PM-1 3 1/2" x 4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

PRECISIONSM RESOLUTION TARGETS

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395 rue Wellington
Ottawa (Ontario)
K1A 0N4

Number: 1-800-953-5696

Chapitre: Notice 10/1/88

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Statistical Modeling of VBR-MPEG Video and its Transmission in ATM Networks

by

M. Riyaz Ismail, B. Eng.

(Gulbarga University, Gulbarga, India)

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for degree of

Master of Engineering

Ottawa-Carleton Institute for Electrical Engineering
Department of Systems and Computer Engineering
Carleton University
1125 Colonel By Drive
Ottawa, Ontario, K1S 5B6
Canada.

July 17, 1995

© 1995 M. Riyaz Ismail

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Author's name

Author's name

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-08958-4

Canada

Name _____

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

1990

U·M·I

SUBJECT TERM

SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS

Architecture 0729
 Art History 0377
 Cinema 0900
 Dance 0378
 Fine Arts 0357
 Information Science 0723
 Journalism 0391
 Library Science 0399
 Mass Communications 0708
 Music 0413
 Speech Communication 0459
 Theater 0465

EDUCATION

General 0515
 Administration 0514
 Adult and Continuing 0516
 Agricultural 0517
 Art 0273
 Bilingual and Multicultural 0282
 Business 0688
 Community College 0275
 Curriculum and Instruction 0727
 Early Childhood 0518
 Elementary 0524
 Finance 0277
 Guidance and Counseling 0519
 Health 0680
 Higher 0745
 History of 0520
 Home Economics 0278
 Industrial 0521
 Language and Literature 0279
 Mathematics 0280
 Music 0522
 Philosophy of 0998
 Physical 0523

Psychology 0525
 Reading 0535
 Religious 0527
 Sciences 0714
 Secondary 0533
 Social Sciences 0534
 Sociology of 0340
 Special 0529
 Teacher Training 0530
 Technology 0710
 Tests and Measurements 0288
 Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS

Language 0679
 General 0289
 Ancient 0290
 Linguistics 0291
 Modern 0401
 Literature 0294
 General 0295
 Classical 0297
 Comparative 0298
 Medieval 0316
 Modern 0591
 African 0305
 American 0352
 Asian 0355
 Canadian (English) 0593
 Canadian (French) 0311
 English 0312
 Germanic 0315
 Latin American 0313
 Middle Eastern 0314
 Romance 0370
 Slavic and East European 0372

PHILOSOPHY, RELIGION AND THEOLOGY

Philosophy 0422
 Religion 0318
 General 0321
 Biblical Studies 0319
 Clergy 0320
 History of 0322
 Philosophy of 0469
 Theology 0323

SOCIAL SCIENCES

American Studies 0323
 Anthropology 0324
 Archaeology 0326
 Cultural 0327
 Physical 0310
 Business Administration 0272
 General 0770
 Accounting 0454
 Banking 0338
 Management 0385
 Marketing 0501
 Canadian Studies 0503
 Economics 0505
 General 0508
 Agricultural 0509
 Commerce-Business 0510
 Finance 0511
 History 0358
 Labor 0366
 Theory 0351
 Folklore 0578
 Geography 0460
 Gerontology 0386
 History 0460
 General 0386

Ancient 0579
 Medieval 0581
 Modern 0582
 Black 0326
 African 0331
 Asia, Australia and Oceania 0332
 Canadian 0334
 European 0335
 Latin American 0336
 Middle Eastern 0333
 United States 0337
 History of Science 0585
 Law 0398
 Political Science 0615
 General 0616
 International Law and Relations 0617
 Public Administration 0814
 Recreation 0452
 Social Work 0626
 Sociology 0627
 General 0938
 Criminology and Penology 0631
 Demography 0628
 Ethnic and Racial Studies 0629
 Individual and Family Studies 0630
 Industrial and Labor Relations 0700
 Public and Social Welfare 0344
 Social Structure and Development 0709
 Theory and Methods 0999
 Transportation 0453
 Urban and Regional Planning 0453
 Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES

Agriculture 0473
 General 0285
 Agronomy 0475
 Animal Culture and Nutrition 0476
 Animal Pathology 0359
 Food Science and Technology 0478
 Forestry and Wildlife 0479
 Plant Culture 0480
 Plant Pathology 0817
 Plant Physiology 0777
 Range Management 0746
 Wood Technology 0306
 Botany 0287
 General 0308
 Anatomy 0309
 Biostatistics 0379
 Botany 0353
 Cell 0369
 Ecology 0793
 Entomology 0410
 Genetics 0307
 Limnology 0416
 Microbiology 0433
 Molecular 0821
 Neuroscience 0778
 Oceanography 0472
 Physiology 0786
 Radiation 0760
 Veterinary Science 0786
 Zoology 0760
 Biophysics 0786
 General 0760
 Medical 0760

EARTH SCIENCES

Biogeochemistry 0425
 Geochemistry 0996

Geodesy 0370
 Geology 0372
 Geophysics 0373
 Hydrology 0388
 Mineralogy 0411
 Paleobotany 0345
 Paleocology 0426
 Paleontology 0418
 Paleozoology 0985
 Polynology 0427
 Physical Geography 0368
 Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES

Environmental Sciences 0768
 Health Sciences 0566
 General 0300
 Audiology 0992
 Chemotherapy 0567
 Dentistry 0350
 Education 0769
 Hospital Management 0758
 Human Development 0982
 Immunology 0564
 Medicine and Surgery 0347
 Mental Health 0569
 Nursing 0570
 Nutrition 0380
 Obstetrics and Gynecology 0354
 Occupational Health and Therapy 0381
 Ophthalmology 0571
 Pathology 0419
 Pharmacology 0572
 Pharmacy 0382
 Physical Therapy 0573
 Public Health 0574
 Radiology 0575
 Recreation 0575

Speech Pathology 0460
 Toxicology 0383
 Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences 0485
 Chemistry 0749
 General 0486
 Agricultural 0487
 Analytical 0488
 Biochemistry 0738
 Inorganic 0490
 Nuclear 0491
 Organic 0494
 Pharmaceutical 0495
 Physical 0754
 Polymer 0405
 Radiation 0605
 Mathematics 0986
 Physics 0606
 General 0608
 Acoustics 0748
 Astronomy and Astrophysics 0607
 Atmospheric Science 0798
 Atomic 0759
 Electronics and Electricity 0609
 Elementary Particles and High Energy 0752
 Fluid and Plasma 0756
 Molecular 0611
 Nuclear 0463
 Optics 0346
 Radiation 0984
 Solid State 0346
 Statistics 0984
 Applied Sciences 0346
 Applied Mechanics 0984
 Computer Science 0984

Engineering 0537
 General 0538
 Aerospace 0539
 Agricultural 0540
 Automotive 0541
 Biomedical 0542
 Chemical 0543
 Civil 0544
 Electronics and Electrical 0348
 Heat and Thermodynamics 0545
 Hydraulic 0546
 Industrial 0547
 Marine 0794
 Materials Science 0548
 Mechanical 0743
 Metallurgy 0551
 Mining 0552
 Nuclear 0549
 Packaging 0765
 Petroleum 0554
 Sanitary and Municipal System Science 0790
 Geotechnology 0428
 Operations Research 0796
 Plastics Technology 0795
 Textile Technology 0994

PSYCHOLOGY

General 0621
 Behavioral 0384
 Clinical 0622
 Developmental 0620
 Experimental 0623
 Industrial 0624
 Personality 0625
 Physiological 0989
 Psychobiology 0349
 Psychometrics 0632
 Social 0451

Carleton University
Ottawa, Canada K1S 5J7

1992-1992

Thesis contains black & white
illustrations which when microfilmed
may lose their significance. The
hardcopy of the thesis is available
upon request from Carleton
University Library.

The undersigned recommend to the Faculty of Graduate Studies and Research
acceptance of the thesis

**Statistical Modeling of VBR-MPEG Video and
its Transmission in ATM Networks**

submitted by M. Riyaz Ismail, B. Eng.

in the partial fulfillment of the requirements for
the degree of Master of Engineering

Thesis Supervisors

Chair, Department of Systems and Computer Engineering

Carleton University

July 17, 1995

Abstract

This thesis presents an efficient transmission mechanism using a traffic shaping scheme called *frame spreading* in ATM networks. The source traffic of concern in this study is a variable bitrate MPEG video source. A priority scheme based on MPEG 2 specification, is implemented in a software encoder that produces traffic in two priority partitions for all three frame types (Intraframe, Predicted and Interpolated) used in MPEG. An ATM multiplexer with a pushout buffer scheme is implemented in order to provide priority scheduling for the two priority partitions. Performance statistics such as the cell loss ratios are studied for various frame spreading scenarios.

Two statistical models are developed for the variable bitrate MPEG video sequence having two levels of priority, using TES (Transform Expand Sample). The first model is matched with the empirical histogram and autocorrelation function of each frame type of the original sequence and the second model is created with an assumption of a Gamma distribution for the sequence of number of bits per frame for each frame type. The above experiment with an ATM multiplexer is repeated using both models and their results are compared.

to my parents ...

•

Acknowledgments

I would like to express my sincere thanks to my thesis supervisors, Prof. A. R. Kaye and Prof. I. E. Lambadaris for their invaluable guidance and helpful suggestions during the course of this thesis. I also wish to thank Dr. M. Devetsikiotis for all the interesting discussions I had with him while working on this thesis.

I would like to thank all the students of the Broadband Networks Laboratory for putting up the nice working atmosphere and the staff of the Department of Systems and Computer Engineering for their help in the development of this thesis.

I gratefully acknowledge the funding support provided by the Telecommunication Research Institute of Ontario and the National Science and Engineering Research Council of Canada.

I would also like to thank my wife Rubina who, being a graduate student herself, provided me with a continuous source of inspiration. Finally, my sincere thanks to my parents for their prayers and continuous encouragement during this research.

Table of Contents

Acceptance Sheet	i
Abstract	ii
Acknowledgments	iv
Table of Contents	v
List of Figures	viii
List of Tables	x
Acronyms	xii
Chapter 1. Introduction	1
1.1 Background	1
1.2 Motivation and Scope	3
1.3 Thesis Outline	5
Chapter 2. Background	6
2.1 Introduction	6
2.2 Digital Pictures	7
2.2.1 Image Formats	7
2.2.2 Linear Color Transformation	8
2.2.3 Subsampling	10
2.3 MPEG	11
2.3.1 Structure of MPEG Video Stream	13
2.3.1.1 Sequence	14
2.3.1.2 Pictures and Group-of-Pictures	14

2.3.1.3	Slices	16
2.3.1.4	Macroblocks and Blocks.....	16
2.3.2	MPEG Video Compression.....	17
2.3.2.1	Frame Reordering	17
2.3.2.2	Motion Compensation.....	18
2.3.2.3	Discrete Cosine Transform	19
2.3.2.4	Quantization and Zig-Zag Scan	21
2.3.2.5	Entropy Coding.....	22
2.3.3	Changes in MPEG-2	23
2.4	TES Processes	25
2.5	ATM	30
2.5.1	The ATM Layer.....	31
2.5.2	ATM Adaptation Layer (AAL)	33
2.5.2.1	Convergence Sublayer	34
2.5.2.2	Segmentation and Reassembly (SAR).....	34
Chapter 3.	Video Modeling	36
3.1	Introduction.....	36
3.2	VBR MPEG Traffic	36
3.3	Two-Layered MPEG.....	39
3.3.1	Priority Requirement.....	39
3.3.2	Priority Implementation	40
3.3.2.1	Data Partitioning	42
3.3.2.2	Decoding process of two-layered MPEG	44
3.3.2.3	Priority Considerations	46
3.3.2.4	Priority Selection	48
3.4	Statistical Video Modeling.....	49

3.4.1	Empirical TES Modeling	52
3.4.1.1	Composite MPEG Model.....	53
3.4.1.2	Implementing scene changes	58
3.4.1.3	Generation of Slices.....	59
3.4.1.4	Implementing Priority in the model.....	62
3.4.2	Gamma-TES Modeling.....	63
Chapter 4.	Performance of an ATM Multiplexer with Video Frame Spreading.....	71
4.1	Introduction.....	71
4.2	ATM Adaptation Layer.....	72
4.2.1	AAL type 2	73
4.2.2	AAL type 5	76
4.3	Frame Spreading	78
4.4	Priority handling	80
4.4.1	Partial buffer sharing.....	81
4.4.2	Push-out scheme.	82
4.5	ATM multiplexer.....	82
4.6	Simulation results	84
4.6.1	Utilization and Total cell loss	84
4.6.2	Applying frame spreading.....	87
4.6.3	Results using MPEG source models.....	92
Chapter 5.	Conclusions and Future Research.....	96
5.1	Conclusions.....	96
5.2	Future Research	99
List of References		101

List of Figures

Figure 2-1	CCIR-601 and SIF formats	11
Figure 2-2	Structure of the MPEG compressed video bitstream.....	13
Figure 2-3	Temporal redundancy reduction using I, P and B frames.....	15
Figure 2-4	Formation of Blocks from a Macroblock.....	17
Figure 2-5	MPEG Video encoder block diagram.	18
Figure 2-6	Applying DCT to 8x8 block to get an 8x8 block of coefficients	19
Figure 2-7	Zig-Zag scanning of the DCT coefficients.....	21
Figure 2-8	An example of TES modeling using TESTool.....	29
Figure 2-9	ATM Protocol Reference Model (from CCITT Rec. I.321)	30
Figure 2-10	An ATM cell structure	31
Figure 2-11	Header format for an ATM cell.....	32
Figure 3-1	A typical bitrate of an MPEG sequence with I, P and B frames.....	37
Figure 3-2	Effect of cell loss in an MPEG frame	38
Figure 3-3	Priority Implementation in a block	41
Figure 3-4	Formation of prioritized slices from prioritized blocks	41
Figure 3-5	Example of Data Partitioning.....	43
Figure 3-6	A segment of bitstream showing the decoding process from two partitions	45
Figure 3-7	Percentage HP per Slice in Flower-garden Sequence (PBP = 12).....	47
Figure 3-8	Percentage HP per slice in Flower-garden sequence at different PBP for the B frames	47
Figure 3-9	Video sequence encoded at PBP = 4, 8, 12, 16, 20, 24, 32, 64.....	50
Figure 3-10	Autocorrelation function of an MPEG sequence (I frame interval = 11 lags).	51
Figure 3-11	Sample paths of I, P and B frame sequences and their models.....	54
Figure 3-12	Matching of the histograms for I, P and B frames with TES Models.....	55
Figure 3-13	Matching of the autocorrelation function for I, P and B frames.....	56
Figure 3-14	Flow Diagram of the Composite MPEG-TES Model.....	57

Figure 3-15	Model for the "Last Action Hero" sequence.....	60
Figure 3-16	Model for the Son.mpg Sequence.....	61
Figure 3-17	Histogram of percentage of bytes per slice in a frame.....	62
Figure 3-18	Histogram of the percentage of bytes in HP partition per slice in B frames.....	63
Figure 3-19	Histogram and Q-Q plot showing the goodness of fit of B frames with gamma distribution (Son.mpg)	64
Figure 3-20	Q-Q plots Showing the goodness of fit with Gamma distribution (Last Action Hero).....	65
Figure 3-21	Inducing Autocorrelation into the iid samples from a Gamma distribution.	68
Figure 3-22	Gamma-TES model and the original sequence (Son.mpg).....	69
Figure 3-23	Normal Probability Plot for the percentage HP per slice in B frames.....	70
Figure 4-1	Video source feeder.....	72
Figure 4-2	From MPEG slice to ATM payload	74
Figure 4-3	ATM cell payload.....	75
Figure 4-4	AAL type 5 CS_PDU.....	77
Figure 4-5	Formation of ATM cells using AAL5.....	78
Figure 4-6	Frame Spreading	79
Figure 4-7	Partial buffer sharing.....	81
Figure 4-8	Push-out scheme	82
Figure 4-9	Statistical multiplexing model	83
Figure 4-10	Multiplexer utilization at various PBPs (using AAL 2).....	85
Figure 4-11	Total cell loss ratio at various PBPs (using AAL 2)	85
Figure 4-12	Total cell loss ratio at various PBPs (using AAL 5)	86
Figure 4-13	Total cell loss ratio at various PBPs (using AAL 5)	87
Figure 4-14	Total cell loss ratio at various frame spreading factors (AAL 2).....	88
Figure 4-15	HP cell loss ratio for various frame spreading factors (AAL 2).....	89
Figure 4-16	LP cell loss ratio for various frame spreading factors (AAL 2)	89
Figure 4-17	Total cell loss ratio at various frame spreading factors (AAL 5).....	90

Figure 4-18 **HP cell loss ratio for various frame spreading factors (AAL 5).....91**

Figure 4-19 **LP cell loss ratio for various frame spreading factors (AAL 5)91**

Figure 4-20 **Number of Multiplexed Sources for a given Frame Spreading92**

Figure 4-21 **Comparison of total cell loss ratios.....93**

Figure 4-22 **Comparison of the HP cell loss ratio94**

Figure 4-23 **Comparison of the LP cell loss ratio.....94**

List of Tables

Table 2-1	Classification of traffic in ATM	33
Table 3-1	MPEG bitrate characteristics.	37
Table 3-2	Percentage overhead associated with priority	46
Table 3-3	Percentage HP traffic in the codec output.	49
Table 3-4	Mean and Variance of Son.mpg sequence	66
Table 4-1	Begin / End tags in the SAR	75
Table 4-2	The Peak-to-average ratio for various frame spreading	80

Acronyms

AAL	ATM Adaptation Layer
ATM	Asynchronous Transfer Mode
B-ISDN	Broadband-ISDN
B frame	Bi-directionally Interpolated frames
B/E tags	Begin/End Tags
Cb	Chrominance component
CBR	Constant bitrate
CLP	Cell Loss Priority
CIF	Common Interchange Format
Cr	Chrominance component
CRC	Cyclic Redundancy Check
CRT	Cathode Ray Tube
CS	Convergence Sub-layer
DCT	Discrete Cosine Transform
EOB	End Of Block
GFC	Generic Flow Control
GOP	Group Of Pictures
GUI	Graphical User Interface
HEC	Header Error Check
HP	High Priority
I frame	Intra frame
ISDN	Integrated Services Digital Networks
JPEG	Joint Pictures Experts Group

LI	Length Indicator
LP	Low Priority
MPEG	Moving Pictures Experts Group
OSI	Open Systems Interconnect
P frame	Predicted frames
PAD	Padding
PBP	Priority Breakpoint
PCM	Pulse Code Modulation
PDU	Protocol Data Unit
PTI	Payload Type Indicator
PVRG-MPEG	Portable Video Research Group-MPEG
Q-Q plot	Quantile-Quantile Plot
RES	Reserved
RGB	Red, Green, Blue
Sr. #	Serial Number
SAR	Segmentation And Reassembly
SEAL	Simple Efficient Adaptation Layer
SDU	Service Data Unit
SIF	Standard Interchange Format
TES	Transform Expand Sample
VBR	Variable Bitrate
VCI	Virtual Channel Identifier
VLC	Variable Length Code
VPI	Virtual Path Identifier
Y	Luminance Component

1

Introduction

1.1 Background

With the availability of high speed—digital transmission facility, switching and signal processing technologies, there is an emerging demand for the development of *broadband-ISDN* (B-ISDN) services [5]. The term *broadband* refers to the use of high bandwidth, to support new and upcoming services. Efforts all over the world are going into the development of a common B-ISDN network standard.

Asynchronous transfer mode (ATM) is accepted as the transfer mode protocol that will implement B-ISDN networks [5]. ATM is designed to be independent of the physical layer protocol. The physical layer is concerned with the actual transmission of bits over a physical network link. ATM is a cell based transmission facility that uses a 53 octet cell for transmitting various types of traffic. It also supports circuit-switched applications with the help of a proper adaptation layer. ATM has the capability to absorb variations in the traffic behavior of input sources with the help of statistical multiplexing. This facilitates the integration of *variable bitrate* (VBR) services [44, 48] in ATM networks.

B-ISDN services can be broadly classified as *interactive* and *distributional* services. Interactive services are usually quite time constrained by their end-to-end delay parameter,

Background

whereas the distributional services usually have greater flexibility in their end-to-end delay requirement. Video applications can belong to any of these two service types, for example video conferencing is interactive and broadcast television is distributional in nature. Video, in its various forms, is expected to be one of the most demanding applications in future B-ISDN networks. Therefore it is essential that video be studied in all its forms and types.

Before considering the transmission of video in a networks it is very important to subject the raw video to some kind of compression, in order to reduce its bitrate. Some of the popular video compression standards are MPEG¹, CCITT Recommendation H.261, and JPEG² for still images. Raw video of size, 720 x 480 pixels, using 24 bits per pixel, at 30 frames per second results in a bitrate of about 248.83 Mbps. MPEG-2, is capable of transmitting this type of video at a bitrate of 10~15 Mbps. Evidently, the compression is of many orders of magnitudes. However, one drawback of a good compression algorithm is that, it renders the video traffic bursty. The size of any video frame (in bytes per frame), after compression is a strong function of the scene contents, motion of objects in the frame, the motion of the camera and the desired resolution.

Earlier compression techniques focused upon providing constant bitrate (CBR) output by utilizing a buffer at the output of the codec (encoder and *decoder*), which smoothed out the variations in the output traffic. The bitrate of the codec is also controlled by implementing an adaptive quantization procedure, which is varied based upon the buffer fullness of the codec output buffer. The codec output bitrate is reduced by increasing the quantization scale; this has an impact of degrading the image quality slightly, each time the quantization scale is increased. By keeping the quantization scale constant, video can be encoded at constant quality [34]. However, this leads to VBR encoding. With the arrival of B-ISDN,

1. Moving Pictures Experts Group

2. Joint Photographic Experts Group.

Motivation and Scope

high quality VBR video will be able to take advantage of the high bandwidth and statistical multiplexing provided by ATM networks.

Of all the video compression algorithms available at present, MPEG provides a very high degree of compression while maintaining a good video quality. MPEG-1 was developed primarily for CBR coding whereas MPEG-2 is designed for VBR applications. We used a public domain MPEG codec called PVRG-MPEG¹ for the purpose of our study which we modified to implement some features of MPEG-2 standard.

1.2 Motivation and Scope

Researchers have studied video and its applications, more in the past few years than ever before [1,9, 11, 26, 34]. Hence, we chose to look at the transmission of VBR MPEG video in ATM networks.

ATM is designed to keep the processing at a network node to a minimum. With the intelligence being pushed to the network end equipments, there is limited buffer management capability available at the network node to tackle congestion. Too many cells could be lost in succession (burst loss) due to buffer overflow. Increasing the intelligence at the node would only lead to an increase in the processing overhead which is quite undesirable in fast packet switching networks [48]. We have suggested a simple traffic shaping strategy for MPEG video, which is quite useful for distributional video and improves the performance of cell loss ratio in an ATM network; even in cases when the network node utilization is quite high.

A major concern with the transmission of video through an ATM network is the cell

1. Private Video Research Group

Motivation and Scope

loss due to buffer overflow [34]. In order to avoid indiscriminate discarding of cells at the network node in the event of congestion, we look at a two-layered MPEG video source with a high priority layer and a low priority layer. This helps in confining the cell losses to the low priority layer as much as possible.

It is a normal practice to use real video sequences to run simulations that are aimed at performance evaluation of a networks [9, 17, 26]. However, the results obtained from such experiments are expensive due to the fact that real video requires a lot of preprocessing, before it can be used in simulation. The applicability of such results can be limited because a short video clip might not include the variations of all video sequences. Therefore, the results from a simulation using a real sequence would be true for only a few similar kinds of video sequences [34]. This is the motivation behind developing video models that represent real video traffic and capture both the first and some second order statistics of a real video sequence. These models can be used to run longer simulations at a much lower cost and critical parameters related with rare events can be evaluated. Various analytical [13,16, 30, 39, 46] and statistical [31,37, 47] models have been proposed in the literature for video traffic. However, none of them discuss the modeling of a two layered video source. As expressed earlier, we investigate the transmission strategy for a two layered MPEG video in ATM networks; we also develop a two layered video model for the MPEG sequences we use in this study. The modeling process uses a technique called *Transform Expand Sample* (TES)[31].

In cases when real video sequence is not available or is too expensive to generate, it may be advantageous to develop video models for MPEG like video sequences based on certain user defined statistical quantities, such as the mean and variance of a video sequence. In the sequel, we develop models that are based on known distributions to represent the video frame sequence. TES is used again in the development of this type of models.

2

Background

2.1 Introduction

This chapter is intended to provide the background information required for understanding the chapters that follow in our discussion. The three main topics of discussion are:

1. MPEG (Moving Pictures Experts Group)
2. TES (Transform Expand Sample)
3. ATM (Asynchronous Transfer Mode)

MPEG is a moving pictures compression standard and is accepted as the *de-facto* standard for most video applications. It has been defined for CBR and VBR mode of operation as MPEG-1 and MPEG-2 respectively. Efforts are underway to come up with a low bitrate video standard called MPEG-4 for video-phone like, low bitrate applications. In Section 2.3 we will describe the MPEG video compression. Video modeling has been an important issue for quite some time now and various statistical and analytical models have been developed so far. However it is very difficult to develop a model for video traffic that can be used to fit all kinds of video data. We use TES for the purpose of modeling MPEG video traffic as it is the only technique that can capture the given marginal distribution and autocorrelation function of a given empirical sample sequence in two unique steps. TES is a statistical modeling technique that decouples the matching of the given marginal from the

Digital Pictures

capturing of the autocorrelation. We use it to match the two layered MPEG source. The modeling of MPEG video using TES will be presented in Chapter 3 however, a brief illustration of TES methodology is presented in Section 2.4. A network node based on ATM is modeled for the purpose of simulating the behavior of multiplexed traffic from similar MPEG video sources. Section 2.5 illustrates the ATM protocol used for our experiment. As our first discussion involves the compression of digital images using MPEG it is necessary that we look at the color representation. The next Section discusses digital pictures and various color space representation.

2.2 Digital Pictures

In order to compress an image, it must be first transformed into digital format from its natural analog form. Likewise, before applying MPEG compression to a sequence of images, we need to have a sequence of such digital images. Images in their digital format are represented as a two dimensional array of pixels; the size of the two dimensional array is a function of the desired resolution at the time of digitization.

2.2.1 Image Formats

Images can be represented in greyscale or in color. Greyscale (monochrome) representation has only one component, called the *luminance* (Y) and each pixel is represented completely using only this component. Color representation requires at least three components to completely represent each pixel. In a color television or a *cathode ray tube* (CRT), almost all colors are generated on the screen using the red, green and blue (RGB) color components. However, controlling colors in this format is extremely difficult [38]. Therefore other color representations are defined that are based on the isolation of the luminance from the *chromi-*

nance (color information, *Cr* and *Ch*) components. This representation is routinely used for transmitting video information for a standard color television. The luminance provides a grey scale version of the image and is the same signal for both, the color as well as the monochrome television. The chrominance components provide the extra information that transforms the grey scale image into a color image.

Luminance-chrominance representations are particularly important for good image compression because the two chrominance information, which are represented as a difference signal from the luminance signal, allow for a more accurate reproduction of colors as opposed to the RGB format [38]. In the next Section we discuss the relationship of this representation with the RGB coordinates.

2.2.2 Linear Color Transformation

The luminance in a CRT display is the sum of the individual contribution to luminance from the three individual color components; red, green and blue phosphors. When all three colors are of equal intensity, i.e., $R = G = B$, the resulting value of the luminance lies on the grey scale. For any particular values of R , G and B , the luminance (Y) can be calculated from the following weighted sum:

$$Y = 0.3R + 0.6G + 0.1B \quad [2-1]$$

The scaling factors of RGB are chosen such that the luminance is also expressed by a relative scale from 0 to 1 and the weights reflect the contribution of the individual primary colors to the total luminance. The chrominance is defined as the difference between a color and a reference white at the same luminance. The chrominance information can then be expressed by a set of color difference values, V and U , where V and U are defined by:

Digital Pictures

$$V = R - Y \quad [2-2]$$

$$U = B - Y$$

These color differences are zero when $R = G = B$, as this condition produces grey, which has no chrominance.

Another color space that is used in North American television is the YIQ color space. The luminance signal in YIQ is the same as the luminance in YUV , as in Equation (2-1). The YIQ color space is related to the YUV color space as follows:

$$I = 0.7V + 0.27U \quad [2-3]$$

$$Q = 0.48V + 0.41U$$

Still another color coordinate system, $YCbCr$, is used extensively in the development of JPEG, MPEG and H.261 standard. This color system is very closely related to YUV and uses the same Y coordinate as the YUV system, whereas U and V are scaled and zero-shifted to produce the variables Cb and Cr , respectively. The relationship between YUV and $YCbCr$ is as follows:

$$Cb = \left(\frac{U}{2} \right) + 0.5 \quad [2-4]$$

$$Cr = \left(\frac{V}{2} \right) + 0.5$$

The scaling and zero shifting in Equation (2-4) ensures that the chrominance values are always in the range 0 to 1. The values of Y , Cb and Cr are then multiplied by 255, so that they can be represented by an 8 bit integer.

In order to use these digital pictures in the MPEG coding process, they have to be

transformed according to the CCIR recommendation 601 specification[34]. The CCIR-601 nominal range is between 16 and 235 for the luminance (Y) and between 16 and 240 for the chrominances (Ch , Cr). Hence the above $YCbCr$ color space is rescaled and shifted accordingly:

$$\begin{aligned} Y' &= \frac{219}{255}Y + 16 \\ U' &= \frac{224}{255}Ch + 16 \\ V &= \frac{224}{255}Cr + 16 \end{aligned} \quad [2-5]$$

This results in the 4:4:4 format, as each pixel is defined using all three components of Equation (2-5). If we use an 8 bit representation of these components we have 24 bits per pixel. The size of the chrominance components is reduced through subsampling as outlined in the following Section leading to fewer bits per pixel.

2.2.3 Subsampling

The human visual system has limitations in perceiving fast changes in the chroma domain. Taking advantage of this fact, the 4:4:4 format can be subsampled to reduce the size of the chrominance component. CCIR-601 defines another format called 4:2:2 at 60 fields per seconds. Two fields combine to form a frame as they are interlaced. These fields are subsampled in the horizontal direction for both the chrominance components to reduce the size by half. The resulting 4:2:2 format is shown in Figure 2-1 (a). The size of the image in CCIR 601 is 720 x 480 pixels, and the size of chrominances is 360 x 480 pixels in 4:2:2. MPEG, has to further convert this format into 4:2:0 *standard interchange format* (SIF) of Figure 2-1 (b). This is done by sampling the odd fields to reduce the horizontal resolution by

MPEG

Figure 2-1 CCIR-601 and SIF formats

Note: Circles represent Luminance and Boxes represent Chrominances

half and then by similarly subsampling all three components in the horizontal direction[20]. In addition the chrominance values are further subsampled vertically. This process results in a size of 360 x 240 pixels for the luminance and a size of 180 x 120 pixels for both the chrominance components. The final image has 12 bits per pixel in the 4:2:0 SIF format. MPEG-1 makes use of only this format, whereas MPEG-2 is capable of encoding from CCIR 601 (4:4:4 and 4:2:2), as well as 4:2:0 SIF format.

For encoding in MPEG an image is supposed to be a multiple of 16 (explained later) in horizontal and vertical direction. Therefore the image in SIF format is reduced to 352x240 pixels by discarding four columns of pixels from the right and left side of the image. This gives us the *common interchange format (CIF)*; 352x240 pixels per frame for luminance and 176x120 pixels per frame for each chrominance components.

2.3 MPEG

MPEG is a motion video compression standard designed to take advantage of the temporal dependence between frames. It achieves about three times better compression than a sequence of JPEG¹ compressed images, when measured in bits/pixel per frame [38]. Spatial

MPEG

correlation is reduced using *discrete cosine transform* (DCT) followed by entropy (Huffman) coding. We will discuss the video compression using the MPEG compression standard in Section 2.3.1. However video is not the only component of an MPEG bitstream. MPEG is defined as a standard for multimedia traffic. Three main components in MPEG are:

1. Systems Layer
2. Audio Layer and
3. Video Layer

Additional information such as text, graphic, etc., could be multiplexed into one MPEG bitstream.

The systems layer is mostly involved in control specific functions. The systems layer multiplexes multiple video, audio and data streams and preserves their timing order and synchronizing information, as time stamps, that are used by the systems layer at the receiving end to demultiplex and forward each stream in the right order [20]. Audio in MPEG-1 has two channels which may be compressed at 64, 128, or 192 kbits/second per channel. In MPEG-2, the audio compression specified is of PCM¹ resolution at a sampling frequency of 48, 44.1 and 32 kbits/second at 16 bits per sample [20]. There are five channels specified for MPEG-2, that are sufficient to encode surround quality sound or provide five multilingual channels. Audio channels are CBR hence they do not contribute to the variability in the output bitrate. Furthermore the bandwidth used by the audio channels is only a small fraction of the total bandwidth required for the transmission of MPEG bitstream, typically about 1.5 Mbps for MPEG-1 and about 6~10 Mbps for MPEG-2. As the systems layer encapsulates video, audio and data it adds only some header information to the bitstream, which represents an insignificant increase in the total bitrate. Therefore, we feel that, in order to get an

1. Joint Picture Experts Group: Still pictures compression standard
1. pulse code modulation

Figure 2-2 Structure of the MPEG compressed video bitstream

insight into the transmission behavior of MPEG traffic in any network, it is sufficient to consider only the video component of the MPEG bitstream.

2.3.1 Structure of MPEG Video Stream

The MPEG video stream is a good example of layered encoding. Its hierarchical structure is illustrated in Figure 2-2 as follows:

1. Sequence Layer -- made up of group of pictures.
2. Group-of-Pictures -- made up of I, P and B pictures.
3. Pictures -- made up of slices.
4. Slices -- made up of macroblocks.
5. Macroblocks -- 4 (2x2) *Y* blocks + 1 *Ch* block + 1 *Cr* block.

MPEG

6. Block -- 8x8, two dimension array of samples.

Four byte start codes are present at the beginning of the sequence layer, group-of-pictures layer, picture layer and the slice layer. All these start codes are unique and are byte aligned in order to be easily accessible by any program, without having to decode the entire bitstream first. Therefore, the smallest denomination of the MPEG stream that can be accessed in the compressed form is the *slice*, and for this reason we deal with slices extensively for video modeling and network simulation. The macroblocks have a header that represents its address increment from the previous encoded macroblock. This being a VLC with a number of bits between 1 and 11, cannot be easily read from the compressed bitstream. A brief description of each layer is given in the following Sections:

2.3.1.1 Sequence

The sequence header initializes the state of the decoder. It is followed by a one or more *Group of Pictures* (GOPs) and is ended by a `sequence_end_code` (see Figure 2-2). Additional sequence headers may appear in a sequence along with all parameters. Repeated sequence headers facilitates random access into the video stream.

2.3.1.2 Pictures and Group-of-Pictures

Four types of pictures (i.e., frames) are used in MPEG video compression. Each one is explained below and the relationship among I, P and B frames is presented in Figure 2-3:

1. I frames (Intra frames) are coded, based on the information within a frame with no temporal dependence from any other frame in the sequence.
2. P frames (Predicted frames) are predicted using forward motion estimation from a previous I or P frame.

Figure 2-3 Temporal redundancy reduction using I, P and B frames.

3. B frames (bidirectional interpolated frames) are derived from the residual frame after prediction from a past and a future frame in the temporal domain.
4. D frames (DC frames) contain only the DC (8x8 block average) component of each block. This is an optional frame type and may not be mixed with the other three frame types.

A GOP is a set of pictures which are contiguous in display order (I B B P B B P...), contain at least one I frame and may be followed by any number of I or P frames. Any number of B frames may be inserted between each pair of I or P frames. We assume the case of two B frames between any two I and P frames or P and P frames in our study. A GOP has the following properties:

Property 1. A GOP begins with an I picture and may be followed by any number of I or P pictures.

MPEG

- Property 2.** A GOP must begin with an I or B picture in display order, and must end with an I or a P picture. The smallest GOP has one I picture but the size of a large GOP is theoretically unlimited.
- Property 3.** A GOP begins with a GOP header and ends at the next GOP header or the next Sequence header whichever comes first.

2.3.1.3 Slices

Each frame is divided into slices. A slice contains an integral number of macroblocks (16x16 samples) in the raster scan order. Slice may be of different sizes and are non-overlapping. Hence we assume each slice to be made up of a row of macroblocks. For a frame with dimensions of 352x240 pixels we get 15 slices with 22 macroblocks in each slice. The minimum number of slices could be unity (equal to the frame size) and the maximum number could be equal to the number of macroblocks in a frame.

A slice header contains information about the vertical as well as the horizontal position of the slice in a frame. As the slice headers are byte aligned, each slice can be easily read and decoded independently from the other slices in the frame. This structure could be easily used to implement parallel encoding and decoding of the slices using multiprocessor codecs. This is also one of the reasons why we chose to model and simulate an MPEG source at the slice level.

2.3.1.4 Macroblocks and Blocks

A macroblock is a 16x16, two dimensional array of samples (pixels) of an image. Each macroblock has samples of the luminance (Y), and the two chrominance (Cr and Cb) components. If the image samples are in 4:4:4 format (the luminance and chrominance have the same size), the chrominance components are subsampled in half, both in the horizontal and vertical direction, to reduce their size to an 8x8 block of samples. Simultaneously, the Y

Figure 2-4 Formation of Blocks from a Macroblock

samples are split into four blocks of 8x8 samples. This way we end up with six blocks for each macroblock. Other formats supported by MPEG-2 (4:2:2 and 4:4:4) are not considered here. In Figure 2-4, the subsampling and formation of blocks from a macroblock is shown.

2.3.2 MPEG Video Compression

2.3.2.1 Frame Reordering

Until now we have discussed about the various building blocks (constituents) of an MPEG compressed bitstream. In Figure 2-3 we saw the MPEG sequence and the dependence of different frames on each other. I frames are the points where the complete image is refreshed. The following frames that are encoded as B frames, require a future P frame for encoding. Therefore the next frame that is encoded after an I frame is the P frame. With the future P frame and the past I frame already encoded, the intermediate B frames are encoded through

MPEG

Figure 2-5 MPEG Video encoder block diagram.

interpolation. This requires reordering of the frames at the encoder such that the frames arrive at the encoder in the following sequence:

I P B B P B B P B B P B B ...
 1 4 2 3 7 5 6 10 8 9 13 11 12 ...

After the frames have been decoded they are presented in their original order as:

I B B P B B P B B P B B P ...
 1 2 3 4 5 6 7 8 9 10 11 12 13 ...

2.3.2.2 Motion Compensation

As shown in Figure 2-5, after frame reordering, the encoder estimates a motion vector for each macroblock in the new frame, using the information from the past and/or the future frame, that are stored in the codec frame buffers (picture store and predictor). When a close match is found an error signal is generated, which is derived as a difference signal between

Figure 2-6 Applying DCT to 8x8 block to get an 8x8 block of coefficients

the input samples (macroblock) and the matched samples (macroblock). If the resulting macroblocks of the error signal has energy lower than the original macroblocks, this macroblocks is encoded along with the motion vectors. If the calculated motion vectors do not bring down the energy of the error macroblock (i.e., the macroblock is not properly predicted), then the macroblock is encoded as intra-macroblock (encoding the original samples and not the sample errors). However, motion vectors are still added to the compressed bitstream even if a macroblock is encoded as intra, to assist decoding in the presence of errors. The motion vectors are encoded as VLCs and follow the macroblock header in the compressed bitstream.

2.3.2.3 Discrete Cosine Transform

By applying a two dimensional DCT to a block of image samples, we can transform the spatial information of an image into the frequency domain resulting in the DCT coefficients. The DCT coefficients are arranged in a progressively increasing order of frequency; the

MPEG

direction of increase is shown by the dashed arrows in Figure 2-6(b). The top left coefficient in Figure 2-6(b) is devoid of any frequency components and represents the average energy of the total block. Therefore, this coefficient is referred to as the DC coefficient of the block. The DC coefficient is also handled separately during the entropy coding as explained later in Section 2.3.2.5. All other coefficients have a frequency component associated with them hence they are called AC coefficients.

The DCT used for this purpose is equivalent to performing a one dimensional DCT on the column and one dimensional DCT on the rows of the block in order to get the DCT coefficients as in Figure 2-6 (b). The equation for the implementation of DCT is of the form:

$$F(u, v) = \left(\frac{1}{4}\right) C(u) C(v) \sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \cos\left((2i+1)u \frac{\pi}{16}\right) \cos\left((2j+1)v \frac{\pi}{16}\right) \quad [2-6]$$

where

$$C(x) = \begin{cases} \frac{1}{\sqrt{2}} & x = 0 \\ 1 & \text{otherwise} \end{cases} \quad [2-7]$$

The inverse of the two dimensional DCT is written as:

$$f(i, j) = \left(\frac{1}{4}\right) \sum_{u=0}^7 \sum_{v=0}^7 C(u) C(v) F(u, v) \cos\left((2i+1)u \frac{\pi}{16}\right) \cos\left((2j+1)v \frac{\pi}{16}\right) \quad [2-8]$$

Figure 2-7 Zig-Zag scanning of the DCT coefficients.

2.3.2.4 Quantization and Zig-Zag Scan

The transform coefficients are quantized in order to reduce their magnitudes, before applying entropy coding. Generally, images have high energy in the slow varying spatial frequencies (low frequencies) and have low energy in the fast varying spatial frequency (higher frequency). We know that the human visual system is not capable to detecting fast changes in the high frequency range. Therefore, after applying DCT, the transform coefficient can be treated separately according to their importance. In the sequel, two types of quantizers are defined; one with dead-zones and one without. The quantizer with dead-zone has a bigger step size near zero, in order to reduce all coefficients in the vicinity of zero, to zero. This quantizer is applied to the high frequency coefficients. The quantizer without dead-zone, actually has a smaller step size near zero, in order to encode as many coefficients near zero as possible. This quantizer is applied to the low frequency coefficients.

The quantization step size is sent along with the encoded coefficients to assist the decoder in the decoding process. All coefficients that have low magnitude are reduced to zero in this step. The next step of the zig-zag scan (see Figure 2-7) transforms the two

MPEG

dimensional block into a one-dimensional array of quantized coefficients, in a progressively increasing order of frequencies. This process help to increase the run-length of zero valued coefficients. Entropy coding exploits this while encoding the quantized coefficients as explained in the next Section.

2.3.2.5 Entropy Coding

Entropy coding is a means of assigning longer code words to symbols that occur rarely and shorter code words to symbols that are quite frequent. Entropy here is related to the amount of information a particular symbol conveys. If the symbol is highly likely then it does not convey much information upon occurrence, and vice-versa.

The DC coefficients are dealt with separately. After quantization the DC coefficient in an I frame is encoded directly using the Huffman code table. However, in P and B frames the DC coefficient is first derived as a differential quantity; taking the difference between the present DC coefficient and the DC coefficient of the previous block. The difference quantity is then encoded using Huffman code.

The one dimensional array of quantized AC coefficients usually contain a number of zero valued coefficients as a result of the quantization. Before entropy coding, these quantized coefficients are converted into symbols of the type (x, y) , where x represents the run-length of consecutive zeros preceding a non-zero coefficient and the y denotes the magnitude of the non-zero coefficient. These symbols are then encoded using Huffman code. The symbols of high probability of occurrence are assigned shorter VLCs while symbols that occur rarely are assigned longer VLCs.

The final compressed bitstream is made up of VLCs. Since the VLCs cannot be recognized until they have been decoded, it is important that there be some code words that can

MPEG

be read by other programs without having to decode the entire sequence. This is the reason for having byte aligned headers that can easily identify the beginning and end of a sequence and the beginning of GOPs, pictures, and slices.

Until now, our discussion regarding MPEG compression was quite general, it applies equally to both standards; MPEG-1 and MPEG-2. MPEG-1 is actually a subset of MPEG-2. In the following Section we bring out some of the major differences between MPEG-1 and MPEG-2.

2.3.3 Changes in MPEG-2

MPEG-2 addresses video applications at higher sample rates such as broadcast television, using the CCIR 601, 4:2:2 format, which has a bitrate of 216 Mbps[3]. MPEG-2 is capable of encoding this video at about 15 Mbps. Broadcast television is interlaced having 60 fields which make up 30 frames per second; such interlaced video is supported by MPEG-2. MPEG-1 can only implement non-interlaced video.

MPEG-2 video has four scalability features. Scalability is referred to as the method of encoding video in a layered structure, where the output traffic is made up of two or more distinct bitstreams (layers). The most important layer is referred to as the *base layer* and can be decoded independently of the other layers. The other layers are called the *enhancement layers* as these layers enhance the image quality of the base layer. Usually, there are only two layers; one base layer and one enhancement layer. In our study we implement one scalability feature, which is known as *data partitioning* in the MPEG-2 standard. Data partitioning will be explained in detail, later in Section 3.3.2. The other forms of scalability features offered by MPEG-2 are known as; SNR scalability, Spatial scalability and Temporal scalability [20].

We have seen that, MPEG-1 includes motion vectors for only the P and B frames. In MPEG-2 however, concealment motion vectors are included for I frames. The function of these concealment motion vectors is to assist the decoder while decoding corrupted I frames. The DC coefficient in MPEG-1 have a fixed precision of 8 bits only while in MPEG-2 this value is variable and can be chosen from 8, 9, 10 and 11 bits.

The size of the input video frame is required to be a multiple of 16 pixels in MPEG-2, both in the vertical and horizontal direction. This ensures that each frame can be easily divided into macroblocks (16x16), without requiring any preprocessing. This is not the case in MPEG-1 where one of the function of preprocessing is to make sure that the input is made a multiple of 16 by adding dummy pixels if required.

The maximum size of the video frame has been increased to 16384x16384 pixels in MPEG-2, from 4096x4096 pixels in MPEG-1. MPEG-2 also supports more image formats, i.e., CCIR 601 (4:4:4 and 4:2:2) and 4:2:0 SIF, where as MPEG-1 supports only 4:2:0 SIF format. The EOB (end of block) uses a four bit code in MPEG-2, whereas MPEG-1 uses a two bit code as EOB.

MPEG-1 is basically designed for CBR applications like CD-ROMs and other digital storage devices. The rate control is implemented using adaptive quantization and bitstuffing at the macroblock level. In MPEG-2 bit stuffing is declared illegal and rate control, if required, is implemented by varying the quantization scale.

2.4 TES Processes

Transform Expand Sample (TES) is a versatile methodology for generating autocorrelated, stationary time series with an arbitrary marginal distribution (a first order statistic), and an

arbitrary autocorrelation function (a second order statistic). TES is able to synthesize data that match closely with any set of given observations of a stationary time series in terms of the marginal distribution and autocorrelation function. It is based on an autoregressive modulo-1 scheme. The advantages of this scheme are;

1. It can generate any marginal distribution or an arbitrarily close approximation.
2. The associated autocorrelation function can be rapidly computed numerically.
3. TES methods for generating time series are fast, in fact computationally they are only a little more complex than the generation of the underlying pseudo random number generator.
4. TES behavior is versatile, ranging from driftless random walks to cyclic random walks.

More details can be found in [10,18, 31, 32].

The development of a TES model is performed in two phases. In the first phase a correlated sequence (also called **background** TES process), $\{U_n^+\}_{n=1}^{\infty}$ or $\{U_n^-\}_{n=1}^{\infty}$ is obtained as follows;

$$\begin{aligned}
 U_n^+ &= \begin{cases} U_0 & n = 0 \\ \langle U_{n-1}^+ + V_n \rangle & n > 0 \end{cases} \\
 U_n^- &= \begin{cases} U_n^+ & n \text{ even} \\ 1 - U_n^+ & n \text{ odd} \end{cases}
 \end{aligned}
 \tag{2-9}$$

where U_0 is a random variable uniformly distributed in $[0, 1)$; $\{V_n\}_{n=1}^{\infty}$ is a sequence of iid (independent and identically distributed) random variables independent of U_0 , called the

innovation sequence (with probability density f_v). The angular brackets, $\langle . \rangle$ in Equation (2-9) denote the modulo-1 operation which is defined as follows; for every real x ,

$$\langle x \rangle = x - \max \{ \text{integer } n: n \leq x \} \quad [2-10]$$

In Equation (2-9), $\{U_n^+\}_{n=1}^{\infty}$ and $\{U_n^-\}_{n=1}^{\infty}$ can generate lag-1 autocorrelation in the range $[0, 1)$ and $[-1, 0)$ respectively, i.e., $\{U_n^+\}_{n=1}^{\infty}$ gives samples that are positively correlated and $\{U_n^-\}_{n=1}^{\infty}$ negatively correlated. Regardless of the probability distribution of the innovation density, f_v , the TES background sequence forms a stationary Markovian process with a uniform marginal distribution in $[0, 1)$ [31]. Intuitively, the modulo-1 arithmetic used in defining TES processes in Equation (2-11) has a simple geometric interpretation as a random walk on a circle of circumference one (unit circle) with a mean step size $E[Vn]$. The walk will have zero, positive or negative drift around the unit circle according to $E[Vn] = 0$, $E[Vn] > 0$ or $E[Vn] < 0$, respectively. It needs to be mentioned that on the unit circle the points 0 and 1 coincide, therefore when we consider a linear mapping of these points to the real line we observe discontinuities at 0 and 1. To compensate for these discontinuities a smoothing operation called the *stitching transform*, parameterized by ξ , ($0 \leq \xi \leq 1$) is applied to the background TES sequence as follows:

$$S_{\xi}(U_n) = \begin{cases} U_n / \xi & (0 \leq U_n \leq \xi) \\ (1 - U_n) / (1 - \xi) & (\xi \leq U_n \leq 1) \end{cases} \quad [2-11]$$

For $0 \leq \xi \leq 1$, the effect of the transform $S_{\xi}(\cdot)$ is to make the sample paths of the background TES sequence have a '*more continuous*' appearance. It should be mentioned

that the transformed sequence $\{S_{\xi}(U_n)\}_{n=1}^{\infty}$ also has a uniform marginal in $[0, 1)$.

In the first phase of the TES model development, the objective is to generate a time series of correlated random variables with uniform marginals in $[0, 1)$. The amount of correlation depends upon the structure of the density function f_v of the innovation sequence $\{V_n\}_{n=1}^{\infty}$, which in general may be quite arbitrary. In the second phase, synthetic data called the *foreground sequence*, that resembles the real samples is derived from the background sequence using the inversion technique [32] that allows us to transform any uniform random variate, to one with an arbitrary distribution. More specifically, a histogram $H(\cdot)$ of the empirical video sequence is built consisting of J cells, where the j th cell is positioned on the interval $[l_j, r_j)$ and is characterized by a probability p_j ($j = 1, 2, \dots, J$). Such a histogram can be inverted as follows;

$$H^{-1}(x) = \sum_{j=1}^J I_{[\hat{C}_{j-1}, \hat{C}_j)}(x) \left[l_j + (x - \hat{C}_{j-1}) \frac{w_j}{\hat{p}_j} \right] \quad [2-12]$$

$$0 \leq x \leq 1$$

where I_A is the indicator function of set A , $w_j = r_j - l_j$ is the width of cell j , and $\hat{C}_j = \sum_{i=1}^j \hat{p}_i$, $1 \leq j \leq J$, $\hat{C}_0 = 0$ and $\hat{C}_J = 1$).

The random variate $\{H^{-1}(S_{\xi}(x))\}_{n=1}^{\infty}$ has a marginal distribution $H(\cdot)$ and autocorrelation that depends on the innovation density f_v . The modeling process is carried out in two steps. The first step essentially captures the autocorrelation of the input time series using an innovation density f_v and stitching parameter ξ . The second step of inverting the back-

ground TES sequence using the empirical histogram always guarantees the matching of the marginals. The task of finding a suitable (f_{ν}, ξ) pair, that approximates the empirical autocorrelation is carried out via a heuristic search using TESTool [10].

TESTool is a visual interactive environment for modeling autocorrelated time series and is a convenient and straightforward tool that reads in the empirical time series and calculates the histogram, autocorrelation and spectral density in textual and graphical representation [18]. In Figure 2-8, the graphical user interface (GUI) of the TESTool is shown. It includes four windows. The top left window in Figure 2-8 displays the sample path, the top right window displays the histogram of the sample path and the bottom left window displays the autocorrelation function of the sample path. The bottom right window in this case is used to select the innovation density based on a visual search for best fit of the model autocorrelation function to the empirical counterparts. Consequently the sample path and the histogram of the developed model are displayed along with their empirical counterparts in their respective windows. The matching of the histogram is guaranteed, while the matching of the autocorrelation is carried out in an interactive manner with visual feedback for the best (f_{ν}, ξ) pair. After a good match is found, the model is saved to a file which is used later by another program to generate a sample sequence of any length.

This concludes the modeling of any given stationary time series using the TES methodology and the visual aid package that is used to facilitate modeling.

2.5 ATM

ATM is selected by CCITT Study Group XVIII, as the target technology [5] to overcome the difficulties facing the existing transfer modes and allowing the creation of a B-ISDN

Figure 2-8 An example of TES modeling using TESTool

ATM

[14,44]. An extremely wide variety of services are expected to be provided over the B-ISDN networks [14]. The resulting traffic types will have a wide range of requirements from the network resources. Therefore it is extremely important that the transmission protocol be sufficiently simple and flexible. For this reason, the seven layer OSI¹ stack found in the conventional networks, which is known for the heavy overhead it offers to the transmission facility, is replaced with the protocol reference model (PRM) of Figure 2-9 in ATM net-

Figure 2-9 ATM Protocol Reference Model (from CCITT Rec. 1.321)

works. The ATM layer is primarily responsible for managing and generating the header portion of the cell. For transmission, the ATM layer accepts data from a higher layer, adds the header and forwards an entire 53 octet cell to the Physical Layer; upon receipt, it accepts a cell from the Physical Layer, strips off the header, and passes the user data to the layer above.

In order to allow the higher layers to make use of the services of the ATM layer, an adaptation layer is defined called the ATM adaptation layer (AAL). The AAL is placed above the ATM layer and it provides service dependent functions to the layers above the AAL. The function of the AAL is mainly to conceal the behavior of the ATM layer from the higher layers and those of the higher layers from the ATM layer. The physical layer is com-

1. Open Systems Interconnect

Figure 2-10 An ATM cell structure

pletely independent of the ATM layer and is primarily concerned with the transmission of bits over any transmission medium (fibre optics, coaxial or copper wire). In the following two Sections we will briefly discuss the ATM and the AAL layer, respectively.

2.5.1 The ATM Layer

ATM is designed to carry a variety of traffic types in its fixed sized containers of 53-octets, called *cells* and the flexibility needed to support variable transmission rates is provided by the transmission of the necessary number of cells. The principal driver for the payload size was the latency for voice traffic. Having a fixed cell helps in developing network equipments that are quite universal and by keeping the protocol simple, transmission speeds are not sacrificed. An ATM cell is made up of a 5 octet header and a 48 octet payload as shown in Figure 2-10.

ATM cells are multiplexed statistically. Figure 2-11 shows the bit allocation structure for the header of an ATM cell. A 16 bit identifier is given to every connection, upon connection establishment on an end to end basis. This identifier is called the virtual channel identifier (VCI). The virtual path identifier (VPI) on other hand is significant only for the length of a particular link. It is represented in 8 bits. When a connection passes through an ATM node the VPI is mapped to a new VPI on the next link. Both these identifiers are called *virtual* because the channel bandwidth is not physically allocated to each connection at the time of connection establishment but is dynamically shared between all established connections.

Figure 2-11 Header format for an ATM cell

The other components in the cell header are; the 4 bit generic flow control (GFC), the 2 bit payload type indicator (PTI) used to identify the type of payload the cell is carrying, i.e., user data, control cell, management cell etc., a 1 bit reserved (RES) field, a 1 bit cell loss priority (CLP) used by ATM to provide two levels of priority handling and the last component in the cell header, the 8 bit header error check (HEC) which is a cyclic redundancy check (CRC) code that is used to protect the cell header from bit errors. HEC is capable of correcting one bit error and detecting multiple bit errors. The RES field is combined with PTI field in AAL5 in order to provide congestion notification [25].

The CLP will be used later for implementing priority in our simulations. If CLP is set the cell is said to be of low priority and in the event of a congestion the network node is allowed to discard such a cell first, if one is available.

2.5.2 ATM Adaptation Layer (AAL)

The main function of AAL is to map the information from the higher layer into the payload field of an ATM cell. While the ATM layer is responsible for the generation of the 5-octet header, the AAL is responsible for formatting the 48-octet ATM cell payload [5].

There are a variety of traffic behaviors, therefore it is important to first classify the different traffic types into classes. Four service classes can be defined based on the bitrate (VBR or CBR), timing requirements (time constrained or not) and the connection modes (connection oriented or connectionless). Table 2-1 presents the various traffic classes as proposed in [5]. Class A services include circuit switched emulation and CBR services. Class B services include VBR services requiring time synchronization between the sender and the receiver, i.e., VBR isochronous traffic such as interactive voice and video. Class C services provide connection oriented data transfer similar to the X.25 packet switching and frame relay. Class D is a connectionless data transfer.

Table 2-1 Classification of traffic in ATM

	Class A	Class B	Class C	Class D
Timing relationship between source and destination	Required	Required	Not Required	Not Required
Bit Rate	Constant	Variable	Variable	Variable
Connection Mode	Connection Oriented	Connection Oriented	Connection Oriented	Connection less

In order to provide an efficient mapping from the higher layers to the ATM layer the AAL layer is further subdivided into [5]:

1. Convergence sublayer (CS)
2. Segmentation and Reassembly (SAR) sublayer

ATM

These two layer are briefly explained in the following Sections. However the adaptation layer protocol that is implemented for our simulation model will be explained in detail later, in Chapter 4.

2.5.2.1 Convergence Sublayer

Convergence sublayer is responsible for providing enhancing service features to the ATM layer. It hides the limitations of the lower layers and makes the higher applications feel that these enhancing services are provided by the underlying B-ISDN layer, such as the transmission of the timing information as part of the CS_PDU¹.

The convergence sublayer accepts the variable length data from the higher layers and forms the CS_PDU by adding the necessary headers and trailers and is then passed down to the SAR sublayer. On receiving it strips off the CS overhead and delivers the original data to the receiving application layer.

2.5.2.2 Segmentation and Reassembly (SAR)

SAR sublayer accepts the CS_PDU from the CS and fragments them into an appropriate payload segment for transmission in an ATM cell. It also reassembles cell payloads at the receiving end to form the CS_PDU.

These two AAL sublayers work hand in hand and provide all the network services leading to the formation of the PDUs (for all four classes of traffic) that are transported in the ATM network. In order to provide service for the four classes of traffics of Table 2-1, four AAL protocols are defined in [5]; AAL1, AAL2, AAL3 and AAL4 respectively. However, later another protocol was suggested by ANSI T1S1.1, called *Simple and Efficient Adaptation Layer* (SEAL). This is now popularly known as the AAL 5.

1. PDU: Protocol Data Unit

ATM

We will discuss the AAL2 and AAL5 in detail, when we set up the simulation model in Chapter 4

3

Video Modeling

3.1 Introduction

ATM networks are expected to carry a wide variety of video applications, characterized by high bandwidth, stringent delay and loss requirements. In order to obtain useful predictions about the network resource requirements and the study of performance of VBR video traffic in B-ISDN networks we need to develop flexible models. The difficulty of preparing and handling large amounts of data means that simulations with actual data are practical only on a small scale. On other hand, if an appropriate model is used to generate data, very accurate predictions can be made regarding the network performance [34].

In this chapter we deal with the modeling of VBR MPEG video with priority encoding. Towards this end, we will first discuss the bitrate characteristics of VBR MPEG compressed video traffic followed by the implementation of priority in the video traffic.

3.2 VBR MPEG Traffic

As explained in the previous Chapter, MPEG encoding is characterized by three types of frames i.e., I, P and B frames. The I frame, which is encoded with information from within a video frame, has the highest bitrate. The bitrate of P frames is lower than I frames since, part

Figure 3-1 A typical bitrate of an MPEG sequence with I, P and B frames

of it is predicted from a previous I or P frame. The B frames are interpolation from a past and a future frame hence, they usually have much lower bitrate than I or P frames. B frames contribute to the maximum compression of an image sequence. In Figure 3-1 an example of the sample path is shown, taken from a real MPEG video sequence (son.mpg), showing the bitrate characteristics of a typical MPEG sequence .

Table 3-1 MPEG bitrate characteristics.

Frame Type	Mean Bytes/ frame	Percentage of total bitrate	ATM cells/ frame	Mean Frames/ second.
I frame	28,000 Bytes	30.3%	637 cells/frame	2.5 frames/sec.
P frame	14,000 Bytes	45.5%	319 cells/frame	7.5 frames/sec.
B frame	2800 Bytes	24.2%	64 cells/frame	20 frames/sec.

As an example, let us consider the average values of the bytes/frame for I, P and B frames from Figure 3-1 as shown in Table 3-1. The GOP in this sequence has 12 frames; one I frame three P frames and eight B frames. If we consider these frames to be encoded at 30 frames per second, we can calculate the output bitrate to be approximately equal to 1.85

Figure 3-2 Effect of cell loss in an MPEG frame

Mbps. The contribution of each frame type to the total bitrate is calculated to be approximately 30.3% (2.5 I frames/sec.), 45.5% (7.5 P frames/sec.) and 24.2% (20 B frames/sec.). The average number of ATM cells (if we assume a payload of 44 octets) that are generated by each frame type is also given in Table 3-1 to be 637, 319 and 64 cells for I, P and B frames respectively. We know that I frames are very important and that they produce a large information content which requires many more cells to transmit than either P or B frame. The network is thus required to provide guaranteed service to an I frames for a long duration in order that we not see any hole in the received images.

In Figure 3-2, we see the effect of an error due to cell loss. Figure 3-2 (a) is a frame with no loss. In Figure 3-2 (b), four consecutive cells are discarded in the sixth slice, to observe the effect of lost data in a frame. We have considered a slice to be made up of one row of macroblocks and the slice headers are byte aligned as mentioned previously. This helps in containing the errors to a single slice. The loss of data in Figure 3-2 (b) caused the decoder to discard all macroblocks until the end of slice while the subsequent slice is

decoded and displayed at the proper location.

3.3 Two-Layered MPEG

A two-layered MPEG encoder is one that provides two traffic streams which are characterized by the importance of the information they carry. Data that have a significant impact on the visual quality are designated as high priority (HP) data while the data that does not have a significant impact on the visual quality are designated as the low priority (LP) data.

3.3.1 Priority Requirement

Since B frames are not used for prediction of future frames, transmission errors in such frames are limited to only one frame. However, errors due to cell loss in either I or P frames propagate until the arrival of the next I frame, when the complete frame is effectively refreshed. Likewise, if the loss of data in Figure 3-2 (b) had occurred in a B frame the effect would not have been very noticeable but, if it were to occur in an I or P frame then the error is visible. This motivates us to consider the usefulness of incorporating some type of priority in the MPEG video traffic.

If we consider assigning priority to the MPEG traffic at the frame level, the B frames would be assigned low priority, P frames, medium priority and I frames high priority. Since ATM networks allow only two priority classes, we could consider both the I and P frames to be of high priority as both of them propagate errors to future frames. The B frames can be assigned low priority as before. From the Table 3-1 we see that I frames generate approximately 637 cells, and if all of them are transmitted as high priority then we have to ensure the proper delivery of all these cells. Similarly P frames that generate approximately 319 cells have to be high priority also. B frames generate only 64 cells and they can be tagged as

low priority without any problem as they do not propagate errors to the future frames. The problem with this type of priority mechanism is that the cells arrive in batches of high priority (HP) and low priority (LP) at network node (buffer). Hence, there is a high probability of cells of the same kind being in the node buffer at a given time. In the event of congestion, if there are only HP cells in the buffer and the arriving cells are also all HP, then the node has no choice but to discard the arriving HP, as long as there is no LP to discard. Therefore, there is a necessity to have a priority scheme that can distribute the cells into the two priority partitions at a much finer time scale, instead of prioritizing at the frame level.

3.3.2 Priority Implementation

Priority is implemented in MPEG video traffic at the DCT level in [36, 40]. In Figure 2-7, we saw how the DCT coefficients are scanned in a zig-zag fashion in order to convert the two dimensional block of coefficients into a one dimensional array of coefficients. This yields coefficients in a progressively increasing order of frequency. The high frequency coefficients usually have smaller amplitudes, and the human eye is not as capable of easily detecting changes in the high frequency range. Therefore, in general, higher frequency coefficients are of less importance. This leads to partitioning at the DCT level by splitting the DCT coefficient array (after zig-zag scanning) at a predetermined point called the priority breakpoint (PBP) as shown in Figure 3-3. We assign high priority to the DC coefficient and the low frequency AC coefficients, and low priority to the high frequency AC coefficients within a block. All headers and motion vectors are always assigned high priority. There is always a fixed number of coefficients in the HP and LP portion of the block based upon the value of the PBP [36, 40].

This type of partitioning of the MPEG traffic results in two output streams. One called the high priority partition and the other the low priority partition. These two streams

Figure 3-3 Priority Implementation in a block

Figure 3-4 Formation of prioritized slices from prioritized blocks

of video data are buffered at the output of the encoder as shown in Figure 3-4, in order to accumulate the macroblocks (made up of blocks), and when a slice is completely encoded for both partitions, the HP portion of each slice is followed by its LP portion on the transmission link. A slice in our case is made up of one row of macroblocks. We have considered the implementation of priority at the slice level because in MPEG this represents the lowest denomination of distinguishable entities.

However, there is a limitation to this type of priority implementation, where the PBP

represents a fixed number of DCT coefficients. The DCT coefficients are all encoded as VLCs. Some coefficients may not be encoded independently after quantization because they are too small and fall into the dead zone of the quantizer. Such DCT coefficients are encoded as part of a non-zero coefficient. Therefore the HP partition may not always have as many encoded coefficients as the PBP. At the receiving end, after the decoder has extracted the coefficients for a block from the HP partition, it must know when to switch to the LP partition for decoding the remaining coefficients to complete the block. Therefore, the PBP must be explicitly encoded in the compressed bitstream for this purpose. It has been calculated in [36] that the amount of overhead associated with this priority implementation is almost 20% of the total bitrate. This high overhead is due to the codewords inserted into the compressed bitstream to highlight the PBP in it.

3.3.2.1 Data Partitioning

The *data partitioning*, as described in MPEG-2 specifications, overcomes the above problem by defining the PBP in a slightly different manner. The PBP here, is defined as the number of encoded coefficients (VLCs) and not the actual number of DCT coefficients, as was the case in [36,40]. The difference between these two schemes is illustrated with the help of an example (Figure 3-5) below;

Consider a 4x4 matrix of DCT coefficients (1 to 16) as shown in Figure 3-5a. In Figure 3-5b the coefficients that have a finite value after quantization are represented with an 'x', and those that are reduced to zero are represented with a '0'. Scanning the coefficients in a zig-zag fashion (Figure 3-5b), results in an array of coefficients as shown in Figure 3-5c. Let us consider a PBP value of five. As per the earlier definition of PBP [36,40], this corresponds to PBP1, which is the absolute location of the fifth coefficient in Figure 3-5c. Implementation of PBP according to MPEG-2's data partitioning corresponds

Figure 3-5 Example of Data Partitioning

to PBP2 in Figure 3-5c, which is the location of the fifth encoded coefficient. The location of PBP2 is not fixed as it depends upon the location of non-zero coefficients in an MPEG block.

Each non-zero coefficient is often encoded along with zero valued coefficients preceding it. This leads to the total number of encoded coefficients being less than the total number of coefficients per block. As the PBP in this scheme represents the number of encoded coefficients, an MPEG video sequence encoded with this scheme may not have the same number of actual coefficients in the HP partition for all blocks, but will always have PBP encoded coefficients. If the total number of coefficients encoded per block is less than the PBP then all encoded coefficients are placed in the HP partitions.

The value of the PBP is included in the slice header. All headers and motion vectors are always placed in the HP partition. The DC coefficient and the low frequency (encoded) AC coefficients are also placed in the HP partition as in the previous scheme. In order to assist the decoder in synchronizing the two partitions, the LP slice header is derived from the HP slice header. This helps the decoder in making a comparison between the two slice headers (HP and LP) from both partitions and picking those two slices that belong to the

Two-Layered MPEG

Figure 3-6 A segment of bitstream showing the decoding process from two partitions

same original slice. When there are no transmission errors, and the slice headers match, the original slice can be generated by combining the two slices from both partitions as explained in Section 3.3.2.2. If the two slice headers do not match then the decoder having the knowledge of the PBP, can use only the HP slice for decoding.

The advantage of this type of encoding is that, the decoder is now capable of interpreting the two partitions based on the knowledge of the PBP alone. We use an example from [20] to explain the decoding process of this scheme. Figure 3-6 shows a portion of non prioritized MPEG bitstream and also the same bitstream when it is encoded using priority as specified by MPEG 2.

3.3.2.2 Decoding process of two-layered MPEG

Upon receiving the two-layered MPEG video stream the video decoder must proceed in the following steps (Figure 3-6):

- Compare the slice headers in the two partitions. If they match then remove the two headers simultaneously.

Two-Layered MPEG

- Begin decoding from the HP slice (quantization scale→macroblock header→blocks). For each block decode (DC coefficient→ encoded HP coefficients) until the encoded coefficient which represents the PBP is reached.
- Upon reaching the PBP in the HP partition the remaining coefficients for that block are decoded from the LP partition until the end-of-block (EOB) code.
- Upon reaching the EOB in the LP partition, the decoder switches back to the HP partition, to begin decoding for the next block until the PBP.
- The switching between the two partitions continues until the complete slice is decoded, at which point the slice headers for the next slice are compared from both partitions and if they match, the whole process mentioned above is repeated again.
- If the number of encoded coefficients is less than PBP, then the EOB appears in the HP partition and the decoder does not switch to the LP partition at all. This information is implicitly transmitted to the decoder, and reduces the amount of overhead required.

The amount of overhead associated with implementing priority according to this scheme was calculated for two video sequences and is shown in Table 3-2. This overhead is quite low when compared with the 20% overhead reported in [36].

Table 3-2 Percentage overhead associated with priority

Sequence	Percentage Overhead
Flower-garden.mpg	7.3%
Bike.mpg	2.41

3.3.2.3 Priority Considerations

In the previous section we discussed the implementation of priority in the MPEG video sequence. It was observed that, if the PBP is defined as the number of encoded coefficients then it has the benefit of a smaller overhead. However it was also pointed out that, the LP

Figure 3-7 Percentage HP per Slice in Flower-garden Sequence (PBP = 12)

partition may or may not have encoded coefficients depending upon the total number of encoded coefficients (EC). The following two cases are possible;

- $PBP < EC$: the HP partition has a total of PBP encoded coefficients and the LP partition has $EC - PBP$ number of encoded coefficients.
- $PBP > EC$: the HP partition has a total of EC encoded coefficients and the LP has zero encoded coefficients.

The percentage of traffic in the HP partition per slice is plotted for the flower-garden sequence in Figure 3-7, which is encoded using the MPEG-2 type of priority assignment. It is evident from Figure 3-7, that the distribution of traffic is not uniform in both partitions. We notice a higher percentage of B frame traffic in the HP partition. This anomaly is due to the fact that, B frames do not generate as many encoded coefficients as I or P frames. Usually the B frames are predicted quite well from the past and future frame and the error infor-

Figure 3-8 Percentage HP per slice in Flower-garden sequence at different PBP for the B frames

mation that is finally encoded does not have enough information content to generate as many non-zero coefficients as generated by an I or a P frame.

As discussed earlier, the least important of the three frame types is the B frame, therefore, it is desirable to have as much of this frame in LP partition as possible. However, we see from Figure 3-7 that the B frames get a better service guarantee by having a higher percentage of its traffic in the HP partition. It is therefore necessary to use a lower PBP value for the B frames in order to reduce its traffic in the HP partition. The value of PBP for the B frames was reduced in steps and each time the percentage HP plot of Figure 3-7 was re-plotted to observe the reduction in traffic in the HP partition for the B frames. Figure 3-8 shows a better distribution of traffic for all frames in both partitions with the B frame PBP equal to half of the I frame PBP value.

For different video sequences we may have to use different values of PBP for

B frames in order to have a similar distribution of traffic in both partitions as in Figure 3-8. However, empirically, a value that is half the value used for I frames is quite acceptable for B frames in most cases. Extending this analogy we used a PBP value for the P frames which is three quarters ($3/4$) of the value used for I frames. The result of using different PBP for the three frame types was to transfer about 5% of the traffic from the HP to the LP partition, which was basically a portion of the traffic from B and P frames. This transfer of traffic was the result of the high frequency coefficients belonging to B and P frames, moving into the LP partition.

3.3.2.4 Priority Selection

The idea behind having priorities in a video stream is to maintain a minimum acceptable image quality at the receiving end at all times. One way to achieve this is to have priority implemented in the coding process and employ a transmission facility that supports the prioritized traffic. In order to have an acceptable image quality at all times we must first arrive at a value for the PBP. We investigated this by encoding the video sequence at various PBP levels and then discarding the whole LP portion of the traffic, to emulate a worst case situation of cell loss during transmission.

Table 3-3 Percentage HP traffic in the codec output

Priority Breakpoint	12	16	20	24	64
Son.mpg	71.90%	80.50%	86.25%	89.90%	100%
Flower_Garden.mpg	44.40%	53.14%	61.25%	68.41%	100%
Tennis.mpg	52.15%	61.36%	69.8%	77.33%	100%

We estimated that a value of 16 for the PBP gives quite good picture quality even with all LP traffic lost. However a value of 12 for the PBP could comfortably be used without too much deterioration in image quality. Our experiment indicated that a value lower

than 12 would not be acceptable at all. As an example, in Figure 3-9 we show the 'Flower-garden' sequence encoded at various priority levels. Close examination of these images on the workstation monitor, helped us in arriving at the above conclusions. We repeated the experiment with other sequences, e.g., 'Tennis.mpg' and 'Son.mpg', and arrived at the same conclusions. Table 3-3 shows the percentage of HP traffic in a two-layered MPEG sequence.

3.4 Statistical Video Modeling

A major drawback of using real video sequences for network designs and performance studies is the enormous amount of data one needs to collect for a decent simulation run length; the disk space limitations and the practically small range of its application [34]. Simulations based on a few real video sequences do not give us enough insight into the behavior of the network, for a greater variety of video applications. Furthermore, it is not always feasible to run simulations based on real sequences. Therefore it is important to develop models, based on statistical estimates of the video sequence on analytical formulas, that are flexible to handle and easy to use in simulations. Models can dramatically speed up the simulations as they do not involve manipulation of real video sequences.

In order to model the arrival process from a prioritized MPEG source, it is required that the model capture the behavior of the traffic arriving from both partitions of the codec. Earlier works concentrated on developing models that captured the frame arrival process [13,15, 37, 42]. This is an important aspect of modeling, as it allows us to capture the autocorrelation within frames in the temporal domain. Figure 3-10 shows the autocorrelation function of a typical MPEG sequence. The autocorrelation function follows a pattern similar to the sample path or the original MPEG sequence of Figure 3-1. This is because of the way

Figure 3-9 Video sequence encoded at PBP = 4, 8, 12, 16, 20, 24, 32, 64

**Figure 3-10 Autocorrelation function of an MPEG sequence
(I frame interval = 11 lags).**

the MPEG is encoded. A typical MPEG sequence may look like:

... I B B P B B P B B P B B I ...

As long as no scene change occurs this sequence is maintained and a very high correlation exists between any frame and its eleventh lag, as in the case of the two I frames shown above. Hence we notice a high correlation coefficient at every 11th and 23rd lag and so on. The negative autocorrelation is due to the switching of the bitrate between high bitrate (I or P) frame and low bitrate (B) frames.

In our study we deal with traffic from two-layer sources. Therefore, the model we develop must generate layered traffic such that the bit arrival process from both partitions resembles the two-layered traffic generated from a real sequence. In other words, the mixing of the two partitions must resemble the traffic from a two-layered encoder. The encoder explained in Figure 3-4 had partitioning at the slice layer with HP partition followed by the LP partition for every slice. Therefore, we are interested in generating a model that first

divides the bytes per frame into bytes per slice. The slices are then split into the HP and LP partitions.

Clearly, the autocorrelation function of MPEG traffic is quite complex and is extremely difficult to capture using a single model. This problem can be overcome by using separate models for the three frame types using the TES process explained in Section 3.4.1 [42]. The traffic statistics collected during the encoding of MPEG video are used as the time series required for building the TES models. Since this model is based on empirical data, it is referred to as the *empirical-TES model*.

3.4.1 Empirical TES Modeling

In Section 2.4 we presented a statistical modeling technique that takes as input the empirical sample sequence and generates from it a model $\{X_n\}_{n=0}^{\infty}$, in two steps. In the first step the autocorrelation of the input samples is captured by a random sequence, $\{U_n\}_{n=1}^{\infty}$, uniformly distributed in $[0,1)$. In the second step this uniform distribution is inverted using the empirical histogram $\{H^{-1}(S_{\xi}(x))\}_{n=1}^{\infty}$ of Equation (2-12). The second step guarantees the matching of the histogram by default. The final challenge remaining is to find the innovation densities f_n that yield the required autocorrelation function. We have discussed earlier the software called TESTool, which is used for implementing the TES process. It helps to calculate the innovation densities in a visually interactive manner. The innovation densities and other parameters (histogram and initial seeds) can be then saved as a model for generating sample sequences which capture both the autocorrelation and the histogram of the original input sequence.

Figure 3-10 shows that the autocorrelation function of the MPEG sequence is a very complex function. This autocorrelation function cannot be modeled using a single TES

Figure 3-14 Flow Diagram of the Composite MPEG-TES Model

model. Hence, we resort to modeling each frame type separately. This required the separation of the traffic corresponding to the three frame types first. In Figure 3-11, we show the sample path of a clip from the movie 'Last Action Hero', for all three frames types, along with the sample paths from their respective models (X_n^I , X_n^P and X_n^B) generated using TES.

The matching of histograms and the autocorrelation functions for the three frame types are shown in Figure 3-12 and Figure 3-13 respectively. We have tried to obtain, as close a match as possible. After the three frame types are matched with appropriate models to our satisfaction, the next step is to generate the composite MPEG sequence (interleaved I, P and B frames) as shown in Figure 3-14.

3.4.1.1 Composite MPEG Model

In the TES models (X_n^I , X_n^P and X_n^B) developed above, we concentrated on capturing the correlation in frames of the same type. However, there is a significant correlation between frames of different types. An I frame with high information content will usually be followed

Figure 3-11 Sample paths of I, P and B frame sequences and their models

Figure 3-12 Matching of the histograms for I, P and B frames with TFS Models.

Figure 3-13 Matching of the autocorrelation function for I, P and B frames.

by P and B frames with high information contents. Moreover the I frames are completely self contained, so we can consider them to be the frame sequence driving the bitrates in the other two frame types as they are both predicted from it [42]. If we are able to incorporate this in our model then we will capture the cross correlation in between frames.

As was explained in Section 2.4 there is a background sequence $\{U_n\}_{n=1}^{\infty}$ behind every TES process. This background sequence is regarded as a random walk on the unit circle. The position on the unit circle at any point in time determines the bin from the empirical histogram that is to be used to generate the next frame. In order to induce cross correlation between I frames and the first P and B frames that follow, the position on the 'I' unit circle (i.e., U_n^I) is copied to the background TES processes U_n^P and U_n^B of the P and B frames, respectively. This method guarantees the generation of a high bitrate P and B frame following a high bitrate I frame and vice-versa. This procedure is implemented each time an I frame is generated. If we let a GOP be the sequence of frames starting with an I frame until the last frame before the next I frame then each GOP will have one I frame followed by several P and B frames. Each GOP begins with the induction of cross-correlation from I to P and B frames, and then the remaining frames that follow the first P and B frame in the GOP are derived from their respective TES models $\{X_n^P\}_{n=0}^{\infty}$ and $\{X_n^B\}_{n=0}^{\infty}$.

3.4.1.2 Implementing scene changes

The sequence of operations in the generation of the composite MPEG sequence is represented graphically by the flow diagram in Figure 3-14. The order of frame generation is:

... I B B P B B P B B P B B I ...

We assume that there are twelve frames in a GOP. An I frame is followed by two B

frames, and every P frame is also followed by two B frames in the GOP. However, at the beginning of the sequence, a P frame is encoded immediately following an I frame instead of two B frames. This is because a B frame which requires a future P (or I) frame for encoding is not yet encoded. This results in only ten frames being encoded for this particular GOP. Scene changes are assumed to occur as a result of independent Bernoulli trials. A Bernoulli random variable is drawn at the beginning of every GOP and a scene change occurs within that particular GOP with probability p_{sc} . If the result of the experiment is a scene change, then its location within the GOP is decided upon, with the help of a discrete uniform random variable U_{sc} , in the interval $[2, 12]$. If no scene change occurs then the whole GOP is encoded before making another Bernoulli trial for a scene change. When a scene change occurs the histogram to distribute the frames into slices is switched as explained in the next Section.

The decision parameters, $I_counter$ and $P_counter$, are incremented by one each time their respective frame is generated and it is assumed that both counters are set to zero in the beginning. When the $I_counter$ is equal to 1, it signifies that the first I frame was generated after the beginning of sequence. This suggests that the next frame has to be a P frame. If the $I_counter$ is greater than 1, then two B frames are generated instead.

The results of this composite TES model are shown in Figure 3-15. We notice a very good match of the autocorrelation function and the histogram of the composite model to that of the original sequence. This concludes the modeling at the frame level. In the following section we discuss the generation of slices and priority. The sequence used for this purpose is 'Son.mpg' as it has a smaller number of scene changes which can be separated and modeled easily. This sequence is also used in the simulation in Chapter 4. A composite TES model for this sequence is built using the same procedure described for the 'Last Action Hero'. The results of modelling of 'Son.mpg' sequence are presented in Figure 3-16. The

Figure 3-15 Model for the "Last Action Hero" sequence

Figure 3-16 Model for the Son.mpg Sequence

Figure 3-17 Histogram of percentage of bytes per slice in a frame

following section discusses the generation of slices and the contents of a scene.

3.4.1.3 Generation of Slices

After the generation of the frame size sequence using the composite-TES model, the next step is the generation of the slice level information. The slice level information is obtained by distributing the bytes per frame into bytes per slice according to a certain rule. We found that, the plot for of percentage of bytes per slice in a frame, plotted for all slices of the frame, as shown in Figure 3-17, is approximately constant over the duration of a particular scene. We selected two scenes from the sequence 'Son.mpg' and extracted the histograms for the percentage of bytes per slice as shown in Figure 3-17. These two histograms are used to distribute the bytes generated for each frame into slices. Whenever a scene change is decided, the model was made to switch to the other histogram. However, without any loss of generality, other arbitrary histograms could be used to create different types of scenes.

3.4.1.4 Implementing Priority in the model

The next step in the modeling effort is to implement priority in the slices generated, as in Section 3.4.1.3. This was achieved with the help of additional TES models. In Figure 3-8, we saw a trace of the percentage of bytes in HP partition per slice. The histogram of such a

Figure 3-18 Histogram of the percentage of bytes in HP partition per slice in B frames

trace is shown in Figure 3-18 for the B frames of 'Son.mpg'. Making use of this histogram and the sample path of the percentage of bytes in HP partition per slice, a TES model is developed for the sequence of the percentage of bytes in HP partition per slice for B frames. Similarly, other models are developed for the sequence of percentage in HP partition per slice for I and P frames. The output of these models are used to distribute the slices generated in Section 3.4.1.3 from the their respective models, into two partitions, i.e., HP and LP partitions.

This concludes our development of a model for VBR MPEG that implements priority in its output bitstream. The statistics of the bitrate from this model will be used in Chapter 4 where the performance of an ATM multiplexer driven with VBR MPEG traffic is studied.

3.4.2 Gamma-TES Modeling

In this Section we look at the development of a video model for VBR MPEG like video based on certain known distribution for the video frame sequences. The reason for investigating such a modeling technique is to develop a strategy for generating MPEG like video traffic from statistical quantities, such as, the mean and variance of a video sequence, and not based on the empirical sample sequence. The motivation behind this approach is to develop a process that would enable the generation of MPEG like sequence by someone who does not have access to real video or lacks capability to generate real video sequence.

In Section 3.4.1, we saw the development of two models, one based on the 'Last Action Hero' and the other on 'Son.mpg'. The 'Last Action Hero' sequence was matched at the frame level, whereas the 'Son.mpg' was taken further, to generate the slices and incorporate priority in the model output. These models were based upon:

- the empirical distributions of each frame type (Figure 3-12)
- the histogram of the percentage of bytes per slice in a frame (Figure 3-17) and
- the histogram of percentage of bytes in HP partition per slice (Figure 3-18).

In order to proceed with the development our model, the distribution of each frame type has to be decided upon first. The histograms for all frame types of the 'Last Action Hero,' shown in Figure 3-12, have a shape similar to the gamma distribution. To confirm the goodness-of-fit of these histograms with the gamma distribution a *quantile-quantile plot* (Q-Q plot) is used. A Q-Q plot is often used for small sample sizes to get a quick insight into the underlying distribution of the observed samples. It is defined as a graph of the observed quantiles (along Y axis), vs. the theoretical quantiles of the underlying distribution (along X axis) [24, 28]. If the observed samples do come from the given theoretical distribution, then

Figure 3-19 Histogram and Q-Q plot showing the goodness of fit of B frames with gamma distribution (Son.mpg)

the Q-Q plot will be a linear plot.

The matching of the histograms of the three frame types of 'Last Action Hero' with a gamma distribution is shown with the help of the Q-Q plots in Figure 3-20. We notice that all three frame types are approximated well by a gamma distribution. The 'Son.mpg', is a relatively short sequence with a total of only 171 frames as compared with the 'Last Action Hero,' with 1835 frames. Son.mpg has 113 B frames, 43 P frames and 15 I frames. Figure 3-19 shows the matching of the B frames of Son.mpg with gamma distribution. Hence, we assume gamma distribution for all frame types for the purpose of this model. We call this modeling methodology *gamma-TES* because of the gamma approximation for all frame type distributions.

The value of the mean and variance (of the bitrate) for each frame type is taken from the real sequence, 'Son.mpg,' so that the gamma-TES model developed here, can be compared with the empirical-TES model, developed in Section 3.4.1. The *mean bitrates* for a particular frame type refers to the bitrate, when that frame type alone is transmitting at 30 frames per second. In Table 3-4 the mean and variance of the bitrate of each frame type are

Statistical Video Modeling

Figure 3-20 Q-Q plots Showing the goodness of fit with Gamma distribution (Last Action Hero)

given.

Table 3-4 Mean and Variance of Son.mpg sequence

	I Frames	P Frames	B Frames
Mean	6.2 Mbps	2.9 Mbps	0.7 Mbps
Variance	1.55e+12	2.6e+13	2.58e+12

However, if the specific mean bitrates of each frame types are not known then we could assume appropriate values for them using the following procedure: Let μ_I , μ_P and μ_B be the mean bit-rates of I, P and B frames in Mbps, respectively, then the mean bitrate, μ_{ave} of the model under development is given by,

$$\mu_{ave} = \frac{(\mu_I + 3\mu_P + 8\mu_B)}{12} Mbps \quad [3-1]$$

where, we consider one I frame, three P frames and eight B frames in a GOP. At the beginning a value for the average bitrate of the model is chosen along with its burstiness (peak-to-average ratio). The peak-to-average ratio is a function of μ_{ave} and μ_I . In order to vary the peak-to-average ratio of the model, minor changes could be made to μ_I along with compensating changes to μ_P and μ_B in order to keep the mean, μ_{ave} , constant.

After the means for the three frames are finalized, the next step is the generation of gamma random variate for each frame types. This has to be done before using TESTool to induce autocorrelation into the samples as the TESTool does not provide inversion with a gamma distribution by default. The gamma *pdf* (probability distribution function) is given by

$$f(t) = \frac{\lambda (\lambda t)^{s-1}}{\Gamma(s)} e^{-\lambda} \quad t \geq 0 \quad [3-2]$$

where, $\frac{s}{\lambda}$ is the mean and $\frac{s}{\lambda^2}$ is the variance of the bitrate. The variances in our case came from the original test sequence (Table 3-4). However, if the variances are not provided, then they could be chosen arbitrarily to proceed with the modeling.

The output of Equation (3-2) is a stream of uncorrelated iid¹ samples from a gamma

distribution. Correlation is induced into the sample sequence using the TES process, with the help of the TESTool. By giving suitable innovation densities (f_i) in the TESTool window we can induce an exponentially decaying autocorrelation function in an interactive manner as shown in Figure 3-21. Autocorrelation is induced in all gamma samples, similarly.

Once the correlated samples for the three frame types have been generated, the composite-TES modeling methodology of Figure 3-14 is used to interleave the three frames in right order. This model produced a bitrate which is shown in Figure 3-22. The dotted line represents the gamma-TES sequence and the solid line the real sequence. We see that the model based on the principle of approximating the histogram of each frame type with a gamma distribution gives a close resemblance with the real sequence.

The frames generated were distributed into slices using the two histograms of Figure 3-17, that were used during the development of the empirical-TES model. The implementation of priority at the slice level is the next task. Towards this end, we developed a model based on the histogram in Figure 3-18 for the percentage of traffic in HP partition. We noticed that this distribution has a bell shape, with a major spike at 100%. We matched this distribution with a normal distribution using Matlab Toolbox [2] which gave the plot of Figure 3-23 for the B frames. The match of B frames for the entire range is quite satisfactory except near 100% where the values in the sample sequence are truncated as the percentage HP cannot have a percentage greater than 100%. Hence, for the purpose of developing a model for distributing slices into two partitions we made use of a truncated normal distribution. We found that this tail of the normal distribution or 'the overflow' beyond 100%, approximately equals the spike at 100% in Figure 3-18. We assume a normal distribution for all three percentages per slice, used TESTool as before to induce some correlation in the

1. independent and identically distributed

Sample Paths for Uncorrelated Gamma Samples and Correlated Gamma Samples

Histogram for Uncorrelated Gamma Samples and Correlated Gamma Samples

Autocorrelation Function for the Uncorrelated and Correlated Gamma Samples

TES MODEL INNOVATION DENSITY FOR bmat data.e

Figure 3-21 Inducing Autocorrelation into the iid samples from a Gamma distribution.

69

Figure 3-23 Normal Probability Plot for the percentage HP per slice in B frames

output samples, and use the resulting samples for prioritizing the gamma TES model. All three sample sequences are used in simulating an ATM multiplexer in Chapter 4. In this Chapter we have shown that visually, both the empirical-TES and gamma-TES models can be seen to match the original MPEG sequence. In Chapter 4 we will compare the results from these three models to validate them further.

4

Performance of an ATM Multiplexer with Video Frame Spreading

4.1 Introduction

In the previous Chapters, we presented topics related to the development of a modeling strategy for VBR MPEG compressed video. We also showed how closely these models approximated the original video sequences, i.e., with respect to their means, variances and autocorrelation functions. The usefulness of such models is in the flexibility of their application in simulations for network performance evaluation [34]. Some ideas regarding ATM as a future B-ISDN networking standard were discussed in Chapter 2.

In this chapter we will use the models from Chapter 3, to run simulations for the study of cell loss in an ATM multiplexer. A new traffic shaping design is presented for MPEG traffic, which we call *frame spreading*. We will also demonstrate the improvement in the cell loss ratios in an ATM multiplexer by this shaping. This frame spreading buffer is positioned between the ATM multiplexer and the MPEG source. The output of the MPEG source is packetized into cells using one of the AAL protocols presented later in this Chap-

Figure 4-1 Video source feeder

ter. The stream of ATM cells is fed into the frame spreading buffer which feeds it to the multiplexer at a controlled rate. The sequence of these operations are shown in Figure 4-1.

Finally, in Section 4.6, the results of our simulations are used to validate the MPEG source models of Chapter 3.

4.2 ATM Adaptation Layer

The ATM standardization is still an ongoing process, and at the moment there are still some topics that are not fully defined [5]. For example the AAL type 2 protocol, which is defined as the ATM adaptation layer protocol for carrying Class B traffic (see Table 2-1) has yet to be standardized. However some guidelines are available in [5], based on which we define a protocol for our purpose. We will also study the cell loss performance using the AAL type 5 protocol.

In Figure 3-4, a two layered MPEG encoder is shown, which produces two output streams (HP and LP). The traffic from this encoder is buffered until a complete slice is col-

lected, before sending it for transmission. As mentioned earlier, this process is carried out in order to maintain the slice as the minimum recognizable entity in the MPEG stream. The MPEG encoding process provides byte aligned start codes for slices and all the layers above it. This enables them to be observable in the compressed bitstream without first decoding them. At the layers below, macroblocks and blocks are only visible after they are decoded. If we wish to make the data within a slice readable by other programs, other than the decoder, then additional markers have to be inserted into the bitstream so that macroblocks and blocks are visible. These extra markers would increase the priority overhead even more. For this reason we designated the slice as the smallest data unit that is passed from the higher layers to the ATM adaptation layer. The structure of the proposed AAL type 2 protocol is given below.

4.2.1 AAL type 2

The MPEG slice, i.e., the MPEG_PDU, is passed to the AAL layer from the encoder (higher layer). The MPEG_PDU is passed through the convergence sub-layer first which is responsible for the enhancement of services provided by the ATM layer, by providing services such as timing information, CRC etc.

Each frame header (picture header) in the MPEG bitstream contains timing fields that encode time stamps for the presentation of that frame at the decoder. Therefore we assume the CS does not have to provide any timing information for its SDUs, i.e., slices. Each slice is encoded independently from other slices in the frame. Each slice header contains the address of the location where that slice is located in the frame. The slice headers provide serial numbers for the slices. This feature satisfies the necessity to provide explicit serial numbers at the CS_PDU level. We also do not consider providing any CRC at the CS layer, as the main cause of errors that we may encounter during transmission would be

Figure 4-2 From MPEG slice to ATM payload

losses due to congestion, in which case a CRC will not help to reconstruct a cell which is discarded due to congestion. A cell loss will always cause the data to be lost until the end of the slice and the next slice will be decoded at the right location as we saw in Figure 3-2. Therefore, in our AAL protocol we do not consider any CS layer overhead at all.

The layer below CS is the SAR sub-layer. The SAR sub-layer is responsible for the formation of cell payload. The sequence of operations towards the generation of an ATM cell is shown in Figure 4-2. The MPEG_PDU is passed to the CS sub-layer, which forwards it, as is, to the SAR sub-layer. The SAR sub-layer accepts the CS_PDU and chops it down into SAR_SDUs, and adds the information required for the headers and the trailers. The proposed structure of the SAR_PDU is shown in Figure 4-3. The headers and trailers in the SAR_PDU are now discussed in further detail:

- **Sr. # (Serial number):** This is an 8 bit field that is used for detecting cell losses during transmission. It is capable of detecting 256 consecutive cell losses. It also performs the

Figure 4-3 ATM cell payload

task of detecting misinserted cells that may arrive due to undetected bit errors in the ATM cell header.

- **B/E Tags (Begin / End tags):** This is a two bit code that is used to inform the receiving node about the beginning of a particular slice, its continuation and its end. If the contents of a slice can fit into one SAR_PDU then such cells have a separate code. Table 4-1 illustrates the codes used in the B/E tag fields. This code can be used to iden-

Table 4-1 Begin / End tags in the SAR

01	Beginning of Slice
00	Continuation of Slice
10	End of Slice
11	One Slice in payload

tify each slice by the receiving node and to forward it to the decoder appropriately.

- **LI (Length indicator):** The cell payload may not always be completely filled, when it is only partly filled, this field helps to inform the receiving node about the useful contents in the payload. It includes 6 bits which can successfully encode a length of 64 octets. However as the ATM payload (ATM_SDU) is 48 octets in size, 6 bits are sufficient.
- **CRC (Cyclic Redundancy Check):** This is a 10 bit codeword that is used to protect the SAR_PDU from bit errors. The CRC calculation is performed on the SAR_PDU headers, SAR_SDU payload field and the LI field, and it is capable of detecting two correlated bits in error. This code may be calculated using the polynomial [14]:

$$G(x) = x^{10} + x^9 + x^5 + x^4 + x + 1 \quad [4-1]$$

- **Res (Reserved):** This 6 bit field is unused, hence reserved.
- **PAD (Padding bits):** This field is used to fill the cell payload and make it 48 octets in size when there are unused octets.

It must be mentioned that the size of the SAR_SDU must be an integral multiple of octets because the LI field is only capable of representing integral number of octets in the SAR_PDU; also, the PAD inserts an integral number of octets in order to make the size of the SAR_PDU equal to 48 octets. In fact no further adaptation is required to achieve this as the slices that are passed down to the AAL from the encoder are already an integral multiple of octets due to the byte aligned nature of all headers.

4.2.2 AAL type 5

The AAL type 5 protocol is aimed at improving the efficiency of the adaptation layer. As we saw, the SAR_PDU was left with only 44 bytes of useful payload while using AAL type 2 protocol. AAL type 5 proposes to append one 32 bit CRC at the CS layer, instead of protecting each SAR_PDU with a CRC in every cell. This is intended to provide a better error protection for the user data [25] and does not use any space from the SAR_PDU. This way the complete payload of 48 octets is available in the SAR sub-layer.

The AAL5 CS_PDU format is shown in Figure 4-4. Each field in the CS_PDU is described below:

- **User Data:** This is the data from the higher layers (CS_SDU). AAL5 specifies that this data be not more than 65,536 octets in length. In our case the user data is an MPEG slice. If the slice happens to be bigger than 65,536 octets then it has to be split into two CS_SDUs. Each CS_SDU is then appended with a CS trailer and padded accordingly.

Figure 4-4 AAL type 5 CS_PDU

- **PAD:** This is used to align the entire CS_PDU to a 48 octet boundary.
- **Control field:** This field is provided for future study; two zero filled octets must be present if not used.
- **LI (Length indicator):** This field is used for representing the actual length of user data in the CS_PDU. The user data is limited to 65,536 octets because the LI uses 2 octets (16 bits) to represent information of a maximum length of 2^{16} (= 65,536) octets.
- **CRC-32:** This cyclic redundancy code provides bit error detection for the CS_PDU which, in most cases, is a complete slice.

There is no SAR header or trailer for AAL5, unlike other AAL types. The SAR_PDU contains a 48 octet segment of the CS_PDU as shown in Figure 4-5. The last cell in this case does not have to be padded as the CS_PDU is already made up of an integral multiple of 48 octets. As the SAR_PDU overhead is not present, the maximum transmission efficiency that AAL5 can attain is about 87.2%, whereas the maximum transmission efficiency that can be reached using AAL2 is only about 80% [25].

Figure 4-5 Formation of ATM cells using AAL5

4.3 Frame Spreading

In this Section we discuss a transmission strategy for cable head-end applications. This is important, whenever an application requires sending isochronous traffic without stringent requirements on the end-to-end delay.

Several previous authors have assumed that the variations in the bitrate within a frame are smoothed by using pre-buffers at the output of the codecs to give a constant value over a frame period [34,17, 42]. We carry this a little further and define a technique called *frame spreading*. The function of the pre-buffer is modified to implement the frame spreading feature, and this buffer will now be called the *frame spreading buffer*. Let us also consider that the frame generation rate is equal to 30 frames per second; frame period is equal to $1/30$ th of second.

The concept of frame spreading is very simple and is useful when the end-to-end delay is not critical, as we assume, is the case at a cable head-end in commercial TV broadcast. Frame spreading preserves the advantages of using VBR video for statistical multiplex-

Frame Spreading

Figure 4-6 Frame Spreading

ing in ATM networks, but at the same time reduces the variability in the video data (i.e., reduces burstiness).

Figure 4-6 illustrates frame spreading. Instead of transmitting each frame strictly in one frame period, the frames are now transmitted at an average transmission rate (R_n) of more than one frame. Figure 4-6 shows the case where the average transmission rate is calculated for three consecutive frame. The average transmission rate of a frame, T_n is calculated at the beginning of each frame when that frame becomes available at the output of the codec as:

$$T_n = \frac{G_n}{Fs \times Fp} \quad [4-2]$$

where;

T_n = Transmission rate of the n th frame in bits/second.

G_n = Size of the n th frame in bits/frame.

Fs = Frame spreading factor (i.e., 1, 2, 3 ...).

2 of/de 2

PM-1 3 1/2" x 4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

Priority handling

Fp = Frame period ($1/30$ th of a second)

The total transmission rate, R_n , is calculated using Equation (4-3). The frame is then transmitted to the ATM multiplexer, from the frame spreading buffer, at R_n bits per second for the next frame period:

$$R_n = T_{n-2} + T_{n-1} + T_n \quad [4-3]$$

R_n is calculated at the beginning of each frame period when a new frame arrives. The transmission rate (R_n) is then maintained constant until the end of the frame period. The reduction in the peak-to-average ratio in the real video sequences, as they come out of the frame spreading buffer is shown in Table 4-2, as the frame spreading factor is increased.

Table 4-2 The Peak-to-average ratio for various frame spreading

Frame Spreading	1	2	3	4	5
Peak-to-ave. ratio	3.95	2.2	2.0	1.77	1.73

In Section 4.5 we will set up an experiment with an ATM multiplexer that uses frame spreading at its input.

4.4 Priority handling

The MPEG sources we discussed earlier in Chapter 3, implement two levels of priority in their output traffic to help guarantee a desired QoS¹. Therefore a priority handling scheme for the source traffic has to be implemented in the multiplexer buffer. In the broad area of selective discarding algorithms for handling two classes of traffic, there are two well known schemes [19] that we consider for handling priority as explained below.

1. Quality of Service.

* if there is an LP in common

Figure 4-7 Partial buffer sharing

4.4.1 Partial buffer sharing

This scheme is equivalent to splitting the multiplexer buffer into two cascaded virtual buffers, called the *reserved* and the *common* buffers, as shown in Figure 4-7. The *common* buffer may have cells of both priority levels in it, whereas the *reserved* buffer is allowed to have only HP cells. Depending on the status of the *reserved* and the *common* buffer, we may have the following cases:

Case 1. *Reserved* is empty and *common* is not full: If a cell arrives (either HP or LP), then it is stored directly in the *common* buffer.

Case 2. *Reserved* is not full and *common* is full: A arriving HP cell is accepted in the *reserved* buffer but an arriving LP is always discarded.

Case 3. *Reserved* is full and *common* is full: All cells are discarded.

This scheme is useful due to its simplicity of implementation. However it results in a low performance for the LP traffic, i.e., if the *common* is full and there are no HP cells in the *reserved* buffer, then the LP still is not able to make use of the reserved buffer [19]. There is also the problem of finding an optimum splitting factor for the buffer for a given type of traffic.

ATM multiplexer

* only if there are no LP cells in the buffer

Figure 4-8 Push-out scheme

4.4.2 Push-out scheme.

This scheme allows the complete usage of the buffer by both priority classes. However, when the buffer is full, then an arriving HP will be accepted as long as there is at least one LP cell in the buffer. The LP cells are always pushed out on a LIFO¹ basis. This scheme provides an optimum performance from a priority point of view. However, the search algorithms can be quite complicated. Figure 4-8 illustrates this scheme.

The push-out scheme was assumed in this work.

4.5 ATM multiplexer

In order to study the performance of MPEG video traffic in ATM networks a simple multiplexing model, such as the one shown in Figure 4-9, is generally used. In this model, N independent VBR MPEG sources are passed through N independent frame spreading buffers. For a frame spreading factor of one, this model is similar to the models described in [9,15, 32, 42].

The output streams from the frame spreading buffers are multiplexed into an ATM

¹ last-in-first-out

Figure 4-9 Statistical multiplexing model

multiplexer in an asynchronous manner (statistical multiplexing). The model of the ATM multiplexer in Figure 4-9 is a first order approximation of an ATM node, i.e., a *single queue single server* system with a finite capacity buffer. The cells from each source accumulate in the common buffer and are served in a FIFO¹ order.

The average bit-rate of the MPEG sources we consider is approximately 1.85 Mbps. This leads us to consider a transmission link speed of 45 Mbps. It seems likely that similar results would be obtained for typical MPEG-2 like traffic (~6 Mbps), on a 150 Mbps link. A buffer capacity of 525 cells was chosen, such that it offers a maximum delay of no more than 5 milliseconds [26]. The transmission time, i.e., service time, of an ATM cell is fixed at 9.42 μ s for a 45 Mbps link.

All N sources in our experiment are similar and have the same mean, variance and autocorrelation function. This is due to the fact that we used the same sample sequence for each one of the input sources. However, as each source becomes active, an arbitrary starting

1. first-in-first-out

Simulation results

point within the sequence is chosen in order to begin transmission [16]. The point in time that each source becomes active was also random within a frame period. We ran each experiment with a fixed value of the frame spreading factor and a fixed number of sources. The number of sources is increased by one and the experiment is repeated. Similarly the frame spreading factor is also increased in steps. In order to have an idea of the performance improvement by implementing the technique of frame spreading we forced the initial conditions be the same for each experimental run. Therefore, the value of the starting point in the video sample sequence and the transmission starting time for each source was saved and used again in the subsequent runs for their respective sources.

4.6 Simulation results

4.6.1 Utilization and Total cell loss

'N' MPEG sources are encoded at a PBP of 12, 16, 20, 24 and 64, and fed to an ATM multiplexer. The frame spreading factor is fixed at one initially and the number of sources are increased from 12 to 25. Using the AAL 2, the utilization of the multiplexer is shown in Figure 4-11. The interesting thing to note is that the utilization of the multiplexer is about 3~4% higher when priority is implemented in the MPEG traffic. This is related to the overhead associated with implementing priority, which was shown in Table 3-2 to be about 3~5%. The higher utilization of the multiplexer has the effect of producing higher cell loss ratios, which is evident in Figure 4-11 for the same experiment using AAL 2.

In Figure 3-2, the effect of having lost a few cells left a hole in the frame (lost image information). If this hole were to appear in an I or a P frame it would have propagated until the arrival of the next I frame. Priority was implemented so that such a situation does not

Simulation results

Figure 4-10 Multiplexer utilization at various PBPs (using AAL 2)

Figure 4-11 Total cell loss ratio at various PBPs (using AAL 2)

Figure 4-12 Total cell loss ratio at various PBPs (using AAL 5)

arise and the viewer always has certain minimum acceptable image quality at all times.

The AAL 5 provides a higher payload capacity (48 octets) due to the absence of the SAR_PDU overhead (headers and trailers). Subsequently, this leads to a higher transmission efficiency for the video traffic. The results of the experiment using AAL 5 are shown in Figure 4-12 and Figure 4-13 for the utilization and total cell losses, respectively. However, the general trend, in the results using AAL 2 and AAL 5, in terms of utilization and total cell loss ratios is quite similar except for a slight shift of the curves in favor of AAL5. The utilization curves are about 5% lower for AAL 5 for almost the whole range of the curves below 95%. Similarly the total cell loss ratio also reflects a better performance (about 3~5% lower) when AAL 5 is used.

Figure 4-13 Total cell loss ratio at various PBPs (using AAL 5)

4.6.2 Applying frame spreading

In this section we will discuss the results of the experiment on frame spreading. The value of frame spreading factor is increased from 1 to 5 and for each value of frame spreading factor we increase the number of sources stepwise from 12 to 25. The initial conditions for each experimental run are fixed. Therefore, if there is an improvement in any of the performance curves, after increasing the frame spreading factor, is the result of frame spreading. The first set of experiments is performed using AAL 2, and subsequently repeated using AAL 5.

In Figure 4-14 the total cell loss is shown. For a frame spreading factor of one, we see that the cell loss becomes noticeable when the number of sources reaches 14. As the frame spreading factor is increased, there is an improvement in the total cell loss ratio. This is the result of reduced burstiness of the input bitrate, as was shown in Table 3-3. However

Figure 4-14 Total cell loss ratio at various frame spreading factors (AAI, 2)

this improvement lasts only till the number of input sources reaches 21. With 21 sources, the bitrate of the sum of all input sources is well over the total transmission capacity (45 Mbps) of the output link and the link utilization is about 99%. As the sources increase beyond 21, there is a linear increase in the cell loss ratio, proportional to the increased number of sources, and it is same for all frame spreading factors.

In Figure 4-15 we see a significant improvement in the HP cell loss ratio as frame spreading is applied. This improvement in cell loss ratio goes beyond the 21 sources (and 99% utilization), unlike the behavior of total loss in Figure 4-14. Therefore frame spreading helps give a better performance to HP traffic, even at high multiplexer utilizations. However the cell loss ratios in this case are actually percentages of the total HP traffic. The percentage of loss is lower when expressed as the percentage of the total traffic lost.

Figure 4-16 presents percentage of LP cell loss, with respect to the total LP traffic.

Simulation results

Figure 4-15 HP cell loss ratio for various frame spreading factors (AAL 2)

Figure 4-16 LP cell loss ratio for various frame spreading factors (AAL 2)

Figure 4-17 Total cell loss ratio at various frame spreading factors (AAL 5)

We notice a crossover point between 19 and 20 sources, in the curves for different frame spreading factors. This crossover indicates the point where the sum of the input traffic bitrate equals the transmission capacity. Beyond this point any cell loss improvement in the HP or total traffic comes at the expense of higher LP loss. This is an advantage of implementing priority, as some traffic is tagged as LP in order to give a better performance to the HP traffic at the expense of LP traffic.

The above experiment is now repeated using the AAL 5 as the adaptation layer protocol. The results of the simulation run are presented in Figure 4-17, Figure 4-18 and Figure 4-19 for the total, HP and LP cell loss ratios. As mentioned earlier in Section 4.6.1, we notice that these figures show a similar trend with experiments using AAL 2. All three Figures show an improvement in their respective cell losses. This is due to lower overhead of the AAL 5 protocol. The crossover point that was mentioned earlier regarding the LP traf-

Figure 4-18 HP cell loss ratio for various frame spreading factors (AAL 5)

Figure 4-19 LP cell loss ratio for various frame spreading factors (AAL 5)

Figure 4-20 Number of Multiplexed Sources for a given Frame Spreading

fic, is now found to have shifted to the interval between 20 and 21 sources in Figure 4-19. This means that we can multiplex one more source using AAL 5, before the sum of all input sources equals the output link capacity.

For an HP cell loss of ratio of less than 10^{-1} , Figure 4-20 shows the total number of multiplexed sources as the frame spreading factor is increased from 1 to 5 for both adaptation layer protocols. We see a definite advantage in using frame spreading in both cases.

4.6.3 Results using MPEG source models

The results we presented until now were all obtained with simulations carried out using the real video sequence 'Son.mpg'. In Chapter 3, we developed two models for our real video sequence that produced statistics of a two layered MPEG video source. We use those two models to repeat the simulations of Section 4.6.2. The underlying assumptions were the

Figure 4-21 Comparison of total cell loss ratios

same for all the simulation runs as before. The results obtained are presented for a frame spreading factor of one. In Figure 4-21, the total cell loss ratio is shown for the three experimental runs, i.e., one with the real sequence, one with the empirical-TES model (Section 3.4.1) and one with the gamma-TES model (Section 3.4.2). The three curves are found to be in fairly close agreement.

A comparison of the HP cell losses from the three sources (real and modeled) is given in Figure 4-22. We find that typically, the HP cell loss from the empirical TES follows the cell loss using the real sequence. However after the number of sources reaches, 21, the real source shows a lower cell loss than the loss shown either model. This behavior of the real sequence may be contributed to the fact that the real sequence was of a short duration (170 frames). However, the gamma TES follows the empirical TES closely, though always underestimating the cell loss ratio. This may be due to the smooth approximation of the

Simulation results

Figure 4-22 Comparison of the HP cell loss ratio

Figure 4-23 Comparison of the LP cell loss ratio

Simulation results

gamma distribution, whereas the real sequence did not exhibit such smoothness. Longer sequences of the Last Action Hero did show better matching with the gamma distribution, though. As the gamma TES model is built using statistical values chosen by the model developer, the model can be changed by varying those values to match the empirical model even closer.

The Figure 4-23 shows the LP cell loss comparison. As both the models and the empirical data had the same mean values, the LP loss curves for the three traffic sources show a common crossover point at 19 sources. The gamma TES model underestimates the empirical sample path slightly, as it had slightly more LP cells than the real sequence. The percentages in the Figure are the ratio of the actual cell loss with respect to the amount of LP traffic.

From the above we could say that the two models could be used to simulate MPEG-like traffic, replacing lengthy sequences that may require large storage devices, and a lot of computing power, in order to collect the video sequence.

5

Conclusions and Future Research

5.1 Conclusions

In this thesis we have addressed two related aspects of MPEG video transmission in B-ISDN networks. One deals with the development of statistical models for MPEG video sources and the other with a video transmission strategy in ATM networks. B-ISDN services are expected to be dominated by applications using video in one form or the other. In order to design networks that will be able to transfer a variety video traffics, it is essential to develop flexible and versatile video models that can replicate real video characteristics.

There have been many studies on the modeling of video traffic based upon analytical and statistical data. Transform Expand Sample (TES) is one promising statistical modelling strategy which is used very successfully in modeling video traffic. It captures the marginal and autocorrelation of the sample sequence in two unique steps and guarantees the matching of the histogram by default. The autocorrelation can be matched using a visual tool called TESool.

The effect of cell loss in an MPEG frame was shown in Figure 3-2. Priority was implemented in the MPEG stream in order to prevent such visible data loss in a video frame

Conclusions

and to maintain a minimum acceptable image quality at the receiving end at all times. The video was prioritized according to MPEG-2 specifications. The advantages of choosing this priority scheme is discussed in Section 3.3.2. The traffic characteristics of the two layered MPEG were studied and it was shown that there is a necessity for implementing dissimilar PBPs for all the three frame types used in MPEG in order to make sure that there is a proportional distribution of traffic in both the partitions for all three frame types. The percentage of traffic in HP partition was shown in Figure 3-7 where the PBP was fixed at 12 and the same sequence was encoded with dissimilar PBP and the results were shown in Figure 3-8. The PBP value for the B frames was chosen in Figure 3-8 by repeatedly plotting the percentage HP and selecting a value when the percentage HP sequence showed the least variation.

The I frames must have the best image quality of the three frame as they are a anchor frames of the sequence. Any errors in the sequence can propagate only until the arrival of an I frame, when the frame is effectively refreshed. It was decided by experiments that discarded the LP completely (Figure 3-9) to represent the worst case in transmission, that the least acceptable value for the PBP can be as low as 12. A PBP value of 16 can give quite a good picture quality even with all LP traffic lost.

MPEG traffic is quite complex due to its three frame types, hence, it is virtually impossible to match the complete MPEG sequence with a single TES model. Therefore, we had to separate the three frame types and model each empirical frame sequence separately. Finally, the three models were interleaved in the proper order to represent a typical MPEG sequence. The cross correlation between the different frames was captured by copying a seed from the background TES sequence of the I frame into the first P and B frame that follow an I frame. In this way the I frames become the frames that drive the bitrates of the P and B frames.

Conclusions

Scene changes in the MPEG sequence were implemented using a geometric random variable at the GOP level and a uniform random variable at the frame level within a GOP. The final results of the above model matches quite well with the original MPEG sequence. It was observed that the histogram of the percentage of bits per slice was approximately constant for the duration of a particular scene. Therefore, such a histogram was used to distribute the frame contents into slices.

None of the statistical models in the literature discuss the implementation of priority in a video sequence. We used a TES model of the percentage HP per slice to implement priority into the slices generated in the previous step. This was called the empirical-TES model. The simulation results have shown the model to give cell loss ratios similar to the original traffic.

Finally, in our work regarding modeling video traffics, we developed a video source model purely from statistical quantities. The statistical quantities, i.e., the mean and the variance were assumed to be the same as that of the empirical sequence in order to facilitate comparison of the analytically developed model with the empirical sequence. The analytical model used a gamma approximation for the three sample sequences of the video frame. An exponentially decaying, arbitrary autocorrelation was induced in the analytically generated frames. It was shown that the final model matched the empirical sequence quite well. This model was called the gamma-TES model. The frames were distributed into slices using the same histogram as used by the empirical-TES model. Priority was implemented in the gamma-TES model using a normal distribution which is truncated at 1.

An experimental model with a single ATM multiplexer was set up which was fed with similar MPEG sources. Priority handling in the ATM multiplexer was assumed to be the push-out scheme.

Future Research

A traffic shaping strategy called frame spreading was introduced, which was inserted between the sources and the multiplexer. There was an improvement in the burstiness of the MPEG traffic as shown in Table 4-2. This strategy is useful for cases where the end-to-end delay is not very critical, such as distributional video. As the frame spreading factor is incremented from one to five it was shown that the burstiness of the MPEG video reduces and more sources can be multiplexed for the given cell loss ratio.

The frame spreading strategy was implemented in simulation with real video sources as well as empirical-TES and the gamma-TES video models, and the results show a good resemblance of the models with the real video sequence.

The adaptation layer protocol used in our simulations was presented. It was based upon the guidelines set for AAL2 [5]. Initial simulation were also carried out using AAL5 and the results show that the AAL5 has a lower cell loss ratio and the multiplexer was able to service more sources using AAL5 than with AAL2 for the same multiplexer utilization and cell loss ratios.

All the simulation results from the three sources (empirical and the two models) showed a close match for all the cell loss ratios that were investigated.

5.2 Future Research

This thesis dealt with the implementation of different PBPs for all frame types in MPEG video. However, the PBP was kept constant for the duration of encoding of the sequence. As an extension the PBP could be varied more dynamically so that the HP loss is kept as low as possible based on some estimate of the frame traffic and the buffer occupancy of the ATM multiplexer. Based on the same estimate, the delay induced in implementing frame spread-

Future Research

ing can be reduced by one frame period if the frame estimation is reliable, as the buffer will not have to wait for the complete frame to be received before calculating the transmission rate for that frame period.

Rate control based on feedback from the frame spreading buffer can be investigated. As the traffic is smoothed using frame spreading, the rate control measures may not have to be as drastic as would be necessary in the light of no frame spreading.

MPEG 2 offers three more scalabilities, i.e., SNR scalability, spatial scalability and temporal scalability which could be investigated for transmission in fast packet networks.

List of References

- [1] K. Anand, A. R. Kaye, T. A. Gulliver and S. A. Mahmoud, "FEC and Priority for VBR Video Distribution over ATM," *Can. J. Elect. and Comp. Eng.*, Vol. 19, No. 3, pp. 123-130, 1994.
- [2] B. Jones, "Statistics TOOLBOX—For use with MATLAB," Jan 1994.
- [3] CCIR: Recommendation 601-1, "Encoding Parameters of Digital Television for Studios," Geneva, 1986.
- [4] CCITT: Recommendation I.121, "Broadband Aspect of ISDN," Geneva, 1991.
- [5] CCITT: Recommendation I.362, "B-ISDN ATM Adaptation Layer (AAL) Functional Description," Geneva, 1991.
- [6] CCITT: Recommendation I.363, "B-ISDN ATM Adaptation Layer (AAL) Specification," Geneva, 1991.
- [7] CCITT: Recommendation H. 261, "Video Codec for Audio Visual Services at p x 64 kbits/s," Geneva, 1990.
- [8] P. F. Chimento Jr., "A Review of Video Sources in ATM Networks,"
- [9] D. M. Cohen and D. P. Heyman, "A Simulated Study of Video Teleconferencing Traffic in ATM Networks," in *Proc. IEEE INFOCOM '93*, San Francisco, CA, 1993, pp. 894-901.
- [10] V. S. Frost and B. Melamed, "Overview of Simulation and Traffic Modeling for Telecommunications Networks," *IEEE Commun. Magazine*, Vol. 32, No. 3, Mar. 1994.
- [11] M. Ghanbari, "Two-Layer Coding of Video Signals for VBR Networks," *IEEE J. Sel. Areas Commun.*, Vol. 7, No. 5, Jun. 1989.
- [12] S. J. Golestani, "A Framing Strategy for Congestion Management," *IEEE J. Sel. Areas Commun.*, Vol. 9, Sept. 1991, pp. 1064-1077.
- [13] R. Grunenfelder, J. P. Cosmas, S. Manthorpe and A. Odinme-Okafor, "Characterization of Video Codecs as Autoregressive Moving Average Processes and Related Queueing Systems Performance," *IEEE J. Sel. Areas Commun.*, Vol. 9, No. 3, Apr. 1991, pp. 284-292.
- [14] R. Händel and M. N. Huber, "Integrated Broadband Networks—An introduction to ATM-Based Networks," *Addison-Wesley Publishing Co.*, 1991.

- [15] D. P. Heyman, A. Tabatabai, and T. V. Lakshman, "Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks," *IEEE Trans. Circuits Syst. for Video Tech.*, Vol. 2, No. 1, Mar. 1992, pp. 49–59.
- [16] D. Heyman, T. V. Lakshman, A. Tabatabai and H. Heeke, "Modeling Teleconference Traffic from VBR Video Coders," in *Conf. Rec. ICC '94*, New Orleans, Louisiana, 1994, pp. 1744–1748.
- [17] D. P. Heyman and T. V. Lakshman, "Source Models for VBR Broadcast Video Traffic," in *Proc. IEEE INFOCOM '94*, Toronto, Ontario, 1994, pp. 664–671.
- [18] J. R. Hill, B. Melamed and C. J. Walsh, "User Guide and Reference Manual for TEstool (Version 1.0)," *NEC USA, Inc.*, Sep. 1993.
- [19] G. Hebuterne and A. Gravey, "A Space Priority Queuing Mechanism for Multiplexing ATM Channels," *Comp. Networks and ISDN systems*, Vol. 20, 1990, pp. 37–43.
- [20] ISO/IEC/JTC1/SC29, "Coding Representation of Pictures, Audio and Multimedia/Hypermedia Information," *Committee Draft ISO/IEC CD 11172*, Japan, Dec. 1991.
- [21] ISO/IEC/JTC1/SC29, "Generic Coding of Moving Pictures and Associated Audio," (*Recommendation H.262*) *Committee Draft ISO/IEC CD 13818*, Seoul, Nov. 1993.
- [22] M. R. Ismail, I. E. Lambadaris, M. Devetsikiotis and A. R. Kaye, "Frame Spreading and a Modified Priority Scheme for VBR MPEG Video Transmission in ATM Networks," in *Proc. Can. Conf. on Elect. and Comp. Eng.*, Halifax, Nova Scotia, Sep. 1994, pp. 526–530.
- [23] M. R. Ismail, I. E. Lambadaris, M. Devetsikiotis and A. R. Kaye, "Modelling Prioritized MPEG Video Using TES and a Frame Spreading Strategy for Transmission in ATM Networks," in *Proc. INFOCOM '95*, Boston, Mass., 1995, to appear.
- [24] R. Jain, "The Art of Computer Systems Performance Analysis," *John Wiley & Sons, Inc.*, 1991.
- [25] G. C. Kessler, "ISDN—Concepts, Facilities, and Services," 2nd ed. *McGraw-Hill Series on Comp. Commun.* 1993.
- [26] H. Kanakia, P. Mishra, A. Reibman, "An Adaptive Congestion Control Scheme for Real-Time Packet Video Transport," in *Proc. ACM SIGCOMM '93*, Ithaca, NY, Sep. 1993, pp. 20–31.
- [27] R. C. Lau, P. E. Fleicher, S. Lei, "Receiver Buffer Control for Variable Bit-Rate Real-Time Video," in *Conf. Rec. ICC '92*, Chicago, MI, 1992, pp. 0544–0550.
- [28] A. M. Law and W. D. Kelton, "Simulation Modeling and Analysis," *McGraw-Hill, Inc.*, 1991.

- [29] D. Le Gall, "MPEG: A Video Compression Standard for Multimedia Applications," *Commun. ACM*, Vol. 34, No. 4, Apr. 1991, pp. 47-58.
- [30] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson and J. D. Robbins, "Performance Models of Statistical Multiplexing in Packet Video Communications," *IEEE Trans. Commun.*, Vol. 36, No. 7, Jul. 1988, pp. 834-843.
- [31] B. Melamed and B. Sengupta, "TES Modeling of Video Traffic," *IEICE Trans. Commun.*, Vol. E75-B, No. 12, Dec. 1992, pp. 1292-1300.
- [32] B. Melamed, D. Raychaudhuri, B. Sengupta and J. Zdepski, "TES-Based Modeling for Performance Evaluation of Integrated Networks," in *Proc. INFOCOM '92*, Florence, Italy, pp. 75-84.
- [33] M. Nomura, T. Fujii and N. Ohta, "Basic Characteristics of Variable Rate Video Coding in ATM Environment," *IEEE J. Sel. Areas Commun.* Vol. 7, No. 5, Jun. 1989, pp. 752-760.
- [34] N. Ohta, "Packet Video - Modeling and Signal Processing," *Artech House, Inc.* 1994.
- [35] P. Pancha, M. El Zarki, "A look at the MPEG Video Coding Standard for Variable Bit Rate Video Transmission," in *Proc. IEEE INFOCOM '92*, Florence, Italy, pp. 85-94.
- [36] P. Pancha and M. El Zarki, "Prioritized Transmission of VBR MPEG Video," in *Proc. IEEE GLOBECOM '92*, Orlando, Florida, Dec. 1992, pp 1135-1139.
- [37] P. Pancha and M. El Zarki, "Bandwidth Allocation Schemes for Variable Bit Rate MPEG Sources in ATM Networks," *IEEE Trans. Circuits Syst. for Video Tech.*, Vol. 3, No. 3, Jun. 1993, pp. 190-198.
- [38] W. B. Pennebaker and J. L. Mitchell, "JPEG: Still Image Data Compression Standard," *Van Nostrand Reinhold*, New York, 1993.
- [39] G. Ramamurthy, B. Sengupta, "Modeling and Analysis of a Variable Bit Rate video Multiplexer," *7th Specialist Seminar, Int. Teletraffic Conf.*, Morristown, NJ, Oct. 1990.
- [40] D. Raychaudhuri, R. Siracusa, J. Zdepski and K. Joseph, "A Flexible and Robust Packet Transport Protocol for Digital HDTV," in *Proc. GLOBECOM '92*, Orlando, Florida, Dec. 1992, pp. 1374-1380.
- [41] D. Reininger, D. Raychaudhuri, B. Melamed, B. Sengupta and J. Hill. "Statistical Multiplexing of VBR MPEG Compressed Video on ATM Networks," in *Proc. INFOCOM '93*, San Francisco, Mar 1993. pp. 919-926.
- [42] D. Reininger, B. Melamed and D. Raychaudhuri, "Variable Bit Rate MPEG Video: Characteristics, Modeling and Multiplexing," in *Proc. of ITC '94*, Antibes, France, 1994.

- [43] D.Reininger and D. Rahchaudhuri, "Bit-Rate Characteristics of a VBR MPEG Video Encoder for ATM Networks," in *Conf. Rec. ICC '93*, pp. 517–521.
- [44] H. Saito, "Teletraffic Technologies in ATM Networks," *Artech House, Inc.* 1994.
- [45] R. J. Siracusa, K. Joseph, J. Zdepski and D. Raychaudhuri, "Flexible and Robust Packet Transport for Digital HDTV," *IEEE J. Sel. Areas Commun.*, Vol. 11, No. 1, Jan. 1993, pp. 88–98.
- [46] P. Sen, B. Maglaris, N. Rikli and D. Anastassiou, "Models for Packet Switching of Variable Bit Rate Video Sources," *IEEE J. Sel. Areas Commun.*, Vol. 7, No. 5, Jun. 1989, pp. 865–869.
- [47] P. Skelly, M. Schwartz and S. Dixit, "A Histogram-Based Model for Video Traffic Behavior in an ATM Multiplexer," *IEEE/ACM Trans. Networking*, Vol. 1, No. 4, Aug. 1993, pp. 446–459.
- [48] W. Stalling, "Advances in ISDN and Broadband ISDN," *IEEE Computer Society Press*, May 1992.
- [49] T. Tanaka, S. Okubo, H. Hashimoto and H. Yasuda, "A Study of Comparison Between VBR and CBR Video Service in ATM Environment," in *Conf. Rec. ICC '92*, Chicago, MI, pp. 0551–0555
- [50] W. Verbiest and L. Pinnoo, "A Variable Bit Rate Video Codec for Asynchronous Transfer Mode Networks," *IEEE J. Sel. Areas Commun.*, Vol. 7, No. 5, Jun. 1989.
- [51] F. Yegenoglu, B. Jabbari, Y. Zhang, "Motion-Classified Autoregressive Modeling of Variable Bit Rate Video", *IEEE Trans. on Circuits Syst. for Video Tech.* Vol. 3, No. 1, Feb. 1993, pp. 42–53.

END

0/4-06-9/6

FIN