



INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.320

(03/2004)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

Infrastructure of audiovisual services – Systems and
terminal equipment for audiovisual services

Narrow-band visual telephone systems and terminal equipment

ITU-T Recommendation H.320

ITU-T H-SERIES RECOMMENDATIONS
AUDIOVISUAL AND MULTIMEDIA SYSTEMS

CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100–H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	
General	H.200–H.219
Transmission multiplexing and synchronization	H.220–H.229
Systems aspects	H.230–H.239
Communication procedures	H.240–H.259
Coding of moving video	H.260–H.279
Related systems aspects	H.280–H.299
Systems and terminal equipment for audiovisual services	H.300–H.349
Directory services architecture for audiovisual and multimedia services	H.350–H.359
Quality of service architecture for audiovisual and multimedia services	H.360–H.369
Supplementary services for multimedia	H.450–H.499
MOBILITY AND COLLABORATION PROCEDURES	
Overview of Mobility and Collaboration, definitions, protocols and procedures	H.500–H.509
Mobility for H-Series multimedia systems and services	H.510–H.519
Mobile multimedia collaboration applications and services	H.520–H.529
Security for mobile multimedia systems and services	H.530–H.539
Security for mobile multimedia collaboration applications and services	H.540–H.549
Mobility interworking procedures	H.550–H.559
Mobile multimedia collaboration inter-working procedures	H.560–H.569
BROADBAND AND TRIPLE-PLAY MULTIMEDIA SERVICES	
Broadband multimedia services over VDSL	H.610–H.619

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation H.320

Narrow-band visual telephone systems and terminal equipment

Summary

This Recommendation specifies technical requirements for narrow-band visual telephone systems and terminal equipment, typically for videoconferencing and videophone services. It describes a generic system configuration consisting of a number of elements which are specified by respective ITU-T Recommendations, definition of communication modes and terminal types, call control arrangements, terminal aspects and interworking requirements.

This revised version of H.320 introduces a number of enhancements and clarifications to the previous version, primarily the description on the usage of ITU-T Recs G.722.1, H.239, H.241, H.264, and ISO/IEC 14496-3 in H.320 systems.

Source

ITU-T Recommendation H.320 was approved on 15 March 2004 by ITU-T Study Group 16 (2001-2004) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

© ITU 2004

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

CONTENTS

Page

1	Scope	1
2	Definitions	1
3	System description.....	2
3.1	Block diagram and identification of elements.....	2
3.2	Signals	3
3.3	Bit rate options and infrastructure	3
3.4	Call control arrangements.....	6
3.5	Optional enhancements	11
4	Terminal requirements.....	11
4.1	Environments.....	11
4.2	Audio and video arrangements.....	12
4.3	Delay compensation in the audio path.....	13
4.4	Control and Indications (C&I).....	14
4.5	Multipoint operation.....	17
5	Intercommunications	17
5.1	Intercommunication between different visual telephone terminal types.....	17
5.2	Intercommunication with telephony.....	17
5.3	Intercommunication with other audiovisual terminals.....	18
6	Maintenance.....	19
7	Human factor aspects.....	19
7.1	Physical section	19
7.2	Logical section.....	19
Annex A	Optional enhanced video algorithms for H.320 systems.....	20
A.1	Introduction to enhanced video algorithms	20
A.2	Use of ITU-T Rec. H.262 in H.320 systems	21
A.3	Use of ITU-T Rec. H.263 in H.320 systems	23
A.4	Enhanced video algorithm hierarchy	24
A.5	Use of ITU-T Rec. H.264 in H.320 systems	26

ITU-T Recommendation H.320

Narrow-band visual telephone systems and terminal equipment

1 Scope

This Recommendation covers the technical requirements for narrow-band visual telephone services defined in H.200/F.720-series Recommendations, where channel rates do not exceed 1920 kbit/s.

NOTE – It is anticipated that this Recommendation will be extended to a number of Recommendations each of which would cover a single videoconferencing or videophone service (narrow-band, broadband, etc.). However, large parts of these Recommendations would have identical wording, while in the points of divergence the actual choices between alternatives have not yet been made; for the time being, therefore, it is convenient to treat all the text in a single Recommendation.

The service requirements for visual telephone services are presented in ITU-T Recs F.720 for videotelephony and F.702 for videoconference; video and audio coding systems and other technical aspects common to audiovisual services are covered in other Recommendations in the H.200/F.700-series.

2 Definitions

This Recommendation defines the following terms:

2.1 bit-rate allocation signal (BAS): eight bits positioned within the frame structure of ITU-T Rec. H.221 to transmit, e.g., commands, control and indication signals, capabilities.

2.2 control and indication (C&I): End-to-end signalling between terminals consisting of control which causes a state change in the receiver and indication which provides for information as to the functioning of the system; see also ITU-T Rec. H.230.

2.3 data port: Input/output gate for the user data transmitted within service channel or sub-channels according to ITU-T Rec. H.221.

2.4 human-machine interface (HMI): Human-machine interface between user and terminal/system which consists of a physical section (electro-acoustic, electro-optic transducer, keys, etc.) and a logical section dealing with functional operation states.

2.5 in-band signalling: Signalling via BAS of the H.221 frame structure.

2.6 lip synchronization: Operation to provide the feeling that the speaking motion of the displayed person is synchronized with that person's voice. Alternatively, 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.

2.7 multipoint control unit (MCU): A piece of equipment located in a node of the network or in a terminal which receives several channels from access ports and, according to certain criteria, processes audiovisual signals and distributes them to the connected channels.

2.8 narrow-band: Bit rates ranging from 64 kbit/s to 1920 kbit/s. This channel capacity may be provided as a single B/H₀/H₁₁/H₁₂-channel or multiple B/H₀-channels in ISDN.

2.9 out-band signalling: Signalling via a channel not being part of the B/H₀/H₁₁/H₁₂-channel (due to I.400-series Recommendations).

2.10 visual telephone services: A group of audiovisual services including videophone defined in ITU-T Rec. F.721 and videoconferencing defined in ITU-T Rec. F.702.

3 System description

3.1 Block diagram and identification of elements

A generic visual telephone system is shown in Figure 1. It consists of terminal equipment, network, Multipoint Control Unit (MCU) and other system operation entities.

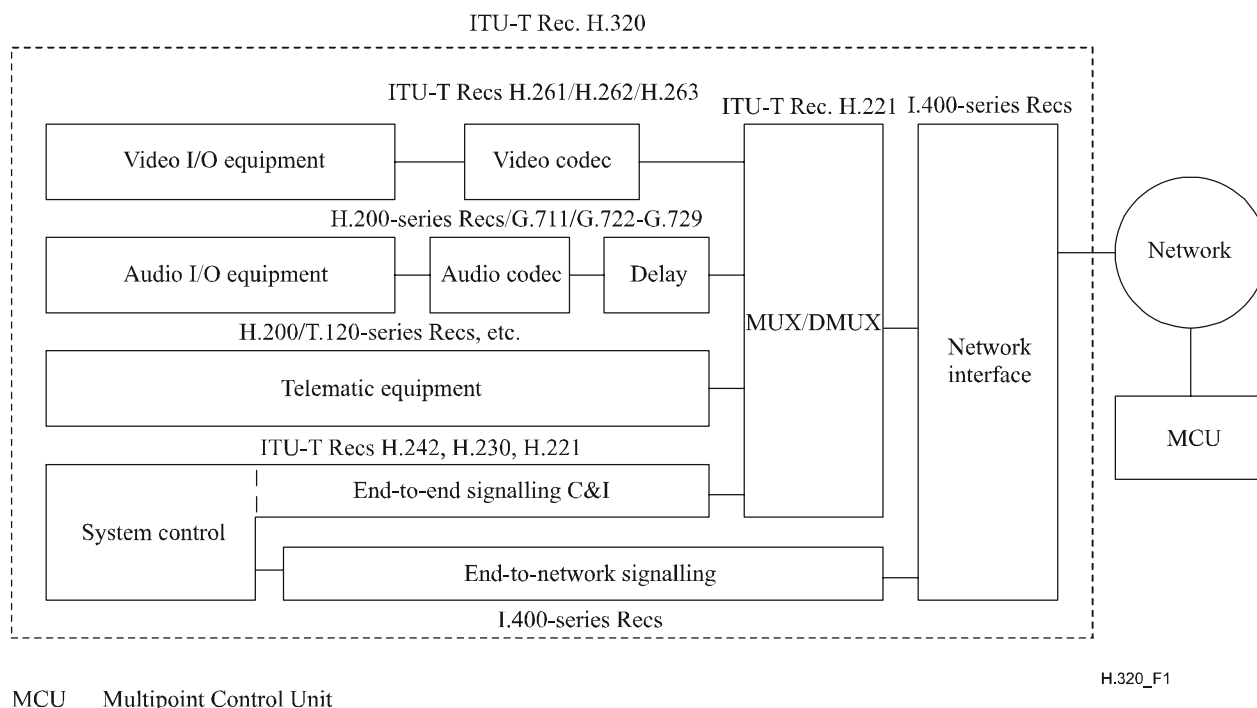


Figure 1/H.320 – Visual telephone system

A configuration of the terminal equipment consisting of several functional units is also shown in Figure 1. Video I/O equipment includes cameras, monitors and video processing units to provide functions such as split-screen scheme. Audio I/O equipment includes microphones, loudspeakers and audio processing units to provide such functions as acoustic echo cancellation (see ITU-T Rec. G.167). Telematic equipment includes visual aids such as electronic blackboard, text conversation facility and still picture transceiver to enhance basic visual telephone communication. The system control unit carries out such functions as network access through end-to-network signalling and end-to-end control to establish common mode of operation and signalling for proper operation of the terminal through end-to-end signalling. The video codec carries out redundancy reduction coding and decoding for video signals, while audio codec does the same thing for audio signals. The delay in the audio path compensates video codec delay to maintain lip synchronization. The mux/dmux unit multiplexes transmitting video, audio, data and control signals into a single bit stream and demultiplexes a received bit stream into constituent multimedia signals. Network interface makes necessary adaptation between the network and the terminal according to the user-network interface requirements defined in the I.400-series Recommendations (see Note).

NOTE – For leased line networks, the network interface is defined in ITU-T Rec. G.703 for bit rates in the range of 64 kbit/s to 2048 kbit/s. An alternative interface is defined in ITU-T Rec. X.21. For $n \times H_0$ channels, timeslot allocation is given in clause 5/G.704 for the G.703 interface. It is stressed that interworking towards ISDN requires synchronous operation of the leased line network.

3.2 Signals

Visual telephone signals are classified into video, audio, data and control as follows:

- Audio signals are continuous traffic and require real-time transmission.
NOTE – In order to reduce the average bit rate of audio signals, voice activation can be introduced (in which case the audio signals are no longer continuous).
- Video signals are also continuous traffic; the bit rate allocated to video signals should be as high as possible, in order to maximize the quality within the available channel capacity.
- Data signals include still pictures, facsimile and documents, or other facilities such as text conversation; this signal may occur only occasionally as required and may temporarily displace all or part of the audiovisual signal content. It should be noted that data signals are associated only with optional enhancements to the basic visual telephone system; therefore, the opening of a path to carry such signals is preceded by negotiation between the terminals.
- Control signals are some system control signals by definition. The path for the terminal-to-network control signals is provided in the D-channel, while the path for the terminal-to-terminal control signals is provided in BAS or service channel only when necessary by the mechanism defined in ITU-T Rec. H.221.

3.3 Bit rate options and infrastructure

3.3.1 Communication modes of visual telephone

Communication modes of visual telephone are defined in Table 1 according to their channel configuration and coding. A particular communication mode of operation is established according to the H.242 procedures.

Table 1/H.320 – Communication modes of visual telephone

Visual telephone mode (Suffix n corresponds to audio coding)	Channel rate [kbit/s]	ISDN channel	ISDN interface		Coding		Notes
			Basic	Primary rate	Audio	Video	
a _n	64	B	Not applicable	Applicable	n Rec. 0 G.711 1 G.722 2 G.728 3 G.723.1 (Note 5) 4 G.729	Rec. H.261 Rec. H.262 Rec. H.263 (Note 6)	1, 2
b _n	128	2B					3
c _n	192	3B					3, 4
d _n	256	4B					3, 4
e _n	320	5B					3, 4
f _n	384	6B					3, 4
g _n	384	H ₀					
h _n	768	2H ₀					3
i _n	1152	3H ₀					3
j _n	1536	4H ₀					3
k _n	1536	H ₁₁					
l _n	1920	5H ₀					3
m _n	1920	H ₁₂					

Table 1/H.320 – Communication modes of visual telephone

Visual telephone mode (Suffix n corresponds to audio coding)	Channel rate [kbit/s]	ISDN channel	ISDN interface		Coding		Notes
			Basic	Primary rate	Audio	Video	
NOTE 1 – If a visual telephone interworks with a wideband speech terminal, G.722 audio (mode a ₁) may be used instead of G.711 audio (mode a ₀).							
NOTE 2 – If two terminals connect at this rate and run G.711 (mode a ₀) or G.722 (mode a ₁) and both have video capability, H.261, H.262 or H.263 may be used. It should be noted, however, that the video performance is limited due to the very low bit rate available for this purpose.							
NOTE 3 – For multiple channels of B/H ₀ , all channels are synchronized at the terminal according to 2.7/H.221. For bit rates higher than 64 kbit/s, channel aggregation of multiple B-channels may be used as specified in ITU-T Rec. H.244 or ISO/IEC 13871.							
NOTE 4 – This mode is applicable to the ISDN basic interface if multiple basic accesses are used.							
NOTE 5 – For G.723.1 audio, a channel shall be allocated as specified in ITU-T Rec. H.221. The audio data and their packet formats are described in ITU-T Recs G.723.1 and H.223. Use of bits not required for G.723.1 audio coding is specified in ITU-T Rec. H.221.							
NOTE 6 – The hierarchy of H.261, H.262 and H.263 is defined in Annex A.							

3.3.2 Terminal types of visual telephone

The terminal type is categorized according to the type of communication channels with which the terminal can communicate; mxB (type X with parameter a-f), n × H₀ (type Y with parameter 1-5; see Note), H₁₁/H₁₂ (type Z with parameter α-β) or their combinations. Table 2 lists terminal types of visual telephone and communication modes which each terminal type is required to support.

NOTE – Type Y terminals shall have the SM-comp or 6B-H₀ compatibility mode defined in ITU-T Rec. H.221 for interworking of evolving networks (see 3.3.2.2).

Table 2/H.320 – Visual telephone terminal types

Terminal type		ISDN channel	Modes to be supported		
			Mandatory	Recommended	Optional
X (Note 1)	X _a	B	a ₀	–	all other a _n
	X _b	2B	a ₀ , b ₀	–	all other a _n , b _n
	X _c	3B	a ₀ , b ₀ , c ₀	–	all other a _n , b _n , c _n
	X _d	4B	a ₀ , b ₀ , c ₀ , d ₀	–	all other a _n , b _n , c _n , d _n
	X _e	5B	a ₀ , b ₀ , c ₀ , d ₀ , e ₀	–	all other a _n , b _n , c _n , d _n , e _n
	X _f	6B	a ₀ , b ₀ , c ₀ , d ₀ , e ₀ , f ₀	–	all other a _n , b _n , c _n , d _n , e _n , f _n

Table 2/H.320 – Visual telephone terminal types

Terminal type		ISDN channel	Modes to be supported		
			Mandatory	Recommended	Optional
Y (Note 2)	Y ₁	H ₀	g ₀	g ₁	all other g _n
	Y ₂	2H ₀	g ₀ , h ₀	g ₁ , h ₁	all other g _n , h _n
	Y ₃	3H ₀	g ₀ , h ₀ , i ₀	g ₁ , h ₁ , i ₁	all other g _n , h _n , i _n
	Y ₄	4H ₀	g ₀ , h ₀ , i ₀ , j ₀	g ₁ , h ₁ , i ₁ , j ₁	all other g _n , h _n , i _n , j _n
	Y ₅	5H ₀	g ₀ , h ₀ , i ₀ , j ₀ , l ₀	g ₁ , h ₁ , i ₁ , j ₁ , l ₁	all other g _n , h _n , i _n , j _n , l _n
Z	Z _α	H ₁₁	k ₀	k ₁	all other k _n
	Z _β	H ₁₂	m ₀	m ₁	all other m _n
NOTE 1 – If a visual telephone interworks with a wideband speech terminal, it should support G.722 audio (mode a ₁) in addition to G.711 audio (mode a ₀).					
NOTE 2 – Terminal of this type shall conform to 3.3.2.2.					

3.3.2.1 Examples

- Type X_b equipped with G.711, G.722 and G.728 audio codecs is a terminal capable of operating at modes a₀, a₁, a₂, b₀, b₁ and b₂ through B or 2 × B-channel;
- Type X_bY₁ equipped with G.711, G.722 and G.728 audio codecs is a terminal capable of operating at modes a₀, a₁, a₂, b₀, b₁, b₂, g₀, g₁ and g₂, through B, 2 × B- or H₀-channel;
- Type X_fY₄Z_α equipped with G.711, G.722 and G.728 audio codecs is a terminal capable of operating at modes x₀, x₁ and x₂ (x = a-k) through (1-6) × B, (1-4) × H₀- or H₁₁-channel.

The type of remote terminal is identified through the transfer rate capability exchange defined in ITU-T Rec. H.242.

3.3.2.2 Conditional requirements for single-channel equipment to operate into H.244 Channel Aggregators

In the case where a single channel-equipment, such as a Type Y₁ or Type Z terminal defined here and a terminal which has sub-primary rate channel access capability, operates as an option into a Channel Aggregation Unit (CAU) conforming to H.244 Mode H₂, the following properties are relevant (see Note):

- The terminal shall be capable of {SM-comp} defined in ITU-T Rec. H.221 – able to vacate bit 8 of the first 16 octets of every timeslot in each frame, in TS₂, ..., and to accept an incoming signal of the same structure.
- The terminal shall be able to accept capability sets containing repeated {null} values as described in ITU-T Rec. H.244.
- The terminal should be able to operate at all multiples of 64 kbit/s up to the maximum rate (optional – able to operate with available channel capacity when one or more channels are lost or unavailable).
- The terminal should be able to obey the command [capex] as described in ITU-T Rec. H.244.
- The terminal should be able to recognize and act upon the command [AggIN]* as described in ITU-T Rec. H.244.