ETSI TECHNICAL REPORT

ETR 160

January 1995

Source: ETSI TC-HF

Reference: DTR/HF-01016

ICS: 33.020, 33.040.40

Key words: Animation, audio, audiovisual, data, database retrieval, graphics, hypermedia, image, information types, media, media quality, moving graphics, multimedia service, quality level, service component, service task, synchronisation, video

Human Factors (HF);
Human Factors aspects of multimedia telecommunications

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) on the human factors aspects of multimedia telecommunications has been prepared 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 Telecommunication Standard (ETS) or Interim European Telecommunication Standard (I-ETS) status. An ETR may be used to publish material which is either of an informative nature, relating to the use or the application of ETSs or I-ETSs, or which is immature and not yet suitable for formal adoption as an ETS or an I-ETS.

Introduction

Multimedia services will provide much needed flexibility to users so that they can more easily carry out a wide range of tasks. Using a multimedia service, a new medium can be added to a "call" in progress to meet an unforeseen requirement. This could allow a user to, for example, show and point to a location on a map during an audio call. Eventually, it may become easier for service providers to provide *all* telecommunications services using what are currently thought of as "multimedia" services. However, combining different media at the user interface brings extra problems which could make these services more difficult to use. It is the aim of this ETR to provide human factors guidance to ensure that multimedia services are easy to access and use by all customers.

The intended users of this ETR (in order of priority) include those shown in table 1.

Table 1: Intended users of this ETR and potential benefits to these users

	User	ETR used for	Potential benefits
1.	Service and terminal designers	User requirements specification.	Improved usability of services and terminals leading to increased uptake.
2.	ETSI STC TE4	Identification of user (usability) requirements for terminals.	Improved terminal specifications.
	ETSI STC NA1	Identification of requirements for multimedia services.	More effective descriptions of (more usable) services.
3.	ETSI STC TE10	Input to the Multimedia Reference Model.	More effective reference model by inclusion of all functions and more complete specification.
4.	Service providers	To assist the identification of end users, tasks and services.	Provision of services more appropriate to users' needs.

Blank page

1 Scope

This ETR provides definitions of media, multimedia and hypermedia as applied to telecommunications services. The main goal of this ETR is to specify human factors recommendations for telecommunications services that support multimedia applications.

Multimedia first gained a foothold in the marketplace in local stand-alone systems. The standards that they use have also been taken into account due to their impact on telecommunications applications.

This ETR examines the options for offering multimedia services using Integrated Services Digital Network (ISDN) and Broadband-ISDN (B-ISDN), and the effects these options will have on the usability of the service offered.

2 References

This ETR incorporates by dated and undated reference, provisions from other publications. These references are cited at the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this ETR only when incorporated in it by amendment or revision. For undated references the latest edition of the publication referred to applies.

[1] ISO 8613-1 (1989): "Information processing - Text and office systems - Office Document Architecture - Part 1: Introduction and general principles".

[2] CCITT Recommendation G.711 (1988): “Pulse code modulation (PCM) of voice frequencies”.

[3] CCIR Recommendation 601: "Encoding parameters of digital television for studios".

[4] ISO/IEC CD 13522-1: “Coded Representation of Multimedia and Hypermedia Information Objects (MHEG)".

[5] ETR 116 (1994): “Human Factors (HF); Human factors guidelines for ISDN terminal equipment design”.

[6] ITU-T Recommendation I.113: “Vocabulary of terms for broadband aspects of ISDN”.

[7] RACE ISSUE project (R1065): “Human Factors Guidelines for Multimedia”.

[8] ISO 9241, Parts 1 to 3 (1992): “Ergonomic requirements for office work with visual display terminals (VDTs)”.

[9] CCITT Recommendation G.722: “7 kHz audio-coding within 64 kbit/s”.

[10] ISO/IEC 9282, Parts 1 and 2: "Information processing - Coded representation of pictures".

[11] ITU-T Recommendation T.101 (1993): "International interworking for videotext services".

[12] ITU-T Recommendation T.4 (1993): "Standardization of group 3 facsimile apparatus for document transmission".

[13] ITU-T Recommendation T.6 (1993): "Facsimile coding schemes and coding control functions for group 4 facsimile apparatus".

[14] ISO/IEC 11544 (1993): "Information technology - Coded representation of picture and audio information - Progressive bi-level image compression".

[15] ISO/IEC 10918-1 (1994): "Information technology - Digital compression and coding of continuous-tone still images - Requirements and guidelines".

[16] ITU-T Recommendation H.261 (1993): "Video codec for audiovisual services at p x 64 kbit/s".

[17] Marsh, A. H.: “Minimum legibility requirements for a high resolution graphics transmission system”, 6th International Symposium. on HF in Telecommunications, 1972.

[18] Blohm W. and Mühlbach L.: “Videotelephony with high resolution transmission of documents”, 13th International Symposium on HF in Telecommunications, Torino, 1990.

[19] Frowein, H. W. et al.: “Video Coding Delay and Audio-visual Synchronisation in ISDN Videotelephony”, Proc. HFT’93, Darmstadt, May 1993.

[20] McQuillan, J.: “An introduction to multimedia networking”, Business Communications Review, November 1991.

[21] Richards, I. A.: “The philosophy of rhetoric”, Oxford University Press, 1936.

[22] Condon, C. & Keuneke, S.: "Metaphors and Layers of Signification: The Consequences of Advanced User Service Interfaces", Proc. 2nd International Conference on Intelligence in Broadband Services and Networks, Aachen, September 1994.

[23] Kurita, T., Lai, S and Kitawaki, N.: “Assessing the effects of transmission delay - interaction of speech and video”, Proc. HFT’93, Darmstadt, Germany, May 1993.

[24] ETS 300 389: “Integrated Services Digital Network (ISDN); Circuit mode multiple-rate unrestricted 8 kHz structured bearer service category; Service description".

[25] ITU-T Recommendation H.221 (1993): “Frame structure for a 64 to 1920 kbit/s channel in audiovisual teleservices”.

[26] ETS 300 144: “Integrated Services Digital Network (ISDN); Audiovisual services; Frame structure for a 64 kbit/s to 1 920 kbit/s channel and associated syntax for inband signalling”.

[27] ITU-T Recommendation H.242 (1993): “System for establishing communication between audiovisual terminals using digital channels up to 2 Mbit/s”.

[28] ETS 300 143: “Integrated Services Digital Network (ISDN); Audiovisual services; Inband signalling procedures for audiovisual terminals using digital channels up to 2 048 kbit/s”.

[29] ETR 170: “Human Factors (HF); Generic user control procedures for telecommunication terminals and services”.

3 Definitions and abbreviations

3.1 Definitions

Multimedia can be defined as the ability to handle combinations of media having fundamentally different properties. Examples of these are text and video, or graphics, audio and video[1)]. A more rigorous definition is required so that multimedia services can be specified. In this subclause, the individual media types are defined, followed by the multimedia objects which result from combinations of these media types.

Hypermedia is the combination of monomedium or multimedia objects which have explicit links between them, such that it is possible to access one object using a link from another object.

This subclause contains the definitions of media types, multimedia and multimedia services adopted by ISO/IEC, ITU-T and by other ETSI STCs.

3.1.1 ISO/IEC definitions of media

The following definitions were agreed by ISO/IEC JTC1/SC29/WG12 MHEG (Multimedia and Hypermedia Experts Group):

medium (plural media): A means by which information is perceived, expressed, stored or transmitted.

NOTE 1: This is a weak definition due to the fact that this term is used with too many meanings depending on the context. Thus, the term is to be avoided in its stand-alone form. To be unambiguous, it should only be used in expressions such as: perception medium, representation medium, presentation medium, storage medium, transmission medium.

perception medium: The nature of the information as perceived by the user.

NOTE 2: Examples of auditory perception: speech, noise, music. Examples of visual perception: text, drawings, moving scene.

presentation medium: The type of physical means which is used to reproduce information to the user (output device) or to acquire information from the user (input device).

NOTE 3: Output devices: screen, paper printer, loudspeaker... Input devices: keyboard, mouse, button, microphone, camera...

NOTE 4: Presentation is not to be confused with the Open Systems Interconnection (OSI) Presentation layer.

NOTE 5: The present definition of output presentation medium is consistent with the one given in ISO 8613-1 [1].

representation medium: The type of interchanged data, which defines the nature of the information as described by its coded form.

NOTE 6: Table 2 provides examples of representation media.

---

1) In commercial products, the term multimedia is often reserved for those involving moving video or still images with text and/or graphics.

Table 2

Nature of Information	Possible coded forms
Characters or text	telex, ASCII, EBCDIC...
Graphics	CEPT, NAPLPS or CAPTAIN Videotex, CGM...
Audio	ITU-T Rec. G.711 [2], MIDI, MPEG Audio, ...
Still picture	Fax Group 3, JBIG, JPEG, ...
Audiovisual sequence	CCIR Rec. 601 [3] plus associated audio, MPEG, ...

NOTE 7: The representation medium is defined independently of the direction of interchange (i.e. to or from the user, or between equipment). Each representation medium can be used for both input or output. For example, character-type representation may be used both for text display and text input from a keyboard; graphics-type representation may be used for both graphic display and graphic input (location) from a mouse. Audio-type or picture-type representations may be used both for reproduction and capture.

interchange medium: The type of means to interchange data; it can be either a storage, a transmission medium, or a combination.

storage medium: The type of physical means to store data. For example, electronic memory, floppy disk, hard disk, optical disk, magnetic tape.

transmission medium: The type of physical means to transmit data. For example, twisted pairs, coaxial cable, optical fibres, radio link.

3.1.2 Multimedia

information type: One is defined for the transmission of each representation medium.

multimedia representation: The property of handling several types of representation media (see ISO/IEC CD 13522-1 [4].

NOTE 1: The term multimedia is an adjective, which when used is attached to a noun which provides the context: multimedia service or application, multimedia terminal, multimedia network, multimedia presentation,...

NOTE 2: In ISO/IEC CD 13522-1 [4] multimedia is used in the sense of multiple representation media. “Multimedia” starts with two types of representation media in a piece of information and thus a remote processing service or Videotex when using only a keyboard and a purely textual display cannot be said to be multimedia.

multimedia terminal: A terminal able to transmit and receive combinations of media simultaneously and communicate separately in each medium (see ETR 116 [5]).

multimedia service: A service in which the interchanged information consists of more than one type, such as text, graphics, sound, image and video (see ITU-T Recommendation I.113 [6]).

multimedia communication: Communication involving at least two different types of information.

multiple media communication: Communication involving at least two different information types *or* instances of the same type of information.

hypermedia: The ability to access monomedium and multimedia information by the use of explicit links (see ITU-T Recommendation I.113 [6]).

3.1.3 Multimedia services

This subclause contains definitions related to multimedia services. These can be described using a number of basic elements:

service component: A service component provides the communication capabilities for a single information type.

A "generic" service component can be provided for each information type as shown in table 3. Information types are characterised by their basic elements. Basic elements are the smallest parts of information that can be processed separately. Table 3 is derived from RACE Common Functional Specification C210 as part of the RACE Issue project RI065 [7].

Table 3: Service components related to information types

Information type	Basic element	Service component
character text	character	text
audio (including speech)	tone	audio
geometric graphics	geometric symbol	graphics
raster image	pixel	picture
moving images	picture	video
binary data	binary digit (bit)	data

To be used as part of a service description, a service component should also have a quality level associated with it. There can be many "specific" service components of each type. For example, specific video service components may be defined with quality levels equivalent to Standard Resolution TV quality (SRTV), Extended Quality TV (EQTV) and High Definition TV quality (HDTV).

service task: A group of service components that are handled together as a unit. Service tasks can be used to represent complex information types where synchronisation is required between media, as described in subclause 7.2.3.

service control element: A set of primitives that allow a user to control a telecommunications service. Examples of service control elements which apply to all services are:

SETUP/RELEASE	(an instance of a Service, i.e. a call);
ACTIVATE/DEACTIVATE	(a service task);
ALLOCATE/DEALLOCATE	(a service component);
MODIFY	(service component attribute value(s));
REPORT	(on failure situations, by the service provider to the user);
INVOKE	(a supplementary service).

Other service control elements could be identified for specific services.

3.2 Abbreviations

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

AGC	Automatic Gain Control
AVT	Audio-Visual Terminal
CD	Compact Disc
DAB	Digital Audio Broadcasting
EQTV	Extended Quality TV
GIF	Graphic Interchange Format
HDTV	High Definition TV quality
ISDN	Integrated Services Digital Network
ITU-T	International Telecommunications Union - Telecommunications Standards Sector
JBIG	Joint Bilevel Image Group
JPEG	Joint Photographic Experts Group
MHEG	Multimedia and Hypermedia information coding Experts Group

MPEG	Motion Pictures Experts Group
NNI	Network-Node Interface
PSTN	Public Switched Telephone Network
RDS	Radio Data System
SRTV	Standard Resolution TV quality
TIFF	Tag Image File Format
TV	Television
UNI	User-Network Interface

4 A Human Factors approach to multimedia

Most human interaction involves more than one medium and "multimedia" is therefore not a completely new problem area. Most multimedia questions can be approached by using traditional human factors methods, for example, by considering task requirements in relation to human capabilities and limitations.

Nevertheless, new issues are raised by technologies which allow users to choose between media and to combine media in different ways. This ETR provides recommendations to help designers of such services to take account of the special human factors issues involved.

4.1 Applications

An application consists of a set of tasks, some of which are communication tasks. Typical applications benefiting from the use of multimedia services include:

- teleshopping
  Current teleshopping applications using Videotex databases are not ideal because users can not identify items easily from their text descriptions alone. An image showing the object, as well as text for confirmatory identification, makes recognition and selection much easier.
- telemedicine
  Patient records are naturally stored by medical staff as multimedia items (text descriptions, X-rays, electrocardiograms, etc.). These records therefore need to be stored, transmitted, viewed and edited as multimedia objects when transferred to a multimedia application. Telemedicine can also include conversation between experts for collaborative diagnosis.
- tele-education
  This is an example of an application requiring a point-to-multipoint service. Teachers can distribute multimedia information (such as videos, spoken text passages, music, graphics, images and virtual blackboards) to students in a number of remote locations.
- multimedia mail
  This is an example of an application involving both sending and retrieval tasks. Mail (including any form of media) can be sent by one user, and retrieved by another at a later date.
- collaborative design
  Interaction between designers using voice and video for conversation and shared graphical and image based tools for joint design work.

Each of the above applications is an instance of one of the generic applications listed in annex A.

4.2 Tasks

Each of the applications listed in annex A involves a number of communication tasks. There are three basic communications tasks as shown in table 4.

Table 4: Communications tasks

Conversing	To interact with another communicating entity in real time (using any medium or combinations of media).
Sending	To send an object or objects (of any media) to another communicating entity.
Retrieving	To retrieve an object (of any media) from another communicating entity.

Other tasks are listed in subclause 5.2.

4.3 Media

The definitions of perception medium, presentation medium and representation medium are similar to those used by ISO/IEC (see subclause 3.1.1). The other media types are also felt to be important from a human factors perspective. These types of media are related as shown in figure 1. An example of the relationship between the media types is shown by the arrows. Here, audio is the representation medium. This is presented to users through a loudspeaker. The sense of hearing is used to receive it, it is perceived as a tone and recognised as a warning.

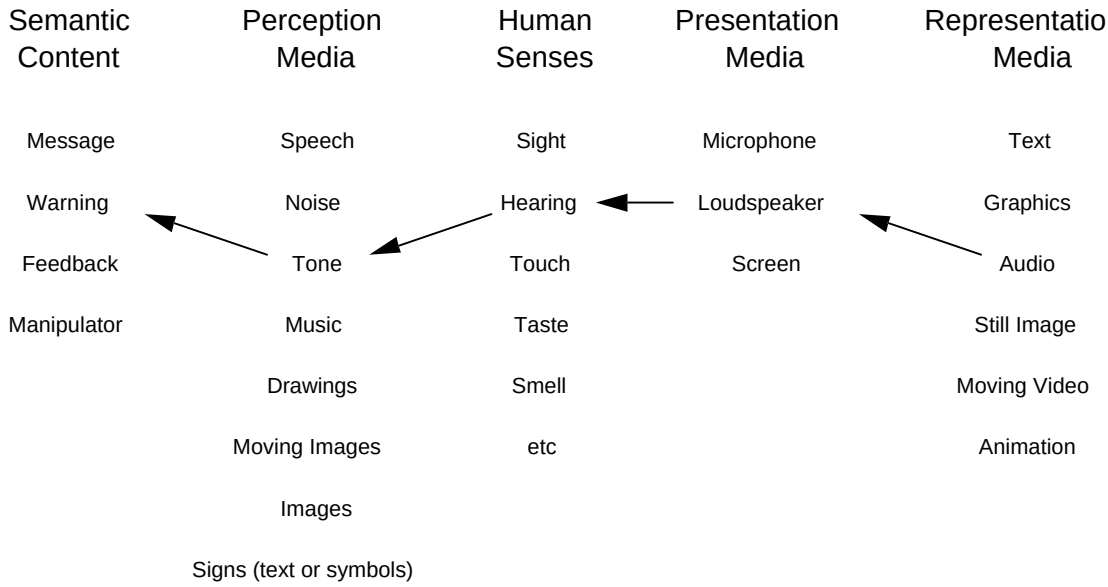

Figure 1: The relationship between types of media from a human factors perspective

Only representation media are considered further in this ETR.

From the human factors perspective it is useful to consider media in terms of their implications for cognitive processing and mental workload. The human factors recommendations take into account the capabilities and limitations of users when processing information from a combination of media. For example, a particular combination of media may lead to non-optimal performance because the information cannot be processed in parallel.

The cognitive implications provide a basis for distinguishing media. For example, animation and still graphics can be distinguished from a human factors perspective because still graphics remain available on the screen, whereas animations change with time. Users can therefore remove their attention from a still graphic for an extended period of time without losing any information. Users can also take as long as necessary to process the information contained in the graphic. In contrast, animations change with time which means that users need to focus their attention continuously on the animation for the duration of the sequence or information may be lost. Thus, the processing of information from still graphics and animations involves different cognitive processes which need to be taken into account when services are designed. The distinction between static and dynamic media as a first level differentiator is illustrated in figure 2.

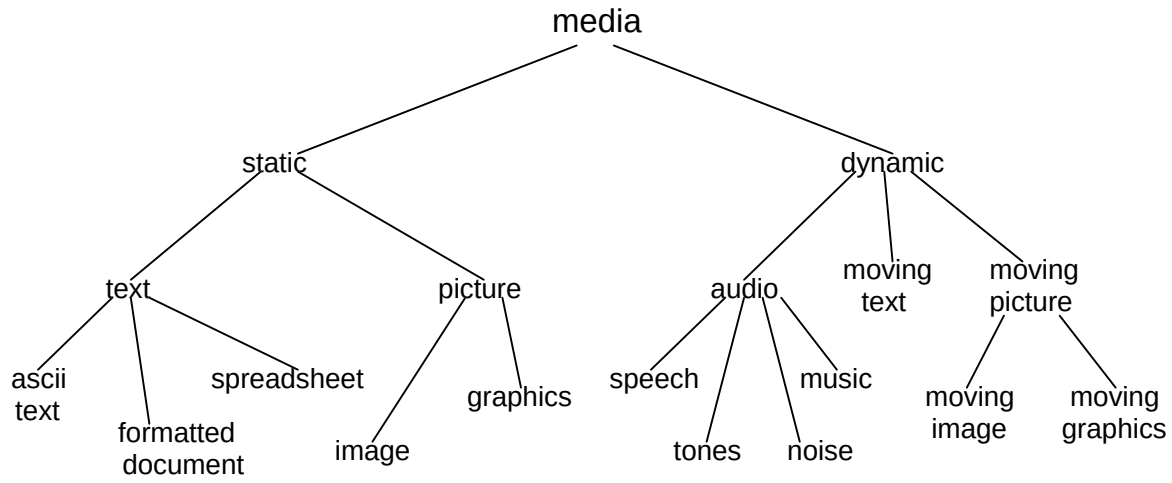

Figure 2: Example classification of media

The fact that these types of media are different has consequences for the amount and type of user control which should be provided. For example, users should, in most cases, be able to start and stop an animation sequence as well as replay it from the beginning. Different controls are necessary for still graphics. Thus, still images and animation require different human factors advice.

Using a human factors approach, the following representation media can be identified as being distinct media within the limits of present day technology:

- text (including characters);
- audio (speech and sound);
- still graphics;
- moving graphics (animation);
- still image;
- moving image.

Data is required as a medium for machine-machine communication, but is not a perception medium, as it can not be perceived directly by a human user. However, machines have their own requirements for representation media in terms of quality levels, response times, etc., and so the ability to communicate data should be provided (see subclause 3.1.3).

4.4 Multimedia services

From a human factors perspective, a multimedia service is one which integrates two or more representation media accessed using the same terminal equipment. Therefore, *from a human factors perspective*:

- multimedia database retrieval clearly is a multimedia service;
- telephony is not a multimedia service;
- broadcast radio is not a multimedia service (unless it incorporates a Radio Data System (RDS));
- broadcast TV is a multimedia service because it incorporates audio and video which *from the user's point of view* can be manipulated separately. The degree of manipulation can vary, depending on the sophistication of the terminal and the service, from control of the volume, brightness and contrast, to the selection of the language spoken and the display of subtitles and other information (e.g. weather forecasts) using teletext facilities;
- videotelephony is a multimedia service because it incorporates audio and video which can be manipulated separately.

From a human factors perspective, there is no essential difference between a weather forecast (incorporating text and image) retrieved from a multimedia database using a multimedia retrieval service, and a weather forecast (incorporating the same media) retrieved from teletext data associated with a broadcast TV service. The differences between these two situations from a human factors perspective are peripheral, and include issues such as the possibly increased response times necessary to retrieve the information, and the lack of flexibility to receive the information in a different form (a multimedia database retrieval service is likely to offer a greater range of information in a wider range of media types).

Broadcast TV and videotelephony services have already been defined and their human factors aspects covered in other recommendations. Therefore this ETR will not include guidelines for these services.

5 Media-specific issues

This clause covers issues specific to individual media. Combinations of media are covered in clause 6. This clause does not cover issues related to services that have already been defined (e.g. videotelephony, Audio-Visual Terminal (AVT)), as the human factors issues have already been specified in appropriate recommendations.

5.1 Media types and quality levels required for different tasks

The most appropriate medium to use depends on the task being carrying out, where 'appropriate' refers to the effectiveness, efficiency and satisfaction of using each medium to carry out that task. Some media place a higher cognitive load on users to carry out the same task.

When choosing media types, it should be remembered that some users may have difficulty using some media or may not be able to use some media at all. However, this subclause does not address the requirements of disabled users.

Media types can be divided into static media and dynamic media. Users should not have to wait while a stored dynamic medium (speech, animation or moving image) runs its course. It should be possible to terminate these sequences and move to the next step as if they had completed normally.

The quality levels required for different media for different tasks are also covered in this subclause. The quality levels used for different perception media should be comparable. For example, high quality video and audio should be used together rather than high quality video with low quality audio or vice versa. This should be the default, but users should be able to override this (e.g. for reasons of cost or user abilities).

5.1.1 Text

Text is useful for:

- conveying large amounts of detailed information;
- providing guidance to users;
- writing reports.

Where possible, text should not be used for providing large amounts of numeric information.

Quality of text is measured in terms of screen resolution, font, and character size and spacing. To ensure good readability of text the guidelines given in RACE ISSUE project R1065 [7] should be used. To display scanned text as an image, there should be no more than 30 lines of text per 1 000 scanning lines. There should be a minimum of 12 scanning lines per line of text, to allow for the inclusion of diacritical marks, etc. SRTV quality is insufficient to display an A4 page. SRTV quality allows a maximum of 24 lines of 40 to 50 characters each.

The colour of text relative to its background also has an important influence on legibility. For instance, red text on a blue background should never be used. Recommendations on the optimum colour combinations to ensure good legibility are given in ISO 9241 [8] and RACE ISSUE project R1065 [7]. See subclause 5.1.6 for more information on transmission of images in legible quality.

5.1.2 Audio

Audio in the form of speech is the natural medium for conversational services.

Audio is useful for:

- adding atmosphere to information being presented on the screen;
- attracting the user's attention to information being presented on the screen;
- elaborating on visual information.

Audio is not appropriate for:

- conveying large amounts of information for later recall;
- providing detailed descriptive information.

Audio is usually the most difficult medium for selection between objects.

It is difficult to remember extended audio messages. For this reason, audio messaging makes the task of sending a message simpler for the sender, but places a higher cognitive load on the receiving user. It is recommended that audio messages should be:

- short (for the convenience of the receiver);
- editable (for the convenience of the sender);
- reviewable (for the convenience of both).

Some kind of volume control (Automatic Gain Control (AGC)) should be provided for audio. It should be possible to mute the microphone.

Audio quality is measured in terms of the parameters defined in CCITT Recommendations G.711 [2] and G.722 [9], such as harmonic distortion. 3,1 kHz analogue quality is acceptable for telephony applications where a handset is used. 7 kHz analogue quality is desirable for conversational tasks using hands free mode (because of higher noise levels), and for audioconferencing to provide enhancement of speaker recognition. 15 kHz analogue quality is necessary to provide FM quality music reception (Digital Audio Broadcasting (DAB) and above), and for applications such as conference interpretation where speech needs to be understood verbatim. 16 kHz is the highest frequency discernible by the human ear and, although not defined as an audio quality, is therefore sufficient for high quality music reception. Compact Disc (CD) quality (44 kHz digital sampling) has been defined to provide high quality music reproduction.

Monophonic audio is adequate for most applications. Stereophonic (2 channel) audio is desirable for music and for multipoint applications to help identify the speaker. Emphasising the voice of a particular participant should be possible for some applications.

5.1.3 Still graphics

Still graphics are useful for:

- expressing ideas;
- showing statistics and trends in numeric information;
- creating and manipulating objects that do not actually exist;
- highlighting desired features of objects (e.g. shape);
- showing relations in space and time.

Graphics can be used in applications where it is necessary to show accurately detailed representations of real or imaginary objects or processes. When transforming detailed information into a simple graphical representation, care should be taken that important information is not lost in the simplification process. Some kind of zooming control should be provided to allow more detail to be seen in complex graphics.

Graphical quality is measured in terms of spatial resolution and colour rendition, which is a function of the colour table (number of colours used) and either monochrome grey scale or colour depth (number of colours available). Spatial resolution depends on screen resolution and image resolution. If screen resolution is too poor lines will appear "jagged". To transmit the same information as a graphic requires a lower bit rate than an image. Graphics take longer to prepare prior to sending.

The following graphical quality standards are defined:

ISO 9282 [10] - for computer graphics.
ITU-T Recommendation T.101 [11] - for Videotex.

Colour should be used with restraint. 5 colours are often sufficient where colours are used for coding purposes (where a separate meaning needs to be associated with each colour). A small colour table is therefore sufficient. However, colour depth is important to ensure accurate representation of colours.

5.1.4 Moving graphics (animation)

Animation is useful for:

- showing trends;
- creating imaginary figures;
- expressing imaginary events;
- exploring viewpoints to determine the relationship between an object and its environments.

It should be possible for users to skip or terminate animation sequences. The ability to manipulate and fast forward graphics is also desirable for some applications.

Moving graphics quality is measured in terms of spatial resolution and colour rendition (as for still graphics), with the addition of the number of frames displayed per second.

5.1.5 Still image

Still images are useful for:

- portraying detailed information well;
- representing pictorial information, e.g. photographs exactly;
- portraying the real world;
- interactive use;
- user guidance (in support of text).

Still images are a static medium and so allow unlimited viewing time.

It is usually easier to recognise a physical object from its image rather than from a textual description. An exception is where users have to choose between a number of similar objects, perhaps without size cues - for example, in selecting between screw sizes. In this case, a textual identifier makes it easier to select between objects. Hence identification of dissimilar objects is best done visually, but identification of similar objects is best done using text.

Image quality is measured in terms of spatial resolution and colour rendition (as for graphics). However, the number of colours that can be displayed simultaneously (colour table) is more important for images than for graphics. An increase in the size of an image may result in a reduction of image quality. A decrease in the size of an image may result in detailed information being lost. Remote document viewing using still images requires a quality higher than that of SRTV quality.

The following image qualities are defined:

ITU-T Recommendation T.4 [12] - group 3 facsimile;

ITU-T Recommendation T.6 [13] - group 4 facsimile;

ISO/IEC 11544 [14] (JBIG) - primarily for monochrome images - designed to replace facsimile group 3/4;

ISO/IEC 10918-1 [15] (JPEG) - lossy compression technique;

Graphic Interchange Format (GIF);

Tag Image File Format (TIFF).

The time to create an image on a screen (for a fixed bit rate) depends on the image size and resolution. An image requires a higher bit rate or longer response time than a graphic to display the same information. Any change in the size of an image may result in a reduction of image quality or detailed information may be lost. Zooming to a more detailed part of a still image should not be so rapid as to cause users to lose context information.

Accuracy may be as important as perceived quality for some applications. Medical applications (e.g. X-rays) are an example where the accuracy of representation is important (for legal reasons) as well as a high perceived quality.

It may be easier to prepare information in the form of an image than a graphic, as video or scanning techniques become available, but still images remain more difficult to manipulate than graphics.

5.1.6 Moving image

Moving images are useful for:

- portraying descriptive information;
- portraying processes;
- portraying cultural and social information;
- portraying action.

Moving images are not appropriate for providing users with large amounts of detailed information.

Moving images can be a help or a hindrance depending on the task being carried out. If the observance of motion patterns is central to the task being carried out, then the inclusion of moving images in the interface is clearly required. If moving images as part of a multimedia presentation are not central to the task being carried out then the image should be frozen. It is sometimes difficult to justify the use of moving image to increase the effectiveness with which users can carry out a task. However, the addition of moving images may increase user satisfaction with the interface and therefore user motivation and performance.

It should be possible for users to skip or terminate moving image sequences.

Moving image quality is measured in terms of spatial resolution and grey scale or colour rendition (as for still images), with the addition of the number of frames displayed per second. Therefore, moving image generally has the highest bit-rate requirements of the six representation media described. The following video qualities are defined:

ITU-T Recommendation H.261 [16] - for p*64 kbit/s;
MPEG1 - up to 1,5 Mbit/s;
MPEG2 - 1,5 Mbit/s to 15 Mbit/s;
MPEG4 - for low bit rate video coding < 64 kbit/s.

The moving image quality required for head and shoulders videotelephony depends on the task being carried out. Standard TV quality is acceptable for a wide range of applications. However, many collaborative activities involve showing or modifying some kind of static image or document. For these, it has been shown by A. H. Marsh [17] that standard TV quality permits the transmission in legible quality only up to A5 format documents. High resolution images, such as of HDTV quality, are needed for larger documents. Higher resolution also improves user satisfaction (see Blohm W. and Mühlbach L. [18]), especially when both an overall view of the whole document and the recognition of small details are required at the same time. In other applications, such as medical diagnosis, a very high image quality is likely to be necessary.

For some tasks the frame update rate can be reduced (e.g. remote supervision). For other tasks, frame rate is critical to the performance of the task (e.g. for signing by deaf users Frowein, H. W. et al [19]). Tasks need to be considered carefully before the frame rate is reduced. The option for users to be able to change the frame rate or resolution should be considered.

For restricted bit rates with highly compressed coding techniques, it is inadvisable to switch between video sources frequently, as the picture may take a long time to build up after each change.

Accuracy is as important as perceived quality for some applications (as for subclause 4.1.5).

5.2 Manipulation of media

The *functions* used to manipulate media are task dependent not media dependent. For example, users will want to browse, select, view, edit, delete and link objects independent of their media types. However, the manner in which these functions are carried out will depend on their media types. Viewing and editing will be fundamentally different actions in different media.

Tasks carried out using multimedia applications include those listed below. Some of these tasks can only be carried out on media once they have been stored.

- Browsing.
  To find out (to create a mental model of) the information that is available (in a multimedia database).

- Searching.
  To look for a particular object or information on a particular subject (in a multimedia database). This could be considered to be a goal-directed form of the browsing task.

- Pointing.
  To indicate the location of an object. This can be part of the selection task, or used as part of a conversational service to bring the attention of other users to an object.

- Selecting.
  To select an object (usually so that something else can be done to it).

- Creating.
  To create an object (of any medium).

- Viewing/reading/listening.
  Tasks that apply to specific media. They are forms of the generic 'taking in information' task which can apply to any medium.

- Editing.
  To modify an (already created) object (of any medium). Editing operations can include transformations such as mixing, frequency shifting and chroma keying.

- Deleting.
  To delete an object (of any medium).

- Linking.
  To link objects (of any medium or combination of media) in a hypermedia database.

- Unlinking.
  To remove a link between objects in a hypermedia database.

Different applications use different subsets of these tasks. Not all of these tasks require communications functions and so may not need a multimedia service. Users often want to carry out a number of tasks at the same time. This can put a higher cognitive load on users which is dependent on the media used to carry out the tasks (see subclause 5.1), and the way in which the media are combined (see clause 6).

6 Combinations of media

One of the advantages of the human factors approach to the definition of media types given in clause 4 is that it provides a useful framework for examining the effect of integrating media. Decisions about which media can be combined can be based on an understanding of the their implications for cognitive processing. Using more than one medium simultaneously leads to an increase in the understandability of the information, as people learn more readily when exposed to information in different modes (see McQuillan J. [20]). However, the effectiveness of multimedia depends more on the chosen combination of media than on the provision of a rich set of media. Overload can easily occur if media are not chosen carefully.

6.1 Using different media for the same task

A combination of media used to provide information to users for the same task can be beneficial and help users to carry out a task. For example, the selection task described in subclause 5.1.5 can best be carried out using a combination of image and text. Dissimilar objects can then be identified easily from their still images, while similar objects can be distinguished using the textual information. Parallel processing by users can also be used to identify the objects in the easiest form. In addition, one medium can reinforce the other, for example, a video with accompanying commentary to reinforce the message portrayed by the

video and give additional information. Sound and graphics can be effectively used together if the information needs to be explicitly noticed (e.g. for warnings).

Combinations of media may also be beneficial when an interface is to be used by a wide range of users. Some users have a greater ability to process some types of media than others, and so a combination of media may mean that an interface is usable by a wider range of users. Where possible, users should be able to select the media used.

6.2 Using different media for different tasks

If the service uses different media to present information to users for different tasks simultaneously, then this is more difficult for users. For example, a conversation and an editing task can be carried out simultaneously (in different areas of the screen). Another example is listening (e.g. for a baby alarm) while carrying out another task (such as editing a document). In these cases it is easier if the media used to provide information to users for the two tasks are different, e.g. it is easier to watch a monitor while talking to someone on the telephone than it is to listen to two different telephone conversations or to carry on a telephone conversation while watching/listening to a TV programme.

6.3 Using the same media for different tasks

If the service uses the same media to present information to users for two (or more) different tasks simultaneously, then this represents the most difficult situation for the user. It would not be advisable to expect a user to process two animation or video sequences at the same time, whereas two graphical images might appear on the screen together, provided other human factors principles are not violated. Two continuous, time dependent media should only be used simultaneously if the information being presented in one medium is complementary to the other. Otherwise cognitive overload may occur.

6.4 Principles for media integration

Integration principles can apply to any combination of media, or be specific to certain combinations of media, or can only be applied in specific applications with specific user groups. Subclause 6.4.1 lists generic principles which apply to any combinations of media used by any users or applications. Subclause 6.4.2 gives specific rules which apply only to certain combinations of media or to specific types of user or application.

6.4.1 Principles which apply to any combination of media

The principles described in subclauses 5.1 to 5.3 can be summarised as follows:

a) it is easier for users to carry out a task if more than one medium is provided. This is an example of the redundancy principle (below);

b) it is easier for users to carry out different tasks if different media are provided for each task.

Other generic principles which can be applied to any combination of media are:

1) Compatibility
   Users should be able to use knowledge they have gained from outside the service.

2) Coherence (consistency)
   Users should be able to generalise their experience from parts of the service to other parts of the service. The same action should always produce the same result.

3) Simplicity
   The number and complexity of necessary actions should be reduced to a minimum. It should take the minimum number of actions to achieve a given task.

4) Redundancy
   Users should be provided with information in several ways. Users should receive information in the most appropriate form for its use.

5) Salience
   Critical information should be presented to users in a sufficiently intrusive way.

6) Transparency
Information about tasks which can be performed, as well as about states of the service, should be stated unambiguously and clearly. Users should know what's going to happen when an action is carried out.

7) Completeness
Information about tasks which can be performed, as well as about the states of the service, should be complete in such a way that users have knowledge of, or are presented with, all available options at the time they are needed.

8) Support orientation
If the information to be presented is too complex or covers more than is possible to present at one time, users should be helped to find the relevant information by giving them support in orientation.

9) Feedback
Users should be informed about the consequences of their actions.

10) Reversibility
Users should be able to restore pre-existing states of the service.

11) Controllability
Users should be able to control actions of the service.

12) Flexibility
Users should be able to choose the modalities of their task performance with respect to the input media available and the experience they have.

6.4.2 Selecting appropriate media for users, tasks and applications

Having considered the properties of each medium and the integration of media in general terms, it may be necessary to select the most appropriate media to use for a specified set of users, tasks and applications. Examples include:

a) for a collaborative task where it is necessary to reach a decision quickly, speech should be used because it is fast but, if there is a need to go back to the information, text should be used;

b) in general for conversational tasks, sound is the dominant channel for communication, but specific users have specific requirements, e.g. for lip readers, the visual channel is dominant although the audio channel can reinforce the information.

Sound and vision are complementary modes of information. One has to look in the appropriate direction to see a visual object, whereas one does not need to face a source of sound to listen to it. In addition, sounds are relatively transient in nature and thus are well suited to conveying changes in the state of a service and for presenting discrete messages.

In general, the sound quality is more important than the image quality for an audiovisual conversational task.

The legibility of text when combined with graphics or images should be taken into account. There are well known techniques for ensuring that text is still legible when superimposed on a graphic or image.

7 Multimedia service issues

This subclause introduces the human factors issues relevant to multimedia and hypermedia services and provides some guidelines on these issues. Issues include requirements for addressing and call set up, bit rates and quality levels, delay and response times, and synchronisation of media. This section is divided into general issues, multimedia issues and hypermedia issues.

7.1 General issues

This subclause contains issues which apply to all services, but have implications which require highlighting for multimedia and hypermedia services.

7.1.1 Metaphors

A metaphor is "the transport to one thing of a name which designates another" (from "The philosophy of rhetoric" [21]. An example is the "cut and paste" metaphor found on word processing systems, relating the operation to the scissors and glue normally used to physically cut and paste paper. Therefore, metaphor can be defined as the presentation of one idea in terms of another. The first idea is transformed. For example, the idea of a window in a room can be transformed to that of a window on a screen. The first idea may also be extended. For example, once a selection is "cut" in a computer document it may then be "pasted" into more than one location.

Metaphors can be useful to help explain the concepts of a new domain (such as a word processing system) in terms of the mental model the user has of an existing domain (such as scissors, paper and glue). However, the match between the domains is rarely perfect and the use of a metaphor can also limit the user's ability to reason about the new domain.

7.1.1.1 Examples of metaphors

Examples of metaphors that can be useful for multimedia services include:

- control panels simulating audio and video recorder controls can be used for the manipulation of audio and video media;

- desktop metaphors can be useful to provide a framework for single-user multimedia applications. However, they can cause problems for multiple-user applications (e.g. groupware), as it is difficult for groups of users to work at the same desk at the same time;

- a meeting room metaphor (perhaps extended by corridors and building maps) may be more suitable for multiple-user applications;

- publishing metaphors such as newspapers or cable television can be particularly suitable where communication is one-to-many or many-to-many;

- a hyperspace navigation metaphor is useful for hypermedia services. This implies that a user can move instantly from one place to another in a "communications space";

- agents ("human" or otherwise) can be used to partition the functionality of the services, such as passing "mail" to a postman or post office.

7.1.1.2 Choosing a metaphor

Which of these metaphors might be most useful deserves careful consideration. Three axes have been identified along which metaphors can be classified:

a) the activity;
b) the interaction;
c) the spatial (and temporal) dimensions.

Research by Condon, C. and Keuneke, S. [22] has shown that these axes are critically important in directing the user's attention to different aspects of the services:

- activity-directed metaphors, such as agents, direct the user's attention to the activities or functions of the system. Evidence shows that users pick up a better understanding of the system functionality with this type of metaphor;

- interactional metaphors, such as publishing, emphasise the interactional nature of the underlying tasks, particularly its social aspects;

  NOTE 1: The task does not need to be the same as the metaphor - publishing metaphors have been used very successfully to support remote co-operation in manufacturing design.

- spatial metaphors tend to direct the user's attention to the user interface, and are therefore often of immediate appeal but of lesser long term value.

NOTE 2: A spatial metaphor does not imply a graphical user interface (c.f. older games software, or the UNIX command "move").

7.1.1.3 Disadvantages of metaphors

There are some disadvantages to using metaphors for multimedia services:

a) the concepts in the existing domain may not apply in the new domain. For example, the provision of a typewriter metaphor to help the explain how to use a word processor may cause difficulty in using operations such as insert and delete, as these do not work in the same way as the equivalent operations on a typewriter;

b) the new domain may contain functionality not present in the existing domain. For example, a word processor contains many other functions not present on a typewriter, such as automatic formatting and the generation of tables of contents. Users may assume that these have to be carried out in the same way as on a typewriter;

c) only parts of the existing domain may apply in the new domain. If the interface contains non-functional aspects this is frustrating to users and it may be necessary to add functionality to the new domain in order to support the implications of the metaphor. For example, a lampshade might make a room look more realistic but should only be shown if a suitable functionality can be provided, such as dimming the screen;

d) it may be necessary to combine 2 or more metaphors to describe a complete system. For example, the cut and paste metaphor works independently of the metaphors used in the underlying applications (typewriter, drawing board, card index, etc.). When the metaphors imply different ways of doing the same thing, then all need to be supported at the user interface;

e) metaphors may be culturally dependent and may not work in different (international) environments. For example, the assumption of the most recently accessed items being located to the right of older items may not apply in environments where writing is from right to left.

7.1.2 Addressing

From a human factors perspective, addressing mechanisms should adopt the following principles:

- simplicity, so that as few operations as possible are required to address an object;
- consistency within and between addresses.

A unique address should be used for a user for all media in a call. It may be preferable for an address to identify a user or a location rather than a terminal, so that the address is always the same independent of the media or service being used. The terminal capabilities used to present the media can then depend on a service identifier or information type field.

Alternatively, if the structure or form of the address is different for different services then it is easier for users to select the service being offered. For example, if all teletext databases used a common range of addresses then it would be easier to recognise these and not select them by mistake.

Using a hypermedia database makes it easy to address objects within a small set (for example the user's most frequently used objects). Mechanisms to point and click will normally be available and a single operation is all that is required within a limited context. However, where a wider address space is required, users many have to point and click many times to get to the local context where the final object to be addressed can be selected. Therefore, other addressing mechanisms may also be required to provide short cuts.

7.1.3 Call set up

Generic user procedures for call set up have been defined for ISDN telecommunications services in ETR 116 [5]. User procedures for videotelephony are currently being drafted as an I-ETS by ETSI TC-HF with a view to publication in Autumn 1995.

Call set up is more complicated for multimedia calls. There are many different media types and many different quality levels possible for each medium. A user setting up a call may not know the capabilities of

the remote user's terminal (common functional mode) or the capabilities available in the intervening network. Therefore, there should be a negotiation phase at the start of each call where the media types and quality levels possible both in the network and at the remote terminal are established, and an optimum set of media types and quality levels established for the call.

A user profile containing information on preferred quality levels can simplify call set up. A user identification procedure is necessary for this.

A hypermedia link may be used so that a single point and click leads to the complete setting up of a call within a limited context.

7.1.4 Call termination

This should be as simple as possible, preferably by a single operation. However, it should not be possible for users inadvertently to lose information as a result of call termination. If the user has entered information during the call that would be lost or discarded at call termination, the user should be alerted and given the opportunity explicitly to save or discard the information, or to complete the operation before call termination.

7.1.5 Response times for communications functions

The response times required for communications functions depend on the users and the tasks they are carrying out. However, for general use, response times for communications functions should not exceed those shown in table 5.

Table 5: Maximum recommended response times for communications functions (see RACE ISSUE project R1065 [7])

Initiation of a simple service (equivalent to a single service task)	2 s
Initiation of a complex service (multiple service task)	5 s
Feedback on an error	4 s
Retrieval of the first page of information from a remote database	10 s
Request for the next page of (similar) information	1 s
Display update rate for real time tracking tasks	0,5 s

From a human factors perspective, response times should be low and not vary between media. However, due to technological limitations, actual response times may be different for different media, and so a maximum should be specified for each. If it is not possible for the service to respond within the maximum response times given above, then feedback should be provided to users as described in subclause 7.2.4.

In order to display (still) images consisting of a large amount of data, an image build-up period may be required during which all data necessary to portray the image is transmitted to the terminal for local presentation. If delays are short then this build up process should be hidden from users, such that only the complete image is presented on the screen when the build up process is complete. If delays are longer then the build-up may be presented to users in one of two ways:

- line by line build up;
- build up in mosaic form.

The most appropriate method is dependent on the user, task and application. Pictorial information may best be built up in mosaic form. Tabular information may best be built up line by line.

Conversational services require shorter response times than retrieval services because turn taking is necessary (see subclause 7.2.3).

7.1.6 Navigation

Navigation is an essential part of any information retrieval task. Navigational support, through such things as indexes, chapters, paragraphs, titles and keywords, has always been important to success in using reference books. It will also be crucial to the usability of multimedia services.

The primary problems encountered by users of a database are:

- users get lost;
- users find it difficult to gain an overview of the material - they do not know what is there;
- users have difficulty finding specific information. Users know what they want but cannot find it, or even when they have found it they cannot find it again.

Navigational support provides answers to questions, such as: “Where am I?”, “How did I get here?”, “Can I get back?”, “Where have I been?”, “Where can I go?” and “How do I get there?”. To answer these, navigational support should provide both the information and the means of access.

Whatever the type and number of media involved in a communication, there are three basic forms of data:

- continuous linear;
- discrete linear;
- non-linear.

A source of information may use combinations of these, as well as special cases (e.g. loops which can be considered as a special case of either linear form). These data forms require different navigational support functions. Continuous linear data (such as audio or video sequences) may require functions based on a tape/video recorder metaphor (e.g. play, stop, fast forward, fast rewind). Discrete linear data (such as books, directories and card indexes) may require functions based on a book or card index metaphor (e.g. next, previous, first, last, ABC). These simple metaphors are not sufficient for data in non-linear forms, and the designer should impose a recognisable structure or methodology to link the data items. For example, hierarchies provide simple link structures which can help users to form a two dimensional model of the information in the database.

Familiar icons should be used for moving between pages or screens of information. A user will learn to recognise standard arrow symbols to move:

- to the next page;
- to the previous page;
- to the first page;
- to the last page;
- up a level;
- to the map or home page.

However, a mental model should be developed to reinforce these concepts, especially where the notion of moving to a lower (more detailed) and higher (more general) level of information is embodied.

Users have additional problems with large information sources containing lengthy linear sequences and/or large numbers of non-linear or linked data items. Users may forget where they came from, may be unable to remember if they have been somewhere before, or may forget how to find a specific item of information. Consequently, users need more support functions. Examples are:

- audit trails and bookmarks to help return to previous items of information;
- anchor points to help users to structure their route through a database;
- “GoTo”s to help users jump to specific items of information;
- “undo” and “redo” functions to allow users to try routes without getting lost.

Explicit maps can also be used to provide aids to navigation. A map provides a symbolic representation of the database with the user’s current location superimposed on it. If the screen size is large enough, the map can remain in place on the screen and be updated as the user moves between pages. However, if the screen size is insufficient to display both the map and the information, a button with a familiar icon should be provided so that users can re-display the map from any point in the database.

7.1.7 User guidance

Guidelines on the provision of user guidance are similar to those for non-multimedia services and include:

- help information should be presented in such a way that users can immediately try out the advice received, so try it out without explicitly leaving "help";

- if it is necessary to leave the help system before the user can continue the task, instructions should be available on how to leave the help system.

7.1.8 Charging

The selection by users of a different medium or quality level within a medium may incur a large increase or decrease in the charging rate, and so users should be provided with sufficient information, both prior to and during use of the service, to enable them to optimise their use of the service.

Users of a multimedia retrieval service may be reluctant to request further items of information without information on the charges that will be incurred. Users may also be reluctant to pay twice for receiving information already received, perhaps as a result of a navigation error.

7.2 Multimedia issues

This subclause contains issues which are central to the provision of multimedia services.

7.2.1 Use of restricted bit rates

The bit rates required by each information type in a multimedia service may not be available from the network when the call is set up. For example, only a limited number of 64 kbit/s channels may be available.

In such situations, it should be possible to optimise the sharing of the total bit rate available between service components based on algorithms taking account both the guidelines given in clause 5 and the quality thresholds defined by users (if the equipment permits this). In worst cases, users could decide to drop some of the service components (video for example) if the quality is insufficient.

7.2.2 Change of media types and quality levels during a call

A useful feature a multimedia service can offer to users is the possibility of modifying dynamically some parameters of the service components initially allocated, or even to select new service components depending on personal preferences, task requirements and cost.

Such a possibility can be used to, for example, improve the video quality level from SRTV to EQTV for a short time to show an A4 document to remote users, since the SRTV quality does not offer sufficient resolution to show A4 documents adequately.

Another possibility for users, if the equipment permits, would be to keep the SRTV quality unchanged and to add temporarily an EQTV component to show documents through another type of device, e.g. a rostrum camera. This is possible if the person-to-person visual contact needs to be maintained and if the mental workload is admissible.

It is often desirable to reduce the frame rate but increase the resolution. Examples of where this may be necessary include signing for deaf users.

It is desirable for the recipient, as well as the sender, to be able to change media types and quality levels during a call. However, this has charging implications.

7.2.3 Synchronisation between media

Synchronisation is often required between objects in a multimedia or hypermedia application. The types of synchronisation defined in ISO/IEC CD 13522-1 [4] are:

Elementary synchronisation: two objects synchronised either both with regard to the same reference time origin (parallel mode) or one with regard to the other (sequential mode). Examples of elementary synchronisation include:

- between audio and video media to provide lip synchronisation during an audiovisual call;
- between a pointer and an image in a shared viewing application;
- between an image and the audio channel to synchronise subtitles with images in a remote learning application;
- between text and speech for voice synthesis.

Chained synchronisation: a set of objects presented one after the other in the form of a chain.

Cyclic synchronisation: one or more objects presented repetitively.

Conditional synchronisation: the presentation of an object linked to the satisfaction of a condition.

It is more difficult to provide adequate synchronisation for conversational services than for retrieval services. Synchronisation between audio and video can be achieved by trading off fixed delay against variable delay. This can achieve acceptable synchronisation at the expense of adding a fixed end-to-end delay to the communications path. For conversational services, this unidirectional end-to-end delay should not be greater than 250 ms. Above 125 ms, the delay becomes noticeable.

For the lip synchronisation of audio and video in a videotelephony application, the differential delay between audio and video should be as small as possible and never greater than 125 ms. If there is a differential delay between the audio and video, then the audio should always be delayed with respect to the video, as stated in a contribution by Kurita T. et al [23]. Loss of lip synchronisation is detectable when the sound is delayed by more than 40 ms relative to the image, but is also detectable when the image is delayed by more than 20 ms relative to the sound. It should be possible to set the audio delay to zero, at the expense of lip synchronisation, for applications such as document display.

A higher delay can be tolerated for a retrieval service than for a conversational service, because turn taking is not necessary. A retrieval service can tolerate an end-to-end delay of 1 second for most operations, and this can be higher provided that users can anticipate this (see subclause 7.1.5).

Multipoint communications require that delays along all paths should be the same and be in the ranges given above. Otherwise it is possible that, depending on the network topology, turn taking sequences will appear overlapped in time from the point of view of a third participant.

7.2.4 Feedback to users

Multimedia services will provide more flexibility for users and so there may be many more options for users to select. It may, therefore, be necessary to provide more feedback to users on the media types available and the quality levels set. There is little point, for example, in a user setting a high quality level for image transmission if the remote terminal is not capable of displaying this image quality. Alternatively, the call may involve interworking between services or between networks (for example, between ISDN and the Public Switched Telephone Network (PSTN)) which may mean that it is not possible for the remote user to adequately receive the quality levels set by the sending user. Feedback is therefore required to the calling user on the quality levels being achieved by the receiving terminal.

It is also more important to provide feedback to users on the charging rate for the call as described in this subclause.

Within a hypermedia database, navigation feedback is especially important to prevent a user getting lost. Status information should be provided in a fixed area of the screen so that users know where to look for it. More detail is provided in the next subclause.

7.3 Hypermedia issues

Hypermedia is a style of providing multimedia based information systems with easy to use, associative links (connections) between composite multimedia documents.

The associative relations between multimedia documents build structures called networks of links. These network structures imposed on a set of multimedia documents allow for easy access to information resources interactively without a need to use formalised query languages required in conventional data base systems. Nevertheless, it can be of advantage to combine the two approaches.

The structures of associative links provide many alternative paths for searching in a set of multimedia documents - it opens new areas in usability of information systems in the telecommunication environments. The levels of the use of hypermedia systems begin at remote access to hypermedia servers and reach to co-operative conference working sessions of viewing and composing hypermedia documents. In order to assure usability of such co-operative conferences a framework for control mechanisms for co-ordination and synchronisation is needed.

The documents retrieved and created in interactive working sessions will be stored in private and public shared data bases.

7.3.1 Links

Links are relations between individual and/or composed multimedia objects. The relations connect (one or more) source monomedium objects or documents with (one or more) destination documents. This leads to one criteria for defining types of links:

- singular monomedium object to singular monomedium object;
- singular monomedium object to composite multimedia object;
- composite multimedia object to singular monomedium object;
- composite multimedia object to composite multimedia object;
- singular monomedium object to a set of monomedium objects;
- singular monomedium object to a set of composite multimedia objects;
- composite multimedia object to a set of monomedium objects;
- composite multimedia object to a set of composite multimedia objects.

Other criteria for classification of links includes their ownership by individuals and groups of users or the strength of association of a link with underlying multimedia information unit.

The important human factors issue of hypermedia systems is the recognition of anchors - source points of a link relation. The perceptible manifestations of anchors are called markers.

It is necessary, that the markers placed in multimedia documents are easily identifiable not only by some conventions in style of presentation (coloured highlights) but also by means of direct manipulative techniques such as immediate interactive feedback. Possible solutions are:

- change of colour of the marker while "touching" the marker;
- change of the shape of the pointing device (mouse arrows shape);
- textual or audio explanation of the "touched" object (e.g. balloon help);
- automatic pop ups of miniatures of the documents.

A combination of the above techniques should be configurable and adaptable to user preferences.

The strongest potential of the hypermedia systems lies in the opportunity to impose many different structures of links on the same set of basic multimedia material.

7.3.2 Navigation

An essential feature of hypermedia navigation is non-linear organisation of underlying multimedia material. The other forms of organisation like linear, or discrete linear forms of data organisation are of less importance in the context of hypermedia. Hierarchical organisation of documents offers advantages for keeping some order in navigation, but is not necessarily required, especially in post edited private views imposed on hypermedia documents (inclusion of loops and jumps between topics).

Hypermedia systems differ from conventional retrieval systems in that they encourage informal, personalised, content-oriented information seeking strategies. These informal strategies based on digressions lead to discovering new, unexpected associations between units of information. This feature is called the principle of serendipity. The process of navigation can influence and potentially redefine the original goals of users.

Nevertheless, a set of complementary tools for supporting navigation are required to support effective work with hypermedia material:

- maps: global views of the structure of the network with possibilities to zoom into a selected part;

- fish eye views: showing surrounding of currently inspected document using, e.g. miniatures of neighbouring documents;

- history trees: graphical representation of the sub-net of hypermedia documents visited in actual or previous sessions;

- guided tours: aids for novice users of a hypermedia database, especially useful in the education domain;

- indexes: list of major key words in forms of lengthy menus;

- tables of contents: textual documents resembling conventional tables of contents with direct links to chapters.

The operational advantages of hypermedia navigation can be summarised as:

- easy tracing of references: machine support for link tracing means that all references are equally easy to follow forward or backward;

- ease of creating new references: users can grow their own networks or combine their documents and/or annotations with someone else's documents;

- information structuring: both hierarchical and non-hierarchical organisations can be imposed on unstructured (or restructured) hypermedia information - even multiple hierarchies can organise the same material;

- global views: browsers provide table of content or maps views on large data bases of hypermedia documents - mixing of views on different levels of hierarchies is of advantage;

- customised documents: multimedia document can be threaded together in many ways allowing the same document to serve multiple functions;

- modularity of information: since the same multimedia information unit can be referenced from several places, ideas can be expressed with less overlap and duplication;

- consistency of information: references can be strongly associated with the media objects - moving of the object to other documents carries the links preserving the direct access to the references;

- collaboration: several authors can collaborate in annotating and linking the documents with private and public hypermedia data bases.

7.3.3 Use of hypermedia in telecommunication services

The typical situation in working with hypermedia databases is that the first basic network of associations between individual documents is provided by the author-originator. Then, other users read the documents, place private electronic book-marks on the information elements, which are of private interest for them. A more advanced and convenient technique of dealing with hypermedia material is imposing private structures of associative links on existing set of documents stored in a private data base. In the next stage, users become authors of new documents tailored for their special purpose or interest.

This process of re-use of information is supported on many levels, starting from placing private links on existing documents for more convenient access to the interesting information material - this technique is called private views.

On the next level, new composite documents are constructed from monomedium objects from different hypermedia data bases - compiling the multimedia material.

Finally, a group of users can introduce interactively in conference sessions multimedia information from local sources (cameras, video recorders, scanners and diverse media editors). Specialists from different domains can work together contributing the highest quality materials from their private resources without being forced to meet personally in one location.

Tele-meetings of joint editing of hypermedia documents supported by conversational services of (video) telephony with placing the produced documents in public data bases will introduce new qualities into working with knowledge materials.

8 Implications for network capabilities

This clause examines the different ways of providing a multimedia service and specifies the network capabilities required to meet the human factors requirements.

From the foregoing clauses, multimedia services have the following requirements:

a) it should be possible to vary the quality levels of individual media, either independently or in combination;

b) the response time from the selection of an item to the display of that item should be short or predictable. This is likely to be more of a problem when high definition or moving images are required to be displayed;

c) a high bit rate may be required at certain times to display high definition images or moving images, while lower bit rates may be sufficient at other times. This implies that a variable bit rate service is required to minimise cost;

d) signalling capabilities should be present to allow users to determine charging rates both before and during a call.

Multimedia and hypermedia services can in principle be provided using either narrowband ISDN or B-ISDN. Use of narrowband ISDN to provide multimedia and hypermedia services is more difficult because the bit rates are not available to provide a large range of media qualities and combinations. The media quality levels are limited to those that can be provided using a number of 64 kbit/s channels with appropriate coding and synchronisation of the media using, for example, the multiple rate bearer service.

There are two broad ways of providing a multimedia service, as described in subclauses 8.1 and 8.2.

8.1 A separate connection per service component

The perceived benefit of providing a multimedia service using ISDN (or B-ISDN) is that a separate connection can be provided with the appropriate transfer capabilities for each service component required. Each connection can be of a different connection type and therefore have a different transfer capability. This is shown in figure 3.

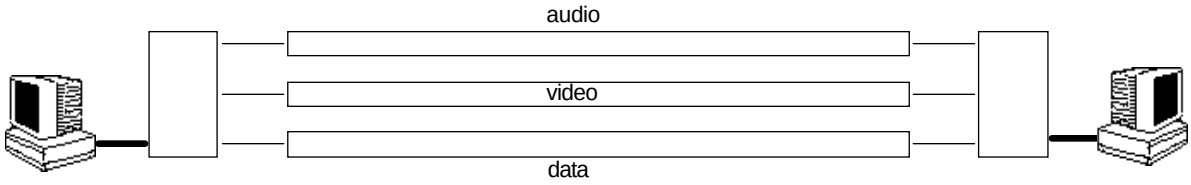

Figure 3: A multimedia service provided using a separate connection for each information type

The quality levels of each information type can be varied independently of the quality levels of the other information types in a multimedia call by varying the service component used. This may result in a change of connection type used for the information type, because the capabilities of that connection type (e.g. maximum bit rate) may be exceeded. For example, in B-ISDN, if users require higher quality video, then a SRTV service component can be de-allocated and an HDTV service component allocated instead. This may require the corresponding connection to be released and a new connection with the appropriate capabilities allocated instead, to ensure that the higher bit rates required are available.

The following problems occur with this approach:

1) synchronisation may be required between media as described in subclause 7.2.3. For example, the audio and video streams may require to be synchronised so that lip synchronisation is maintained. If two independent connections are set up for the two media then they may take different routes to the destination and delays may be considerably different;

2) when, following a change of required media quality, a connection is deallocated and a new one allocated, a new connection with the correct capabilities may not actually be available, or may have a different delay to the previously allocated connection.

These problems make synchronisation more difficult to maintain in separately allocated connections. The preferred method of providing a multimedia service using currently available network technology is therefore as described in subclause 8.2.

8.2 A single connection for all service components

A multimedia service can be provided using a single connection for all information types, as shown in figure 4.

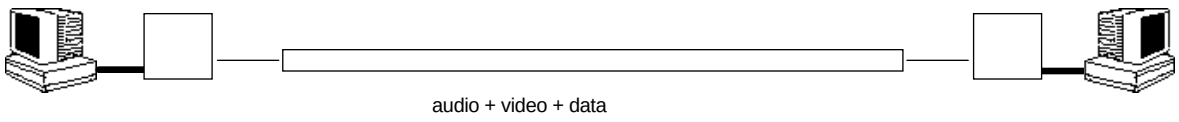

Figure 4: A multimedia service provided using an appropriate bearer service (a single connection for all information types)

An appropriate bearer service is provided from User-Network Interface to User-Network Interface. The information types are multiplexed and demultiplexed in the terminals or the terminal adapters. The network may provide dynamic bit rate allocation and de-allocation, e.g. on a per service component basis where each service component can be associated with a virtual channel identification in the same end-to-end connection. In the simplest case, the network provides only the capacity to support the maximum bit rate required by the multiple information type bundle with the response times required.

To support the change of media types and quality levels during a multimedia call, the network may provide:

- a fixed bit rate connection that can accommodate any likely maximum combination of media and media quality requirements. This is likely to be more expensive for users;

- a variable bit-rate connection that will allow incremental increases in bit rate (and hence media quality). This is not possible with the service described in ETS 300 389 [24], as the number of 64 kbit/s channels needs to be specified at the start of a call and cannot be changed subsequently;

- a maximum bit rate for all media, but allow trade-offs between the bit rate allocated to each medium. This is possible using the capabilities specified in ITU-T Recommendation H.221 [25] (also ETS 300 144 [26]) and ITU-T Recommendation H.242 [27] (also ETS 300 143 [28]). Allocation of bit rates to media is described in ITU-T Recommendation H.221 [25].

9 Extensions to generic user procedures

ETSI TC-HF has defined a set of general rules for user control procedures and a set of generic user control procedures for basic call handling in telecommunication terminals and services, see ETR 170 [29]. The objective is to minimise the risk of inconsistency between user control procedures developed for different services and terminals. These general rules and generic user control procedures are applicable to multimedia telecommunications. The generic procedures defined are:

- Generic Call Set-up;
- Generic Incoming Call;
- Generic Payment;
- Generic Identification;
- Generic Change;
- Generic Call In Progress;
- Generic Termination.

The three procedures covering payment, identification and change can be considered as sub-routines in that they can be accessed from the other four procedures as required.

Provision has therefore been made in the generic user procedures for:

- adding or changing a medium within a call;

- changing the quality level of a medium or a combination of media during a call;

- negotiating on the media and quality levels acceptable to the receiving terminal on call set up and when changes are requested during a call;

- accommodating new occasions for payment and/or identity checking required by changes made to the media or the quality levels.

Annex A: Generic applications

1) Single task applications

Table A.1: Single task applications

C	R	S
Conversation (human-human) Interaction (human-machine) Exchange (machine-machine)	Retrieve	Send

2) Two-task applications

Table A.2: Two-task applications

Collaborative retrieval	People in conversation, and retrieving information.
Interactive retrieval	One or more persons interacting with a machine, and retrieving information.
Collaborative interaction	People in conversation, and together interacting with a machine.
Retrieval and forwarding	Retrieving information from a machine, and forwarding it on to other(s).
Collaborative forwarding	People in conversation, and together forwarding information on to other(s).

3) Three-task applications

Table A.3: Three-task applications

Collaborative retrieval and forwarding	People in conversation, retrieving information from a machine, and forwarding it on to other(s).
Interactive retrieval and forwarding	Interacting with a machine, retrieving information, and forwarding it on to other(s).
Collaborative interaction and forwarding	People in conversation, interacting with machine(s), and forwarding information on to other(s).

History

Document history	
January 1995	First Edition
February 1996	Converted into Adobe Acrobat Portable Document Format (PDF)