ETSI

ETSI TECHNICAL REPORT

ETR 297

July 1996

Source: ETSI TC-HF

Reference: DTR/HF-01006B

ICS: 33.020

Key words: MMI, telephony, video

Human Factors (HF); Human Factors in Videotelephony

ETSI

European Telecommunications Standards Institute

ETSI Secretariat

Postal address: F-06921 Sophia Antipolis CEDEX - FRANCE
Office address: 650 Route des Lucioles - Sophia Antipolis - Valbonne - FRANCE
X.400: c=fr, a=atlas, p=etsi, s=secretariat - **Internet:** secretariat@etsi.fr

Tel.: +33 92 94 42 00 - Fax: +33 93 65 47 16

Contents

Foreword

This ETSI Technical Report (ETR) was produced by the Human Factors (HF) Technical Committee of the European Telecommunications Standards Institute (ETSI).

ETRs are informative documents resulting from ETSI studies which are not appropriate for European Telecommunications Standards (ETS) or Interim European Telecommunications Standards (I-ETS) status. An ETR may be used to publish material which is either of an informative nature, relating to the use of the applications of ETSs or I-ETSs, or which is immature and not yet suitable for formal adoption as an ETS or I-ETS.

The intended users of this ETR include:

	User	ETR used for	Potential Benefit
1	Service providers	Guidance on user issues for the design of services.	
2	Terminal designers	Guidance on user issues for the design of terminals. To assist the approval of terminals.	Increased usability, acceptability, and uptake of services and products.
3	Network operators	Guidance on user issues for transmission provision.	
4	Regulatory bodies	To assist the approval of terminals.	
5	ETSI TCs	Development of standards. Supporting user requirements.	Harmonisation of Recommendations.

Introduction

This ETR is intended primarily for use by designers and manufacturers of videotelephones, as a guide to improving the usability of their products. The report discusses Human Factors (HF) issues related to the design of videotelephony terminals and the man-machine interface necessary for the access to services and the control of tasks involved in videotelephony. Poor design may induce errors in use, and the concept of usability is introduced.

Videotelephony devices should have natural user interfaces, be easy to learn and use, and be less prone to error. Human physiology, skills of different users, including the disabled and the elderly people, and expectations with ergonomics of the videotelephone system should be considered in the design and implementation see RACE 1065 [1].

It is acknowledged that it may be necessary at times for manufacturers of terminals to compromise usability to some degree, where technical or production aspects of design are taken into account. One of the aims of this ETR is to ensure that any such compromise favours the user. Unless indicated explicitly, the Recommendations made in this ETR apply to ISDN and broadband videotelephony (in some cases also to Public Switched Telephone Network (PSTN videotelephony), and may also in principle to Internet videocommunications.

Wherever possible, Recommendations are based on human sciences research and behavioural considerations related to tasks in telecommunications services and are aimed at guidance and not standardization.

Clause 4 briefly presents an overview of user requirements and usability issues, including a definition of usability, its specification and evaluation.

Clause 5 covers HF Recommendations for a range of videotelephone functions, including call control and control of service modes, together with issues related to video and audio quality. Clause 6 addresses user procedures for videotelephony.

Clause 7 covers general HF issues pertinent to the controls and indications in videotelephony. In subclause 7.1, the Indicate-Control-Indicate (I-C-I) model is presented, which describes the minimum information which should be provided for the user about system and control states prior to and following a control action. The remaining subclauses deal with issues of indication, identification, feedback, display, control, implementation, handset, hands-free, and loudspeaking modes.

Clause 8 covers issues related to videotelephony usage by elderly and disabled persons.

1 Scope

This ETR identifies some of the important Human Factors (HF) usability issues in telecommunications terminals and services, particularly as applied to videotelephony. The ETR contains Recommendations based on good HF practice in the design of man-machine interfaces, including providing for the needs of disabled and elderly users.

The ETR is intended for designers and manufacturers of services and terminals and those engaged in the preparation of standards. This ETR is not a standard, but will supplement existing or future standards in the subject area. As such it contains references to original source material for further reading by interested parties.

The videotelephone functions and user procedures proposed provide for most of the basic functionality needs for point-to-point videotelephony. Videoconferencing and multi-point videotelephony are not dealt with in this ETR, and reference should be made to ETR 175 [2].

This ETR makes no Recommendations concerning the type of feedback indications or text messages that might be employed for inform the user about call progress, as outlined in the procedures. Neither are any specific Recommendations made about controls, indications, or their layout, as these are dependent on implementation by the terminal manufacturer.

2 References

For the purposes of this ETR, the following references apply:

[1] RACE Project 1065: "ISSUE Videotelephony Guidelines".

[2] ETR 175 (1995): "User procedures for multipoint videotelephony".

[3] ETR 116 (1993) (ETSI Handbook): "Human Factors Guidelines for ISDN Terminal Equipment Design".

[4] Lindgaard, G (1993): WOW - 568 "Smart features on your PABX: What really determines the uptake of technology?"; Proceedings of the 14th International Symposium on Human Factors in Telecommunications. R. v. Decker's Verlag.

[5] ISO 9241-11: "Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability".

[6] ETR 095 (1993): "Guide for usability evaluation of telecommunications systems and services".

[7] ETS 300 145 (1994): "Integrated Services Digital Network (ISDN); Audiovisual services. Videotelephony systems and terminal equipment operating on one or two 64 kbit/s channels".

[8] ETS 300 264 (1994): "Integrated Services Digital Network (ISDN): Video telephony teleservice, Service description".

[9] Mühlbach, L, Böcker, M, & Prussog, A (1993): "Communicative presence in videoconferencing"; Proceedings of the 14th International Symposium on Human Factors in Telecommunications. R v Decker's Verlag.

[10] Flohrer, W & Weikinnis, H (1988): "Man-machine aspects of an office videophone"; Proceedings of the 12th International Symposium on Human Factors in Telecommunications.

[11] Frowein, HW, Smoorenburg, GF, Pyfers, I and Schinkel D (1991): "Improved Speech Recognition Through Videotelephony: Experiments with the Hard of Hearing"; IEEE Jnl. on Selected Areas in Telecommunications, Vol. 9, No 4, May 1991.

[12] Kellner, B, Mühlbach, L. & Romahn, G (1988): "On the audio quality in a videotelephone service"; Proceedings of the 12th International Symp. on Human Factors in Telecommunications.

[13] CCITT Recommendation P.34 (1989): "Transmission characteristics of hands-free telephones". CCITT Blue Book, Vol. V.

[14] ETR 198 (1995): "User trials of user control procedures in ISDN videotelephony".

[15] CCITT Recommendation P.31 (1989): "Transmission characteristics for digital telephones". Blue Book, Vol. V.

[16] Böcker, M. & Anderson, D (1993): "Remote conference interpreting using ISDN videotelephony: A requirements analysis and feasibility study"; Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting.

[17] Mühlbach, L (1988): "User-friendly videotelephone terminals"; Proceedings of the 12th International Symp. on Human Factors in Telecommunications.

[18] CCITT Blue Book, Vol. V, Supplement 16 (1989).

[19] Romahn, G & Mühlbach, L (1988): "Multipoint teleconferencing: Empirical studies with different variants"; Proceedings of the 12th International Symp. on Human Factors in Telecommunications.

[20] CCITT Recommendation G.114 (1989): "Mean one-way propagation time"; Blue Book, Fasc. III.1.

[21] CCITT Recommendation G.131 (1989): "Stability and echo"; Blue Book, Fasc. III.1.

[22] CCITT Recommendation P.30 (1989): "Transmission performance of group audio terminals (GATs)"; Blue Book, Vol. V.

[23] ETS 300 738: "Human Factors (HF); Minimum Man-machine Interface (MMI) to public network based supplementary services".

[24] ETR 170 (1995): "Generic user control procedures for telecommunications terminals and services".

[25] RACE Project 1065 - ISSUE: "Human Factors Guidelines for Multimedia".

[26] Nuclear Regulatory Commission: NUREG-0700. "Human Factor Guidelines for Control Room".

[27] ISO 3864: "Safety colours and safety signs".

[28] IEC 73: "Coding of indicating devices and actuators by colours and supplementary means".

[29] HFES 100: "Human Factors Engineering of Visual Display Terminal Workstations"; ANSI.

[30] Böcker, M (1993): "A Multiple Index Approach for the Evaluation of Pictograms"; Proceedings of the 14th International Symposium on Human Factors in Telecommunications. R v Decker's Verlag.

[31] ETR 070 (1993): "The Multiple Index Approach (MIA) for the evaluation of pictograms".

[32] ETS 300 375 (1994): "Pictograms for point-to-point videotelephony".

[33] ETR 113 (1995): "Results of an evaluation study of pictograms for point-to-point videotelephony".

[34] CCITT Recommendation E.180 (1995): "Technical characteristics of tones for the telephone service"; Blue Book, Fasc. II.2.

[35] CCITT Recommendation E.181 (1989): "Customer recognition of foreign tones"; Blue Book, Fasc. II.2.

[36] CCITT Recommendation E.184 (1989): "Indications to users of ISDN terminals"; Blue Book, Fasc. II.2.

[37] ETR 187 (1995): "Recommendation on characteristics of telephone service tones when locally generated in telephony terminals".

[38] ISO/IEC 13714: "Information Technology, Document processing and related communications; User interface to telephone-based services; voice messaging applications".

[39] ETR 096 (1995): "Human Factors (HF): Guidelines for the design of minimum phone-based interfaces for computer-based services".

[40] Gili, J & Clarke, A (1991): "Icons in Multipoint Videotelephony an International Study"; Deliverable no. 65 /TID /HFG /DS /A /003 /B1.

[41] Gili, J, Vivas, A & Reillo, Ch: "Methodology for Algorithm Evaluation"; TRANSIT: RACE-2055.

[42] Lindström, JI (1993): "Picture Communication for People with Disabilities: Promising Results from Pilot Projects in Sweden"; Proceedings of the 14th International Symposium on Human Factors in Telecommunications. R v Decker's Verlag.

[43] Pereira, LM & Lindström, J-I (1994): "Videotelephony for disabled and elderly people"; COST 219. Commission of the European Communities, Directorate-General Telecommunications, Information Industries and Innovation.

[44] COST 219 (1991): "Issues in Telecommunications and Disabilities"; von Tezchner, S (ed.), Commission of the European Communities, Directorate-General Telecommunications, Information Industries and Innovation.

[45] Wilson, F, Martin, M, Osman-Allu, N, & Stephens, A. (1993): "Specifying acceptable service quality in videotelephony for deaf people"; Proceedings of the 14th International Symposium on Human Factors in Telecommunications. R v Decker's Verlag.

[46] Frowein, HW, Smoorenburg, GF, Pyfers, L & Schinkel, D (1991): "Improved speech reception through videotelephony: An experiment with the hard of hearing"; IEEE Journal on Selected Areas in Communications, 9, 611-616.

[47] Klaus, H. & Fellbaum, K (1993): "Design Aspects of Man-Machine Interfaces for Voice-controlled Applications with Respect to Physical Disabilities"; Proceedings of the 14th International Symposium on Human Factors in Telecommunications. R v Decker's Verlag.

[48] IEC 417: "Graphical Symbols for use on equipment, survey and compilation of the single sheets".

3 Definitions and abbreviations

3.1 Definitions

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

control: Any system component (e.g. switch, key, mouse, voice input control) used to produce a change in the videotelephone system state.

effectiveness: The accuracy and completeness with which specified users can achieve specified goals in particular environments.

efficiency: The resources expended in relation to the accuracy and completeness of goals achieved.

indication: Any system component (e.g. LED, LCD text, pictogram, acoustic signal, etc.) used to indicate a videotelephone control state or system state.

interface design principles: Generic rules and concepts that are required for an effective use of any interface. The principles should consider the system's requirements, as well as an effective user model (ETR 116 [3]).

NOTE: Effectiveness and efficiency are also referred to in combination as "performance measures".

satisfaction: The comfort and acceptability of the videotelephone system to its users and other people affected by its use.

usability attributes: The features and characteristics of a product which influence its usability.

usability measures: The specific metrics that quantify the usability components of effectiveness, efficiency, and satisfaction, comprising performance and attitude measures.

usability: The effectiveness, efficiency, and satisfaction with which specified users can achieve specified goals in particular environments.

3.2 Abbreviations

For the purposes of this ETR, the following abbreviations apply:

CRT	Cathode Ray Tube
ISDN	Integrated Services Digital Network
ISO	International Standards Organisation
LCD	Liquid Crystal Display
LED	Light Emitting Diode
PABX	Private Automatic Branch Exchange
PSTN	Public Switched Telephone Network
TELR	Talker Echo Loudness Rating

4 User requirements for videotelephony

4.1 General requirements

As with all telecommunications services, the functionality of videotelephony is available only as a result of user actions. The user, in turn, requires a set of controls, appropriate feedback, and an easily learnable man-machine dialogue. An example is the placing of a normal telephone call: the user lifts the handset, gets feedback if a line is available, dials a number digit-by-digit (and gets visual or audio tone feedback on some phones), with final feedback via the ringing tone before being connected to called party.

This simple sequence is easily learned and remembered and the only basic variations are for number unobtainable or busy. Extra functions and facilities, such as conference calls, short code dialling, stored numbers etc. require more complex access procedures and are often under-utilised by the average user, because the procedures are difficult to learn and remember. In particular, many of the large numbers of features available in PABXs are complex, little understood and under-utilised as a result in (Lingaard,G) [4].

A videotelephone has many more added functions, such as the video picture itself or self-view and therefore each function requires a procedure plus controls and indications for its use. These procedures should be based on the users' expected knowledge, reactions, and skills, and be consistent with other telecommunications procedures.

The video function carries a number of options, including that of the user himself deciding whether to use video when making or receiving calls. In addition, he can view himself (self-view), use video for external or document viewing and may be able to change the quality of audio and video.

These diverse, and as yet relatively unfamiliar functions to most users, require simple, logical dialogues and rational layout for controls and indications, supported by carefully compiled user guidance, both built-in to terminals and in written form.

4.2 Usability

Usability is one of the main concerns of HF in telecommunications. The aim of providing HF guidelines is to improve performance and user satisfaction by reducing errors and improving efficiency and safety.

Error has a general effect on the users of telecommunications services in wasted time and increased cost, as well as irritation and frustration, directed mainly at the provider of the service.

The overall aim of this ETR is to maximise and promote the usability of videotelephone systems. This clause is included to promote the concept of usability and to encourage the use of usability specification and measurement, as defined by ISO 9241-11 [5] and described in ETR 095 [6].

4.2.1 Usability specification

When specifying or measuring the usability of a videotelephone system, the following information should be provided (see also ISO 9241-11 [5]):

- a precise and verifiable description of the relevant characteristics of the context of use, including users, equipment, environments and user tasks. This may be a description of an existing context, or a specification of a required context.

The tasks should be described in terms of the goals or results the user can achieve, and not in terms of the functions provided by products or the activities or steps that should be performed to achieve these goals. Other relevant characteristics of the task, such as frequency or duration, should also be described:

- a specification of the quality of interaction. This should consist of measurable values of performance and acceptability, primarily in terms of effectiveness, efficiency, and satisfaction.

4.2.2 Usability evaluation

Finally, the importance of evaluation should be stressed as part of the design process, in order to assess usability and give directions for improvement.

As a result of the evaluation of the usability of videotelephone systems and services, it is recommended that the manufacturer, supplier, or service provider collect data from laboratory tests and user trials that assess the level of the usability components of effectiveness, efficiency, and satisfaction (see ETR 095 [6]). Typically, usability measures would include performance time and errors as well as subjective measures such as satisfaction, acceptability, aesthetic appeal, stress, etc. Some index of learning, such as the number of trials to reach an acceptable performance, may also be included.

In particular, it is recommended that this data be collected for all usability attributes relevant for all goal tasks (the task the user actually wishes to accomplish, i.e. videotelephone communications) and enabling tasks (the steps the user has to perform prior to being able to perform his goal task) the videotelephone equipment is intended to support. For example:

- installing and setting up the videotelephone;
- making videotelephone and telephone calls;
- receiving videotelephone and telephone calls;
- completing videotelephone and telephone calls;
- converting between videotelephone and telephone calls;
- making use of ISDN network services such as multipoint calls, message transfer, call forwarding, and call back.

As usability measures are more often comparative than absolute, it is desirable for the data collected to be compared with similar data from other systems, so that improvements and deficiencies can be recognised, e.g. making a telephone call on a videotelephone vs. ISDN business telephone vs. conventional analogue telephone.

It is desirable that manufacturers, suppliers, or service providers keep a database of user performance and preference for their products and services, covering the parameters of performance and attitudes, as a reference against which future products and systems can be assessed.

5 HF Recommendations of aspects of videotelephone functions

The basic set-up for videotelephone communications requires at least one monitor, a camera, an audio connection, and some form of input device all of which can be integrated in the form of a dedicated terminal, a PC, or a multimedia terminal. Essential video features are given in RACE Project 1065 [1], ETS 300 145 [7] and are also under investigation in ETSI STC-TE4[1)]:

- call status feedback, which indicates when an incoming call is in video mode;
- camera on/off (video pause, camera freeze), which inhibits the transmission of outgoing video;
- self-view function, which allows users to view their own image;
- screen adjustability (e.g. colour, brightness, and contrast).

A generic videotelephone system description can be found in ETS 300 145 [7]. Basic audio functions include handset and hands-free modes, which will be reviewed separately, and the possibility of adjusting the terminals audio output by the user. The terminal should be equipped with a microphone on/off function (mute, suppression of outgoing audio) for reasons of privacy (see RACE 1065 [1]).

5.1 Call control functions

The following control functions are necessary for the safe and simple operation of the videotelephone service:

- on/off hook and dialling function (i.e. call set-up function);
- control of service mode.

1) Work Item DTR/TE-04113 on a videotelephone reference terminal.

5.1.1 On/off hook and dialling functions (call set-up)

The procedures for these functions, i.e. call set-up procedures, should be consistent with audio telephony today and with future ISDN procedures. The user procedures for ISDN videotelephony, including hands-free operation and en-bloc dialling, are the subject of a draft I-ETS within TC HF[2)].

5.1.2 Control of service mode

Broadband, ISDN and PSTN videotelephony differ in terms of the bandwidth available for audio and video transmission. Broadband videotelephony (> 2 Mbit/s) approaches the video and audio quality of broadcast television.

In PSTN videotelephony, three operating modes are possible: audio plus video, audio only (telephony mode) and video only. However, in the absence of agreed standards, different coding algorithms are employed so that different manufacturers videotelephone terminals may not be compatible with each other and they are not compatible with ISDN-videotelephony terminals.

ISDN videotelephony terminals employ internationally agreed coding algorithms making terminals of different manufacturers compatible with one another, and are capable of interworking with PSTN tele-phony. ISDN videotelephony allows for a number of pre-defined combinations of bandwidth allocations for audio and video called service modes (ETS 300 264 [8]). The important HF issue in ISDN-narrowband videotelephony with two B-channels (each with 64 kbit/s) relates to call set-up and quality-selection procedures.

5.1.2.1 Description of service mode in ISDN videotelephony

In a call establishment procedure, when ISDN communication is set up (one B-channel connected), a terminal-to-terminal negotiation procedure starts. This negotiation ends up with choosing a common service mode in which to operate initially (details on the negotiation procedure and the parameters for each service mode are defined in ETS 300 264 [8]).

Service mode is a difficult concept to convey to the user. It is a function of subjective quality of audio and video separately and combined. It also defines total transmission capacity (one or two B-channels). A user may wish to use two B-channels to get the best possible quality. Or he may wish to use one channel only to keep the cost of the call as low as possible or to keep the second B-channel free for incoming calls.

5.1.2.2 Principles for procedures

The considerations in the previous subclause have led to some principles for the user procedures for control of service mode:

- the user should be able to select a service mode from a limited set with not more than five or six modes; this simplifies the dialogue;

- the active mode should be indicated to both calling and called user;

- the user should be guided to set the default service mode for incoming and outgoing calls to be the same. There is evidence that users get confused if, for example, the default service mode on incoming calls is audio-only (to ensure privacy) and on outgoing calls is audio-visual;

- for the preservation of privacy, users should be able to recognise an incoming call as being in video mode before they answer the call;

- both the calling and called user may change the service mode at any time during a call. Any increase in cost should be charged to the person requiring the service mode change;

- an indication that the selected mode is not possible should be given if the user selects a mode which is not supported by the other side, or if the called user selects a mode which would increase the cost of the call. The same indication should be given in both cases to make the procedure appear symmetrical with respect to calling and called user;

---

2) Work Item DI/HF-01018.

- it should be possible to change service mode (e.g. upgrading from one to two channels to improve video quality) during a call without having to terminate the call;

- it should be possible to select one of the modes in the set as a default mode to be active whenever a call starts. In this way, consistency with telephony call set-up procedures can be kept as the user does not have to take an active part in the initial service mode negotiation procedure (see previous subclause), i.e. the user does not have to select service mode in every call set-up phase.

5.2 Video related functions

This subclause deals with the video related HF aspects of videotelephony. The special cases of document camera and external camera source control are not addressed but it is acknowledged that they have important HF aspects.

5.2.1 Incoming video indication

One of the main principles of the user procedures is to ensure the user's privacy. An indication should therefore be given to the B-party (called party) when an incoming call is in video mode. This indication should be acoustic (e.g. an alerting signal with a clear difference between an audio-only ring signal and an audio-visual ring signal) and/or visual (e.g. a screen message) or optical (e.g. an LED, see also subclause 7.2).

NOTE 1: This will enable the called party to decide whether to accept the call in video mode or not. The camera on/off and/or service mode control functions can be used to control this.

NOTE 2: There is strong evidence that the incoming video indication should NOT be a flashing LED associated with a service mode change key, especially if the default service mode for incoming calls is audio-only and the user action of cancelling the flashing LED by pressing the service mode change key accepts the call in audio-visual mode.

5.2.2 Camera on/off (video pause, one way; video cut)

There are three options for controlling visual privacy during a videotelephone call: to conduct an audio-only telephone call, to use a mechanical shutter in front of the camera, and to use the camera off function. The camera off function (video pause) inhibits outgoing video transmission as long as it is activated. It should not have any effect on the active service mode or call connection, i.e. when it is selected the call will continue in the same service mode as before the function was activated.

A simple and ergonomic solution to implementing the camera on/off function consists in a mechanical shutter pushed by the user in front of the optic of the camera. This way, the user can be certain that he cannot be seen by his interlocutor.

The camera off function has no effect on incoming video, i.e. it works one way only. The user on the other side should be given an indication whenever the camera-off function is active.

It should be possible to pre-set a default state for this function, i.e. whether it is normally on or off. Which default state is preferred is a matter for further study. It is likely that until videotelephones are widely used and accepted, some users may want camera transmission initially off at the start of a call. As users get used to being seen, a shift towards having camera transmission on at the start of a call may occur. One should note that having camera off as a default requires an additional action by the user to set up a full videotelephone call, and will on that point differ from call set-up procedures for telephony.

5.2.3 One-way videotelephony (single-point videotelephony)

Apart from camera-off working as a security function for the user (the user can control whether he is to be seen or not), there are many possible applications of one-way videotelephony. For instance, when entering a public video database, the user wishes to receive video and maybe audio, but not to send. One-way videotelephones could also be used to authorise entry to secure premises, e.g. a company with buildings distributed geographically could have a central manned security post for identification. In this case, only one-way video is used, but maybe two-way audio.

One-way videotelephony offers a wide range of uses for videotelephones, apart from straightforward face-to-face conversations, and may be important in determining how widespread the use of videotelephones will be.

5.2.4 Self view function

The self-view facility enables users to view themselves to ensure they have a presentable image. It may also be used to view documents or other objects if a separate document camera or other camera is used.

The main principles governing the usage of this function are:

- self-view may be selected at any time, on or off-hook or during a call;
- self-view should have no effect on outgoing video.

In videotelephones with a single fixed camera system, self-view should be displayed in mirror image form to comply with user expectations of movement direction, and in coded form to enable checking of the transmitted image quality. If manual focus is provided, it is recommended that focusing is available on the uncoded image to remove the codec delay and to optimise sharp edges to achieve best focus.

In videotelephones with removable or external cameras, it is recommended that automatic, or exceptionally, manual, control be provided over the self-view image format depending its source as follows:

- external document or object camera: non-mirror view, coded and non-coded image;
- removable camera: user control over mirror/non-mirror, coded and non-coded image.

It is also recommended that:

- self-view should be displayed continuously on a separate screen or as a window on a main screen (picture in picture);

- if the self-view picture fills the whole screen where the incoming picture will appear, then it should be interrupted by the incoming picture on completion of call establishment;

  NOTE 1: This procedure is to ensure that the subscriber has an immediate indication that a videotelephone call is established. It will always be possible to revert immediately to self-view if required (see main principle above).

- self-view video should be processed by the video codec unless the delay before the whole screen is fully updated exceeds 2 seconds (in the case of full screen self-view on the same screen as incoming video).

  NOTE 2: This is because the procedure of using self-view, when presented on the whole screen, interrupts the incoming picture. If the user wants to quickly check position or lighting, the delay before the incoming picture is updated again might be intolerable and would discourage the use of the self-view function. In this case it will be better to use uncoded self-view.

5.2.5 Other video hardware considerations

5.2.5.1 Camera position and adjustments

It is desirable for the videotelephone to have parallax-free construction for face-to-face communication, where the camera system's focal plane is coincidental to the eyes on the imaged face. This can be achieved by employing a half-silvered mirror.

If a parallax difference is unavoidable, then it is recommended that the camera system's focal plane is situated centrally above the display (see [9], and [43]), and that the parallax differences between the camera axis and the image display's eye-level axis is kept to a minimum. It is recommended that the difference does not exceed 8° at the furthest point of the preferred viewing range/distance.

5.2.5.2 Camera field of view and viewing distance

It is recommended that a head and torso image be captured. The camera field of view at the nearest point of the preferred viewing range should provide at least a head and shoulders image.

To accommodate the normal range of different eye heights of the sitting and standing users within the camera's field of view, it is recommended that the camera system's focal plane be adjustable vertically. No preference is made between vertical linear or tilt adjustment, except that significant distortion should not be introduced.

Available technology will invariably dictate screen size, but to maintain a preferred viewing distance of between 50 cm and 120 cm, see table 1.

Table 1: CRT screen sizes and viewing distances

Diagonal Screen Size	Height of Screen	Viewing Distance
15 cm	10 cm	60 cm
23 cm	15 cm	90 cm
30 cm	20 cm	120 cm

NOTE: Although individual differences for viewing distance exist, the optimal distance can be computed using the following formula:

$$d = \frac{h}{2 \tan \frac{p}{4 f_{\max}}}$$

where d is the optimum distance, h is height of the display, p is display resolution, and f_{max} is the maximum spatial frequency perceived by the human visual system (50 cycles/degree, i.e. f_{max} = 50 implies that no more than 50 pixels (100 with alternate black and white ones) can be distinguished when seen within one degree angle of vision).

Whether these dimensions apply to other screen technologies, such as LCD, is unclear. There is no evidence to provide Recommendations at the time of publication.

Any documentation associated with the videotelephone should reflect the preferred viewing distance.

5.2.5.3 Camera lens

A fixed lens appropriate for capturing the standard head and torso section of the user is recommended. In some situations, the user may want a field of view other than head and torso. If lenses of different types are required, such as zoom, wide angle, etc. this will have some impact on the codec working and hence picture quality. This should be studied during usability evaluation. Any degradation of normal picture quality should be stated in documentation.

5.2.5.4 Camera iris control

To cope with a possible wide range of lighting conditions, the camera should have an iris control. It is recommended that the control of the iris be automatic. It must be noted that the location of the terminal with respect to bright light sources will influence the quality of the picture. For videotelephones with automatic iris control, the impact on the codec of different rates of changes of lighting is for further investigation.

5.2.5.5 Camera focus control

It is recommended that the camera lens used have provision for focusing. It is recommended that focusing be controlled by the user from the normal viewing distance. It is further recommended that an automatic focusing device, if fitted, be activated for only one focusing action at a time, at the user's request.

5.2.5.6 Ambient lighting and background

The ambient lighting provided for videotelephone users can have a marked effect on the performance of the videotelephone and on the quality of the transmitted image. It may also affect the "legibility" of the screen. The lower the transmission rate, the more important lighting becomes.

The general requirements are for diffuse ambient lighting within the range 100 lux to 10 000 lux. Direct and strong reflected light in the camera's or the user's field of view should be avoided. Direct lighting onto the screen surface is also to be avoided, to minimise viewing problems. "White" light in preference to "warm white" or "incandescent" light should always be used to avoid colour distortion of the transmitted image.

Movements in the users background should always be avoided as it occupies processing time in the codec and reduces the overall video quality. The background should avoid patterns, be non-reflective, and provide good contrast to the user.

It is recommended that simple, clear instructions on basic lighting and background requirements and on diagnosing and correcting typical lighting errors be included within the user documentation.

5.2.5.7 Colour correction

To ensure the optimum colour rendition of the transmitted image, the camera system needs to accommodate the colour temperature of the ambient light source (normally automatic white balance adjustment).

In exceptional circumstances and for some special environmental conditions or user tasks, it may be desirable for the user to be able to set the "white balance" or adjust the relative colour values. If such a provision is made, it is recommended that provision be also made so that the user can make a rapid return to the default or original settings.

5.2.5.8 Screen adjustability

To accommodate the range of eye heights of the users and to enable them to optimise the screen viewing angle to avoid screen reflections, it is recommended that the positioning of the screen be user adjustable. Adjustment may be made in three planes: Tilt (inclination of the surface of the screen with respect to the horizontal axis), swivel (inclination of the surface of the screen with respect to the vertical axis), and vertical linear. It is recommended that tilt be provided, swivel and vertical adjustment are desirable.

Theoretically, the normal viewing angle for a seated user is about 15° below the horizontal, but in practice the average preferred viewing angle is about 35°. For most applications, the normal viewing angle can assume just about any value between 0° and 50°.

5.2.5.9 Display refresh and frame rate

If progressive CRTs are used, the display refresh rate should be at least 50 fields per second, in order to minimise flicker effects. In interlaced CRTs operating at 90 Hz, flicker effects are avoided for over 90 % of the population.

There is good evidence to recommend a minimum frame rate of at least 15 frames per second, to optimise intelligibility (including lip reading) when a videotelephone is used by the hard of hearing (see [11]).

5.3 Audio related functions

5.3.1 Audio modes

Table 2 shows the alternatives foreseen for the audio transmission in ISDN videotelephony.

Table 2: Alternatives for coding and bandwidths

Audio Mode	kbit/s	Coding Law	Audio Band (kHz)
0	64	G.711	0,3 to 3,4
1	64	G.722	0,1 to 7,0
2	56	G.722	0,1 to 7,0
3	48	G.722	0,1 to 7,0
5	32	G.721	0,3 to 3,4
7	16	(note)	0,3 to 3,4
9	64		No audio (video/data only)
NOTE: Several under study in ITU-T			

Mode 0 is also needed because a videotelephone also should be able to function, in a default mode, as an ordinary telephone connected to ISDN. Modes 1, 2, 3 and 5 are recommended because of the higher sound quality offered. Mode 7 is needed because the videotelephone should always be able to operate on one B-channel for both video and audio. Mode 9 is included because for some important applications no sound transmission is needed. At present, no ITU-T Recommendation exists for a stereophonic transmission suitable for narrow-band videotelephones.

5.3.2 Audio functions

Three audio functions should normally be provided:

1) hands-free mode;
2) normal handset mode;
3) loudspeaking, i.e. handset mode augmented by loudspeaker listening.

Note that the coding and bandwidths requirements in table 2 apply for all alternatives.

Alternative 1, i.e. hands-free, is in general to be preferred provided that echo cancellation is adequate and a satisfactory audio quality can be obtained with reasonable cost (see [12]). However, because difficult acoustic conditions may occur (high room noise, high reverberation, acoustic echoes) alternative 2 is required and alternative 3 may be offered. Choosing one mode or another depends on individual preferences and prior experiences. The use of hands-free mode is advantageous in situations where documents or small objects have to be shown or explained to other participants. However, hands-free mode can produce interference with other people's work, specially when several people share the same workplace.

Transmission characteristics of hands-free telephony such as sending and receiving sensitivity, frequency response curves, voice switching, and measuring conditions are specified in ITU-T Recommendation P.34 [13]. The hands-free mode should operate in rooms with a background noise level of at least 50 dB(A) and 400 ms reverberation time. (RACE PROJECT 1065 [1] p.15).

In a terminal with handset and loudspeaking modes, the procedure for switching between the two modes would simply be to switch the loudspeaker on or off. If all three modes are offered in the same terminal, there are several solutions for the user procedures. One is to keep the procedures for switching between loudspeaking and handset and between hands-free and handset separate and to provide separate controls. Another solution is to provide a procedure for switching between the three modes, most often implemented by using the hook-switch and a hands-free control. The latter solution is known to cause some confusion among users as to how one switches between modes (ETR 198 [14]). As some of the existing implementations of this solution have been rejected by users because the user procedures were unclear, particular care should be taken in designing the user procedures. The procedures should also be carefully tested.

5.3.3 Audio transmission on/off

For reasons of privacy, it is desirable that videotelephones be equipped with a control for switching on and off the transmission of the local microphone ("audio mute"). The transmission of the incoming audio is not affected by this function.

5.3.4 Telephone channel bandwidth (0,3 - 3,4 kHz)

For hands-free telephones, ITU-T Recommendation P.34 [13] applies.

NOTE 1: The relative sensitivities mentioned in clause 1 of ITU-T Recommendation P.34 [13] are to be related to the digital handset specification.

NOTE 2: In general, it takes a determined design effort to achieve a good sound quality in normal hands-free, audio-only, telephony. The addition of a video channel increases the difficulties still more:

- even a moderate initial syllable-clipping is perceived and is annoying;
- the necessary delay increases the disturbance from echoes.

For handset telephones, ITU-T Recommendation P.31 [15] applies.

For loudspeaking telephones which do not provide hands-free operation, the relevant parts of ITU-T Recommendation P.34 [13] apply.

Telephone channel bandwidth when shared with video (as in PSTN videotelephony) is sufficient for general conversations but may not meet ITU-T Recommendation P.34 [13]. There is, however, evidence that a higher bandwidth is required for some professional applications see [16].

5.3.5 Wide-band audio (0,1 - 7,0 kHz)

Videoconferencing experiments confirm that users appreciate the increased audio quality obtained by a wide-band audio channel which is free from noticeable noise and distortion see [17]. In particular, naturalness and intelligibility are increased in comparison with the normal telephone channel bandwidth.

Note, however, that under certain difficult circumstances it has proved beneficial for the overall sound quality to gradually attenuate frequencies below 500 Hz. In this way, the degrading effects of reverberations may be reduced.

The methods for specifying electro-acoustic parameters for the wide-band audio channel is under study in ITU-T. However, similar considerations apply as for telephone channel bandwidth.

5.3.6 Other audio considerations

Restricting the microphone pickup angle may diminish the acoustic echo problems. However, the microphone pickup should at least cover the pickup angle of the camera.

Guidance in acoustic matters can be found in the CCITT Blue Book, Vol. V, Supplement 16 [18]. (In particular, refer to clause 3, ambient noise level considerations, and clause 4, reverberation considerations) It is recommended that the hands-free telephone should be able to operate at least up to the limits 50 dB(A) noise, 400 ms reverberation time. When providing user guidance for videotelephone applications, the above should be taken into account by manufacturers.

Note that, when listening to a speaker, the videotelephone listener is subjected to a certain fundamental acoustic disadvantage compared to a listener in the same room as the speaker. This is because, with the monophonic systems of today, only the latter listener can use his binaural hearing ability to concentrate his attention on a particular speaker, excluding extraneous noises from for instance air conditioning fans, traffic, other persons talking etc. This is a well known phenomenon, sometimes called "the cocktail party effect".

A cleverly designed stereophonic audio system might overcome this difficulty of monophonic systems, improving the perceived audio quality under some difficult acoustic conditions like disturbing noises, multipoint connections, etc. (see in [12] and [19]).

Another important aspect to consider is the interaction between the visual and the aural impressions in person-to-person communication. In general, the combination of visual and audio tele-transmission makes the participants more sensitive to deficiencies in the audio quality because, unconsciously, a comparison is made to a real face-to-face situation (examples of noticeable effects are clipping of initial speech, non-synchronisation between audio and video, non-linear distortion, speech-level inconsistent with apparent person-to-person distance).

In PSTN videotelephony in audio-visual mode, the subjective audio quality is perceived as being significantly below standard telephone quality.

5.4 Audio and video synchronisation

Processing the video signal in the codec to reduce its redundancy introduces a delay with respect to the audio signal. This delay, called differential delay, is variable, depending on the amount of movement in the scene. Since it is not possible to compensate for this delay economically, an average fixed delay of the audio signal is introduced, which may result in a positive or negative differential delay.

Long delays cause difficulties for the users in two ways:

First, by making the personal communication more prone to confusion. It is recommended that the total round-trip delay should not exceed 800 ms unless under the most exceptional circumstances (compare with ITU-T Recommendation G.114 [20]). Several Recommendations have been addressed to this issue. To preserve lip synchronism, the delay difference between sound and vision channels should not exceed:

- 40 ms when the sound arrives after the vision;
- 20 ms when the sound arrives before the vision.

If the quality of the picture is good but audio is low quality, there may be an annoying imbalance between both qualities.

Secondly, by making echoes much more noticeable. For instance, the talker echo loss should be increased by 30 dB if the one-way delay time is increased from 10 ms to 300 ms for the subjective annoyance to be the same. Further details concerning Recommendations for TELR are given in ITU-T Recommendation G.131 [21] and evaluation methods in ITU-T Recommendation P.30 [22].

6 User Procedures

The user control procedures required for basic call point-to-point connections for ISDN videotelephones are being defined in an interim standards being drafted by TC-HF[3)]. These will cover the call set-up, incoming call, and call termination procedures for three types of ISDN videotelephones. These are:

- ISDN videotelephone in audio default mode;
- ISDN videotelephone in audio-visual default mode, with fallback to audio mode available; and
- ISDN videotelephone in audio-visual default mode, without fallback to audio mode available.

The interim standard will define a set of user states necessary to accommodate basic calls, i.e. outgoing call set-up, incoming call acceptance, and call termination, for ISDN videotelephones. The user's transition between these states is defined by one or more control procedures which specify the necessary sequence/s of control actions, prompts, and feedback indications. The interim standard will address the normal conditions; no procedures are to be offered for exceptional conditions, e.g. to provide accidental disconnection protection.

The interim standard specifically will not define which states any particular ISDN videotelephone terminal should support, nor will it define how to implement the user control procedures specified to progress between the states that are supported. These are for the ISDN videotelephone terminal manufacturers and service providers to define.

The user control procedures for point-to-multipoint connections for videotelephony have been considered in ETR 175 [2]. Recommendations are included but no specific procedures have been defined. The user control procedures for supplementary services have not been considered separately for videotelephony (PSTN or ISDN), but they are addressed generically in ETS 300 738 [23].

7 Controls and indications

7.1 Feedback

Users should be given feedback indicating the consequences of a controlling action (e.g. which system states change or which system options become available as a result of a control action). Feedback can be provided in many forms (e.g. tones, visual signals, messages and text) each of which have specific requirements which are dealt with in subclause 7.2.

The I-C-I Model (Indicate-Control-Indicate) of feedback (ETR 170 [24]) requires four items of information to be provided to the user: the status of the system prior to the controlling action, the status of the control device prior to the controlling action, the fact that the control device status has changed, and the knowledge of how the system responds to the controlling action (see figure 1).

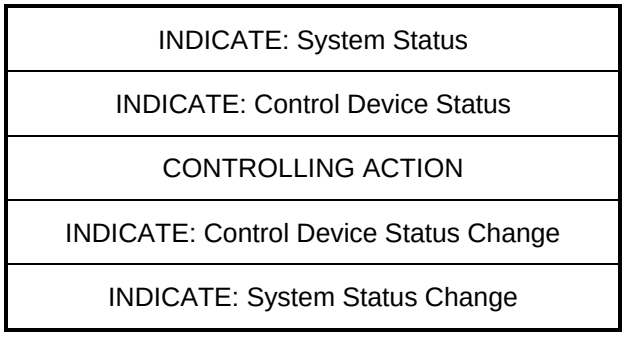


Figure 1: A basic I-C-I model for controls and indications (from ETR 170 [24])

3) Work Item DI/HF-01018.

EXAMPLE: A telephone with a replaced receiver indicates an idle system status. The position of the receiver (placed on the cradle) indicates its current status (on hook). The controlling action of lifting the receiver off the cradle leads to the indication of a change in the status of the control device (off-hook) and in the status of the system (dial tone indicating a prompt for an address). For a telephone with en-bloc dialling and loudspeaking features, however, the receiver placed on the cradle is not an unambiguous indication of an idle system state. In such cases, different indications are required.

In many cases, the indication giving feedback on one controlling action may form the status information for the subsequent controlling action. The requirement, therefore, is to provide the user with indications of all four types of information in an appropriate form and at the appropriate time.

7.2 Indication issues

Indications can be used in videotelephones as a prompt for a user action, as feedback on the correctness (or otherwise) of a user action, as information about an error condition, as ancillary help information, as a redundant information supporting other indications, etc. An indication can (inter alia) take the form of labels, pictograms, visual indications and messages, and acoustic indications. Recommended indications are listed in ETS 300 145 [7] and include indications for camera off and microphone off states.

7.2.1 Naming of functions and labelling

Labels are text indication elements which are unambiguous descriptions for structures and functions presented on control and indication elements of terminals. Labels should be used primarily to describe item functions and names, or object states. The main concern for proper labelling is legibility: the ability to distinguish single characters from each other. A number of Recommendations e.g. [25] are available to ensure the legibility of labels. These Recommendations address issues such as font style and size, spacing, abbreviations, and location.

Labels should be comprehensible by all expected users, this is, words for names should be chosen on the basis of user familiarity. Abbreviations may be used only when they are familiar to the user. Control and display labels should convey verbal meaning in the most direct manner by using simple words and phrases, and should indicate the functional result of control actions.

Issues specific to on-screen labels such as highlighting and coding have been researched specifically for videotelephony see [25]. The use of colour in displays and labelling is covered in [26] [27] [28] and [29].

7.2.2 Graphical symbols

Pictograms are a form of graphical symbols. It is recommended to provide pictograms together with textual labels to increase the chances of user recognition of items. Compared with written commands or labels, the main advantages of pictograms are shown in [1] and [30] that:

- they are more distinctive than labels;
- they require less learning time and effort than labels or any other text-based alternatives;
- they require less space and their syntax and semantics are simpler than text;
- they are language independent; and that
- there is some evidence that they reduce the likelihood of errors.

However, some disadvantages have also been found see [30], namely that:

- they have less capacity to convey abstract and/or detailed information;
- there is a risk of misinterpretation, which can be dangerous when applications are critical;
- pictograms may require prior learning and there is a limit to the number of icons which can appear simultaneously in one application; and
- they may be less adequate than text for expert users.

Pictograms can be used with great flexibility and freedom by the designer, but some human factors Recommendations should be addressed e.g.[1]:

- it is important to maintain meaning consistency among pictograms;
- each pictogram should bear a close perceptual similarity to what it represents, i.e. pictograms that are meant to act as buttons should look like buttons;
- It is important to provide visual feedback of the pictogram being selected.

The design of pictograms should comply with the following basic principles:

- clarity: to represent the function or state without ambiguity;
- simplicity: they should not contain more than two or three ideas;
- univalency: they should represent only one object or action.

It is recommended that pictograms be always empirically tested in different countries and linguistic communities, in order to ascertain that they have the same meaning for different people. ETR 113 [31] describes the Multiple Index Approach to the Evaluation of Pictograms (MIA) as the pictogram evaluation method recommended by ETSI.

7.2.2.1 Application of pictograms for videotelephony

The eight pictograms in annex A for point-to-point videotelephone controls and indications have been empirically evaluated by ETSI. They are published in ETS 300 375 [32] and the empirical evaluation study of these pictograms is described in ETR 113 [33].

The functions represented by the eight pictograms are:

- videotelephone/telephone (for switching between videotelephone and telephone modes);
- videotelephone camera on/off;
- videotelephone microphone on/off;
- videotelephone self-view on/off;
- videotelephone still picture on/off;
- videotelephone document camera on/off;
- videotelephone hands-free on/off;
- videotelephone loudspeaking on/off.

In addition, a number of videotelephone status icons have been empirically tested in an international study see [1] including functions like call alert, dial prompt, busy signal, and remote camera selection (see also subclause 7.2.4).

7.2.3 Visual indication and messages

Visual indicators and messages may be used in videotelephones as a prompt for a user action, as feedback (error or acknowledgement of control action, or change of system state), as ancillary help information, or as redundant information supporting another indication medium, e.g. flashing lamp flashing at the same cadence as an acoustic ring signal. They rely on the user seeing the intended information. Such displays are of considerable importance on videotelephone terminals used by hearing-impaired users, for normal terminal feedback indications.

Visual indicators may take the form of illuminated lamps (sometimes referred to as optical indicators), such as LEDs or filament lamps, or may be illuminated windows or annunciators, or LCD displays or they can take the form of electro-mechanical flags or indicators.

Lamps and LEDs may be information-coded by colour, shape or brightness, see under a variety of cross references in ETR 116 [3]. The rules regarding spatial layout and choice of display for different purposes are common to most user interfaces are also well documented in this reference.

Be aware that flashing lamps and LEDs should be used sparingly and preferably reserved for alarm states only. There is often a strong user tendency to cancel the flashing alarm first and then assess the status of the system. This can lead to conflicts between the user's intended goal (e.g. to answer a call in audio-only, ensuring privacy) and the actual outcome.

Annunciators, whether illuminated optical versions or mechanical, rely on presenting text or symbolic information, and may be used when instructions need to be presented to the user, or for warnings containing specific information.

Text messages may also be presented on an LCD screen, or on the main display screen, where they may be restricted to reserved parts of the screen to indicate spatial relationships, as with "soft" labels associated with soft keys, where informative text may change depending on system state.

ETR 116 [3] contains details of design Recommendations under "Character displays", "Screen messages", "Screen formatting", etc.

7.2.4 Acoustic indications

Sound is commonly used in telecommunications for indication and feedback purposes (e.g. tones used for alerting system states such as dial tone and ring tone).

Acoustic indications may be coded to distinguish sounds from one another by frequency, duration, and intensity. Telephony employs standardised tones used in public networks consisting of dial tone, special dial tone, ringing tone, busy tone, congestion tone, special announcement tone, call waiting tone, and warning tone (ITU-T Recommendations E.180 [34], E.181 [35], E.184 [36], and ETS 300 295 [37]). Additional tones should not be specified without research and user testing to prevent confusion with existing tones and to ensure high association with the chosen function or service.

A particular case is the possible use of a unique tone to indicate an incoming videotelephone call to distinguish it from an audio-only call: videotelephones that can receive both audio and audio-visual incoming calls should provide two distinctive ringing signals, one for incoming audio calls and the other for incoming audio-visual calls. This will enable the user to distinguish between the two types of incoming call without the need to look at or interact with the terminal.

Acoustic indications have advantages over visual indications because they are omnidirectional and can alert users to critical changes in the system state. Disadvantages of acoustic indications are that they are distracting during a videotelephone conversation and that they may distract others in a work place.

Speech and auditory announcements are increasingly used in teleservices e.g. "The number you have dialled is unobtainable". Such acoustic indications should preferably be used in conjunction with a warning tone e.g. the special information tone specified in ITU-T Recommendation E.180 [34]. New services are likely to make increasing use of speech announcements and reference should be made to ISO/IEC 13714 [38], ETR 096 [39], etc.[4)]

Music is a special case where a "comfort tone" is desired to cover waiting time and to indicate that the connection is unbroken see ETR 116 [3]. TC-HF investigations show that the widespread use of music should be avoided.

Dual coding with simultaneous acoustic and visual indication improves alerting and feedback to the user (see subclause 7.2.2).

See also ETR 116 [3] for further information on acoustic indications.

7.3 Control issues

7.3.1 Input hardware

Input hardware includes all elements of a terminal which allow the user to carry out the procedures associated with particular operations or functions (ETR 116 [3]). Design criteria for input devices are adequacy for the task (i.e. mouse or pointing devices for graphical displays, or keyboards for discrete data items) and human error rates in using that technology within the range of tasks expected.

Nowadays, a considerable number of input devices is available, and their use in telecommunications terminal equipment is discussed in ETR 116 [3].

---

4) For example, Work Item DTR/HF-01029.

7.3.2 Function keys

Function keys are keys dedicated to the direct selection of procedures and features. Functions may also be operated by programmable keys at the appropriate point in the dialogue under program control. The function should be consistently labelled, and appropriate feedback should be given for its use.

The following functions are usually implemented in videotelephones with function keys, and a set of pictograms have been internationally tested (see [40]):

- picture-in-picture;
- rotate;
- portrait/document camera;
- multipoint mode;
- camera on/off.

The following Recommendations about the use of function keys are extracted from RACE 1065 [1] and ETR 116 [3]:

- the layout should be organised in a left to right progression, which is the usual order of use;

- function keys and controls should be logically clustered according to their function. The arrangement should reflect the importance, frequency of use and function;

- the number of keys should be kept as low as possible. If shifted functions are employed, they should be limited to two (preferred) or three (maximum) levels;

- the operation of dedicated function keys should be consistent, and not dependent on the state of the terminal;

- dedicated function and soft keys should be labelled and highlighted to reduce the risk of errors that might have irreversible unwanted effects, i.e. terminating a call while still transmitting data; labels should be aligned with soft keys;

- function keys can enable or disable functions of the video telephone, and feedback to the user can be provided with LEDs, by acoustic indicators or screen messages.

7.3.3 Control dialogues

Control dialogues are interactions with the system about a function. The type of dialogue is critical for usability depending on task complexity. Dialogues can be based on dedicated or programmable function keys or by application software currently running on the terminal. Main criteria for choosing between these alternatives are user skills and task complexity. In any case, the basic principles of feedback and consistency should be followed.

Recommendations on control dialogue design can be found in ETR 116 [3] p.106.

7.3.4 Pictograms and icons in control elements

Where pictograms or icons are used as part of control elements, their design, specification, and evaluation should follow the Recommendations in subclause 7.2.2.

7.4 Display issues

7.4.1 Screen layout

The format of information displayed on the screen is very important for the success of a users interaction with a videotelephone. Recommendations concerning information presentation for display screens in general also apply to information presentation on display screens for videotelephones. The following Recommendations are particularly important in videotelephony see ETR 116 [3]:

- studies about the spatial/temporal arrangement during call set up recommend that the initial screen should consist of a large window of fixed size with a plain background and function icons, centred on the screen, and a small self-view at the top right corner. After call establishment, the remote image covers the large window while the self-view is unchanged;

- important information should be displayed in a prominent place to catch the users eye, however, never obscuring the image of the interlocutor;

- status information should be displayed close to the relevant window;

- search times may be shorter for information displayed in the upper half of the window than in the lower half and for windows displayed in the upper half of the screen, but the situation is less clear when the horizontal dimension is considered;

- in multipoint videotelephony, a maximum of nine windows is recommended, since search time increases exponentially with the number of windows.

In addition, in face-to-face videotelephony, it may be advantageous to place on-screen labels, indicators, and messages in a fixed position adjacent to the control area.

7.4.2 Image quality

Subjective image quality has been found to be dependent on the following transmitted image characteristics (see [41]):

- colour;
- brightness;
- background stability;
- speed in image reassembling;
- outline definition;
- dirty window effect;
- mosaic effect.

Of these, the two most important are speed in image reassembling and outline definition.

Broadband videotelephony should have a picture quality comparable with domestic TV (625 line or 525 line) under favourable conditions. Videotelephony systems can incorporate a double video channel between the terminals. This configuration makes it possible to send and receive both personal images and document information simultaneously. This option, if available, should be provided to each partner. When double video channel is implemented (i.e. to transmit documents simultaneously, the use of high resolution (1 249 lines or A4 display) is recommended for the presentation of whole pages in readable quality (ETR 116 [3] p.25).

8 Videotelephony issues pertinent to disabled and elderly people

At this stage of the development of the ISDN market aimed at use by the general public, it is important to aim at supporting the least skilled and least technical members of the user community. The more complex dialogues and powerful commands should be reserved for the skilled users of sophisticated tasks (ETR 116 [3] p.222).

Attention should be paid to the use of Recommendations summarised above for disabled and elderly users. For instance, when designing coloured displays, a trade-off should be made between the use of colour for information coding and the effective use of the displays by colour-blind people. The solution should be the selection of colours with highly differentiated contrast, permitting disabled people to distinguish among them. (RACE 1065 [25]).

In [42] proposes the following classification of disabled and elderly people:

- visually impaired;
- deaf;
- hearing impaired;
- speech impaired;
- having learning disabilities;
- motor impaired;
- elderly people showing a combination of impairments.

The videotelephone holds many promising opportunities for disabled and elderly people but also presents specific design challenges to ensure that the terminals are usable for these groups of users as well.

If the picture quality allows, and the screen refresh rate is not too low, sign language communication offers itself as the most important user area (several studies showing the feasibility of this are reported in [43] and [44]). The special requirements for videotelephony for deaf and hard of hearing people are also addressed in [45] and [46].

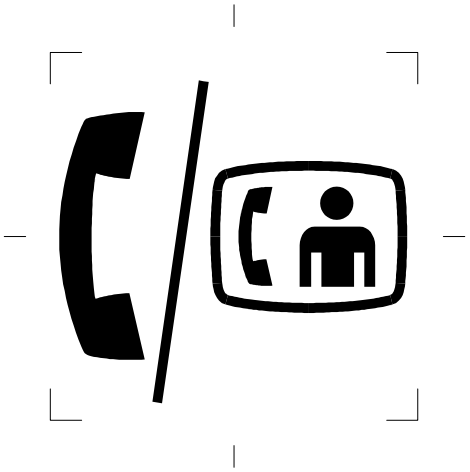
For people with learning disabilities several studies have demonstrated the positive effects of video communication see [43] and the effective use of videotelephony with other disabilities has been demonstrated in a number of studies[5)] (e.g. in [43]).

Some designs have specially taken account of these issues. In one of these studies see [47], speech control systems issues for persons with physical disabilities have been reviewed. In particular, the following issues were covered:

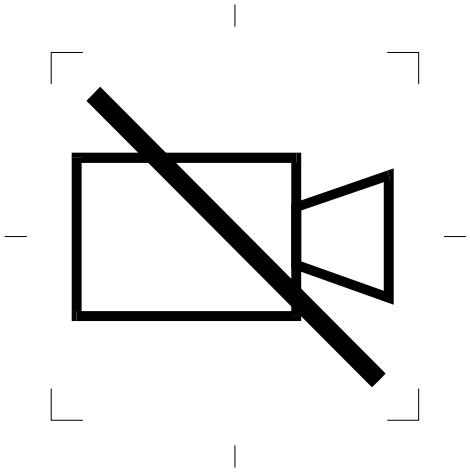
- integration of speech components for control standard office applications;
- speech components for acoustical control of terminals;
- speech control in critical situations.

5) See also Work Item DTR/HF-02003.

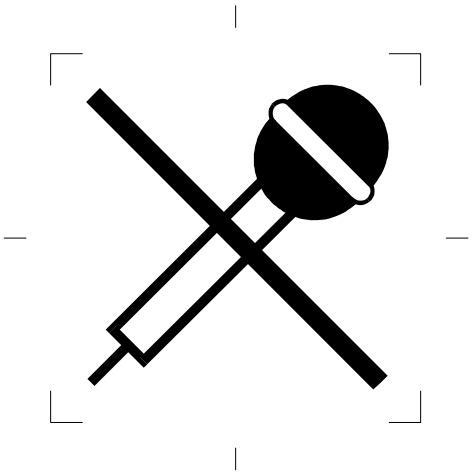
Annex A: Pictograms for point-to-point videotelephony (from ETS 300 375)



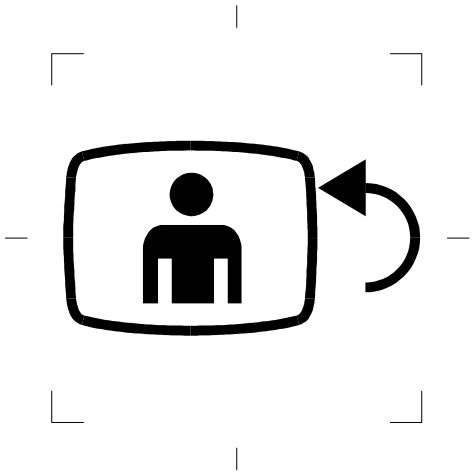
Videotelephone/telephone:
In videotelephony, for switching between videotelephone (sound and picture) and telephone (sound only) modes.



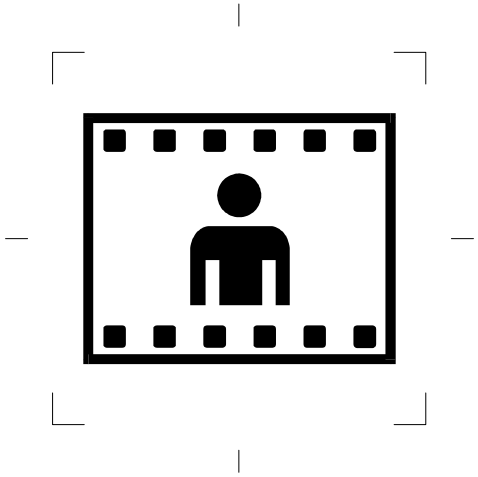
Videotelephone camera on/off:
In videotelephony, for switching on and off the transmission of the camera signal.
NOTE: The diagonal bar may be omitted when the normal system state is one in which the camera is turned off.



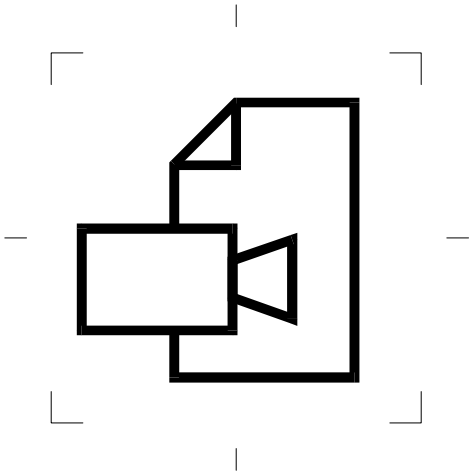
Videotelephone microphone on/off:
In videotelephony, for switching on and off the transmission of the microphone signal.
NOTE: The diagonal bar may be omitted when the normal system state is one in which the microphone is turned off).



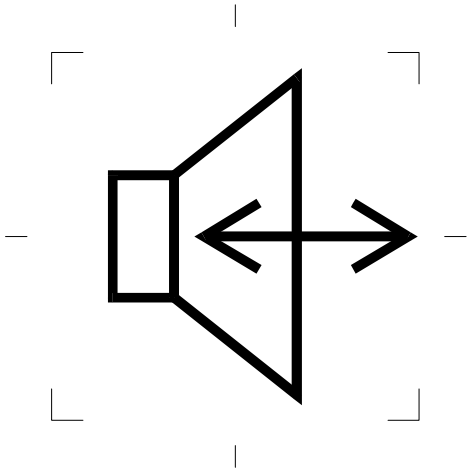
Videotelephone self-view on/off:
In videotelephony, for switching on and off the self-view function.



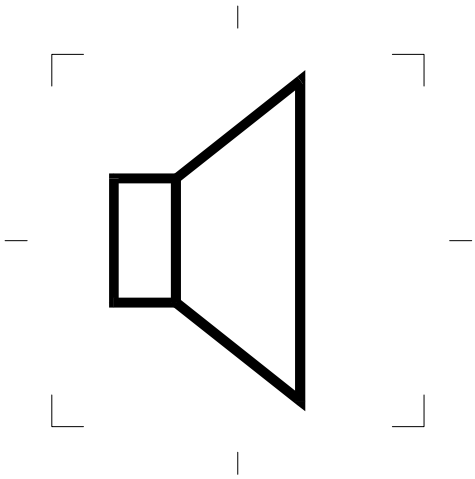
Videotelephone still picture on/off:
In videotelephony, for switching on and off the still picture (screen freeze) function.



Videotelephone document camera on/off:
In videotelephony, for switching on and off the document camera.



Videotelephone hands-free on/off:
In videotelephony, for switching on and off the hands-free mode.
This symbol is identical with the IEC symbol for "Loudspeaker/Microphone" 417-IEC-5081 (IEC 417 [48]).



Videotelephone loudspeaking on/off:
In videotelephony, for switching on and off the loudspeaking mode.
This symbol is identical with the IEC symbol for "Loudspeaker" 417-IEC-5080 (IEC 417 [48]).

History

Document history	
July 1996	First Edition

ISBN 2-7437-0684-8
Dépôt légal : Juillet 1996