the number N needed to satisfy these conditions as 63 packets. If equation (12) was used instead to give the most conservative value for N assuming that the video packet jitter was uniformly distributed between ($-J_{max}$ and $+J_{max}$), N would be 333 packets. Figure 48 shows how N varies depending on the desired probability of error, *P(e)*, while using the same tolerance interval (0.1 J_{max}) for the experimental jitter distribution and the above assumed uniform jitter distribution. For the same example, Figure 49 shows the effect of choosing different tolerance interval on N while keeping *P(e)*= 0.1.

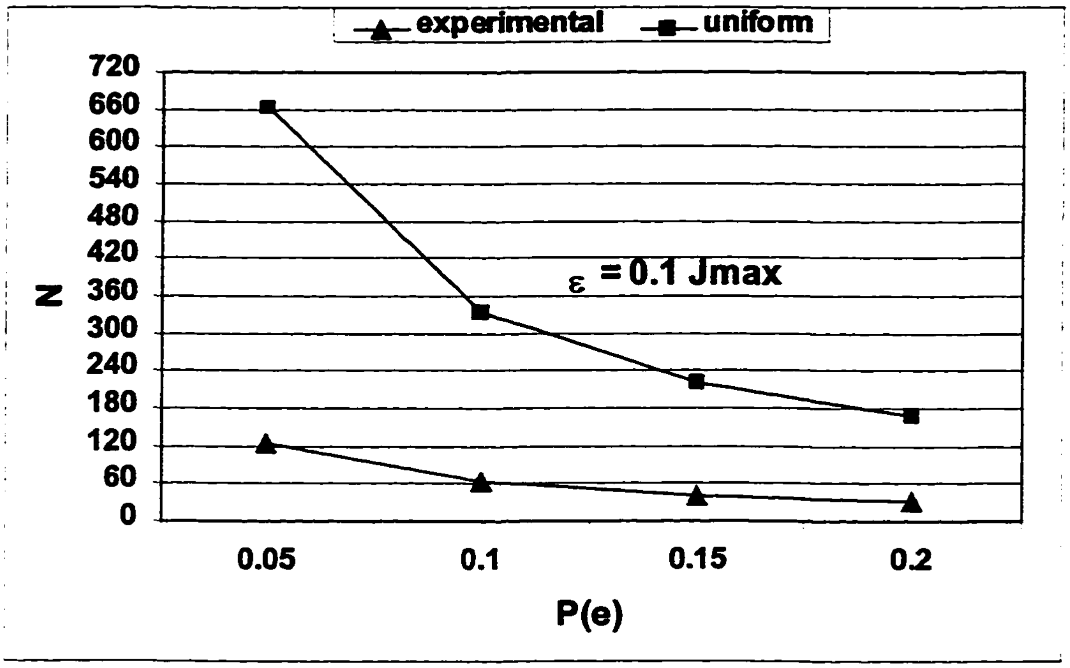


Figure 48: An Example of How the Choice of P(e) Affect the Value of N, Number of Video Packets Used in the Sample Mean Jitter Calculation.

From Figure 49, its shown that if the tolerance interval is 0.1 J_{max} and *P(e)* is 0.1, N can be as low as 63 packets for the experimental jitter distribution.

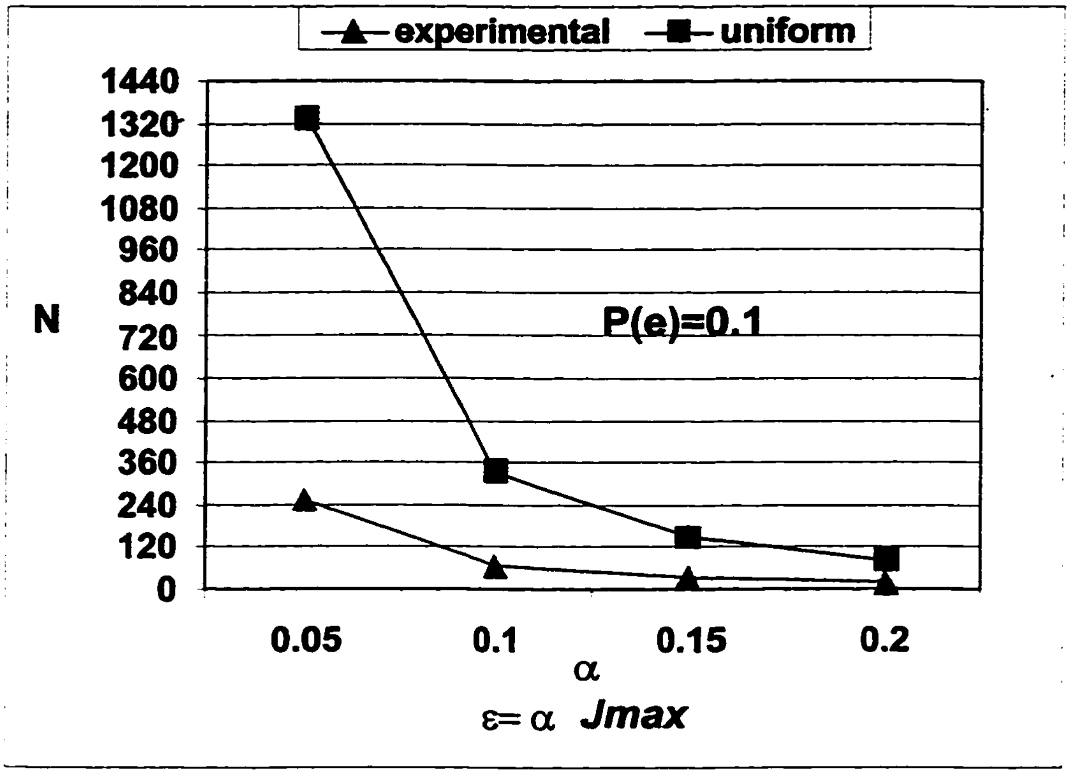


Figure 49: An example of How the Choice of Tolerance Interval Affect the Value of N, Number of Video Packets Used in the Sample Mean Jitter Calculation.

8 Conclusion and Future Work

8.1 Summary and Conclusion.

The work in the thesis can be summarized as follow:

- We have conducted simulations and experiments on several MPEG-2 transport stream clips with different bit rates and encoding sequences to investigate the sensitivity of MPEG-2 transport stream to bit errors and cell losses. To improve the received video quality when transported over ATM networks we proposed a cell loss padding technique. The proposed scheme could be incorporated as part of a Service Specific Convergence Sublayer (SSCS) of the AAL used or as part of the MPEG-2 decoder when decoding a TS. The results presented could be used to selectively apply FEC to minimize the overhead involved while minimizing degradation in the received video quality. Results obtained showed that passing MPEG video transport stream packets to the decoder layer rather than discarding them improves the perceptually received video quality and therefore, the enabling of CRC of the AAL5 should be optional based on the user application. In the case of data loss, simple padding techniques for the lost data were shown to result in a video quality better than discarding uncompleted AAL5 PDUs or passing them to the upper decoding layer.

- We have conducted an intensive and comprehensive empirical study on the transport of ATM CBR traffic under diverse operating conditions. Our results revealed the effect of changing several network parameters such as the applied traffic load and number of switching nodes on the performance and characteristics of the transported CBR traffic. Under the most extreme network conditions presented in this study the end-to-end delay of a cell was below (350μs) and the maximum peak-to-peak CDV was under (240μs). From the overall observations and analysis of CDV and end-to-end delay of CBR traffic, the failure of standard distributions (Normal, Rayleigh etc) in adequately and accurately modeling their behavior was shown. An example of how distinct and different the distributions of CDV and CTD could be under different operating conditions is shown in Figure 50. Quantifying the end-to-end delay and CDV in a WAN ATM network is essential to building the next generation of multimedia services.

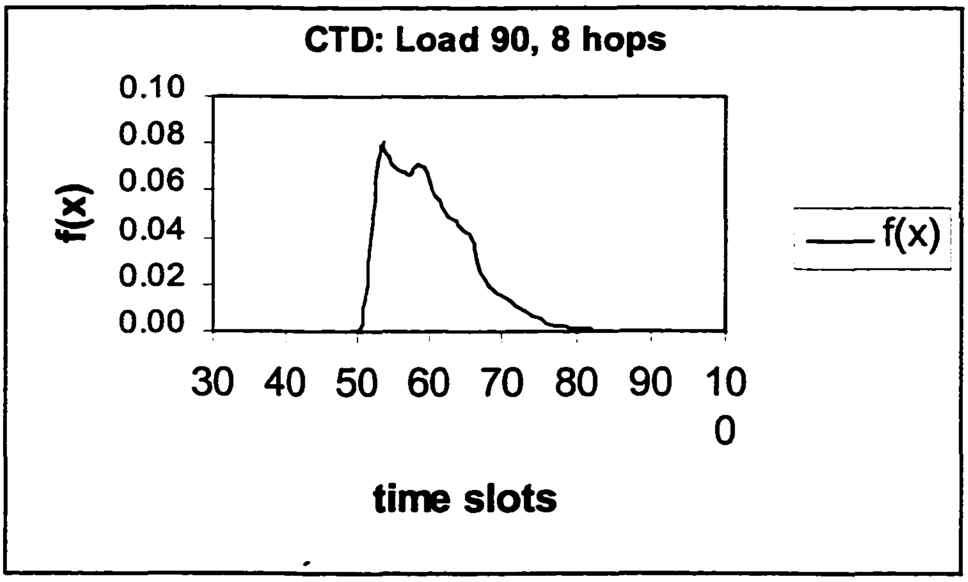


Figure 50: CTD Distribution for Load 90% and 8 Hops

- In this work an overview of the MPEG-4 system and timing model was provided. Clock reference values, time stamps and a PLL are used to reconstruct the system time clock at the decoder and achieve clock synchronization. For this to work, the end-to-end delay between the decoder and encoder must be constant. Variable channel delay introduces error and distortion in the STC recovery and therefore, a robust de-jittering scheme is desired to make an MPEG system decoder exhibit good performance even without the best designed PLL. This work present a novel de-jittering scheme that can applied to preserve the temporal relationship between video objects and therefore preserve the desired quality of service.

8.2 Intended Future Work.

Our intended future work fall under the umbrella of variable bit rate (VBR) multimedia traffic transported over ATM and IP networks and will consist of the following:

- Study the characteristics of delay and delay jitter of VBR MPEG when transported over an ATM network using the VBR service category for transmission. In Transporting VBR MPEG over an ATM network, several issues rise such as, traffic shaping, video smoothing, video scheduling, and preserving the user desired quality of service. The delay and jitter analysis is affected by such issues and become more complicated. Our future work will be centered on building simulation tools to enable us perform our intended VBR video delay and jitter analysis under diverse network operational conditions. We will use pre-encoded MPEG VBR video streams and develop a scheduling mechanism to deliver such streams over an ATM based network. Our objective will be to design our scheduling technique in a way to ensure the preservation of video quality while minimizing the use of network resources. Using such scheduling technique, the inter-departure times of ATM cells carrying VBR MPEG encoded video can be registered. The second step is to model a multi-hop ATM network with several queuing mechanisms and study the distribution of delay and jitter when the pre-encoded VBR MPEG streams are scheduled to pass through the modeled ATM network.

- Our second future work objective is to extend our de-jittering scheme to the VBR case. Also, to test the performance of the proposed de-jittering scheme using MPEG-4 video data rather than MPEG-2. Another aspect of our intended future work will be to use the de-jittering scheme in an IP environment. MPEG video can be transported over IP using the real time protocol (RTP) for encapsulation. Once MPEG video is encapsulated using RTP packets, it is transported using UDP/IP. Since the characteristics of an IP network is different from those of an ATM network, the de-jittering scheme must be adjusted and modified to be used in such an environment.

9 References.

[1] MPEG-1 International Standard ISO/IEC 11172.

[2] ITU-T, I.363, “AAL type 5, Draft Recommendation for section 6 of I.363”, January 1993.

[3] S. Dixit and P. Skelly, "MPEG-2 over ATM for video dial tone networks: issues and strategies," IEEE network magazine, September 1995, pp. 30-40, vol. 9.

[4] ftp://ftp.tektronix.com/pub/streams/video/525.

[5] B. Furht, S. Smoliar and H. Zhang, "Video and image processing in multimedia systems," Kluwer Academic Publishers, 1995.

[6] ISO/IEC 13818-2/1, "Information technology-coding of moving pictures and associated audio", video, system 1996.

[7] ITU-T Draft H.310, line transmission of non-telephone signals broadband audiovisual communication systems and terminals, June 1996.

[8] ATM Theory and Applications, D. McDysan and D. Spohn, McGraw-Hill Series, 1994.

[9] H. Naser, and Leon Garcia, “ A simulation study of delay and delay variation in ATM networks, Part 1: CBR traffic, “ Proc. of IEEE INFOCOM 96, March 1996, pp. 393-400.

[10] EL-Henaoui S., Coelho R. and Tohme S. “ Inter-flows jitter analysis for multimedia traffic in ATM networks”, Proc. of IEEE Symposium on Computers and Communications, 1995, pp. 377-383.

[11] M. Borden, " Properties of CDV and its accumulation", ATM Forum Technical Committee Contribution, August 1995.

[12] S. Banerjee, D. Tipper, B. Weiss and A. Khalil," "Traffic Experiments on the vBNS wide area ATM networks", IEEE communication magazine, August 1997, pp. 126-133.

[13] R. Mandeville," Product evaluation, ATM switches", Data communication, March 1995, pp. 22-30.

[14] K. Shuaib, T. Saadawi M. Lee S. Gringeri, and B. Basch "A Cell Loss Padding Technique for the Transport of MPEG-2 Transport Stream over ATM/AAL" Proc. of the IASTED conference in the Bahamas Dec 7-10 1997, pp. 450-454.

[15] K. Shuaib, H. Elsayed, T. Saadawi M. Lee, S. Gringeri, B. Khasnabish and B. Basch, "Experimental and simulation results in the transport of MPEG-2 video over ATM" Proc. of the SPIE conference on voice, video, and data, Dallas TX, Nov. 2-5 1997, pp. 168-176

[16] K. Shuaib, S. Nawrot, H. Esayed, M.Lee, and T. Saadawi,"Empirical evaluation of ATM CBR traffic transported under diverse operating conditions", Proc. of the ISCC conference, Athens, Greece June 1998, pp. 450-454.

[17] S. Gringeri, K. Shuaib, R. Egorov, A. Lewis, B. Khasnabish and B. Basch, "Traffic Shaping, Bandwidth Allocation and Quality Assessment for the Transmission of VBR MPEG Video over Broadband Networks:" IEEE Network Magazine, Nov./Dec 1998 issue, pp. 94-107.

[18] K. Shuaib, T. Saadawi, M.Lee, and B. Basch, "MPEG-4 System: Timing Model and a De-Jittering Scheme" Proc. of the third ATIRP conference, Maryland, Feb. 3, 1999, pp.323-327.

[19] S. Nawrot, H. Esayed, K. Shuaib, M.Lee, and T. Saadawi "Experimental Evaluation of a Hybrid ATM-Satellite Network: CBR Traffic, CTD and CDV QoS parameters" Military Communications Conference, 1998. MILCOM 98. Proceedings., IEEE Volume: 2 , pp. 554 -559

[20] S. Gringeri, K. Shuaib, R. Egorov, A. Lewis, and B. Basch "Customizing the MPEG Standards for Military Applications" Proc. of the second ATIRP conference, Maryland, Feb. 3, 1998, pp. 330-334

[21] S. Gringeri, A. Lewis, K. Shuaib, R. Egorov, B. Khasnabish and B. Basch, "Transmission of MPEG-2 Video Streams over ATM Networks" spring issue of IEEE Multimedia Magazine 1998, pp. 58-71.

[22] C. Ohta, K. Shinagawa and Y. Onozato, " Cell loss properties for multiplexing of MPEG video sources considering picture coding types in ATM networks" IEEE JSAC, April 1996, pp. 1396 -1400 vol.3.

[23] T. Saadawi, M. Ammar, and A. EL Hakeem, "Fundamental of Telecommunication Networks", Wiley Series in Telecommunications and Signal Processing 1994.

[24] P. Pancha and M. El Zarki ," MPEG coding for variable bit rate video transmission." IEEE Communication Magazine, May 1994, pp. 54 -66.

[25] Krunz, M.; Sass, R.; Hughes, H, "Statistical characteristics and multiplexing of MPEG streams", Proc. of INFOCOM '95. Fourteenth Annual Joint Conference of the IEEE Computer and Communications Societies. 'Bringing Information to People', pp. 455-462 vol.2.

[26] S. Wee, M. Polley, and W. Schreiber, " A generalized framework for scalable video coding", Multimedia Communication and Video Coding, edited by Y. Wang, Plenum Press, New York, 1996, pp. 483- 491.

[27] K. Chandra, and A. Reibman, " Modeling two-layer MPEG-2 video traffic, Multimedia Communication and Video Coding, edited by Y. Wang, Plenum Press, New York, 1996, pp. 419- 426.

[28] Y.Xie, C. Liu, M.Lee, and T.Saadawi, Multipoint multimedia teleconference system with adaptive synchronization, IEEE Journal on Selected Areas in Communications, September 1996, pp.1422-1435.

[29] P. Zarros, M. Lee, T. Saadawi, "Interparticipant Synchronization in Real-Time Multimedia Conference Using Feedback," IEEE/ACM Transactions on Networking, 4(2):173-180, April 1996, pp. 173 -180.

[30] MPEG-4 Systems ISO/IEC JTC1/SC29/WG11 N2201, 1998

[31] MPEG-4 Video ISO/IEC JTC1/SC29/WG11 N 2202, 1998

[32] MPEG-4 DMIF ISO/IEC JTC1/SC29/WG11 N 2206, 1998

[33] A. Puri and A. Eleftheriadis, "MPEG-4: An object-based multimedia coding standard supporting mobile applications", ACM, Journal of Mobile Network and Applications, August 1998, pp. 5-31.

[34] I. Akyildiz, S. Hrastar, H. Uzunalioglu and W. Yen, "Comparison and Evaluation of Packing Schemes for MPEG-2 over ATM using AAL5", Proc. of IEEE, ICC June,1996 Dallas, Texas, pp. 1411-1415.

[35] L. Chariglione, "MPEG and Multimedia Communications," IEEE Trans. On Circuits and Systems for Video Technology, Vol. 7, No. 1, February 1997, pp. 5-18.

[36] ATM Forum, "ATM Traffic Management Specifications 4.0," 1996.

[37] J. M. Pitts and J. A. Schormans, "Introduction to ATM: Design and Performance," John Wiley and Sons Ltd., England, 1996.

[38] ATM Forum, "Audio Visual Multimedia Services: Video on Demand Implementation Agreement Phase 1.0", December 1995.

[39] ATM Forum Technical Committee UNI 3.1 Specification, September 1994.

[40] R. Steinmetz, " Synchronization properties in multimedia systems", IEEE Journal on Selected Areas of Communications, April 1990, pp. 401 -412

[41] M. Perez-Luque, and T. Little, "A temporal reference framework for multimedia syn- chronization", IEEE Journal on Selected Areas on Communications, Vol. 14, No. 1, Jan. 1996, pp. 36 -51.

[42] P. Rangan, S. Kumar, and S. Rajan, "Continuity and synchronization in MPEG" , IEEE J. on Selected Areas on Communications, Vol. 14, No, 1, 1996, pp. 52 -60.

[43] K. Shuaib, S. Gringeri, A. Lewis, R. Egorov and B. Basch "Delivery of VBR MPEG compliant Video Streams" Proc. of the third ATIRP conference, Maryland, Feb. 3, 1999, pp. 335-339.

[44] Principles of Communication Engineering, Wozencraft and Jacobs, John Wiley and Sons, Inc, 1965.

[45] Probabilty Random Variables, and Stochastic Processes, Athanasios Papoulis,McGraw-Hill, 1984.

[46] K. Shuaib, T. Saadawi, and M.Lee "MPEG-4 System: Timing Model and a De-Jittering Scheme" submitted to the Journal of ACM Multimedia System.

[47] K. Shuaib, T. Saadawi, M.Lee and B. Basch "De-jittering in the Transport of MPEG-2 and MPEG-4 over ATM" accepted for presentation at the ICATM, Colmar, France, June, 1999.

Biography

Khaled A. Shuaib was born in 1967 in the town of Rammallah, Palestine. He received his B.S. and M.S. degrees in electrical engineering in 1991 and 1993, respectively from City College of New York (MAGNA CUM LAUDE). He was on the Dean's List from 1987 to 1991. Khaled has been an IEEE member since 1993, and he is also a member of the communication society. From September 1997 to May 1998, Khaled earned a City University of New York Teaching Assistant Fellowship to teach at the Electrical Engineering department at the City College of New York.

Khaled worked at GTE Laboratories during the summer of 1996 as a Member of Technical Staff (MTS). In July of 1997, he started as a full time MTS at GTE Laboratories. In June 1998 he was promoted to his current position as a Senior MTS. Khaled holds the GTE Individual Excellence Awards, April 1998, and December 1998.

Recent publication include:

"MPEG-4 System: Timing Model and a De-Jittering Scheme" submitted to the Journal of ACM Multimedia System.

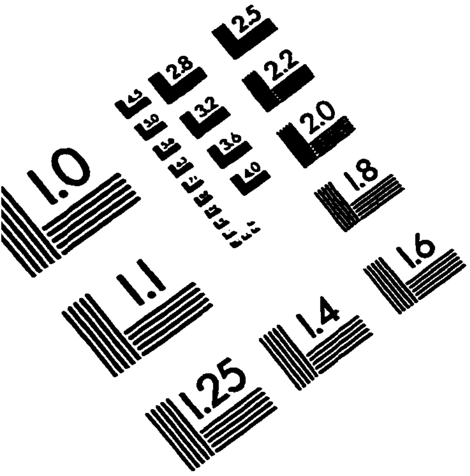
"Traffic Shaping, Bandwidth Allocation, and Quality Assessment for MPEG Video Distribution over Broadband Networks" IEEE network magazine, Nov/Dec. 1998.

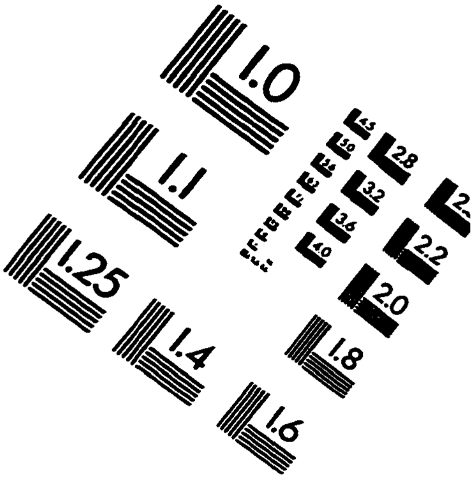
“Empirical Evaluation of CBR ATM Traffic Transported under Diverse Operating Conditions” presented at the ISCC conference, Athens, Greece, June, 1998.

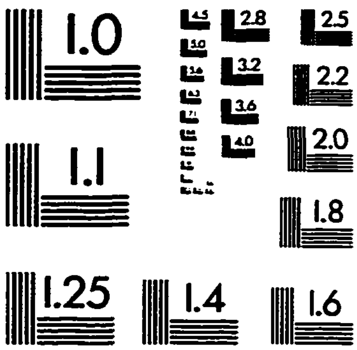
U.S. Patent Applications:

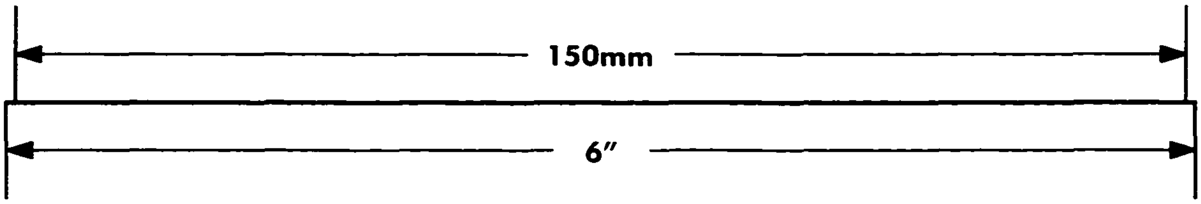
"System and Method for Analyzing and Transmitting Video over A Switched Network", filed to the United States Patent Application Office, Jan. 1999.

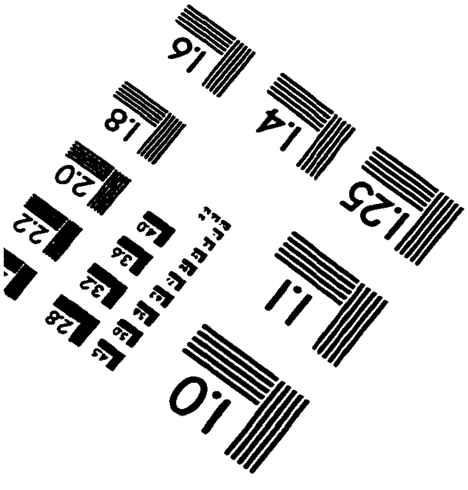
IMAGE EVALUATION TEST TARGET (QA-3)











APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989