

Multimedia Communications

Directions &
Innovations



Jerry D. Gibson, Editor

Multimedia Communications

**ACADEMIC PRESS SERIES IN COMMUNICATIONS,
NETWORKING, AND MULTIMEDIA**

Editor-in-Chief
Jerry D. Gibson
Southern Methodist University

This series has been established to bring together a variety of publications that cover the latest in applications and cutting-edge research in the fields of communications and networking. The series will include professional handbooks, technical books on communications systems and networks, books on communications and network standards, research books for engineers, and tutorial treatments of technical topics for non-engineers and managers in the worldwide communications industry. The series is intended to bring the latest in communications, networking, and multimedia to the widest possible audience.

Books in the Series:

Handbook of Image and Video Processing, Al Bovik, editor
Nonlinear Image Processing, Sanjit Mitra, Giovanni Sicuranza, editors
The E-Commerce Book, Second Edition, Steffano Korper and Juanita Ellis

Multimedia Communications

Directions and Innovations

JERRY D. GIBSON, EDITOR

*Department of Electrical Engineering
Southern Methodist University
Dallas, Texas*



ACADEMIC PRESS

A Harcourt Science and Technology Company

SAN DIEGO / SAN FRANCISCO / NEW YORK / BOSTON / LONDON / SYDNEY / TOKYO

This book is printed on acid-free paper.

Copyright © 2001 by Academic Press

The chapter "Multimedia Conferencing Standards" by David Lindbergh is reprinted from *Digital Compression for Multimedia: Principles and Standards* by Jerry Gibson, Toby Berger, Tom Lookabaugh, and David Lindbergh. Copyright © 1998 by Morgan Kaufmann Publishers, San Francisco. <http://www.mkp.com>. Reprinted with permission.

All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval

permissions

ACADEMIC PRESS, INC., 525 B STREET, SUITE 1900, SAN DIEGO, CA 92101-4495, USA

ACADEMIC PRESS

A Harcourt Science and Technology Company

525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

<http://www.academicpress.com>

Academic Press

Harcourt Place, 32 Jamestown Road, London NW1 7BY, UK

Library of Congress Catalog Card Number: 00-104377

International Standard Book Number: 0-12-282160-2

Printed in the United States of America

00 01 02 03 04 C W 9 8 7 6 5 4 3 2 1

Contents

Preface	xv
List of Contributors	xvii
 Chapter 1: Multimedia Communications: Source Representations, Networks, and Applications	 1
<i>Jerry D. Gibson</i>	
1.1 Introduction	1
1.2 Networks and Network Services	3
1.3 Multimedia Sources	6
1.4 Source and Destination Terminals	8
1.5 Applications of Multimedia Communications Networks	9
1.5.1 Video Streaming to Multiple Users	10
1.5.2 Videoconferencing	11
1.6 Conclusions	12
1.7 For Further Reading	12
 Chapter 2: Future Telecommunication Networks: Traffic and Technologies	 13
<i>Leonid G. Kazovsky, Giok-Djan Khoe, and M. Oskar van Deventer</i>	
2.1 Key Technologies	14
2.2 Impact of Competition	16
2.2.1 Market Penetration	17
2.2.2 Market Development	17
2.2.3 Hypotheses 3 and 4: The Digital Video Age	18
2.3.4 HDTV in the United States	20
2.3.5 Traffic Attributes	20
2.4 Synergy: Future Projections	21
2.5 Summary and Conclusions	22
2.6 Bibliography	22
 Chapter 3: Speech Coding Standards	 25
<i>Andreas S. Spanias</i>	
Abstract	25
3.1 Introduction	25
	 vii

3.2	Speech Analysis-Synthesis and Linear Prediction	27
3.2.1	Long-Term Prediction (LTP)	29
3.3	Linear Prediction and Speech Coding Standards	29
3.3.1	Open-Loop Linear Prediction	29
3.3.2	Standards Based on Analysis-by-Synthesis Linear Prediction	32
3.4	Standards Based on Subband and Transform Coders	39
3.4.1	The ITU G.722 Subband Coder	39
3.4.2	Sinusoidal Transform Coding	40
3.4.3	The Multiband Excitation Coder and the Imrasat-M Standard	40
3.5	Summary and Emerging Standards	41
3.6	References	42

Chapter 4: Audio Coding Standards **45**

Chi-Min Liu and Wen-Whoi Chang

4.1	Introduction	45
4.2	ISO/MPEG Audio Coding Standards	45
4.2.1	MPEG-1	46
4.2.2	MPEG-2	48
4.2.3	MPEG-4	49
4.3	Other Audio Coding Standards	50
4.3.1	Philips PASC	50
4.3.2	Sony ATRAC	51
4.3.3	Dolby AC-3	52
4.4	Architectural Overview	53
4.4.1	Psychoacoustic Modeling	53
4.4.2	Time-Frequency Mapping	54
4.4.3	Quantization	54
4.4.4	Variable-Length Coding	56
4.4.5	Multichannel Correlation and Irrelevancy	57
4.4.6	Long-Term Correlation	57
4.4.7	Pre-echo Control	58
4.4.8	Bit Allocation	59
4.5	Conclusions	59
4.6	Definitions of Key Terms	59
4.7	References	60
4.8	Bibliography	60

61

Michael W. Hoffman and Khalid Sayood

5.1	Introduction	61
5.2	Lossy Compression	62
5.2.1	JPEG	62
5.2.2	JPEG2000	68
5.3	Lossless Compression	71
5.3.1	JPEG	71
5.3.2	JPEG-LS	71
5.4	Bilevel Image Compression	73
5.4.1	JBIG	73
5.4.2	JBIG2	78
5.5	Definitions of Key Terms	79
5.6	References	80
5.7	Bibliography	80

Chapter 6: Multimedia Conferencing Standards	81
<i>David Lindbergh</i>	
6.1 Introduction	81
6.2 H.320 for ISDN Videoconferencing	82
6.2.1 The H.320 Standards Suite	83
6.2.2 Multiplex	84
6.2.3 System Control Protocol	84
6.2.4 Audio Coding	85
6.2.5 Video Coding	86
6.2.6 H.231 and H.243: Multipoint	87
6.2.7 H.232 and H.234: Encryption	89
6.2.8 H.331 Broadcast	89
6.3 H.320 Network Adaptation Standards: H.321 and H.322	89
6.3.1 H.321: Adaptation of H.320 to ATM and B-ISDN	90
6.3.2 H.322: Adaptation of H.320 to IsoEthernet	90
6.4 A New Generation: H.323, H.324, and H.310	90
6.4.1 H.241 Control Protocol	91
6.4.2 Audio and Video Codecs	91
6.4.3 H.323 for Packet Switched Networks	93
6.4.4 H.324 for Low-Bit-Rate Circuit Switched Networks	96
6.4.5 H.310 for ATM and B-ISDN Networks	98
6.5 T.120 for Data Conferencing and Conference Control	98
6.6 Summary	98
6.7 References	99
 Chapter 7: MPEG-1 and -2 Compression	101
<i>Tom Lookabaugh</i>	
7.1 Introduction	101
7.2 The MPEG Model	101
7.2.1 Key Applications and Problems	102
7.2.2 Strategy for Standardization	102
7.3 MPEG Video	103
7.3.1 The Basic Algorithm	103
7.3.2 Temporal Prediction	106
7.3.3 Frequency Domain Decomposition	110
7.3.4 Quantization	111
7.3.5 Variable-L	112
	113
	114
	116
 Chapter 8: MPEG-4 and MPEG-7	117
<i>Jerry D. Gibson</i>	
8.1 Introduction	117
8.2 MPEG-4	118
8.2.1 MPEG-4 Systems Model	120
8.2.2 Natural Video Coding	124
8.2.3 Audio and Speech Coding	125
8.3 MPEG-7	127
8.4 Summary	128
8.5 References	128

Chapter 9: ATM Network Technology	129
<i>Yoichi Maeda and Koichi Asatani</i>	
9.1 Introduction	129
9.2 Overview	130
9.2.1 Background	130
9.2.2 Basic ATM Concepts	131
9.2.3 ATM Network Protocol Structure	131
9.2.4 International Standardization and Recommendations	132
9.3 Physical Layer Specifications	133
9.3.1 Basic Characteristics of the TC Sublayer	134
9.3.2 Interface Bit Rates	134
9.4 ATM Layer Specifications	134
9.5 ATM Adaptation Layer (AAL) Specifications	135
9.6 Network Aspects of B-ISDN	135
9.6.1 Traffic Control	135
9.6.2 ATM Layer Performance	137
9.6.3 OAM Functions	138
9.6.4 Signaling Procedure	138
9.6.5 VBS Interfaces	139
9.7 Other ATM Network Technologies	140
9.7.1 IP Over ATM	140
9.7.2 MPEG2 Over ATM	141
9.8 Concluding Remarks	141
9.9 Definitions of Key Terms	141
9.10 Bibliography	142
9.11 For Further Information	142
Chapter 10: ISDN	143
<i>Koichi Asatani and Toshinori Tsuboi</i>	
10.1 Introduction	143
10.1.1 General Features of ISDN	143
10.1.2 Service Aspects of ISDN	144
10.1.3 Access Features	146
10.2 ISDN User-Network Interface	146
10.2.1 Basic Interface	147
10.2.3 Interface Features	148
10.3 Layers 1, 2, and 3 Specifications of UNI	151
10.3.1 Layered Structure	151
10.3.2 Basic Interface Layer 1	151
10.3.3 Primary Rate Interface Layer 1	158
10.3.4 Layer 2 Specification	162
10.3.5 Layer 3 Specification	168
10.4 Access Transmission Line Systems	171
10.4.1 Outline of Transmission Line System	171
10.4.2 Metallic Transmission Line System for Basic Access	172
10.4.3 Primary Rate Transmission System	176
10.5 References	177

Chapter 11: Video-on-Demand Broadcasting Protocols	179
<i>Steven W. Carter, Dariell D. E. Long, and Jehan-François Pâris</i>	
11.1 Introduction	179
11.2 Common Terms and Concepts	180
11.3 Staggered Broadcasting Protocols	180
11.4 Pyramid Broadcasting Protocols	181
11.5 Harmonic Broadcasting Protocols	184
11.6 Summary	186
11.7 Definitions of Key Terms	187
11.8 References	188
11.9 For Further Information	189
Chapter 12: Internet Telephony Technology and Standards Overview	191
<i>Bernard S. Ku</i>	
12.1 Introduction	191
12.2 Internet Telephony Architecture Overview	192
12.3 Related Internet Telephony Standards	194
12.3.1 IETF	195
12.3.2 ETSI Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON)	195
12.3.3 ITU-T	196
12.3.4 TISI	198
12.4 Current and Developing Internet Telephony Protocols	198
12.4.1 H.323	198
12.4.2 Session Initiation Protocol (SIP)	200
12.4.3 Media Gateway Control Protocol (MGCP)	202
12.4.4 MEGACO/H.248 (H.GCP)	203
12.5 How Voice Over Internet Protocol (VoIP) Works	205
12.5.1 PSTN Gateways	205
12.5.2 VoIP Gateways	206
12.5.3 IFTel Gateways	207
12.6 Open Issues in Internet Telephony	209
12.7 IN/IP Integration	210
12.7.1 New Elements/Functions Required	211
12.7.2 Special Extensions Required	212
12.7.3 New IN/IP Interworking Interfaces	213
12.7.4 Information Transfer from Click-to-Dial (CTD) to IP	214
12.8.1 Future of IP/SS7	215
12.8.2 Future of IP/SS7	216
12.8.3 Future of IP/SS7	216
12.8.3 Future of IP/SS7	217
12.9 Concluding Remarks	217
12.10 Glossary	217
12.11 Definitions of Key Terms	218
12.12 Acknowledgments	219
12.13 Bibliography	219
Chapter 13: Wideband Wireless Packet Data Access	221
<i>Jusin Chuang, Leonard J. Cimini, Jr., and Nelson Sollenberger</i>	
13.1 Introduction	221
13.1.1 The Wireless Data Opportunity	221
13.1.2 Current Wireless Data Systems	222

xii CONTENTS

13.1.3	Emerging and Future Wireless Data Options	223
13.1.4	Summary and Outline of the Chapter	225
13.2	Packet Data Access Using WCDMA	225
13.2.1	Variable-Rate Packet Data	228
13.3	Packet Data Access Using EDGE	228
13.3.1	Link Adaptation and Incremental Redundancy	229
13.4	Packet Data Access Using Wideband OFDM	232
13.4.1	Physical-Layer Techniques	232
13.4.2	Physical-Layer Solutions	232
13.4.3	Frequency Reuse and Spectral Efficiency	234
13.4.4	Dynamic Packet Assignment Protocol	235
13.4.5	Dynamic Packet Assignment Performance	235
13.4.6	Radio Link Resource Organization	236
13.4.7	Frame Structure for Dynamic Packet Assignment	239
13.4.8	Simulation Model	240
13.4.9	Simulation Performance Results	241
13.5	Conclusions	244
13.6	References	244
Chapter 14: Internet Protocols Over Wireless Networks		247
<i>George C. Polyzos and George Xylomenos</i>		
	Abstract	247
14.1	Introduction	247
14.2	Internet Protocols and Wireless Links	248
14.2.1	Internet Transport Layer Protocols	248
14.2.2	Protocol Performance Over a Single Wireless Link	249
14.2.3	Protocol Performance Over Multiple Links	251
14.3	Performance Enhancements for Internet Protocols	253
14.3.1	Approaches at the Transport Layer	253
14.3.2	Approaches Below the Transport Layer	254
14.4	The Future: Challenges and Opportunities	256
14.4.1	Wireless System Evolution	256
14.4.2	Goals for Protocol Evolution	257
14.5	Summary	258
14.6	References	258
Chapter 15: Transcoding for Multimedia		261
15.1	Introduction	261
15.1.1	Adapting to Bandwidth Heterogeneity	262
15.1.2	Adapting to Client Heterogeneity	265
15.2	End-to-End vs. Proxy-Based Transcoding Designs	266
15.3	Architecture of a Transcoding Proxy	271
15.4	To Transcode or Not to Transcode	274
15.4.1	A Store-and-Forward Image Transcoding Proxy	274
15.4.2	A Streamed Image Transcoding Proxy	277
15.5	Transcoding Policies for Selecting Content	280
15.5.1	Optimal Policies for Offline Pretranscoding	281
15.5.2	Policies for Real-Time Transcoding	284
15.6	A Sample Set of Transcoding Policies	289

CONTENTS	xiii
15.7 Related Issues	291
15.8 Acknowledgments	293
15.9 References	293
Chapter 16: Multicasting: Issues and Networking Support	297
<i>Utpal Vorshney</i>	
16.1 Introduction	297
16.2 Multicasting Support	298
16.3 Multicasting in IP-Based Networks	299
16.3.1 Routing Protocols for IP Multicast	301
16.3.2 Multimedia Support and IP Multicasting	301
16.3.3 Multimedia Multicasting Applications on the MBone	302
16.4 Multicasting in ATM Networks	302
16.4.1 Multicasting Schemes for ATM Networks	303
16.5 IP Multicasting Over ATM	305
16.5.1 Problems in RSVP Over ATM	305
16.5.2 IP Multicast Over ATM in VBNS	306
16.6 Reliable Multicast Transport Protocols	306
16.7 Multicasting in Wireless Networks	307
16.7.1 Issues in IP Multicasting Over Wireless	308
16.7.2 Multicast Support in Wireless ATM	308
16.8 Summary and the Future of Multicasting	308
16.9 Definitions of Key Terms	309
16.10 References	309
16.11 For Further Reading	310
Index	311

Table 1.4 Selected Videoconferencing Standards (Basic Modes)

Standard	Network	Video	Audio
H.320 (1990)	ISDN	H.261	G.711
H.323 (1996)	LANs/Internet	H.261	G.711
H.324 (1995)	PSTN	H.263	G.723.1
H.310 (1996)	ATM/B-ISDN	H.262	MPEG-1

Abbreviations: ATMB-ISDN, Asynchronous Transfer Mode/Broadband ISDN; MPEG-1, standard for videotape quality video and high quality audio on a CD; PSTN, Public Switched Telephone Network.

quality since MOS for any coder varies according to test conditions. This is why a range of MOS values is given for G.711. The complexity measure is millions of instructions per second (MIPs), and is also approximate. In the bit rate column, a range of bit rates is shown for G.726, which indicates that this coder has selectable rates at 16, 24, 32, and 40 Kbit/s. In contrast, IS-96 and IS-127 show a range of rates, but these rates are adaptively varied by the coder based upon the input speech. These coders are therefore designated variable rate coders. Some of the coders shown are for wireline telephony, some are for digital cellular, and several are used in videoconferencing and multimedia applications. Key information presented in Table 1.3 is typical bit rates that are possible and the substantial level of complexity that has become common for these telephone bandwidth speech coders.

There are also wideband speech compression methods, such as G.722, and newly evolving standards in this area (G.722.1), plus wideband audio, such as MP3. We leave further discussion of these topics to later chapters. For completeness, Table 1.4 lists important videoconferencing standards. The video and audio codecs listed in Table 1.4 are only for the basic modes, and several alternatives are part of the more complete standard and are often implemented—see Chapter 6. Note that the standards listed in Table 1.4 are systems standards, and as such, have multiplexing and controls part of the specification. These are not shown in the table, but are left for discussion in Chapter 6. Considerable effort is also being expended in developing the Motion Picture Experts Group standards MPEG-4 (for object-based audio-visual representation) and MPEG-7 (for multimedia content description interface), which will be important to multimedia communications over networks, and these are discussed in Chapter 8.

1.4 SOURCE AND DESTINATION TERMINALS

In this chapter, we use the word *terminal* to refer to any device that connects a user to the network. For voice communications over the PSTN, the terminal may simply be a telephone handset, or for the Internet, it may be a desktop computer. However, today and certainly for the future, terminals are going to take a host of shapes and sizes and be asked to accommodate a full range of tasks. Terminals will thus be classified according to characteristics such as: sources handled (messaging, voice, data, images, video), size and weight, battery power and battery life, input devices (keyboards, microphone, or handset), output devices (handset audio, low-resolution black-and-white display, high-resolution color display), input/output (I/O) and processor speeds, special signal processing capabilities, mobility, and portability. One or more of these characteristics can dictate what can and cannot be done in a particular multimedia application.

At the time of this writing, there are several trends in evidence. First, central processing unit (CPU) speeds for desktop machines are at 850 MHz, and 1.2 GHz speeds are on the near horizon. In fact, Intel projects that by 2011, chips will have 1 billion transistors, 10 GHz clock speeds, and an additional 10 times increase in performance. The increases in performance will be due to innovations such as the use of increasing parallelism. Therefore, the 10-GHz chips will be 100 times more powerful than the 1-GHz chips soon to be available. Second, mobile, highly portable terminals such as laptop personal computers (PCs) and palm computing devices are evolving rapidly in terms of weight, power dissipation, battery life, and display capabilities, with palm computing devices certain to be important terminals of the future. Third, it is evident that special-purpose signal processing will be a significant part of tomorrow's terminals, and considering the compression needs as outlined in Tables 1.2–1.4, Digital Signal Processors (DSPs) are going to be a major component in these terminals. This trend will be accelerated for wireless connectivity such as in third-generation digital cellular and evolving wireless LAN standards, since considerable signal processing will be needed to mitigate the effects of the time-varying wireless channels and to achieve reliable data throughput. DSP processor speeds are increasing at a rate that tracks rates of increases in CPU speeds, and DSP designs are also exploiting techniques to extract more MIPS per MHz of processor speeds. There is also substantial effort being expended to develop low-power DSP designs in response to the need for DSPs in wireless devices.

Many different types of terminals will be connected to a network at any one time; furthermore, several kinds of terminals will often be involved in a single multimedia communications session. It is easy to imagine such a scenario for streaming applications. For example, a presentation by a chief executive officer (CEO), a seminar by a researcher, or a training session for a piece of equipment may all be available via streamed audio and video from a server. Users interested in any of these may be spread throughout their organization, out of the office on a business trip, or simply working at home. The streamed session would then be accessed by desktop PCs, laptop PCs, possibly PDAs, or perhaps, through a wireless phone with audio-only capability. Because of the great popularity of the Internet, the issues involved with transcoding multimedia content for access via a variety of different terminals is under way.

The widely varying types of terminals and the diverse types of physical layer channels will create an even more heterogeneous environment than we have today in the Internet. This will keep the pressure on the development of new protocols and network interfaces to maintain interoperability at the high level expected.

WORKS

The heterogeneity mentioned in the previous section becomes explicitly visible when one considers the Multimedia Communications Network shown diagrammatically in Figure 1.3. Network connection speeds range from a few tens of Kilobits per second to more than 100 Mbit/s, and the media that comprise these channels range from optical fiber and coaxial cable through copper wire pairs and free space. The terminals may be high-end workstations with large displays, desktop personal computers, battery-powered laptops, and personal digital assistants (shown as personal communicators in the figure) that have small black-and-white displays.

The two most common multimedia communications applications today and that are foreseen in the near future are video streaming to multiple users and party-to-party or multiparty videoconferencing. In this section, we present typical scenarios for these two applications and highlight some of the issues and challenges.

1.5.1 Video Streaming to Multiple Users

For this application, we assume that the Multimedia Server in the upper left-hand corner of Figure 1.3 wishes to stream a video lecture to any user on the network. The video may be stored in MPEG-2 format [standard for movie quality video and audio], which is very high quality, but direct streaming of MPEG-2 video requires a variable rate of from 4 to 10 Mbit/s. Such rates would be incompatible with many of the network connections shown in Figure 1.3, so the first inclination might be to transcode this information down to a rate that is commensurate with all, or a majority of, the network links. The video compression method of choice to do this might be H.263, which offers a wide range of rates and frame sizes, and is widely supported. The result of the transcoding, however, would be that we would have a least common denominator encoding, so that even those users with higher rate network connections would be forced to accept the quality produced for the low-rate users.

One approach to working around the lowest common denominator limitation would be to use a layered coder with multicasting. That is, you would choose a video coder that allows multiple compression rates that can be obtained by incrementally improving the base layer. Coders that have this capability are sometimes said to be *scalable*. MPEG-2 has several scalability options, including signal-to-noise ratio (SNR), spatial, and frame rate scalability. One or more of these options could be used and combined with multicasting to create (say) three multicast groups. The first group would be the baseline coded layer, and the other two would use scalability to create incremental improvements in output quality as users join the remaining two multicast groups. There might be another good reason to use multicast in this application. Specifically, if the desired number of viewers is large, unicast transmissions to each of them could flood various links in the network. If multicast is employed, users on congested links could reduce the rate by reducing the number of multicast groups that they join.

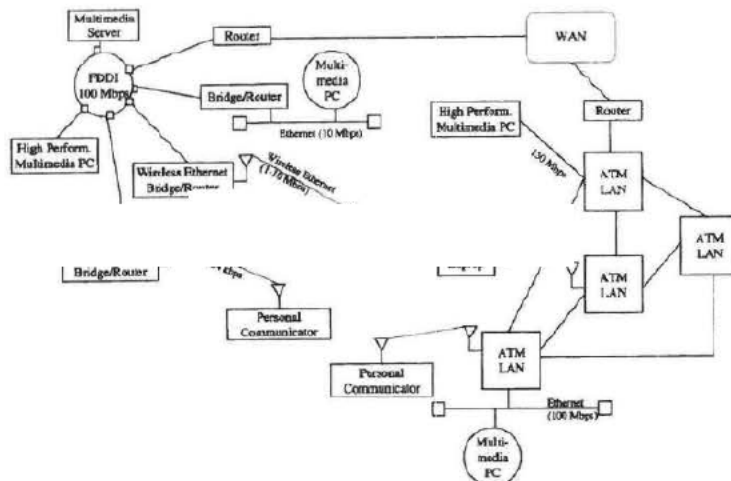


FIGURE 1.3
Multimedia communications network. (ATM = Asynchronous Transfer Mode.)

Another approach to establishing interoperability between networks would be to transcode at network gateways. There are three disadvantages usually cited for transcoding: (1) complexity, (2) delay, and (3) added distortion. For video streaming, because it is one way, delay is not a serious concern. Complexity would be an issue at some network interfaces, and in those cases, the streamed video might not be transcoded, thus yielding degraded network performance and poor delivered video quality. If complexity does not preclude transcoding, then the remaining issue is the distortion added during the transcoding process. Of course, transcoding to a lower rate will yield lower quality, smaller frame size, and/or slower frame rate, so it is key to add as little additional distortion as possible. This implies that we would prefer to not completely decode back to video and then re-encode the video at a lower rate (notice that this has implications in terms of complexity and delay, too). We would add less distortion if we could directly map the encoded stream, or at least the decoded parameters, directly into a lower-rate-coded version. As you can imagine, there are numerous possible options corresponding to any given network interface and the particular compression method, and so we do not go into further detail here.

1.5.2 Videoconferencing

One-to-one videoconferences (single party to single party) are relatively easy to deal with. Compared to video streaming, the serious new issue that arises is *latency* or round trip delay. During call setup, the users can negotiate the particular compression methods to be used and the desired transmitted data rates. Usually the two participants agree on these issues and the conference can be initiated. Each user thus sees and hears the other participant, subject to distortions or breakup in the video or audio due to packet losses. In the situation where the total video and audio transmitted data rate cannot be sustained without packet losses, it is often desirable to give priority to the audio signal. This is because participants are more forgiving of breakups in video than in audio.

Notice in this case that it does not make sense for a participant with a high data rate connection to request or send high-rate video if the other participant has a much lower rate channel. The lower-rate channel cannot send or receive data at the higher rate. Thus, one-to-one videoconferences are negotiated to preferences that reflect the lowest common denominator in transmitted data rate. In Internet videoconferencing, the principal question to be answered before attempting a conference is whether each user supports a common videoconferencing tool. Compared to the familiar PSTN videoconferencing applications, Internet videoconferencing tools offer more diversity and less standardization.

For multicasts with a large number of participants, it is desirable that all participants receive audio from all of the other participants (MCU), and then retransmitting the combined audio to each participant. The drawbacks of such an approach are that all of the audio has to be decoded, combined, and re-encoded, resulting in high complexity and possible performance loss. Another alternative is for each participant to transmit audio to all of the other participants. This approach requires that all participants either use the same audio coding method, so that an appropriate decoder is available at every location, or that all participants be able to decode all compression schemes in the videoconference. Note that the bitrate over the links does not increase linearly with the number of conference participants since there is usually only a single speaker at any time instant.

The question of what video is displayed at each location is also important. Preferably each participant should be able to see all other participants at the same time. As long as bandwidth is not a problem, multiple received video streams can be presented on the computer display in what is often called a "Hollywood Squares" arrangement. With reference to Figure 1.3, notice that this

may not be possible for all participants because of their individual bandwidth limitations or because of the resolution of their terminal's display. This problem can be simplified by adopting the approach that participants are provided with only the video of the current speaker. This approach makes sense because there should be only a single speaker at any one time instant; however, this speaker can be at any of the locations, and the speaker may change fairly often. There are three standard approaches to accommodating the need for switching the video supplied to participants. The first is what is called *director control*, where one participant in the videoconference is designated the director and manually switches between locations as a participant begins to speak. For this to work well, it is desirable that the director have access to all of the audio. A second scenario is where the video is switched automatically by a volume cue that selects the location with the loudest audio. To prevent rapid, but inadvertent, switching due to coughing, dropped objects, doors slamming, etc., there is usually a requirement that the loudest audio be present for some minimum time interval. A third alternative is for each participant to choose which video stream is displayed on his or her individual terminal. All three of these approaches are in use today.

As in video streaming, network heterogeneity and the variety of devices serving as user terminals present challenges to videoconferencing. Since there is already considerable delay due to video compression and possible audio combining, additional delay due to transcoding at network gateways becomes difficult to tolerate. At the same time, it is much preferred for each participant to be able to choose the bit rate over its local network link so that the quality of the videoconference can be kept as high as possible.

Multicast transmission can be particularly useful in this environment. Of course, the situation where all participants are multicasting could also aggravate traffic congestion. However, it would be unusual for all participants to be able to multicast their audio and video, either because they lack multicast capability or because their local link bandwidth precludes sending more than a single data stream. Multicasting combined with layered video compression methods presents the best videoconferencing quality to participants, since each participant can choose the quality received in proportion to either available link bandwidth or the processing and display capabilities of the participant's terminal.

1.6 CONCLUSIONS

The stated goal of "ubiquitous delivery of multimedia content via seamless network connectivity" is becoming a reality. Certainly there are many technical challenges, but new solutions are being developed every day, and the commercial demand for the resulting services is large and growing. This book develops key topics in greater detail and describes many of the particular applications of multimedia communications over networks that are available today and will be available in the near future.

1.7 FOR FURTHER READING

The chapters in the remainder of this book elaborate on many of the issues raised here. Two additional references of interest are:

1. J. D. Gibson, T. Berger, T. Lookabaugh, D. Lindbergh, and R. L. Baker, *Digital Compression for Multimedia: Principles and Standards*, Morgan Kaufmann Publishers, San Francisco, CA, 1998.
2. L. L. Peterson and B. S. Davie, *Computer Networks: A Systems Approach*, 2nd ed., Morgan Kaufmann Publishers, San Francisco, CA, 2000.