INTERNATIONAL TELECOMMUNICATION UNION

ITU-T H.320

(03/96)

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

SERIES H: TRANSMISSION OF NON-TELEPHONE SIGNALS

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

---

Narrow-band visual telephone systems and terminal equipment

ITU-T Recommendation H.320

(Previously “CCITT Recommendation”)

ITU-T -SERIES RECOMMENDATIONS H

TRANSMISSION OF NON-TELEPHONE SIGNALS

Characteristics of transmission channels used for other than telephone purposes	H.10-H.19
Use of telephone-type circuits for voice-frequency telegraphy	H.20-H.29
Telephone circuits or cables used for various types of telegraph transmission or simultaneous transmission	H.30-H.39
Telephone-type circuits used for facsimile telegraphy	H.40-H.49
Characteristics of data signals	H.50-H.99
CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100-H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	**H.200-H.399**
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.399**

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 revision reflects the progress in relevant H-Series Recommendations such as H.233, H.234, H.242, H.243, H.244, etc.

Source

ITU-T Recommendation H.320 was revised by ITU-T Study Group 15 (1993-1996) and was approved under the WTSC Resolution No. 1 procedure on the 10th of April 1995.

FOREWORD

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. Some 179 member countries, 84 telecom operating entities, 145 scientific and industrial organizations and 38 international organizations participate in ITU-T which is the body which sets world telecommunications standards (Recommendations).

The approval of Recommendations by the Members of ITU-T is covered by the procedure laid down in WTSC Resolution No. 1 (Helsinki, 1993). In addition, the World Telecommunication Standardization Conference (WTSC), which meets every four years, approves Recommendations submitted to it and establishes the study programme for the following period.

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.

CONTENTS

Recommendation H.320

NARROW-BAND VISUAL TELEPHONE SYSTEMS AND TERMINAL EQUIPMENT

(revised 1996)

1 Scope

This Recommendation covers the technical requirements for narrow-band visual telephone services defined in H.200/AV.120-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 Recommendations F.720 for videotelephony and F.730 for videoconference; video and audio coding systems and other technical aspects common to audiovisual services are covered in other Recommendations in the H.200/AV.200-Series.

2 Definitions

For the purposes of this Recommendation, the following definitions apply:

2.1 bit-rate allocation signal (BAS): 8 bits positioned within the frame structure of Recommendation 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 Recommendation H.230.

2.3 data port: Input/output gate for the user data transmitted within service channel or sub-channels according to Recommendation 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 $B/H_0/H_{11}/H_{12}$-channel or multiple B/H_0-channels in ISDN.

2.9 out-band signalling: Signalling via a channel not being part of the $B/H_0/H_{11}/H_{12}$-channel (due to I.400-Series Recommendations).

2.10 visual telephone services: A group of audiovisual services including videophone defined in Recommendation F.721 and videoconferencing defined in Recommendation F.730.

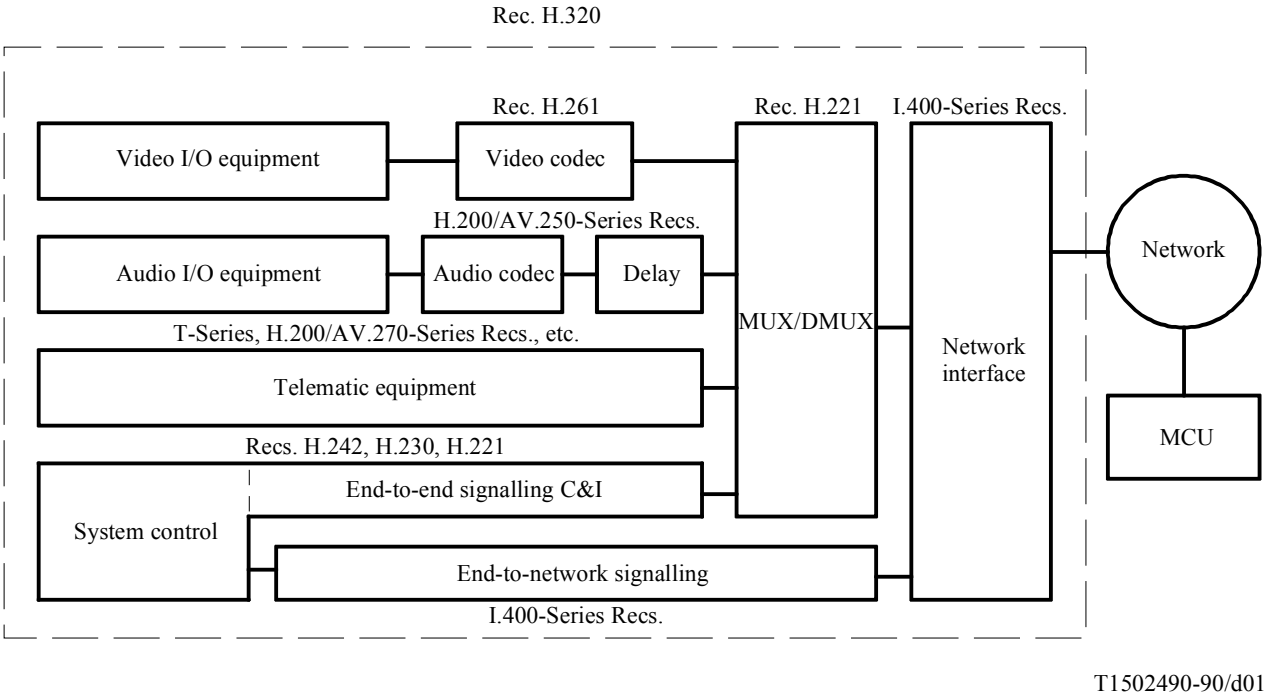
3 System description

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.

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 Recommendation G.167). Telematic equipment includes visual aids such as electronic blackboard, 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 Recommendation G.703 for bit rates in the range of 64 kbit/s to 2048 kbit/s. An alternative interface is defined in Recommendation X.21. For nxH_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.



MCU Multipoint Control Unit

FIGURE 1/H.320

Visual telephone system

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; 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 Recommendation 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.

TABLE 1/H.320

Communication modes of visual telephone

<table>
<tr><td colspan="2" rowspan="2">Visual telephone mode</td><td rowspan="2">Channel rate (kbit/s)</td><td rowspan="2">ISDN channel (Note 2)</td><td colspan="2">ISDN interface</td><td colspan="2">Coding</td></tr>
<tr><td>Basic</td><td>Primary rate</td><td>Audio</td><td>Video</td></tr>
<tr><td rowspan="2">a</td><td>a_0</td><td rowspan="2">64</td><td rowspan="2">B</td><td rowspan="10"></td><td rowspan="10">Applicable</td><td>Rec. G.711 (Note 4)</td><td>Rec. H.261 (Note 6)</td></tr>
<tr><td>a_1</td><td>Rec. G.728</td><td rowspan="11">Rec. H.261</td></tr>
<tr><td rowspan="3">b</td><td>b_1</td><td rowspan="3">128</td><td rowspan="3">2B</td><td>Rec. G.711</td></tr>
<tr><td>b_2</td><td>Rec. G.722</td></tr>
<tr><td>b_3</td><td>Rec. G.728</td></tr>
<tr><td rowspan="3">q
(Note 3)</td><td>q_1</td><td rowspan="3">n x 64</td><td rowspan="3">nB</td><td>Rec. G.711</td></tr>
<tr><td>q_2</td><td>Rec. G.722</td></tr>
<tr><td>q_3</td><td>Rec. G.728</td></tr>
<tr><td colspan="2">g</td><td>384</td><td>H_0</td><td rowspan="7">Not applicable</td><td>Rec. G.722

(Note 5)</td></tr>
<tr><td colspan="2">h</td><td>768</td><td>$2H_0$</td></tr>
<tr><td colspan="2">i</td><td>1152</td><td>$3H_0$</td></tr>
<tr><td colspan="2">j</td><td>1536</td><td>$4H_0$</td></tr>
<tr><td colspan="2">k</td><td>1536</td><td>H_{11}</td></tr>
<tr><td colspan="2">l</td><td>1920</td><td>$5H_0$</td></tr>
<tr><td colspan="2">m</td><td>1920</td><td>H_{12}</td></tr>
<tr><td colspan="8">NOTES
1 (Audio coding of mode b_3). In addition to G.728, higher quality audio coding such as H.200/AV.253 may be used for this mode.
2 For multiple channels of B/H_0, all channels are synchronized at the terminal according to 2.7/H.221.
3 q = c/d/e/f corresponds to n = 3/4/5/6, respectively. This mode is applicable to the ISDN basic interface if multiple basic accesses are used.
4 If a visual telephone interworks with a wideband speech terminal, G.722 audio may be used instead of G.711 audio.
5 Modes (G.711 and G.728 audio) other than this recommended mode may be invoked by H.242 procedures.
6 If two terminals connect at this rate and run Recommendation G.711 and both have video capability, Recommendation H.261 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.</td></tr>
</table>

3.3.2 Terminal types of visual telephone

Table 2 lists terminal types of visual telephone. The terminal type is categorized according to the communication modes and the type of communication channels with which the terminal can communicate; mxB (type X with parameter a-f), nxH_0 (type Y with parameter 1-5; see Note), H_{11}/H_{12} (type Z with parameter α-ß) or their combinations.

NOTE – Type Y terminals must have the SM-comp or 6B-H_0 compatibility mode defined in Recommendation H.221 for interworking of evolving networks - see 3.3.2.2.

TABLE 2/H.320

Visual telephone terminal types

Mode			Type X (Note 2)									Type Y (Note 3)					Type Z	
Transfer rate		Audio coding (Rec.)	a	b_1	$b_{2/3}$	b_4	b_5	q_1	$q_{2/3}$	q_4	q_5	1	2	3	4	5	α	β
a_0	B	G.711 (Note 4)	X	X	X	X	X	X	X	X	X							
a_1	B	G.728	X	X	X			X	X									
b_1	2B	G.711		X	X	X	X	X	X	X	X							
b_2	2B	G.722			X		X		X		X							
b_3	2B	G.728		X	X			X	X									
q_1	nB	G.711 (Note 5)						X	X	X	X							
q_2	nB	G.722 (Note 5)							X		X							
q_3	nB	G.728 (Note 5)						X	X									
g	H_0	G.722										X	X	X	X	X		
h	$2H_0$	G.722											X	X	X	X		
i	$3H_0$	G.722												X	X	X		
j	$4H_0$	G.722													X	X		
k	H_{11}	G.722															X	
l	$5H_0$	G.722														X		
m	H_{12}	G.722																X

NOTES

1 "X" means the terminal of the given type is able to work in the given mode.

2 Types Xb_4 and Xb_5 are defined to take into account that Recommendation G.728 had not yet been approved when this Recommendation was first established.

3 Terminal of this type must conform to 3.3.2.2.

4 If a visual telephone interworks with a wideband speech terminal, G.722 audio may be used instead of G.711 audio.

5 q = c/d/e/f corresponds to n = 3/4/5/6, respectively. Since transfer rates of multiple B are defined hierarchically, Type X_{f1}, for example, supports all of (a_1, b_1, c_1, d_1, e_1, f_1) and (b_3, c_3, d_3, e_3, f_3) modes.

3.3.2.1 Examples

a) type $Xb_{2/3}$ is a terminal capable of operating at modes a_0, a_1, b_1, b_2 and b_3 through B or 2xB-channel;

b) type $Xb_{2/3}Y1$ is a terminal capable of operating at modes a_0, a_1, b_1, b_2, b_3 and g through B, 2xB- or H_0-channel.

c) type $Xf_{2/3}Y4Z\alpha$ is a terminal capable of operating at modes a_0-k through (1-6)xB, (1-4)xH_0- or H_{11}-channel.

For MxB and NxH_0 categories, the terminal should be able to operate at all the values of *m* and *n* not higher than M and N in principle. The type of remote terminal is identified through the transfer rate capability exchange defined in Recommendation 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 Y 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 (Note):

1) the terminal shall be capable of {SM-comp} defined in Recommendation 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;
2) the terminal shall be able to accept capability sets containing repeated {null} values as described in Recommendation H.244;
3) 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);
4) the terminal should be able to obey the command [capex] as described in Recommendation H.244;
5) the terminal should be able to recognize and act upon the command [AggIN]* as described in Recommendation H.244.

NOTE – The 1993 and prior versions of Recommendation H.320 stipulated {6B-H0-comp} rather than {SM-comp} as the capability to be implemented, and there were no provisions for {null}, [capex] and [AggIN]*. A terminal constructed to one of those versions will interoperate with a Type Xf terminal via a CAU at 384 kbit/s only, provided that it can accept capsets containing repeated {null} values.

3.3.3 Video codec

As per Recommendation H.261.

3.3.4 Audio codec

As per Recommendations G.711, G.722, G.728, H.200/AV.253 (see Table 1).

3.3.5 Frame structure

As per Recommendation H.221.

3.3.6 Control and Indication (C&I)

Identified subset of Recommendation H.230 is used (see 4.4).

3.3.7 Communication procedure

As per Recommendation H.242.

3.4 Call control arrangements

To establish intercommunication between various audiovisual terminals, it is necessary to carry out in-band and out-band procedures according to Recommendation H.242 and other relevant Recommendations.

The different stages of the call are referred according to a point-to-point configuration where terminal X is the calling terminal and Y the called terminal.

3.4.1 Establishment of a visual telephone call – Normal procedure

The provision of the communication is made in the main following steps:

– phase A: call set-up, out-band signalling;

– phase B1: mode initialization on initial channel;
– phase CA: call set-up of additional channel(s), if relevant;
– phase CB1: initialization on additional channel(s);
– phase B2 (or CB2): establishment of common parameters;
– phase C: visual telephone communication;
– phase D: termination phase;
– phase E: call release.

3.4.1.1 Phase A – Call set-up

After user initialization, the terminal X performs a call set-up procedure (this procedure is defined in Recommendation Q.939). As soon as the terminal receives an indication from the network that the connection is established, a bidirectional channel is opened from end to end, and it overlays H.221 framing on the channel.

Following the connection establishment, all the terminals will start to work in Mode 0F (A-law or μ-law) as defined in Recommendation H.221.

The in-band procedure is activated.

3.4.1.2 Phase B1 – Mode initialization

3.4.1.2.1 Phase B1-1

Using the procedures provided in Recommendation H.242, framed PCM audio is transmitted in both directions - see 3.4.5.1; after frame and multiframe alignment is gained, terminal capabilities are exchanged.

3.4.1.2.2 Phase B1-2 (terminal procedure)

Determination of the appropriate mode to be transmitted: this will normally be the highest common mode (see Table 3 for the case using a B or 2xB-channel), but a lower compatible mode could be chosen instead.

In the case that both terminals have announced the capability to work on additional channel(s), terminal X initiates the request for additional call set-up.

Visual telephone terminals using multiple connections need to know what number(s) to dial for additional connection(s). In many cases this information is stored, or can be deduced from the number dialled for the initial channel. Where the information is not available or deducible, the terminal can acquire the missing number(s) by the procedure defined in Clause 18/H.242.

NOTE – If the user at either terminal does not wish the call to proceed to two or more channels, even though the terminal has this capability, the user must set the terminal such that only single-channel capability is declared in phase B1-1. In this case, we distinguish the active capability, wished by the users, from the inherent capability of the terminal. See Clause 5/H.242.

TABLE 3/H.320

Common mode (default) for communication between different types of visual telephones using a B or 2xB-channel

Xa	Xb_1	$Xb_{2/3}$	Xb_4	Xb_5	Terminal type
a_1	a_1	a_1	a_0	a_0	Xa
	b_3	b_3	b_1	b_1	Xb_1
		b_2/b_3	b_1	b_2	$Xb_{2/3}$
			b_1	b_1	Xb_4
				b_2	Xb_5

NOTES

1 "b_2/b_3" indicates that either b_2 or b_3 can be used. b_2 gives better audio, while b_3 gives better video.

2 The communication modes indicated in the table above include the possibility to use the CIF format as well as the QCIF format. The choice of the operating format is left to the terminal procedure according to the exchanged capabilities: symmetry in two directions is not essential.

Each terminal may use a minimum decodable picture interval in its sending direction which makes best use of the capability indicated by the other terminal.

This table does not include interworking situations between visual telephones and telephone terminals. If visual telephone terminals are connected to telephones, mode a_0 is used for the communication.

3.4.1.2.3 Phase B1-3 (mode switching)

If the terminal has {G.728} capability and is connected to a 64/56 kbit/s connection, and if the capset received includes {1B}, {G.728} and {H.261/QCIF or CIF}, identifying the remote endpoint as Type Xa, it shall switch to Mode a_1. Likewise if the terminal is itself Type Xa and the received capset includes {G.728} and {H.261/QCIF or CIF}, it shall switch to Mode a_1.

In all other cases a higher capacity than 64/56 kbit/s is available, and the terminal may switch to a higher mode than a1 in Table 1, according to the application.

If the terminal is of Type Xb, Xc or higher and the received capset includes {2B} or higher, it is a matter for local choice whether to mode-switch video on and/or change to another audio coding mode while waiting for additional connections to be set up.

The switching is made using the procedure of Recommendation H.242. If the terminals have not both adopted the common mode, an asymmetric communication may result (this is not precluded) - see 3.4.1.5.

3.4.1.3 Phase CA – Call set-up of the additional channel(s)

Following phases B1-3 and B2 if relevant, the communication phase C proceeds on that channel. If additional channels have been requested these go also through phase A (hence the nomenclature "Phase CA"), exactly as in phase A above, and additional call set-ups are performed by the terminals. On each of the established channels H.221 framing is overlaid (see Note).

NOTE – During phase CA an intermediate audiovisual mode could be offered on the initial channel used for initialization, until full completion of initialization phase.

3.4.1.4 Phase CB1 – Mode initialization on additional channel(s)

3.4.1.4.1 Phase CB1-11

Using the procedure provided in Recommendation H.242, frame and multiframe alignments are gained.

3.4.1.4.2 Phase CB1-12

Synchronization of the channels is achieved.

3.4.1.4.3 Phase CB1-2 (terminal procedure)

Determination of the appropriate mode to be transmitted. This phase will normally be skipped, because the appropriate mode of operation is determined in Phase B1-2.

3.4.1.4.4 Phase CB1-3 (mode switching)

Both terminals switch to the mode they have identified in phase B1-2 using the procedure of Recommendation H.242.

It may be noted that G.711 modes should be avoided after initialization if Recommendation G.728 or G.722 is available, since both these can leave more capacity for video and Recommendation G.722 provides better quality.

Total symmetry is not required: that is, it is not essential that the audio and video transmissions be identical in the two directions, and 8.2/H.242 warns against implementations that automatically select the outgoing mode to be the same as the incoming, listing better methods of mode control. Visual telephone applications require symmetry of transfer rate, but choice of transmitted audio coding and of QCIF versus CIF video settings are made by the user, or preset in the terminal.

3.4.1.5 Phase B2 (or CB2) – Establishment of common parameters

This phase establishes common operational parameters specific to visual telephone (e.g. encryption) after phase B1 process is finished. Capabilities or requirements of the receiving side are first indicated then the sending side decides operational parameters and controls the receiving side. BAS codes for this purpose are defined in Recommendation H.221.

Sometimes users may wish to indicate preference of receiving modes (e.g. choice between mode b_2 with better audio quality and mode b_3 with better video quality as in Table 3). This can be achieved by using the "mode-preference" indications defined in 9.5/H.242. It is recommended that in general visual telephones be equipped to send and respond to these codes.

3.4.1.6 Phase C – Visual telephone communication

In the case where more than one channel is used, there will be intermediate phases CA, CB1, CB2 as described in this subclause. Likewise, if additional channels are dropped during the call there will be intermediate phases CD, CE as described in 3.4.4. The provisions of this subclause apply to any channel, initial or additional, for which phases B1 and B2 have been completed and phase D not yet started.

3.4.1.6.1 Mode switching

According to action by either user (for example, starting a facsimile machine) a different mode from the highest common mode may become more appropriate. Switching to this mode is made according to the procedure of Recommendation H.242.

3.4.1.6.2 Capability change

The user may change the capability of his terminal during the call (for example, by connecting or switching-on auxiliary telematic equipment); the terminal must initiate the capability exchange procedure defined in Recommendation H.242.

3.4.1.7 Phase D – Termination phase

3.4.1.7.1 Phase D1 (terminal procedure)

When one of the users hangs up, the terminal invokes phase D2 directly.

3.4.1.7.2 Phase D2 (mode switching)

Mode 0F is forced according to Recommendation H.242 (or taking into account the result of phase D1 if different; for further study).

In case the call is disconnected for some reason, it may take a long time if it keeps awaiting reaction from the remote terminal, and reception or initiation of the next call may be blocked. ISDN disconnection should be given priority over the H.242 call termination procedure. A time-out of two seconds is suggested.

3.4.1.8 Phase E – Call termination (release)

The terminal which has initiated the hang up sends messages over the D-channel with respect to all channels and idles all of them (that means no more information is sent over).

At the other terminal, the actual disconnection occurs at reception of the other disconnect message(s).

3.4.2 Exceptional procedures for phases A and B

In case of unsuccessful outcome during phases A and B (due to many causes), exceptional procedures may be provided in order to ensure a suitable service. The matter is for further study.

3.4.3 Exceptional procedures during phase C

During the actual exchange of audiovisual data, problems may occur in some channels. Fallback procedures managed by the terminal are activated. The description of the procedures and the appropriate indications are for further study.

3.4.4 Addition and dropping of channels during a visual telephone call

3.4.4.1 Addition

According to action by a user (for example the activation of auxiliary equipment) one or more additional channels are requested. The procedure follows those described for phases CA and CB1.

3.4.4.2 Dropping

Two phases are envisaged.

3.4.4.2.1 Phase CD1

The common mode, appropriate to the channel(s) which remains, is selected.

3.4.4.2.2 Phase CD2

The mode switching procedure of Recommendation H.242 is applied to invoke the mode identified in phase CD1; the remaining channel is the channel used for initialization (see phase A). It supports an appropriate fallback mode. The matter is for further study.

3.4.5 Transmission and presentation of audiovisual information at the start of a visual telephone call

3.4.5.1 Audio

Transmitted audio shall be according to 9.1.1.2/H.242. For videotelephony audio shall not be muted at the transmitting side.

As soon as the initial channel is established and the received audio coding law (G.711 A-law or μ-law) is known according to the procedure defined in 9.1.1.3/H.242, the audio signal should be presented to the user so that audio conversation can promptly be started as in the ordinary telephone communication.

3.4.5.2 Video

According to the chosen terminal procedures, pictures may or may not be visible to both users as soon as initialization is complete. In the case that either phase B1-3 or phase CB1-3 has activated a common mode, including video, mutual visibility of the users is possible.

The following items collect alternative procedures which can be used to suspend picture display until user intervention (by mutual agreement or otherwise) causes pictures to be displayed.

1) *No video transmitted* – In phase B1-2 and (if relevant) phase CB1-2 the mode selected includes video OFF. During phase C either user may unilaterally switch to video ON, alternatively, the terminal may send the C&I BAS code VIR (Video Indicate Ready-to-activate), but not switch to video-ON until video or VIR is received from the other terminal. While the incoming video-OFF state remains, the visual telephone screen should display a symbol or message indicating this (i.e. there is no fault).

   As already noted in 3.4.1, phase B1-2, the request for additional channel may, according to terminal procedure, be delayed while video-OFF is maintained; user action to activate video would then result in procedure phases CA1, CB1 (CB2 if required).

2) *Video pattern transmitted* – An electronically generated or other pattern is transmitted instead of the signal from a normal camera. The C&I BAS code VIS (Video Indicate Suppressed) is used to indicate the situation to the remote party.

3) *Video transmitted but not displayed* – Terminal procedures simply involve local action to display not the incoming signal but an explanatory symbol or message. User action would cause the incoming signal to be displayed, but if this should depend on mutual action by both users then a new C&I BAS code VRD (Video Ready-to-Display) must be defined. This point is for further study.

3.5 Optional enhancements

3.5.1 Data ports

Data ports as physical I/O ports of the terminal for telematic and other equipment are activated/deactivated by BAS commands. Depending on the transmission capability of a connection, e.g. multiples of B/H_0 channels, etc., various bit rates are available at these ports. Allocation of bit streams to the port(s) is performed by in-band signalling. Data conveyed at the port(s) is transparent, data rates being listed in Annex A/H.221.

For the data enhancement of the conformant visual telephone terminals, the following applies for increased level of interworking:

1) Terminals having at least one "audiographic" application such as T.81 still picture exchange, application sharing, on-screen annotation, etc. shall make use of the standardized applications provided by the T.120-Series (if such applications have been recommended by the ITU-T) making use of the MLP/H-MLP channels. A terminal that provides far-end camera control using Recommendations H.281 and H.224 operating in both the LSD and the MLP channels is not required to also support a T.120 far end camera control protocol.

2) Preferred MLP channel rates are 6.4 kbit/s (5.6 kbit/s if ECS is in use), 14.4 kbit/s (13.6 kbit/s if ECS is in use), 32 kbit/s and 40 kbit/s with 6.4 kbit/s (5.6 kbit/s if ECS is in use) as default.

3) If a terminal has an application which uses LSD to obtain critical performance, it should also be able to transmit the requisite data in the MLP channel when in multipoint calls via an MCU.

3.5.2 Encryption

Encryption may be applied on audio and video signals multiplexed; the encryption shall be according to Recommendations H.233, and, where a key-management system is required, one of the schemes in Recommendation H.234 should be used. Switching-on and off the encryption process has to be signalled between the terminals (or terminal and MCU, respectively) via in-band signalling.

3.5.3 Restricted networks

Some networks are restricted in transfer characteristics. For communications among terminals accommodated in restricted networks and those accommodated in non-restricted networks, both in point-to-point and multipoint cases, optional procedures are defined in Clause 13/H.242.

4 Terminal requirements

4.1 Environments

Under study.

4.2 Audio and video arrangements

4.2.1 Audio arrangements

A terminal can have one or more of three different arrangements:

– handset function;

– handsfree function for a small group of users (up to three users);

– handsfree function for more than three users (conference terminal).

The audio characteristics are defined for each of these functions. Furthermore, the bandwidth of the transmitted speech is taken into consideration.

The principles used are identical with those for telephony terminals. That is, the sensitivity for handset function and handsfree function designed for personal use/a small group of users is specified in loudness ratings, and the sensitivity for conference terminals is specified as output levels.

4.2.1.1 Test principles

4.2.1.1.1 Handset function

The sensitivity measurement of a terminal when a handset is used shall be based on the principles described in Recommendation P.64. The loudness rating shall be calculated as described in Recommendation P.79.

4.2.1.1.2 Handsfree function for a small group of users

The sensitivity measurement of the handsfree function of a terminal designed for a small group of users shall be based on the principles described in Recommendation P.34. The applied test signal level at the digital input when measuring receive sensitivity shall be -30 dBm0.

The user position for a visual telephone terminal depends on the design of the terminal. The real user position as recommended by the supplier might be different compared with the position used for measurements. A correction factor shall be used. The correction factor is:

$$F(dB) = 20\log_{10}\left\{\frac{D_s}{D_o}\right\}$$

where D_s is the distance between the recommended user position and the terminal and D_o is the reference distance of 50 cm.

The loudness rating shall be calculated as described in Recommendation P.79.

4.2.1.1.3 Handsfree function for a conference terminal

The principles described in Recommendation P.30 shall be used.

4.2.1.2 Sensitivity

4.2.1.2.1 General

For handset terminals and handsfree terminals designed for a small group of users the sensitivity shall be specified as loudness ratings; Send Loudness Rating (SLR) and Receive Loudness Rating (RLR). Definition of SLR and RLR is found in Recommendation P.10.

For conference terminals the sensitivity shall be specified in terms of input and output levels.

4.2.1.2.2 Receive volume control

For handsfree and loudspeaking terminals a volume control shall be provided.

Where a manual receive volume control is provided, the minimum control range shall be to -15 dB from the test position.

Where an automatic receive volume control is provided, the RLR value obtained with a line level of -15 dBm0 shall not exceed that RLR value which is obtained with a line level of -30 dBm0 by more than 15 dB.

4.2.1.2.3 Handset function

The requirements of Table 4 shall be met.

TABLE 4/H.320

Sensitivity of the handset function

	3.1 kHz bandwidth (Note)	7 kHz bandwidth
SLR	8	8
RLR	2	7
Note - 3.1 kHz bandwidth includes both G.711 and G.728 coding.		

The manufacturing tolerances are ± 3 dB.

4.2.1.2.4 Handsfree function

The requirements of Table 5 shall be met.

TABLE 5/H.320

Sensitivity of the handsfree function

	3.1 kHz bandwidth (Note)	7 kHz bandwidth
SLR	13 - F	13 - F
RLR	-7 - F	-5 - F
Note - 3.1 kHz bandwidth includes both G.711 and G.728 coding.		

The receive RLR requirement shall be met when the receive volume control is in its maximum position. The manufacturing tolerances are ± 4 dB.

4.2.1.2.5 Conference terminals

The procedures and values specified in Recommendation P.30 shall be used.

4.2.2 Video arrangements

Under study.

4.3 Delay compensation in the audio path

The H.261 video codecs require some processing delay, while the H.200/AV.250-Series audio codecs involve much less delay. Hence, unless the audio is further delayed, the sound will be presented significantly in advance of the lip movements of the speaker. If lip synchronization is to be maintained, the video processing delay must be compensated in the audio path. Since video coder and decoder delays may vary according to implementation, delay compensation must be carried out individually at the coder and decoder. A reference measurement method of video coder and decoder delays is defined in Recommendation H.261.

The insertion of lip-synch delay is not mandatory: some users may prefer the immediate speech for its better interactivity, while others prefer lip synchronization, accepting the delay in interaction it causes. If one user selects delay insertion, the terminal should send the code ACE (see Recommendation H.230) to request that the remote end-point do likewise; if zero-delay is selected, ACZ should be sent.

When delay is inserted at the transmitting end, care should be taken not to insert a disturbing noise into the encoded audio path, since the receiving end will not be able to mute this out, not knowing exactly when it will occur. When delay is inserted at the receiving end, similar care should be taken to avoid disagreeable noise presentation.

If delay is inserted in the transmission direction it should also be inserted at the same time in the receiving direction. The insertion could be made:

- at the very start of Phase B (note that if the remote end is not a videophone, the delay is then inappropriate);
- when incoming frame structure is detected (the remote end is then more likely to be a videophone, but could also be an audiographic terminal, for which the delay is then inappropriate);
- when switching video ON [note that this choice may disturb (twice over) the user's initial greetings, and so should be avoided];

- later in the call, or not at all.

4.4 Control and Indications (C&I)

C&Is are chosen from the general audiovisual set contained in Recommendation H.230. For visual telephone systems, those signals in Table 6 are mandatory, where their source, sink, synchronization with picture, transmission channel and codewords are indicated. Additionally support of C&I signals listed in Table 7 is highly desirable for multipoint communication.

All visual telephone terminals have a video source providing a picture of participants, and some terminals may have additional video sources; the participant-picture source is designated #1, having the associated symbol VIA. When incoming video is ON (BAS command (010) [1 or 2]) and VIA, VIA2, VIA3 have not been transmitted, source #1 is assumed.

Some supplementary services provided by some conference bridges, servers, etc. may require terminals to send out the symbols on their dialling keyboards. It is highly desirable that all visual terminals be able to translate these symbols into BAS codes representing the Alphabet as provided for in Recommendation H.230. This functionality is essential for displaying texts input by the user in such applications as videophone for handicapped people.

TABLE 6/H.320

C&I signals mandatory for visual telephone

	C&I signal	C/I	Source	Sink	Sync. with picture	Transmission channel	Codeword definition (Rec.)
Video	Picture format	I	Decoder	Coder	No	BAS	H.221
	Picture format	C	Coder	Decoder	Yes	Embedded in video	H.261
	Minimum decodable picture interval	I	Decoder	Coder	No	BAS	H.221
	Freeze picture request control, VCF	C	Coder or MCU	Decoder	No	BAS	H.221
	Fast update request control, VCU	C	Decoder or MCU	Coder	No	BAS	H.221
	Freeze picture release control	C	Coder	Decoder	Yes	Embedded in video	H.261
MCU	Multipoint command conference, MCC and cancel-MCC	C	MCU	Terminal	No	BAS	H.230
	Multipoint command symmetrical data transmission, MCS	C	MCU	Terminal	No	BAS	H.230
	Multipoint command negating MCS, MCN	C	MCU	Terminal	No	BAS	H.230
Maintenance	Video loop request control, LCV	C	Terminal	Terminal	No	BAS	H.221
	Digital loop request control, LCD	C	Terminal	Terminal	No	BAS	H.221
	Loop off request, LCO	C	Terminal	Terminal	No	BAS	H.221
Conference	Split screen indication (Note)	I	Sending terminal	Receiving terminal	Yes	Embedded in video	H.261
Terminal	Document camera indication (Note)	I	Sending terminal	Receiving terminal	Yes	Embedded in video	H.261
	Audio active/muted indication, AIA/AIM	I	Sending terminal	Receiving terminal	No	BAS	H.230
	Video active indication VIA	I	Sending terminal	Receiving terminal	No	BAS	H.230
	Video suppressed indication VIS	I	Sending terminal	Receiving terminal	No	BAS	H.230
NOTE – The procedures to use these indication signals are under study. However, H.261 decoders shall be able to decode and ignore these indication signals.							

TABLE 7/H.320

Optional C&I signals which should be supported

	C&I signal	C/I	Source	Sink	Sync. with picture	Transmission channel	Codeword definition
MCU	Multipoint indication zero-communication, MIZ	I	MCU	Terminal	No	BAS	Rec. H.230
	Multipoint indication secondary-status, MIS	I	MCU	Terminal	No	BAS	Rec. H.230
Video	Video command reject, VCR	C	MCU	Terminal or MCU	No	BAS	Rec. H.230
Audio	Audio command equalize, ACE	C	Terminal	Terminal	No	BAS	Rec. H.230
	Audio command zero-delay, ACZ	C	Terminal	Terminal	No	BAS	Rec. H.230
Mode preference	Mode preference indicators (A-law 0F, μ-law 0F, G.722-m2, G.722-m3, Rec. G.728, Rec. H.261/QCIF, Rec. H.261/CIF)	I	Terminal	Terminal	No	BAS	Rec. H.230

4.5 Multipoint operation

It is highly desirable that terminals which do not have any auxiliary data equipment should nevertheless be able to open LSD and/or HSD as well as MLP and/or H-MLP channels so that they could participate in multipoint communications without losing video. All conformant terminals should be able to declare the "Nil_Data" capability defined in Recommendation H.221 and use the procedure defined in 12.5/H.242 for this purpose.

Other aspects of multipoint operation of the terminal are under study.

5 Intercommunications

The mechanisms for intercommunication with other services are described in the H.200/AV.240-Series Recommendations.

5.1 Intercommunication between different visual telephone terminal types

A common mode of operation is determined as described in 3.4.1. D-channel signalling should include new LLC and HLC which are appropriate for audiovisual services, but this subclause is for further study.

5.2 Intercommunication with telephony

NOTE – Description of this subclause is for communications using a B-channel.

5.2.1 Intercommunication with ISDN telephones

A call from a visual telephone to an ISDN telephone is first placed as an audiovisual call, but the ISDN telephone returns incompatible destination or the network returns recovery on timer expiry in case of no responses from the called side, then the visual telephone may switch to a speech or 7 kHz audio bearer service call.

A call from ISDN telephone to a visual telephone is accepted by the visual telephone because every audiovisual terminal is equipped with this telephone capability as a minimum function.

For both of the above cases, the operational mode of communication is G.711 speech or G.722 audio.

5.2.2 Intercommunication with PSTN telephones

A call from visual telephone to a PSTN telephone may be initiated as an audiovisual call, but the network returns no route to destination, then the visual telephone may switch to a speech or 3.1 kHz audio bearer service call. The operational mode of communication is G.711 audio coding.

A call from a PSTN telephone is routed into the ISDN as a 3.1 kHz audio call which can be responded by the visual telephone for the same reason as described in 5.2.1. The operational mode of communication is 3.1 kHz audio.

5.3 Intercommunication with other audiovisual terminals

A common mode of operation is determined according to the H.200-Series Recommendations.

6 Maintenance

Some loop-back functions are envisaged to allow verification of the functional aspects of the terminal in order to ensure correct operation of the system and satisfactory quality of the service to the remote party. The following loop-back functions (see Figure 2) are envisaged:

a) *Loop at terminal-network interface (towards network)*

Upon receiving the digital loop-back BAS, loop-back is activated at the digital interface of the terminal toward the network side. In case of a multiple B/H_0 channel arrangement, loop-back is activated in each connection.

b) *Loop at terminal-network interface (towards terminal)*

The procedure is for further study.

c) *Loop at analogue I/O interface*

Upon receiving the video loop-back or audio loop-back BAS, loop-back is activated at the analogue interface of the video/audio codec towards the video/audio codec.

The opportunity of having a self-checking procedure at terminal stage is for further study.

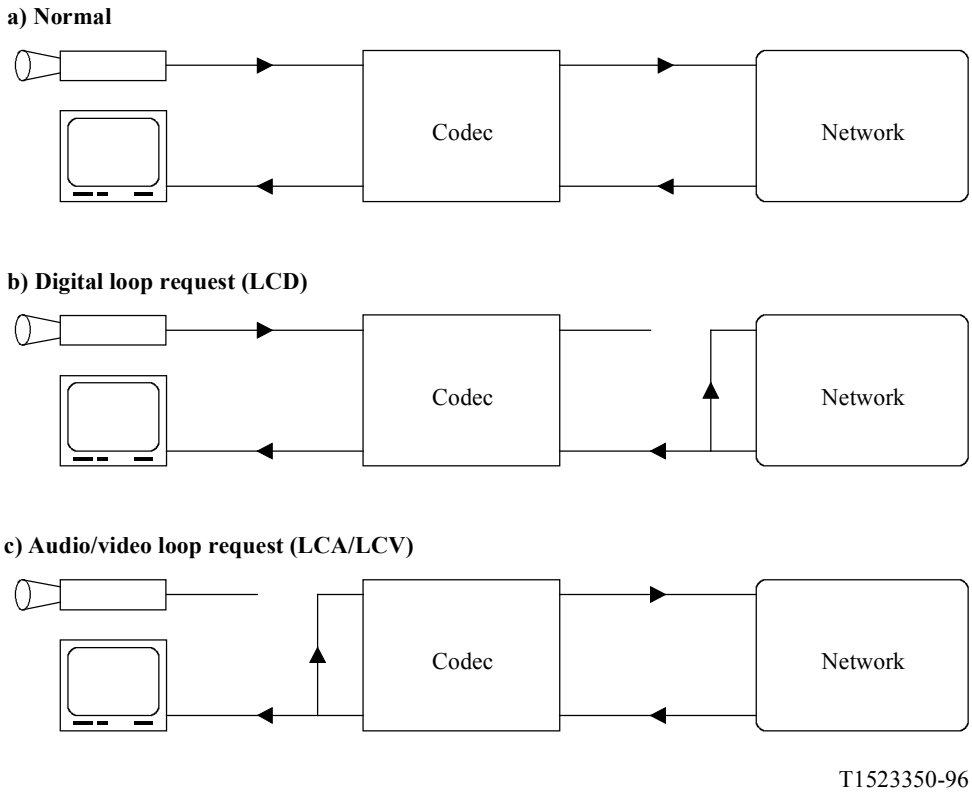


FIGURE 2/H.230

Loop back

7 Human factor aspects

To achieve error free and uncomplicated utilization of terminal equipment and service from the users standpoint, human factor related aspects have to be studied and recommended. These aspects deal with the flow of information between user and terminal/network. This information can be divided into a physical section and a logical section of the HMI.

7.1 Physical section

– Figures and properties of transducers (camera, microphone, etc.).

– Signals particularly related to the service, keys, pictograms.

7.2 Logical section

– Procedures, e.g. for call establishment/release, during communication phase.

– Consistency between the MMIs of visual telephone and terminals of other teleservices.

ITU-T RECOMMENDATIONS SERIES

Series A	Organization of the work of the ITU-T
Series B	Means of expression
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Telephone network and ISDN
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media
Series H	**Transmission of non-telephone signals**
Series I	integrated services digital network
Series J	Transmission of sound-programme and television signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	Maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound-programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminal equipments and protocols for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communication
Series Z	Programming languages

Printed in Switzerland

Geneva, 1996