



ATIS-0100803.1998(R2013)

Overview and Reference for GSTN Multimedia Terminals

AMERICAN NATIONAL STANDARD FOR TELECOMMUNICATIONS



As a leading technology and solutions development organization, ATIS brings together the top global ICT companies to advance the industry's most-pressing business priorities. Through ATIS committees and forums, nearly 200 companies address cloud services, device solutions, emergency services, M2M communications, cyber security, ehealth, network evolution, quality of service, billing support, operations, and more. These priorities follow a fast-track development lifecycle – from design and innovation through solutions that include standards, specifications, requirements, business use cases, software toolkits, and interoperability testing.

ATIS is accredited by the American National Standards Institute (ANSI). ATIS is the North American Organizational Partner for the 3rd Generation Partnership Project (3GPP), a founding Partner of oneM2M, a member and major U.S. contributor to the International Telecommunication Union (ITU) Radio and Telecommunications sectors, and a member of the Inter-American Telecommunication Commission (CITEL). For more information, visit www.atis.org.

AMERICAN NATIONAL STANDARD

Approval of an American National Standard requires review by ANSI that the requirements for due process, consensus, and other criteria for approval have been met by the standards developer.

Consensus is established when, in the judgment of the ANSI Board of Standards Review, substantial agreement has been reached by directly and materially affected interests. Substantial agreement means much more than a simple majority, but not necessarily unanimity. Consensus requires that all views and objections be considered, and that a concerted effort be made towards their resolution.

The use of American National Standards is completely voluntary; their existence does not in any respect preclude anyone, whether he has approved the standards or not, from manufacturing, marketing, purchasing, or using products, processes, or procedures not conforming to the standards.

The American National Standards Institute does not develop standards and will in no circumstances give an interpretation of any American National Standard. Moreover, no person shall have the right or authority to issue an interpretation of an American National Standard in the name of the American National Standards Institute. Requests for interpretations should be addressed to the secretariat or sponsor whose name appears on the title page of this standard.

CAUTION NOTICE: This American National Standard may be revised or withdrawn at any time. The procedures of the American National Standards Institute require that action be taken periodically to reaffirm, revise, or withdraw this standard. Purchasers of American National Standards may receive current information on all standards by calling or writing the American National Standards Institute.

Notice of Disclaimer & Limitation of Liability

The information provided in this document is directed solely to professionals who have the appropriate degree of experience to understand and interpret its contents in accordance with generally accepted engineering or other professional standards and applicable regulations. No recommendation as to products or vendors is made or should be implied.

NO REPRESENTATION OR WARRANTY IS MADE THAT THE INFORMATION IS TECHNICALLY ACCURATE OR SUFFICIENT OR CONFORMS TO ANY STATUTE, GOVERNMENTAL RULE OR REGULATION, AND FURTHER, NO REPRESENTATION OR WARRANTY IS MADE OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OR AGAINST INFRINGEMENT OF INTELLECTUAL PROPERTY RIGHTS. ATIS SHALL NOT BE LIABLE, BEYOND THE AMOUNT OF ANY SUM RECEIVED IN PAYMENT BY ATIS FOR THIS DOCUMENT, AND IN NO EVENT SHALL ATIS BE LIABLE FOR LOST PROFITS OR OTHER INCIDENTAL OR CONSEQUENTIAL DAMAGES. ATIS EXPRESSLY ADVISES THAT ANY AND ALL USE OF OR RELIANCE UPON THE INFORMATION PROVIDED IN THIS DOCUMENT IS AT THE RISK OF THE USER.

NOTE - The user's attention is called to the possibility that compliance with this standard may require use of an invention covered by patent rights. By publication of this standard, no position is taken with respect to whether use of an invention covered by patent rights will be required, and if any such use is required no position is taken regarding the validity of this claim or any patent rights in connection therewith. Please refer to [<http://www.atis.org/legal/patentinfo.asp>] to determine if any statement has been filed by a patent holder indicating a willingness to grant a license either without compensation or on reasonable and non-discriminatory terms and conditions to applicants desiring to obtain a license.

ATIS-0100803.1998(R2013), *Overview and Reference of GSTN Multimedia Terminals*

Is an American National Standard developed by the **ATIS Network Performance, Reliability and Quality of Service Committee (PRQC)**.

Published by

**Alliance for Telecommunications Industry Solutions
1200 G Street, NW, Suite 500
Washington, DC 20005**

Copyright © 2013 by Alliance for Telecommunications Industry Solutions
All rights reserved.

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher. For information contact ATIS at 202.628.6380. ATIS is online at <http://www.atis.org>.

Printed in the United States of America.

American National Standard
for Telecommunications –
**Overview and Reference for
GSTN Multimedia Terminals**

Secretariat

Alliance for Telecommunications Industry Solutions

Approved November 13, 1998

American National Standards Institute, Inc.

American National Standard

Approval of an American National Standard requires review by ANSI that the requirements for due process, consensus, and other criteria for approval have been met by the standards developer.

Consensus is established when, in the judgement of the ANSI Board of Standards Review, substantial agreement has been reached by directly and materially affected interests. Substantial agreement means much more than a simple majority, but not necessarily unanimity. Consensus requires that all views and objections be considered, and that a concerted effort be made towards their resolution.

The use of American National Standards is completely voluntary; their existence does not in any respect preclude anyone, whether he has approved the standards or not, from manufacturing, marketing, purchasing, or using products, processes, or procedures not conforming to the standards.

The American National Standards Institute does not develop standards and will in no circumstances give an interpretation of any American National Standard. Moreover, no person shall have the right or authority to issue an interpretation of an American National Standard in the name of the American National Standards Institute. Requests for interpretations should be addressed to the secretariat or sponsor whose name appears on the title page of this standard.

CAUTION NOTICE: This American National Standard may be revised or withdrawn at any time. The procedures of the American National Standards Institute require that action be taken periodically to reaffirm, revise, or withdraw this standard. Purchasers of American National Standards may receive current information on all standards by calling or writing the American National Standards Institute.

Published by

**American National Standards Institute, Inc.
11 West 42nd Street, New York, NY 10036**

Copyright © 1999 by Alliance for Telecommunications Industry Solutions
All rights reserved.

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without prior written permission of the publisher.

Printed in the United States of America

Contents

	Page
Foreword	ii
1 Scope, purpose and application	1
2 Normative references	1
3 Definitions	3
4 Abbreviations	6
5 Conventions	8
6 Arrangements and architecture	8
7 Control messaging and indication	16
8 Multiplex	20
9 Video codec	21
10 Audio codec	24
11 Interoperability	25
Tables	
1 Basic features of multi-media terminal with baseline components (Implementing ITU-T Recommendations H.263 & G.723.1)	8
2 Picture formats for video terminals	14
3 Number of pixels per line and number of lines for each of the ITU-T Recommendation H.263 picture formats	23
Figures	
1 Elements of ITU-T Recommendation H.324 terminal	11
2 Protocol stack for ITU-T Recommendation H.223	20
3 Information field example	21
4 Simplified block diagram of video codec	21
5 Simplified block diagram of ITU-T Recommendation G.723.1 Encoder	24
6 Simplified block diagram of ITU-T Recommendation G.723.1 Decoder	25
Annex	
A Bibliography	27

Foreword (This foreword is not part of American National Standard T1.803-1998.)

This standard was developed by Technical Subcommittee T1A1 under project T1A1-06 which establishes standard specifications for VideoPhone signals transmitted using voiceband data techniques over the Public Switched Telephone Network (PSTN). These specifications will allow video telephony of digitally encoded signals to communicate voice and imagery in a full duplex mode.

This standard is an overview of the suite of ITU-T Recommendations for GSTN Multimedia Communications Terminals defined in ITU-T Recommendation H.324 and its associated recommendations.

Because the ITU-T Recommendations comprise a detailed suite of standards for compatible terminals for this application, T1A1 has chosen to provide a useful overview document with appropriate normative references to ITU-T Recommendations which should be utilized in developing and operating a standard product. In all cases, the referenced ITU-T Recommendations will be definitive.

Every effort (where practicable) has been made to utilize syntax in accordance with the ANSI Style; however, since some terminology and syntax derives from approved ITU-T texts, in the interest of enabling the reader to correspondingly refer to appropriate sections in the normative references, the ITU-T syntax and terminology has been adopted here.

This standard has one informative annex, which is not considered part of the standard.

Suggestions for improvement of this standard are welcome. They should be sent to the Alliance for Telecommunications Industry Solutions, 1200 G Street, NW, Suite 500, Washington, DC 20005.

This standard was processed and approved for submittal to ANSI by Accredited Standards Committee on Telecommunications, T1. Committee approval of the standard does not necessarily imply that all committee members voted for its approval. At the time it approved this standard, the T1 committee had the following members:

- G. H. Peterson, Chairman
- E. R. Hapeman, Vice-Chairman
- A. Lai, Secretary

- B. Lerich, Senior Editor

<i>Organization Represented</i>	<i>Name of Representative</i>
EXCHANGE CARRIERS	
Ameritech	L. Richard Wood Larry A. Young (Alt.)
AT&T Wireless Services, Inc.	David Holmes
Bell Atlantic.....	Roger Nucho James F. Baskin (Alt.)
Bellcore	James C. Staats Cliff Halevi (Alt.)
BellSouth Telecommunications, Inc.	Malcolm Threlkeld, Jr. John Spencer (Alt.)
GTE Telephone Operations	Bernard J. Harris Richard L. Cochran (Alt.)
Pacific Bell.....	Adrian Eriksen
SBC Communications, Inc.	C. C. Bailey Robert J. Hall (Alt.)
Sprint - Local Telecommunications Division.....	Leroy D. Kellogg
US Telephone Association (USTA)	Paul Hart Anthony Pupek (Alt.)
US WEST	James L. Eitel Darryl Debault (Alt.)

<i>Organization Represented</i>	<i>Name of Representative</i>
INTEREXCHANGE CARRIERS	
AT&T	Charles A. Dvorak
	Jeffrey George (Alt.)
AT&T Canada Long Distance Services.....	David H. Whyte
Comsat Corporation	Mark T. Neibert
	Prakash Chitre (Alt.)
General Communication, Inc.	Derek L. Welton
	C. R. Baugh (Alt.)
MCI Telecommunications Corporation.....	Laszlo I. Szerenyi
	J. Martin Carroll (Alt.)
Sprint - Long Distance Division	Thomas G. Croda
	James Lord (Alt.)
Stentor Resource Centre, Inc.	B. Sambasivan
MANUFACTURERS	
3Com.....	Richard L. Stuart
	Dale Walsh (Alt.)
ADC Telecommunications Inc.	Don Berryman
	Cliff Davidow (Alt.)
Alcatel Network Systems (ANS).....	Ken Biholar
	Bill Powell (Alt.)
AMP, Inc.	Ben Bennett
Analog Devices	Rupert J. Baines
	Sunder Kidambi (Alt.)
Apple Computer, Inc.	Wanda Cox
	David Michael (Alt.)
Ascom Enterprise Networks.....	Z. Putnins
Aware, Inc.	Michael Tzannes
	Marcos Tzannes (Alt.)
DSC Communications Corporation	Pete Waal
	Allen Adams (Alt.)
ECI Telecommunications, Inc.	Ron Murphy
	Danny Etz-Hadar (Alt.)
Ericsson, Inc.	Linda Troy
	Al Way (Alt.)
Fujitsu America, Inc.	Kenneth T. Coit
	Hirohiko Yamamoto (Alt.)
General Datacom, Inc.	Fred Cronin
	William Conway (Alt.)
Globespan Technologies, Inc.	Massimo Sorbara
	Clete Gardenhour (Alt.)
Harris Corporation.....	Yogi Mistry
Hekimian Laboratories	William H. Duncan
Hewlett-Packard.....	Karen Higginbottom
	Richard van Gelder (Alt.)
Hughes Network Systems, Inc.	Leonard Golding
	Enrique Laborde (Alt.)
Lucent Technologies	John H. Bobsin
	Dave R. Andersen (Alt.)
Motorola, Inc.	Ken Skurnak
	Dan Grossman (Alt.)
NEC America, Inc.	Donovan Nak
	Hajime Koto (Alt.)
Nokia Telecommunications, Inc.	Chris Wallace
	Teuvo Jarvela (Alt.)
Northern Telecom, Inc.	Mel N. Woinsky
	John Pugh (Alt.)
Oki America, Inc.	Henri Suyderhoud
	Hisao Fujikawa (Alt.)
Orckit Communications	James Szeliga
	Nigel Cole (Alt.)
Paradyne Corporation	Richard K. Smith
	Phil Kyees (Alt.)
Picturetel Corporation	Marshall Schachtman
	David Lindbergh (Alt.)

<i>Organization Represented</i>	<i>Name of Representative</i>
Pirelli TSG	John McDonough T. C. Nie (Alt.)
Qualcomm, Inc.	Mark Epstein Ed Tiedemann (Alt.)
Reltec Corporation.....	Mark Scott Leroy Baker (Alt.)
Rockwell International	Quentin C. Cassen Tim Werve (Alt.)
Siemens Telecom Networks (STN)	David E. Francisco Dennis Edinger (Alt.)
Telecommunications Solutions.....	M. J. Narasimha Don Chislow (Alt.)
Telecommunications Techniques	Bernard E. Worne John Paul Williams (Alt.)
Tellabs Operations, Inc.	Jim Orme Tom Rarick (Alt.)
Texas Instruments	James T. Carlo Peter J. Melsa (Alt.)
Transwitch Corporation	Jitender Vij Edwin Soltysiak (Alt.)

GENERAL INTEREST

ABC, Inc.	Warner W. Johnston
BellSouth Mobility DCS	Don Zelmer Scott Fox (Alt.)
C. S. I. Telecommunications.....	Michael S. Newman William J. Buckley (Alt.)
Defense Information Systems Agency	C. Joe Pasquariello Don Choi (Alt.)
Gemplus	Leo Legaspi Jennie Ong (Alt.)
Microcell Connexions	Charles Despins
National Telephone Coop Association.....	Paul M. Johnson Scott Reiter (Alt.)
National Communications System.....	Dennis Bodson Marshall Cain (Alt.)
National Telecommunications and Information Administration/Institute for Telecommunications Sciences (NTIA/ITS)	William F. Utlaut Neal B. Seitz (Alt.)
Pacific Bell Mobile Services.....	Mark Younge
Rural Utilities Services.....	Orren E. Cameron III Norberto Esteves (Alt.)

At the time this standard was approved, Technical Subcommittee T1A1 on Performance and Signal Processing had the following members:

N. B. Seitz, Chairman
E. G. Hauch, Vice-Chairman

<i>Organization Represented</i>	<i>Name of Representative</i>
ABC, Inc.	Warner W. Johnston
Alcatel Network Systems (ANS)	Ken Biholar Albert Azzam (Alt.)
Ameritech	Larry F. Brown
AT&T	Rick Canaday Charles A. Dvorak (Alt.)
Bell Atlantic.....	Max Roesch Lita B. Gwinn (Alt.)
Bellcore.....	Ralph E. Jensen W. Garry Couch (Alt.)
BellSouth Telecommunications, Inc.	David M. Brady Glenn Scheyd (Alt.)
C. S. I. Telecommunications.....	Michael S. Newman William J. Buckley (Alt.)

<i>Organization Represented</i>	<i>Name of Representative</i>
Comsat Corporation	Mark T. Neibert Faris Faris (Alt.)
Defense Information Systems Agency	Gary L. Koerner
ECI Telecommunications, Inc.	Ron Murphy David Shraier (Alt.)
Ericsson, Inc.	Mustafa Kocaturk Sangamesh Vinayagamurthy (Alt.)
GTE Telephone Operations	Richard L. Cochran Gary E. McAninch (Alt.)
Lucent Technologies	Robert B. Waller Carl R. Posthuma (Alt.)
MCI Telecommunications Corporation	Laszlo I. Szerenyi James Liou (Alt.)
Microcell Connexions	Charles Despina
National Communications System	An Nguyen Stephen Perschau (Alt.)
National Telecommunications and Information Administration/Institute for Telecommunications Sciences (NTIA/ITS).....	Neal B. Seitz Randy Bloomfield (Alt.)
NEC America, Inc.	Donovan Nak
Newbridge Networks Corporation	Michael Craig Bakri Aboukarr (Alt.)
Northern Telecom, Inc.	Mel N. Woinsky John Pugh (Alt.)
Oki America, Inc.	Henri Suyderhoud Hisao Fujikawa (Alt.)
Pacific Bell.....	Adrian Eriksen
Paradyne Corporation	Richard K. Smith Phil Kyees (Alt.)
Picturetel Corporation	Tony Crossman Marshall Schachtman (Alt.)
Rockwell International	Tom Geary Glen R. Griffith (Alt.)
SBC Communications, Inc.	John E. Roquet Robert J. Hall (Alt.)
Siemens Telecom Networks (STN)	Gilbert W. Hassell Suhass S. Gandhi (Alt.)
Sprint - Long Distance Division	Cannon Hwu James Lord (Alt.)
Stentor Resource Centre Inc.	J. A. Zebarth Dave Milne (Alt.)
Tektronix	Bozidar Janko David Fibush (Alt.)
Telecom Solutions.....	M. J. Narasimha Kishan Sheno (Alt.)
Telecommunications Techniques.....	Bernard E. Worne John Paul Williams (Alt.)
Tellabs Operations, Inc.	Maurice Givens Ricardo Davila (Alt.)
US Telephone Association (USTA).....	Robert Creighton
U S WEST	Bill Wycoff John Grigg (Alt.)
Vtel Corporation	Smita Gupta

At the time it developed this standard, working group T1A1.5 had the following participants:

D. J. Atkinson	Mustafa Kocaturk
Jyotin Basrur	Gary Koerner
Paul Berger	John Libert
Stephen Bersey	Dave Lindbergh
Bhavesh Bhalt	Jeffrey Lubin
Dick Bobilin	Michael E. Lukas
Carl Bushe	Doug Lung
Leigh Chang	Patrick Luthi
Christian J. van den Branden Lambrecht	John Mailhot
Martin H. Brill	Alfred C. Morton
Lorence F. Brown	Wallace Murray
Rick Canaday	Donovan Nak
EMel C. Celi	Mark Neibert
Gregory W. Cermak	An Nguyen
Richard L. Cochran	Chong Ong
Bill Cotton	Mark Perkins
Bill Coufal	Soren Pihlman
P. A. Dare	Michel Poulin
Jim Durant	Neil Randall
Chuck Dvorak	James R. (Rick) Redford
David Fibush	Max J. Roesch Jr.
Inette Furey	John Roquet
Godfrey Garner	Marshall G. Schachtman
Tom Geary	R. Schaphorst
Alan Godber	Ernest Schmidt
John Grigg	Tom Soon
Richard Grinnell	Richard Streeter
David L. Hanna	Henri Suyderhoud
Eric G. Hauch	Laszlo I. Szerenyi
Robert Hicks	Pat Tweedy
Katrina Hopkins	Jim Washura
Ralph E. Jensen	Arthur Webster
Warner W. Johnston	Stephen Wolf
John Jordan	Bernard Worne
Jim Jou	Joe Zebarth
Leslie B. Klein	Charles Ziegler

American National Standard for Telecommunications –

Overview and Reference for GSTN Multimedia Terminals

1 Scope, purpose and application

1.1 Scope

This document is to be a general overview of the implementation of multimedia terminals targeting audiovisual conferencing applications on the GSTN. This document also provides elaboration of implementation details in areas which the ITU-T documents have been found to be vague or unclear. Detailed standards for implementation of compatible terminals are provided by normative reference to approved ITU-T Recommendations. Multimedia terminals according to this standard which comprise combined video and audio (Videophones) shall include means for conventional telephony interworking with telephone terminals on the GSTN. Means for interworking with other modem or FAX terminals which may utilize the GSTN are optional. Such additional arrangements are outside the scope of this standard. Wireless operation is beyond the scope of this standard.

1.2 Purpose

The purpose of the document is to provide the reader a general understanding of the implementation of multimedia terminals targeting audiovisual conferencing applications on the GSTN. Its technical content is purposely left as an overview and technical implementation details are to be found in the referenced ITU-T recommendations. A detailed normative reference list of all applicable ITU-T and ISO/IEC recommendations is included.

1.3 Application

The multimedia terminals discussed in this standard provide a means for users to perform compatible audio and video communication, optionally supporting data sharing, over the GSTN. The terminals described by this and the referenced ITU-T recommendations will allow users to communicate domestically and internationally.

2 Normative references

The following standards contain provisions, which through reference in this text, constitute provisions of this American National Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision and parties to agreements based on this American National Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below.

ISO/IEC 3309 (1991), *Information Technology – Telecommunications and information exchange between systems – High-level data link control (HDLC) procedures – Frame Structure*¹⁾

ISO/IEC TR 9577 (1993), *Information Technology – Telecommunications information exchange between systems – Protocol identification in the network layer*¹⁾

CCIR Recommendation 601-3 1992, *Encoding Parameters of Digital Television for Studios*¹⁾

¹⁾ Available from American National Standards Institute, 11 West 42nd Street, New York, NY 10036.

ITU-T Recommendation G.711 (1988), *Pulse code modulation (PCM) of voice frequencies*; (Paragraph 3.6, Conversion to and from uniform PCM)¹⁾

ITU-T Recommendation G.723.1 (1996), *Dual rate speech coder for multimedia communication transmitting at 5.3 & 6.4 kbit/s.*¹⁾

ITU-T Recommendation G.728 (1992), *Coding of speech at 16 kbit/s using low-delay code excited linear prediction.*¹⁾

ITU-T Recommendation H.223 (1998), *Multiplexing protocol for low bitrate multimedia communication.*¹⁾

ITU-T Recommendation H.224 (1994), *A real time control protocol for simplex applications using the H.221 LSD/HSD/MLP channels* (Normative when the H.324 terminal in question is supporting Far End Camera Control).¹⁾

ITU-T Recommendation H.230 (1995), *Frame-synchronous control and indication signals for audiovisual systems* (Section 3, Definition of C & I signals).¹⁾

ITU-T Recommendation H.233 (1995), *Confidentiality system for audiovisual services* (Normative when the H.324 terminal in question is supporting the option of encryption).¹⁾

ITU-T Recommendation H.234 (1994), *Encryption key management and authentication system for audiovisual services* (Normative when the H.324 terminal in question is supporting the option of encryption).¹⁾

ITU-T Recommendation H.245 (1998), *Control protocol for multimedia communication.*¹⁾

ITU-T Recommendation H.246 (1998), *Interworking of h-series multimedia terminals with h-series multimedia terminals and voice/voiceband terminals on GSTN and ISDN*¹⁾

ITU-T Recommendation H.261 (1993), *Video codec for audiovisual services at px64 kbit/s.*¹⁾

ITU-T Recommendation H.263 (1998), *Video coding for low bitrate communication.*¹⁾

ITU-T Recommendation H.281 (1994), *A far end camera control protocol for video conferences using H.224* (Normative when the H.324 terminal in question is supporting Far End Camera Control).¹⁾

ITU-T Recommendation H.324 (1998), *Terminals for low bitrate multimedia communication.*¹⁾

ITU-T Recommendation T.30 (1996), *Procedures for document facsimile transmission in the general switched telephone network.*¹⁾

ITU-T Recommendation T.35 (1991), *Procedure for the allocation of CCITT defined codes for non-standard facilities* (Normative when the H.324 terminal in question is encoding NSF in its capabilities. It shall utilize the country coding as specified in T.35).¹⁾

ITU-T Recommendation T.120 (1996), *Transmission protocols for multimedia data* (Normative if the H.324 terminal in question is supporting the optional T.120 capability).¹⁾

ITU-T Recommendation T.123 (1996), *Protocol stack for multimedia conferencing applications* (Normative if the H.324 terminal in question is supporting the optional T.120 capability).¹⁾

ITU-T Recommendation T.434 (1996), *Binary file transfer format for the telematic services* (Normative if the H.324 terminal in question is supporting the optional point-to-point file transfer per T.434).¹⁾

ITU-T Recommendation V.8 (1998), *Procedures for starting sessions of data transmission over the GSTN* (Section 6.1 Call Function and table 3/V.8).¹⁾

ITU-T Recommendation V.8bis (1996), *Procedures for the identification and selection of common modes of operation between data circuit terminating equipment (DCE) and between data terminal equipment*

(DTE) over the general switched telephone network and on leased point-to-point telephone-type circuits (Section 8 and all sub paragraphs and tables).¹⁾

ITU-T Recommendation V.14 (1993), *Transmission of start-stop characters over synchronous bearer channels* (Normative if the H.324 terminal in question is supporting data channels).¹⁾

ITU-T Recommendation V.25ter, *Serial asynchronous dialing and control*¹⁾

ITU-T Recommendation V.80 (1997), *In-Band DCE Control and Synchronous Data Modes for Asynchronous DTE*.¹⁾

ITU-T Recommendation V.250 (1998), *Serial asynchronous automatic dialing and control* (also Annex A (1996), *Procedure for DTE-controlled call negotiation*).¹⁾

ITU-T Recommendation V.42 (1996), *Error-correcting procedures for DCEs using asynchronous-to-synchronous conversion*.¹⁾

ITU-T Recommendation V.42bis (1990), *Data compression procedures for data circuit terminating equipment (DCE) using error correction procedures*¹⁾

ITU-T Recommendation V.34 (1996), *A modem operating at data signaling rates of up to 28 800 bit/s for use on the GSTN and on leased point-to-point 2-wire telephone-type circuits*.¹⁾

ITU-T Recommendation X.691 (1995), *Information Technology – ASN.1 Encoding Rules – Specification of Packed Encoding Rules (PER)*¹⁾

3 Definitions

3.1 AL-SDU. The logical unit of information exchanged between the ITU-T Recommendation H.223. multiplex and the audio codec, video codec, or data protocol above.

3.2 audiographic. Audio plus still image and graphics.

3.3 audiographic conference. Conference in which audio plus still image and graphics is employed.

3.4 bi-directional logical channel pair. A bi-directional logical channel pair consists of a pair of associated transmission paths between two terminals, one in each direction of transmission.

3.5 capability. A terminal has a particular capability if it is able to encode and transmit or receive and decode that particular signal.

3.6 channel. A unidirectional link between two end-points.

3.7 codec. Coder/decoder, used to convert audio or video signals to/from digital format.

3.8 command. A command is a message that requires action by the remote terminal but no explicit response.

3.9 connection. A bi-directional link between two end-points.

3.10 control channel. Dedicated logical channel number 0 carrying system control protocol per ITU-T Recommendation H.245.

3.11 data. Information streams other than control, audio and video, carried in a logical data channel (see ITU-T Recommendation H.223).

3.12 elementary stream. Elementary stream is a generic term for a coded video, coded audio or other coded bitstream.

3.13 entry. The word entry is used to refer to elements in sets or tables, such as capability sets and multiplex tables.

3.14 in-band. In-band messages are those that are transported within the channel or logical channel to which they refer.

3.15 in-band signalling. Control signals sent within a specific logical channel other than the control channel, carrying information applicable only to that logical channel.

3.16 indication. An indication is a message that contains information but does not require action or response.

3.17 interworking adapter. A device connected to terminals or multipoint conference units (MCUs) working according to two or more recommendations, which translates the content of one or more logical channels to allow interoperation between otherwise incompatible equipment, sometimes known as "gateway".

3.18 lip synchronization. Operation to provide the feeling that the speaking motion of the displayed person is synchronized with that person's voice. The minimization of the relative delay between the visual display of a person speaking and the audio of the voice of the person speaking. The objective is to achieve a natural relationship between the visual image and the aural message for the viewer/listener.

3.19 logical channel. One of several logically distinct channels carried over a single bitstream.

3.20 logical channel number. A logical channel number is a number that identifies a single logical channel. It is unique for each direction of transmission.

3.21 logical channel signalling. Logical channel signalling is a set of procedures that are used to open and close logical channels.

3.22 master/slave. Master/Slave procedures are provided to resolve the very rare circumstances where there is contention between the two terminals which are usually operated with peer control status; to resolve a control message conflict, there is a random determination at the beginning of any call in which one of the two terminals is given precedence in cases of control message contention. Therefore the notion of Master/Slave does not imply a controlling function, but simply an explicit contention resolution process.

NOTE - The terminology Master/Slave can be interpreted to have unfortunate connotations in current usage. This terminology has been used in the approved ITU-T References and is being retained solely to assist the reader in understanding the principles and when referring to the referenced text.

3.23 master terminal. (See note in 3.22) A master terminal is the terminal that is determined as being the master terminal by the master-slave determination procedure defined in this standard.

3.24 media. One or more of audio, video, or data.

3.25 medium type. A medium type is a single form of information that is presented to a user or the data representing that information: audio, video and text are example Medium Types.

3.26 mode. A mode is a set of elementary streams that a terminal is transmitting, intends to transmit, or would like to receive.

3.27 multimedia communication. Multimedia communication refers to the transmission and/or reception of signals of two or more Medium Types simultaneously.

3.28 multimedia communication system. A system that handles more than one media stream in a synchronized way from the user's point of view. The system may allow interconnection of multiple parties, multiple connections, and the addition or deletion of resources and users within a single communication session.

3.29 MultimediaSystemControlmessage. A choice of message types. Messages defined in the system defined by this standard are classified as request, response, command and indication messages.

3.30 multilink. The use of more than one physical connection to obtain a larger aggregate bit rate.

3.31 multipoint. The simultaneous interconnection of three or more terminals to allow communication among several sites through the use of MCUs (bridges) that centrally direct the flow of information.

3.32 MUX-PDU. The logical unit of information exchanged between the multiplex layer discussed in ITU-T Recommendation H.223 and the underlying physical layer. It is a packet framed by HDLC flags and using HDLC zero-bit insertion for transparency.

3.33 non-segmentable. The mode of operation defined in ITU-T Recommendation H.223 in which AL-SDUs must be sent as consecutive octets in a single MUX-PDU. See ITU-T Recommendation H.223.

3.34 non-standard. Not conforming to a national or international standard referenced in this standard or in ITU-T Recommendation H.324.

3.35 pictures. In total, ITU-T Recommendation H.263 has seven basic picture types (of which only the first two are mandatory) which are defined primarily in terms of their prediction structure: As used herein, a "reference" or "anchor" picture is a picture that contains data which can be used by reference as a basis for the decoding of another picture. This use by reference is also known as "prediction", although it sometimes may actually indicate use in a reverse-temporal direction.

1. Intra: A picture having no reference picture(s) for prediction (also called an I picture);
2. Inter: A picture using a temporally-previous reference picture (also called a P picture);
3. PB: A frame representing two pictures and having a temporally-previous reference picture (see Annex G of ITU-T Recommendation H.263);
4. Improved PB: A frame functionally similar but normally better than a PB frame (see Annex M of ITU-T Recommendation H.263);
5. B: A picture having two reference pictures, one of which temporally precedes the B picture and one of which temporally succeeds the B picture (see Annex O of ITU-T Recommendation H.263);
6. EI: A picture having a temporally simultaneous reference picture (see Annex O of ITU-T Recommendation H.263); and
7. EP: A picture having two reference pictures, one of which temporally precedes the EP picture and one of which is temporally simultaneous (see Annex O of ITU-T Recommendation H.263).

3.36 remote terminal. A terminal endpoint other than the local terminal.

3.37 request. A request is a message that requires a specific action by the receiver, including an immediate response.

3.38 response. A response is a message that is the reply to a request.

3.39 segmentable. The mode of operation defined in ITU-T Recommendation H.223 in which AL-SDUs may be sent in separate multiplex slots carried on one or more MUX-PDUs. See ITU-T Recommendation H.223.

3.40 session. A session is a period of communication between two terminals which may be conversational or non-conversational (for example retrieval from a database).

3.41 signal-to-noise ratio (SNR). Three exact definitions are in common use in connection with image coding. All yield the same nominal value when taking the difference between signal-to-noise measurements in order to compare two different ways of encoding the same picture (because the differences are only in the calculation of the numerator of the argument of the logarithm, which is the same for all representations of the same picture).

1. $PSNR = 10 \log_{10} (255 \cdot 255 / MSE)$ dB

This is the most common use in image processing, and is more properly called "Peak-SNR", but many just call it SNR. The $255 \cdot 255$ come from the square of the range of possible values in an 8-bit pixel (0 to 255).

2. $SNR = 10 \log_{10} (ENERGY / MSE)$ dB

Where ENERGY is the average of the squared values of the pixels in the original picture. This is used by those who take the meaning of the term SNR literally.

$$3. \text{ SNR} = 10 \log_{10} (\text{ENERGY} - \text{MEAN} * \text{MEAN}) / \text{MSE} \text{ dB}$$

Where MEAN is the average of the pixel values in a picture. This is used by those who take the meaning of the term SNR literally and think the overall brightness of the picture should not change their measurement number.

3.42 slave terminal. (See note in 3.22) A slave terminal is the terminal that is determined as being the slave terminal by the master/slave determination procedure defined in this standard.

3.43 support. The ability to operate in a given mode. However, a requirement to “support” a mode does not mean that the mode must actually be used at all times. Unless prohibited, other modes may be used by mutual negotiation.

3.44 terminal. A terminal is any endpoint and may be a user's terminal or some other communication system such as an MCU or an information server.

3.45 uni-directional logical channel. A uni-directional logical channel is a path for the transmission of a single elementary stream from one terminal to another.

3.46 UserInputIndication message. An available message for transport of user input alphanumeric characters from a keypad or keyboard, equivalent to the DTMF signals used in analog telephony.

3.47 videophone. A terminal capable of sending and receiving audio and video information simultaneously.

3.48 Mean Squared Error. The Mean Square Error calculation compares video frames on a pixel-by-pixel basis, and summarizes the differences over all pixels of interest. Thus, for a specific pair of frames (e.g., one from the input sequence and one from the output sequence) the Mean Square Error (MSE) is

$$\text{MSE} = \frac{1}{K_s} \sum_{j=J_{\min}}^{J_{\max}} \sum_{i=I_{\min}}^{I_{\max}} [V'(i, j) - V(i, j)]^2$$

where $V'(i, j)$ is the value of pixel i, j in the output frame, and $V(i, j)$ is the value of pixel i, j in the input frame. K_s is the total number of pixels in the rectangular frame of interest, given by

$$K_s = (I_{\max} - I_{\min} + 1) \times (J_{\max} - J_{\min} + 1)$$

4 Abbreviations

For the purposes of this standard, the following symbols and abbreviations apply.

AAL	ATM Adaptation Layer
AL-SDU	Adaptation Layer Service Data Unit (see ITU-T Recommendation H.223).
AL1,2,3	Adaptation Layers 1, 2 and 3 as defined in ITU-T Recommendation H.223.
ASN.1	Abstract Syntax Notation 1
B-LCSE	Bi-directional Logical Channel Signalling Entity
CESE	Capability Exchange Signalling Entity
CLCSE	Close Logical Channel Signalling Entity
CIF	Common Intermediate Format (of a video picture: refer to ITU-T Recommendations H.261 and H.263)
CRC	Cyclic Redundancy Check

DTMF	Dual tone multi-frequency
DSM-CC	Digital Storage Media - command and control
EIV	Encryption Initialization Vector
GSTN	General Switched Telephone Network (= PSTN)
HDLC	High-level Data Link Control, per ISO/IEC 3309
ISDN	Integrated Services Digital Network
ISO/IEC	International Organization for Standardization/International Electrotechnical Commission
ITU-T	International Telecommunications Union – Telecommunications Standardization Sector
IV	Initialization Vector (used for encryption: refer to ITU-T Recommendations H.233 and H.234)
LAPM	Link Access Procedures for Modems (per ITU-T Recommendation V.42)
LCN	Logical Channel Number (per ITU-T Recommendation H.223)
LCSE	Logical Channel Signalling Entity
MCU	Multipoint Control Unit
MLSE	Maintenance Loop Signalling Entity
MPI	Minimum Picture Interval
MRSE	Mode Request Signalling Entity
MSDSE	Master/Slave Determination Signalling Entity
MSE	Mean Squared Error
MTSE	Multiplex Table Signalling Entity
NLPID	Network Layer Protocol Identifier (per ISO/IEC TR 9577)
PSTN	Public Switched Telephone Network (= GSTN)
QCIF	Quarter CIF
RMESE	Request Multiplex Entry Signalling Entity
RTDSE	Round Trip Delay Signalling Entity
SDL	Specification and Description Language
SDU	Service Data Unit
SE	Session Exchange message (used for encryption: refer to ITU-T Recommendations H.233 and H.234)
SQCIF	Sub QCIF
SNR	Signal-to-Noise Ratio (see definitions in §3)
SPIFF	Still Picture Interchange File Format
SRP	Simple Retransmission Protocol (see Annex A of ITU-T Recommendation H.324)

5 Conventions

The word “shall” is used in this and referenced recommendations to specify a mandatory requirement.

The word “should” is used in this and referenced recommendations to specify a suggested, but not required, course of action.

In this standard elements or operations are identified which provide supporting features or functionality but are not “shall” or “should” requirements. Such identified items are noted as “is optional.”

Throughout this document, GSTN and PSTN are used interchangeably.

6 Arrangements and architecture

Basic Features of the GSTN Multimedia Terminal are listed in table 1:

Table 1 - Basic features of multi-media terminal with baseline components (Implementing ITU-T Recommendations H.263 & G.723.1)

Multi-media Applications		
Video	Variable with content, compression-dependent moving-color video	From less than 1 frame/sec to 30 frames/sec. maximum
Audio	Speech grade comparable to analog telephony	
Data	Flexible data applications constrained to rates consistent with a maximum compatibility and mix of audio and video applications as in ITU-T Recommendation V.34	
Video Input (When provided - Video Capability is Optional)		
Non-interlaced		
525-line system accommodated		
625-line system accommodated		
Frame rate: 30000/1001—>29.97 frame/sec		
Video comprises luminance (Y) and two color difference signals C_B and C_R per CCIR Rec. 601-3		
Picture Formats – pels		
	luminance	chrominance
Sub-QCIF	128x96	64x48
QCIF	176x144	88x72
CIF	352x288	176x144
4CIF	704x576	352x288
16CIF	1408x704	704x576

(continued)

Table 1 (continued)

Pixel Aspect Ratio		Pixel Width : Pixel Height
Square		1:1
CIF		12:11
525-type for 4:3 picture		10:11
CIF for 16:9 picture		16:11
525-type for 16:9 picture		40:33
Extended PAR		m:n, m & n are relatively prime
Video Timing		
Latency: Variable – For Further Study – 300 ms max. suggested		
Audio Input (When provided - Audio Capability is Optional)		
Voice frequency signal	300 – 3400 Hz – Codec dependent – G.722 wideband is allowed	
Input range	-30 – +3 dBm0	
Audio Latency		
Nominally a budget of 150.5 ms (for G.723.1) plus optional Receive path delay for Audio/Visual “Lip Synchronization”		
Audio Bitrates with Required G.723.1 Codec:		
5.3 kb/s	Mandatory	
6.3 kb/s	Mandatory	
Frames	30 ms	
Look Ahead	7.5 ms	
Latency (for a bit rate of 6.4kbps):		
a) Encoder	88 ms	
Comprising:		
Algorithmic	37.5 ms due to:	
Frame size:	30 ms (Time to accumulate Audio Samples)	
Look Ahead	7.5 ms	
Processing	25 ms	
Multiplex	1 ms (Time coded audio waits before transmission)	
Transmission	7 ms (Bits in coded frame divided by bitrate)	
Transmit Modulation	17.5ms (For a V.34 modem @ 28800bps)	

(continued)

Table 1 (concluded)

b) Decoder	62.5 ms
Comprising:	
Receive demodulation	17.5 ms (For a V.34 modem @ 28800bps)
Processing	5 ms
Buffering	40 ms
Algorithmic delay	97.5 ms
a) Encoder	
Comprising:	
Encoder buffering	30 ms (Time to acquire one frame)
Look ahead	7.5 ms
b) Decoder	
Comprising:	
Receive one frame	30 ms
Decode	30 ms
Change bitrates at any frame boundary	
Other ITU approved codecs are Optional; proprietary codecs may be signaled	
Data Input - Data input arrangements are optional	
Data carried in logical channels with required input interfacing according to ITU-T Recommendation H.223	
Multiple logical channels permitted consistent with multimedia trade-off and throughput capacity of transport specified in ITU-T Recommendation V.34	
Logical channels can be opened or closed during a call with an appropriate control mechanism consistent with ITU-T Recommendation H.245	

The ITU-T H.324 suite of recommendations are adopted and shall provide a standardized arrangement for multi-media communication on the GSTN, utilizing signals compatible with the analog interface to the GSTN. It is assumed that this interface will not be more limiting than that provided for conventional analog telephony.

For transmission of multi-media signals, digital concepts are most reliable and provide for a flexible architecture for combining communication signals from various media. Standardization of high-speed modems for digital signal communication over the GSTN, such as in ITU-T Recommendation V.34, provides the underlying transport for multi-media arrangements. Certain multi-media signals such as video and analog speech and music signals must be encoded to utilize a digital architecture although music and other audio signals are not reproduced as faithfully as speech. Through use of optional delay in the received speech path, "Lip Synchronization" should be controlled. While computer-originated data streams are in a digital format, such streams may be synchronous or asynchronous and comprise different formats and data rates.

The bit rates that result from non-compressed encoding of analog video and voice/music signals are much larger than the throughput rates achievable with modems specified in ITU-T Recommendation V.34.

Therefore, a standardized multi-media terminal for GSTN applications must include extensive compression-based encoding algorithms and balanced architectural arrangements for efficient and flexible combining of resultant digital streams consistent with capabilities in ITU-T Recommendation V.34. The ITU-T H.324 suite of recommendations achieves this goal and is adopted by Normative Reference and shall be the ANSI T1 standard multi-media terminal arrangement for the GSTN.

The scope of the elements of ITU-T Recommendation H.324 is shown in figure 1. The elements that are subject to this standard are shown within the box denoted by dotted lines.

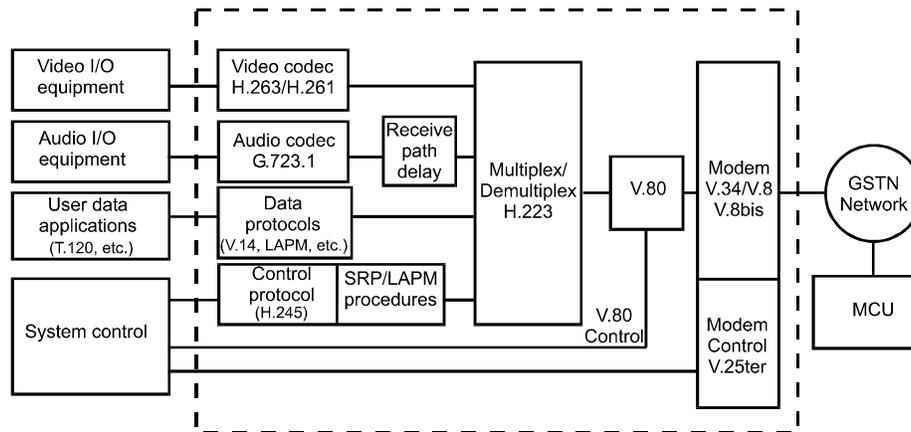


Figure 1 - Elements of ITU-T Recommendation H.324 terminal

NOTE - Other interfaces to telecommunications networks such as ISDN, wireless, etc. are planned. (See clause 11.)

6.1 Required functional elements

ITU-T Recommendation H.324 implementations are not required to have each functional element, except for the modem (ITU-T Recommendation V.34), multiplex (ITU-T Recommendation H.223), and system control protocol (ITU-T Recommendation H.245), all of which shall be supported by all terminals described in ITU-T Recommendation H.324.

Terminals described in ITU-T Recommendation H.324 offering audio communication shall support the audio codec specified in ITU-T Recommendation G.723.1. Those offering video communication shall support the video codecs of ITU-T Recommendations H.263 and H.261. Those offering real-time audiographic conferencing should support the protocol suite in ITU-T Recommendation T.120. In addition, other video and audio codecs, and other data protocols, may optionally be used via negotiation over the control channel described in ITU-T Recommendation H.245.

If a modem external to the ITU-T Recommendation H.324 terminal is used, terminal/modem control shall conform to ITU-T Recommendation V.25ter.

The presence of optional facilities is signaled via the control channel described in ITU-T Recommendation H.245. If both ends support an optional facility, and choose to make use of it, the opening of a path to carry such information streams shall be negotiated according to the procedures of ITU-T Recommendation H.245.

NOTE - This standard does not specify a particular implementation. Any implementation that provides the required functionality, and that conforms to the bitstream format ultimately described by this standard, is considered compliant.

Multimedia Terminals per this standard shall support both the ITU-T Recommendations H.263 and H.261 video codecs in the modes shown in Table 2. Where the terminal is a videophone with both audio and video capabilities, the terminal shall contain arrangements permitting conventional telephone calling over the GSTN network. Details and requirements for such capability are not within the scope of this standard.

6.2 Information streams

Multimedia information streams are classified into video, audio, data, and control as follows:

1. Audio streams are real-time, but it is optional to provide delay in the receiver processing path to maintain synchronization with the video streams. In order to reduce the average bitrate of audio streams, voice activation is optional in the transmitter but receivers shall process silence frames correctly when voice activation is provided. Voice activation is optional in the transmitters.
2. Control streams pass control commands and indications between remote counterparts. Terminal-to-modem control shall conform to ITU-T Recommendation V.25ter for terminals using external modems connected by a separate physical interface. Terminal-to-terminal control shall be according to ITU-T Recommendation H.245.
3. Data streams may represent still pictures, facsimile, documents, computer files, computer application data, undefined user data, and other data streams.
4. Video streams are continuous traffic-carrying moving color pictures. When video is used, it is optional for the bitrate available for video streams to vary according to the needs of the audio and data channels.

6.3 Modem

Modems used for terminals specified in ITU-T Recommendation H.324 shall operate in full duplex, synchronous mode and conform to ITU-T Recommendations V.34 and V.8. Support of ITU-T Recommendation V.8bis is optional. The output of the multiplex specified in ITU-T Recommendation H.223 shall be applied directly to the synchronous data pump of ITU-T Recommendation V.34. When an external, non-integrated modem of ITU-T Recommendation V.34 is used, control between the modem and the terminal shall be via ITU-T Recommendation V.25ter; if external asynchronous modems are used, the principles of ITU-T Recommendation V.80 shall be implemented. In such cases the physical interface is implementation specific. The use of the optional auxiliary channel from ITU-T Recommendation V.34 is reserved for further study.

6.4 GSTN network Interface

Terminals defined by this suite of recommendations which will connect to the North American GSTN shall meet all applicable US FCC Part 15 and Part 68 requirements and/or Canadian DOC and CSA specifications as appropriate.

6.5 Multiplex

The multiplex of ITU-T Recommendation H.223 consists of a multiplex layer, which mixes the various logical channels into a single bitstream, and an adaptation layer, which handles error control and sequence numbering, as appropriate to each information stream. The multiplex layer transfers logical channel information in packets called MUX-PDUs, delimited by HDLC flags and using HDLC zero-bit insertion for transparency. Each MUX-PDU contains a one-octet header followed by a variable number of information field octets. The header octet includes a multiplex code, which specifies, by reference to a multiplex table, the mapping of the information field octets to various logical channels. Each MUX-PDU may contain a different multiplex code, and therefore a different mix of logical channels.

6.6 Control channel

The control channel carries end-to-end control messages governing operation of the system in ITU-T Recommendation H.324, including capabilities exchange, opening and closing of logical channels, mode preference requests, multiplex table entry transmission, flow control messages, and general commands and indications.

There shall be exactly one control channel in each direction within terminals of ITU-T Recommendation H.324. The control channels shall use the messages and procedures of ITU-T Recommendation H.245. The control channel shall be carried on logical channel 0. The control channel shall be considered to be permanently open from the establishment of digital communication until the termination of digital commu-

nication; the normal procedures for opening and closing logical channels shall not apply to the control channel.

General commands and indications shall be chosen from the message set contained in ITU-T Recommendation H.245. In addition other optional command and indication signals which have been specifically defined to be transferred in-band within video, audio or data streams (see the appropriate Recommendation to determine if such signals have been defined) and whether they are mandatory or optional in the context of those Recommendations.

Messages in ITU-T Recommendation H.245 fall into four categories: Request, Response, Command, and Indication. Request messages require a specific action by the receiver, including an immediate response. Response messages respond to a corresponding request. Command messages require a specific action, but do not require a response. Indication messages are informative only, and do not require any action or response. Terminals of ITU-T Recommendation H.324 shall respond to all supported commands and requests specified in ITU-T Recommendation H.245, and shall transmit accurate indications reflecting the state of the terminal.

NOTE - All control channel messages are sent over a link layer protocol which acknowledges correct receipt. This acknowledgment is distinct from the response messages, which convey content beyond that of correct receipt of the message.

Terminals specified in ITU-T Recommendation H.324 shall be capable of parsing all **MultimediaSystemControlPDU** messages described in ITU-T Recommendation H.245, and shall send and receive all messages needed to implement functions required by ITU-T Recommendation H.324 and those optional functions which are supported by the terminal. All messages and procedures of ITU-T Recommendation H.245 related to functions required by ITU-T Recommendation H.324 are required, except for those explicitly described as optional, or which are related to defined optional capabilities the terminal does not support. Terminals specified in ITU-T Recommendation H.324 shall send the **FunctionNotSupported** message in response to unrecognized request, response, or command messages.

NOTE - The semantics of Recommendation H.245 shall be followed.

A control channel indication, **UserInputIndication**, shall be available for transport of user input alphanumeric characters from a keypad or keyboard, equivalent to the DTMF signals used in analog telephony. This may be used to manually operate remote equipment such as voice mail or video mail systems, menu-driven information services, etc. Terminals described in ITU-T Recommendation H.324 shall support the transmission of user input characters 0-9, '*', and '#'. Support for the transmission of other characters is optional.

NOTE - If the encryption procedures of this standard are in use, the control channel will not be encrypted. Users are therefore cautioned regarding the carriage of user data in the control channel, the use of non-standard messages, and the confidentiality risk from traffic analysis of the control channel.

6.7 Video Channel

All terminals described in ITU-T Recommendation H.324 offering video communication shall support both of the video codecs specified in ITU-T Recommendations H.263 and H.261, except that H.320 Interworking Adapters specified in ITU-T Recommendation H.246 Annex B (which are not terminals) do not have to support ITU-T Recommendation H.263 (see section 8.2 of ITU-T Recommendation H.324). The codecs of ITU-T Recommendations H.261 and H.263 shall be used without BCH error correction and without error correction framing. There are five standardized image formats: 16CIF, 4CIF, CIF, QCIF, and SQCIF.

CIF and QCIF are defined in ITU-T Recommendation H.261. For the algorithm of ITU-T Recommendation H.263, SQCIF, 4CIF and 16CIF are defined within the Recommendation. For the algorithm of ITU-T Recommendation H.261, SQCIF is any active picture size less than QCIF, filled out by a black border, and coded in the QCIF format. For all these formats, the pixel aspect ratio is the same as that of the CIF format.

NOTE - The resulting picture aspect ratio for SQCIF in ITU-T Recommendation H.263 is different from the other formats.

Table 2 shows which picture formats are required, and which are optional for ITU-T Recommendation H.324 terminals which support video.

Table 2 - Picture formats for video terminals

Picture Format	Luminance Pixels	Encoder		Decoder	
		H.261	H.263	H.261	H.263
SQCIF	128 x 96 for H.263 ³	Optional ³	Required ^{1,2}	Optional ³	Required ¹
QCIF	176 x 144	Required	Required ^{1,2}	Required	Required ¹
CIF	352 x 288	Optional	Optional	Optional	Optional
4CIF	704 x 576	Not defined	Optional	Not defined	Optional
16CIF	1408 x 1152	Not defined	Optional	Not defined	Optional
<p>NOTE 1 - Optional for interworking adapters of ITU-T Recommendation H.320.</p> <p>NOTE 2 - Mandatory to encode one of the picture formats QCIF and SQCIF; optional to encode both formats.</p> <p>NOTE 3 - SQCIF of ITU-T Recommendation H.261 is any active size less than QCIF, filled out by a black border, coded in QCIF format.</p>					

All video decoders shall be capable of processing video bitstreams of the maximum bitrate which can be received by the implementation of the multiplex in ITU-T Recommendation H.223 (maximum ITU-T Recommendation V.34 rate for single link, 2 x ITU-T Recommendation V.34 rate for double link, etc.).

Identification of particular picture formats, minimum number of skipped pictures, and algorithm options that can be accepted by the decoder is determined during the capability exchange using ITU-T Recommendation H.245. After that, the encoder is free to transmit anything consistent with the decoder's capability. Decoders that indicate capability for a particular algorithm option shall also be capable of accepting video bitstreams representing other options accepted during negotiation even when a particular bitstream does not use a particular decoder option.

When each video logical channel is opened, the maximum operating mode to be used on that channel is signaled to the receiver. The picture header within the video bitstream indicates which mode is actually used for each picture, within the stated maximum. Receivers should signal, via ITU-T Recommendation H.245, a preference for a certain mode.

NOTE - The maximum mode signaled includes maximum picture format, algorithm options, etc. For example, a video logical channel opened for CIF format may transmit CIF, QCIF, or SQCIF pictures, but not 4CIF or 16CIF. A video logical channel opened with only the unrestrictedVector and arithmetic Coding options may use neither, either, or both options, but shall not use options which were not signaled. Implementers are cautioned to review the latest approved versions of ITU-T Recommendation H.263 for information regarding implementation options.

Other video codecs, and other picture formats, are optional via negotiation described in ITU-T Recommendation H.245. More than one video channel is optional, as negotiated via the control channel of ITU-T Recommendation H.245.

NOTE - The method of continuous presence multipoint operation, in which a single picture is divided into multiple sub-pictures, should not be used by terminals specified in ITU-T Recommendation H.324. Instead, multiple logical channels of video should be used.

6.8 Audio channel

All terminals specified in ITU-T Recommendation H.324 offering audio communication shall support both the high and low rates of the audio codec in ITU-T Recommendation G.723.1. Receivers of ITU-T Recommendation G.723.1 shall be capable of decoding silence frames. The choice of high rate, low rate, or silence is made by the transmitter, and is signaled to the receiver in-band in the audio channel, as part of the syntax of each audio frame. Transmitters have the capability to switch rates given in ITU-T Recommendation G.723.1 on a frame-by-frame basis, based on bitrate, audio quality, or other preferences. Receivers should signal, via ITU-T Recommendation H.245, a preference for a particular audio rate or mode.

Alternative audio codecs are optional, via negotiation described in ITU-T Recommendation H.245. Coders may omit sending audio signals during silent periods after sending a single frame of silence, or may send silence background fill frames if such techniques are specified by the audio codec Recommendation in use.

More than one audio channel is optional, as negotiated via the control channel of ITU-T Recommendation H.245.

NOTE - Each audio channel is independent. Grouping of audio channels into stereo pairs or other synchronized groups is for further study.

6.9 Data channel

All data channels are optional. Standardized optional data applications include:

1. ITU-T Recommendation T.120 series for point-to-point and multipoint audiographic teleconferencing including database access, still image transfer and annotation, application sharing, real-time file transfer, etc.
2. ITU-T Recommendation T.84 (SPIFF) point-to-point still image transfer cutting across application borders.
3. ITU-T Recommendation T.123 (1996), *Protocol stack for multimedia conferencing applications* (Normative if the H.324 terminal in question is supporting the optional T.120 capability).
4. ITU-T Recommendation T.434 (1992), *Binary file transfer format for the telematic services* (Normative if the ITU-T Recommendation H.324 terminal in question is supporting the optional point-to-point file transfer cutting across application borders).
5. ITU-T Recommendation H.224 (1994), *A real time control protocol for simplex applications using the ITU-T Recommendation H.221 LSD/HSD/MLP channels* (Normative when the ITU-T Recommendation H.324 terminal in question is supporting real-time control of simplex applications, including ITU-T Recommendation H.281 Far End Camera Control).
6. Network link layer, per ISO/IEC TR 9577 (supports IP and PPP network layers, among others).
7. Unspecified user data from external data ports
8. T.30 facsimile transfer
9. T.140 text conversation protocol

The mechanism for data applications are implementation dependent. These data applications may reside in an external computer or other dedicated device attached to the ITU-T Recommendation H.324 terminal through an ITU-T Recommendation V.24 or equivalent interface (implementation dependent), or may be integrated into the ITU-T Recommendation H.324 terminal itself. Each data application makes use of an underlying data protocol for link layer transport. For each optional data application supported by the ITU-T Recommendation H.324 terminal, this recommendation requires support for the corresponding appropriate underlying data protocol to ensure correct interworking of data applications.

NOTE - The control channel described in ITU-T Recommendation H.245 is not considered a data channel.

Standardized link layer data protocols used by data applications include:

1. Buffered ITU-T Recommendation V.14 mode for transfer of asynchronous characters, without error control.
2. LAPM/ITU-T Recommendation V.42 for error-corrected transfer of asynchronous characters. Additionally, depending on application, ITU-T Recommendation V.42bis data compression may be used.
3. HDLC frame tunneling for transfer of HDLC frames.
4. Transparent data mode for direct access by unframed or self-framed protocols.

ITU-T Recommendation H.324 terminals offering real-time audiographic conferencing should support the T.120 protocol suite.

Data protocols shall operate within logical channels as specified in ITU-T Recommendation H.223. Protocol procedures referring to link establishment or link termination (including setup and disconnection of physical channels particular to a given protocol) shall be interpreted as referring to opening and closing of logical channels, and shall not affect the physical link of ITU-T Recommendation H.324. For protocol procedures which distinguish between an originator and an answerer, the master terminal of ITU-T Recommendation H.324, determined according to the **MasterSlaveDetermination** procedure of ITU-T Recommendation H.245, shall be the originator, and the slave terminal shall be the answerer.

More than one data channel, or more than one protocol may be used at the same time (each in a separate logical channel), as negotiated via the control channel of ITU-T Recommendation H.245. Other data protocols and applications are optional via negotiation as specified in ITU-T Recommendation H.245.

7 Control messaging and indication

Procedures are defined to allow the exchange of audiovisual and data capabilities; to request the transmission of a particular audiovisual and data mode; to manage the logical channels used to transport the audiovisual and data information; to establish which terminal is the master terminal and which is the slave terminal for the purposes of managing logical channels; to carry various control and indication signals; to control the bit rate of individual logical channels and the whole multiplex; and to measure the round trip delay, from one terminal to the other and back. These procedures are explained in more detail below.

Following this general introduction, there are sections detailing the message syntax and semantics and the procedures. The syntax has been defined using ASN.1 notation and the semantics define the meaning of syntax elements as well as providing syntactic constraints that are not specified in the ASN.1 syntax. The procedures section defines the protocols that use the messages defined in the other sections.

Not all of the messages and procedures defined in ITU-T Recommendation H.245 will be applicable to all terminals, and no indication of such restrictions is given here. These restrictions are individual to the applications and the governing recommendations that utilize ITU-T Recommendation H.245.

ITU-T Recommendation H.245 has been defined to be independent of the underlying transport mechanism, but is intended to be used with a reliable transport layer, that is, one that provides guaranteed delivery of correct data.

7.1 Semantic definitions

This section provides semantic definitions and constraints on the syntax elements introduced in the previous section.

MultimediaSystemControlMessage: is a choice of message types. Messages defined in ITU-T Recommendation H.245 are classified as request, response, command and indication messages.

RequestMessage: a request message results in an action by the remote terminal and requires an immediate response from it. The nonStandard message may be used to send non-standard requests.

ResponseMessage: a response message is the response to a request message. The nonStandard message may be used to send non-standard responses.

CommandMessage: a command message requires action but no explicit response. The nonStandard message may be used to send non-standard commands.

IndicationMessage: an indication contains information that does not require action or response. The nonStandard message may be used to send non-standard indications.

NonStandardParameter: this may be used to indicate a non-standard parameter. It consists of an identity and the actual parameters, which are coded as an octet string.

NonStandardIdentifier: is used to identify the type of non-standard parameter. It is either an object identifier, or a type of identifier described in ITU-T Recommendation H.221 that is an octet string consisting of exactly four octets which are country code (octet 1 as in T.35; octet 2*), manufacturer code (next two octets*), *=assigned nationally. The manufacturer codes shall be the same as those assigned for use in ITU-T Recommendation H.320, Contact the Secretariat of Accredited Standards Committee T1 for information concerning assignment in the United States.

7.1.1 Master-slave determination

Conflicts may arise when two terminals involved in a call initiate similar events simultaneously and only one such event is possible or desired, for example, when resources are available for only one occurrence of the event. Normally each terminal is considered to have peer precedence with respect to control messages; however, there are rare instances when the terminals are in contention (a "race" condition). To resolve such contention conflicts, one terminal shall act as a master and the other terminal shall act as a slave terminal. Rules specify how the master and slave terminal shall respond at times of conflict.

The master-slave determination procedure allow terminals in a call to predetermine which terminal is the master and which terminal is the slave in a contention situation which may arise during a call. The terminal status, once determined, remains constant for the call duration.

7.1.2 Capability exchange

The capability exchange procedures are intended to ensure that the only multimedia signals to be transmitted are those that can be received and treated appropriately by the receiving terminal. This requires that the ability of each terminal to receive and decode as well as the ability to transmit shall be known to the other terminal. It is not necessary that a terminal understand or store all in-coming capabilities; those that are not understood, or cannot be used, shall be ignored, and no fault shall be considered to have occurred.

The total capability of a terminal to receive and decode various signals is made known to the other terminal by transmission of its capability set.

Receive capabilities describe the terminal's ability to receive and process in-coming information streams. Transmitters shall limit the content of their transmitted information to that which the receiver has indicated

it is capable of receiving. The absence of a receive capability indicates that the terminal cannot receive (is a transmitter only).

Transmit capabilities describe the terminal's ability to transmit information streams. Transmit capabilities serve to offer receivers a choice of possible modes of operation, so that the receiver should request the mode which it prefers to receive. The absence of a transmit capability indicates that the terminal is not offering a choice of preferred modes to the receiver (but it may still transmit appropriate to the capability of the receiver).

These capability sets provide for more than one stream of a given medium type to be sent simultaneously. For example, a terminal may declare its ability to receive (or send) two independent H.262 video streams and two independent G.723.1 audio streams at the same time. Capability messages have been defined to allow a terminal to indicate that it does not have fixed capabilities, but that they depend on which other modes are being used simultaneously. For example, it is possible to indicate that higher resolution video can be decoded when a simpler audio algorithm is used; or that either two low resolution video sequences can be decoded or a single high resolution one. It is also possible to indicate trade-offs between the capability to transmit and the capability to receive.

Non-standard capabilities and control messages are optional and are issued using the NonStandard-Parameter structure. Note that while the meaning of non-standard messages is defined by individual organizations, equipment built by any manufacturer may signal any non-standard message, if the meaning is known.

Terminals may reissue modified capability sets at any time.

7.1.3 Logical channel signalling procedures

An acknowledged protocol is defined for the opening and closing of logical channels which carry the audiovisual and data information. The aim of these procedures is to ensure that a terminal is capable of receiving and decoding the data that will be transmitted on a logical channel at the time the logical channel is opened rather than at the time the first data is transmitted on it; and to ensure that the receive terminal is ready to receive and decode the data that will be transmitted on the logical channel before that transmission starts. The **OpenLogicalChannel** message includes a description of the data to be transported; for example, logical channels should only be opened when there is sufficient capability to receive data on all open logical channels simultaneously except for use of the new "**ReplacementFor**" parameter.

A part of this protocol is concerned with the opening of bi-directional channels. To avoid conflicts during a call, which may arise when two terminals initiate similar events simultaneously, at the beginning of a call one terminal is randomly defined as the master terminal, and the other as the slave terminal. A protocol defined in this standard shall determine which terminal is the master and which is the slave.

7.1.4 Receive terminal close logical channel request

A logical channel is always opened and closed from the transmitter side. A mechanism is defined which allows a receive terminal to request either the opening or the closure of an incoming logical channel. The transmit terminal shall accept or reject the logical channel closure request. A terminal should, for example, use these procedures to request the closure of an incoming logical channel which, for whatever reason, cannot be decoded. These procedures should also be used to request the closure of a bi-directional logical channel by the terminal that did not open the channel.

7.1.5 ITU-T Recommendation H.223 multiplex table entry modification

The multiplex table described in ITU-T Recommendation H.223 associates each octet within a MUX message with a particular logical channel number. The multiplex table can have up to 15 entries. A mechanism is provided that allows the transmit terminal to specify and inform the receiver of new multiplex table entries. A receive terminal can also request the retransmission of a multiplex table entry PDU.

7.1.6 Audiovisual and data mode request

When the capability exchange protocol has been completed, both terminals will be aware of each other's capability to transmit and receive as specified in the capability descriptors that have been exchanged. It is not mandatory for a terminal to declare all its capabilities; it need only declare those that it wishes to be used.

As noted above a terminal can indicate its capabilities to transmit. A terminal that receives transmission capabilities from the remote terminal can request a particular mode to be transmitted to it. A terminal indicates that it does not want its transmission mode to be requested by the remote terminal by sending no transmission capabilities.

7.1.7 Round trip delay determination

It may be useful in some applications to have knowledge of the round trip delay between a transmit terminal and a receive terminal. A mechanism is provided to measure this round trip delay. Additional guidelines are provided in H.324 Annex E. This mechanism may also be useful as a means to detect whether the remote terminal is still functioning.

7.1.8 Maintenance loops

Procedures are specified to establish maintenance loops. It is possible to specify the loop of a single logical channel either as a digital loop or decoded loop, and the loop of the whole multiplex is for further study.

7.1.9 Commands and indications

Commands and indications are provided for various purposes; some examples are: video/audio active/inactive signals to inform the user; fast update request for source switching in multipoint applications. Neither commands nor indications elicit response messages from the remote terminal. Commands force an action at the remote terminal whilst indications merely provide information and do not force any action.

A specific command, Flow Control Command, is defined to allow the bit rate of logical channels and of the whole multiplex to be controlled from the remote terminal. This has a number of purposes; for example: interworking with terminals using multiplexes in which only a finite number of bit rates are available; multipoint applications where the rates from different sources should be matched; and flow control in congested networks.

7.2 Messaging

7.2.1 Master/slave determination messages

This set of messages is used by a protocol to determine which terminal is the master terminal and which is the slave terminal.

7.2.2 Terminal capability messages

This set of messages is for the exchange of capabilities between the two terminals.

7.2.3 Logical channel signalling messages

This set of messages is for logical channel signalling. The same set of messages is used for uni-directional and bi-directional logical channel pair signalling; however, some parameters are only present in the case of bi-directional logical channel pair signalling.

'Forward' is used to refer to transmission in the direction from the terminal making the original request for a logical channel to the other terminal, and 'reverse' is used to refer to the opposite direction of transmission, in the case of a bi-directional channel request.

7.2.4 Multiplex table signalling messages

This set of messages is for the secure transmission of multiplex table entries (described in ITU-T Recommendation H.223) from the transmitter to the receiver.

7.2.5 Request multiplex table signalling messages

This set of messages is for the secure request of retransmission of one or more **MultiplexEntryDescriptors** from the transmitter to the receiver.

7.2.6 Request mode messages

This set of messages is used by a receive terminal to request particular modes of transmission from the transmit terminal.

7.2.7 Round trip delay messages

This set of messages is used by a terminal to determine the round trip delay between two communicating terminals. It also enables a user described in ITU-T Recommendation H.245 to determine whether the peer protocol entity is alive.

7.2.8 Maintenance loop messages

This set of messages is used by a terminal to perform maintenance loop functions.

7.2.9 Commands

A command message requires action but no explicit response.

7.3 Indication

7.3.1 Indications

An indication contains information that does not require action or response.

8 Multiplex

ITU-T Recommendation H.223 specifies the frame structure, format of fields and procedures of the multiplexing protocol for low bitrate multimedia communication. This protocol can be used between two low bitrate multimedia terminals, or between a low bitrate multimedia terminal and a multipoint control unit (MCU) or an interworking adapter (IWA). The control procedures necessary to implement this multiplexing protocol are specified in ITU-T Recommendation H.245.

The general structure of the multiplexer is shown in figure 2. The multiplexer consists of two distinct layers: a Multiplex (MUX) layer and an Adaptation Layer (AL).

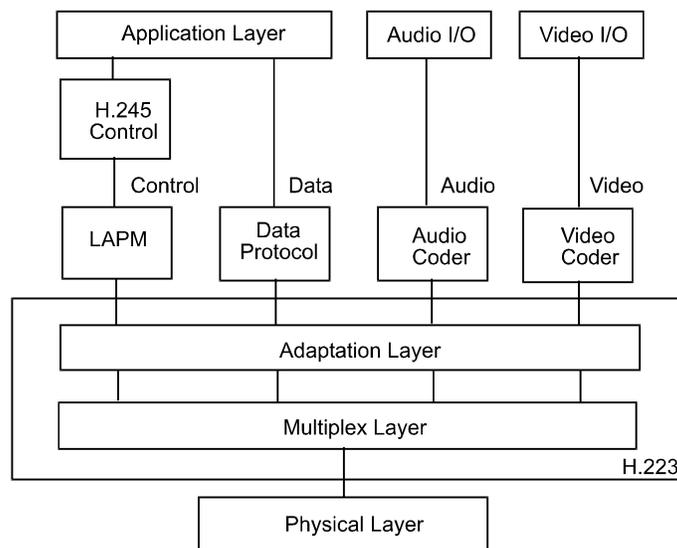


Figure 2 - Protocol stack for ITU-T Recommendation H.223

As an example of how a Packet Data Unit (PDU) would be made up, the following shows how a short audio MUX-SDU, (possibly representing background noise information sent during a silence period), may be mixed using a repeating pattern of data and video. Figure 3 shows an example of the construction of the information field from data obtained from the AL layer.

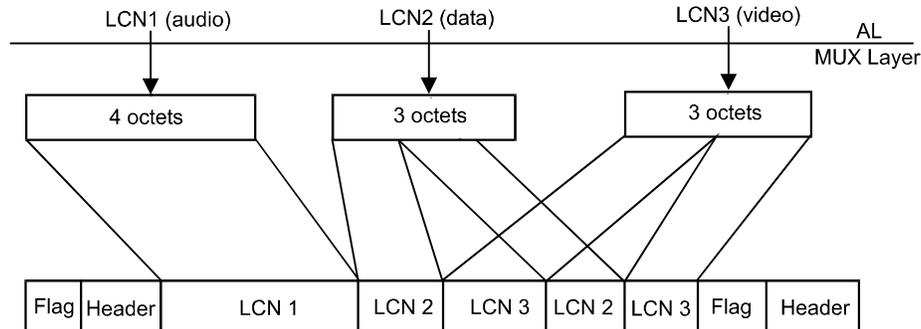


Figure 3 – Information field example

In ITU-T Recommendation H.223, communication between different protocol layers is modeled as a set of abstract primitives, which represent a logical exchange of information. The decomposition of functionality into (sub) layers, as well as the description of the primitives, does not imply a particular method of implementation. In particular, layers may, and usually should, exchange the contents of a logical unit (an SDU) in a "streaming" mode where information exchange may start before the transferring layer has the complete unit in its possession.

9 Video codec

ITU-T Recommendation H.263 specifies a coded representation that can be used for compressing the moving picture component of audio-visual services at low bitrates. The basic configuration of the video source-coding algorithm is based on ITU-T Recommendation H.261. Four negotiable coding options are included for improved performance. These are identified below. Details are available in ITU-T Recommendation H.263 and Annexes.

An outline block diagram of the codec is given in figure 4.

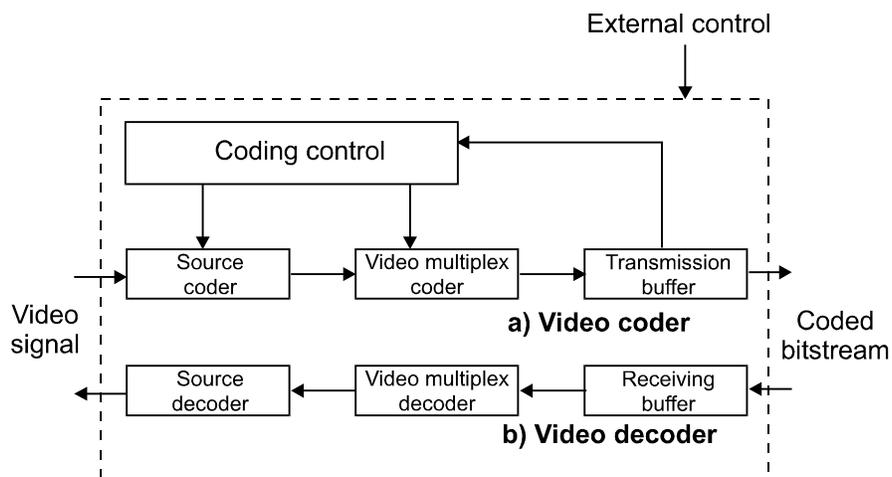


Figure 4 – Simplified block diagram of video codec

9.1 Source coding algorithm and modes

A hybrid of inter-picture prediction to utilize temporal redundancy and transform coding of the remaining signal to reduce spatial redundancy is specified in ITU-T Recommendation H.263. The decoder has motion compensation capability, allowing optional incorporation of this technique in the coder. Half pixel precision is used for the motion compensation, as opposed to ITU-T Recommendation H.261 where full pixel precision and a loop filter are used. Variable length coding is used for the symbols to be transmitted.

In addition to the core coding algorithm of ITU-T Recommendation H.263, four negotiable coding options are included that will be identified in the subsequent subsections. All these options can be used together or separately.

9.1.1 Unrestricted motion vector mode

In this optional mode motion vectors are allowed to point outside the picture. The edge pixels are used as prediction for the “not existing” pixels. With this mode a significant gain is achieved if there is movement across the edges of the picture, especially for the smaller picture formats. Additionally, this mode includes an extension of the motion vector range so that larger motion vectors can be used. This is especially useful in case of camera movement.

9.1.2 Syntax-based arithmetic coding mode

In this optional mode arithmetic coding is used instead of variable length coding. The SNR and reconstructed pictures will be the same, but somewhat fewer bits will be produced.

9.1.3 Advanced prediction mode

In this optional mode overlapped block motion compensation (OBMC) is used for the luminance part of P-pictures (predicted from the previously decoded P-Picture). Four 8x8 vectors instead of one 16x16 vector are used for some of the macroblocks in the picture. The encoder has to decide which type of vectors to use. Four vectors use more bits, but give better prediction. The use of this mode generally gives a considerable improvement. A subjective gain is achieved because OBMC results in less blocking artifacts.

9.1.4 PB-frames mode

A PB-frame consists of two pictures being coded as one unit. The name PB comes from the name of picture types in ITU-T Recommendation H.262 where there are P-pictures and B-pictures. Thus a PB-frame consists of one P-picture which is predicted from the previous decoded P-picture and one B-picture which is predicted from both the previous decoded P-picture and the P-picture currently being decoded. The name B-picture was chosen because parts of B-pictures may be bidirectionally predicted from the past and future pictures. With this coding option, the picture rate can be increased considerably without a major increase in the bitrate, but it adds one frame of delay.

9.1.5 ITU-T Recommendation H.263 reference picture selection mode support

ITU-T Recommendation H.263, Annex N, reference picture selection mode, may optionally be supported. In this mode, video backchannel messages may be mixed with the video data in the opposite direction according to ITU-T Recommendation H.263, or video backchannel messages may be carried on an additional separate logical channel.

In the case where ITU-T Recommendation H.263, Annex N video backchannel messages are carried on a separate LC, the procedure to set up the LC for the video backward channel is different depending on whether the video communication is uni-directional or bi-directional.

In the case of bi-directional video communication, the LCs for original video data shall first be opened as bi-directional LCs which shall support ITU-T Recommendation H.223 AL3. The LCs for the video backchannel messages shall next be opened by the terminal which originated the video LCs. The backchannel LCs shall be opened as bi-directional LCs with the LC dependency parameters indicating dependency on the corresponding video LC flowing in the same direction. The backchannel LCs shall support ITU-T Recommendation H.223 AL2. Until the backchannel LCs are established, the terminal shall not send any video data which requests backchannel messages.

In the case of uni-directional video communication, a single bi-directional LC shall be opened, which shall support ITU-T Recommendation H.223 AL3. Variable length Stuffing (BSTUF) defined in ITU-T Recommendation H.263 shall be used to make all the backward channel messages whole numbers of octets in length.

9.2 Input and output

9.2.1 Video input and output

To permit a single standard to cover use in and between regions using 625- and 525-line television standards respectively, the source coder operates on pictures based on a common intermediate format (CIF). The standards of the input and output television signals, which may, for example, be composite or component, analog or digital and the methods of performing any necessary conversion to and from the source coding format are outside the scope of this standard.

9.2.2 Digital output and input

The video coder produces a self-contained digital bit stream output which may be combined with other bit streams (for example as defined in ITU-T Recommendation H.223). The video decoder performs the reverse process on an input bitstream resulting from a corresponding demultiplexing process.

9.3 Source format

The source coder operates on non-interlaced pictures occurring 30 000/1001 (approximately 29.97) times per second. The tolerance on picture frequency is ± 50 ppm.

Pictures are coded as luminance and two color difference components (Y , C_B and C_R). These components and the codes representing their sampled values are as defined in CCIR Recommendation 601.

1. Black = 16
2. White = 235
3. Zero color difference = 128
4. Peak color difference = 16 and 240.
5. These values are nominal ones and the coding algorithm functions with input values of 1 through to 254.

There are five standardized picture formats: sub-QCIF, QCIF, CIF, 4CIF and 16CIF. For each of these picture formats, the luminance sampling structure is dx pixels per line, dy lines per picture in an orthogonal arrangement. Sampling of each of the two color difference components is at $dx/2$ pixels per line, $dy/2$ lines per picture, orthogonal. The values of dx , dy , $dx/2$ and $dy/2$ are given in Table 3 for each of the picture formats.

Table 3 – Number of pixels per line and number of lines for each of the ITU-T Recommendation H.263 picture formats

Picture format	Number of pixels for luminance (dx)	Number of lines for luminance (dy)	Number of pixels for chrominance (dx/2)	Number of lines for chrominance (dy/2)
sub-QCIF	128	96	64	48
QCIF	176	144	88	72
CIF	352	288	176	144
4CIF	704	576	352	288
16CIF	1408	1152	704	576

10 Audio codec

ITU-T Recommendation G.723.1 specifies a codec that can be used for compressing the speech signal component of multimedia services at a very low bit rate. In the design of this coder, the principal application considered was very low bit rate visual telephony as part of the overall family of standards in ITU-T Recommendation H.324.

10.1 Speech coder description

The description of the speech coding algorithm of ITU-T Recommendation G.723.1 is made in terms of bit-exact, fixed-point mathematical operations. The ANSI C code indicated in Section 5 of that Recommendation, constitutes an integral part of the Recommendation and reflects this bit-exact, fixed-point description approach. The mathematical descriptions of the encoder and decoder, given respectively in Sections 2 and 3 of ITU-T Recommendation G.723.1, can be implemented in several other fashions, possibly not sufficient and leading to a non-compliant codec implementation. Therefore, the algorithm description of the C code of Section 5 shall be definitive as compared to the mathematical descriptions of Sections 2 and 3. A non-exhaustive set of test sequences which can be used in conjunction with the C code are available from the ITU.

10.2 Bit rates

This coder has two bit rates associated with it. These are 5.3 and 6.3 kbit/s. The higher bit rate has greater quality. The lower bit rate gives good quality and provides system designers with additional flexibility. Both rates shall be a mandatory part of the encoder and decoder. It shall be possible to switch between the two rates at any 30 ms frame boundary. Variable rate operation (as a result of silence suppression) using discontinuous transmission and noise fill during non-speech intervals is optional in ITU-T Recommendation G.723.1 but mandatory when implemented in ITU-T Recommendation H.324.

10.3 Delay

10.3.1 Algorithmic delay

See Table 1, Algorithmic Delay.

10.3.2 Latency

Latency is defined to be the Algorithmic delay plus delay due to system aspects, which is nominally budgeted to be 150.5 ms.

See Table 1, Audio Latency.

10.3.3 Block diagrams

10.3.3.1 Encoder

A block diagram of the encoder is shown in figure 5.

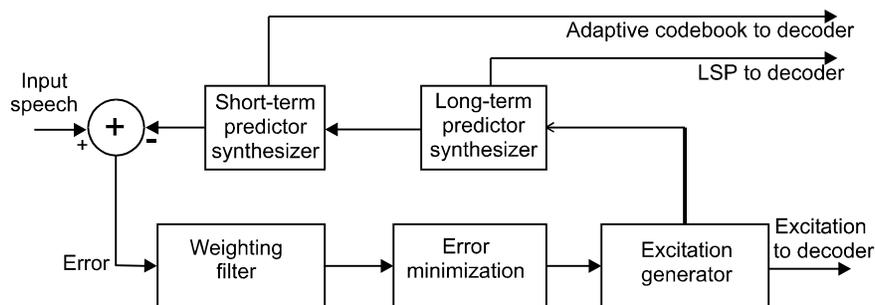


Figure 5 - Simplified block diagram of ITU-T Recommendation G.723.1 Encoder

10.3.3.2 Decoder

A block diagram of the decoder is shown in figure 6.

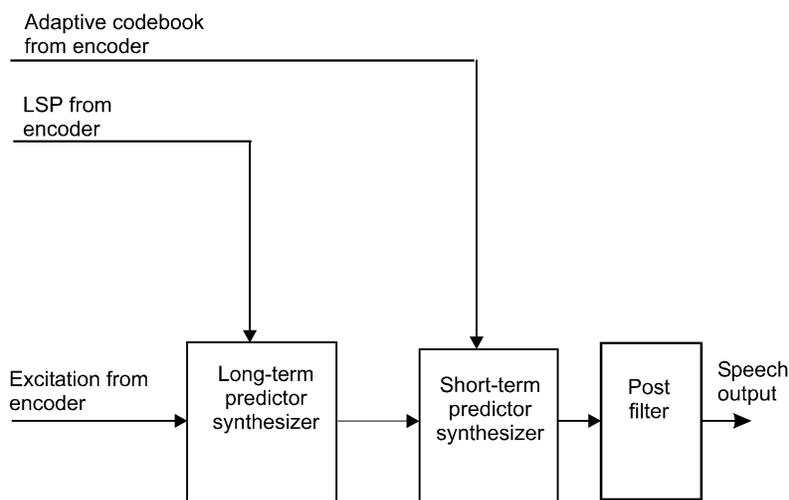


Figure 6 - Simplified block diagram of ITU-T Recommendation G.723.1 Decoder

11 Interoperability

11.1 Speech-only terminals and other analog GSTN terminals

Multi-media terminals described in ITU-T Recommendation H.324 shall support interoperation with analog speech-only telephones over the GSTN. Means included in terminals to interwork with other modem or FAX terminals used on the GSTN are optional. These means are outside the scope of this standard.

11.2 Multimedia telephone terminals over the ISDN

Interoperation with multimedia telephone terminals over the ISDN (ITU-T Recommendation H.320) can be provided by:

1. using an interworking adapter on the ISDN, or
2. using dual-mode (ISDN and GSTN) terminals on the ISDN.

An interworking adapter appropriate for terminals according to ITU-T Recommendations H.324 and H.320 is located at the interface between ISDN and GSTN signals. It transcodes the multiplexes of ITU-T Recommendations H.223 and H.221, and the content of control, audio, and data logical channels between the protocols of ITU-T Recommendations H.324 and H.320.

In order to ease communication between terminals of ITU-T Recommendations H.324 and H.320 via interworking adapters, terminals of ITU-T Recommendation H.324 which support video shall support the video codec of ITU-T Recommendation H.261 in the QCIF picture format so that the additional delay of video transcoding can be avoided. When this mode is in use, interworking adapters shall insert and remove BCH error correction and error correction framing (described in ITU-T Recommendations H.261 and H.263) as appropriate for each terminal type. ITU-T Recommendation H.324 terminals shall respond to the **FlowControlCommand** of ITU-T Recommendation H.245, so that transmitted video streams of ITU-T Recommendation H.324 can be matched to the video bit rate of ITU-T Recommendation H.320 in use by the multiplex of ITU-T Recommendation H.221.

Dual-mode (ITU-T Recommendations H.320 and H.324) terminals on the ISDN shall send GSTN signals of ITU-T Recommendation H.324 by the use of a "virtual modem," which generates and receives an ITU-T

Recommendation V.34 analog signal encoded as an ITU-T Recommendation G.711 audio bitstream over the ISDN.

Annex D/ITU-T Recommendation H.324 (approved February 1998) specifies operation of the H.324 protocol on ISDN and similar switched digital circuits at bit rates from 56 kbps to 1920 kbps. This operation mode is called "H.324/I" and is intended by ITU-T as a second-generation replacement for H.320.

In the H.324/I mode, the bitstream format of the ITU-T Recommendation H.324 protocol on the digital channel is identical to that carried on the synchronous V.34 modem in the case of PSTN operation. The V.34 modem is simply replaced with an I.400 series user-network interface to the digital channel.

Operation is generally identical with the PSTN mode of H.324, except that during call setup, the role of V.8/V.8bis is replaced by V.140, which probes the digital circuit to automatically determine 56 vs. 64 kbps operation, and to determine byte alignment. V.140 then executes a simple capability exchange to determine if the call should proceed in H.324/I, H.320, voice, or another mode.

For purposes of backward compatibility, terminals conforming to Annex D/ITU-T Recommendation H.324 are required to also support ITU-T Recommendation H.320 on ISDN, interoperation with PSTN H.324 systems via V.34 modulation, and ordinary ISDN voice telephony. Selection among these (and other possible) modes is automatically accomplished during call setup according to the procedures of ITU-T Recommendation V.140.

Multiple channel operation (using multiple PSTN or ISDN channels or a combination) is allowed according to Annex F/ITU-T Recommendation H.324 (scheduled for approval September 1998). In this case the digital channel carrying the ITU-T Recommendation H.324 bitstream is the single aggregated channel, provided as described in Annex F/ITU-T Recommendation H.324.

11.3 Multimedia telephone terminals over mobile radio

It is expected that multimedia telephone terminals will also be used on mobile radio and satellite networks. Rate matching between wireless terminals and GSTN terminals can be achieved by the use of the **Flow-Control Command** from ITU-T Recommendation H.245.

Implementation of wireless and mobile systems is now found in Annex C of ITU-T Recommendation H.324 (1998) and ITU-T Recommendation H.223 Annex A (1998), ITU-T Recommendation H.223 Annex B (1998) and ITU-T Recommendation H.223 Annex C (1998).

ITU-T Recommendation H.223 Annexes A, B and C define four levels of multiplexing to address the more error-prone environments of wireless and mobile. These four levels define ever more error robust strategies of multiplexing starting with Level 0 being defined as use of the current ITU-T Recommendation H.223, and Levels 1, 2 and 3 implementing ITU-T Recommendation H.223 Annexes A, B and C respectively. (Level 1 addresses low error-prone channels, Level 2 moderately error-prone channels, and Level 3 addresses highly error prone channels.)

Annex A
(informative)

Bibliography

ITU-T Recommendation H.262 (1995): *Generic coding of moving pictures and associated audio: video – ISO/IEC 13818-2*¹⁾

ITU-T Recommendation H.320 (1993): *Narrow-band ISDN visual telephone systems and terminal equipment*¹⁾

ITU-T Recommendation T.51 (1993): *Latin based coded character sets for telematic services*¹⁾

Draft ITU-T Recommendation T.84|ISO/IEC 10918-3 (199x): *Digital compression and coding of continuous tone still images – Extensions*¹⁾

ITU-T Recommendation X.680 (1994): *Information Technology – Abstract Syntax Notation One (ASN.1) – Specification of basic notation* (Reference document for ITU-T Recommendation X.691)¹⁾