INTERNATIONAL STANDARD

ISO 14915-3

First edition
2002-10-15

Software ergonomics for multimedia user interfaces —

Part 3: Media selection and combination

Ergonomie des logiciels pour les interfaces utilisateur multimédias —

Partie 3: Sélection et combinaison des médias

Reference number
ISO 14915-3:2002(E)

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.ch
Web www.iso.ch

Printed in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this part of ISO 14915 may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 14915-3 was prepared by Technical Committee ISO/TC 159, *Ergonomics*, Subcommittee SC 4, *Ergonomics of human-system interaction*.

ISO 14915 consists of the following parts, under the general title *Software ergonomics for multimedia user interfaces*:

— *Part 1: Design principles and framework*

— *Part 2: Multimedia navigation and control*

— *Part 3: Media selection and combination*

Annexes A to D of this part of ISO 14915 are for information only.

Introduction

The design of user interfaces for multimedia applications typically involves a much wider range of design and evaluation issues than that of conventional user interfaces based only in textual and graphical format. Many different techniques and design options are available. Multimedia user interfaces incorporate, integrate and synchronize different media (static media such as text, graphics, images, and dynamic media such as audio, animation, video or other sensory modalities). Within each medium, further distinctions can be made. Graphics, for instance, can be presented either in two- or three-dimensional format and audio can be further categorized according to the level of sound quality or with respect to mono, stereo or surround sound.

Ergonomic design enhances the ability of users to operate multimedia applications effectively, efficiently and with satisfaction (see ISO 9241-11). This can be achieved by careful design of multimedia applications with respect to the tasks they are intended to fulfil (e.g. for work, education or performance support), user characteristics and the environment in which the system will be used. Multimedia applications are often used for communicative purposes. An ergonomic design of multimedia user interfaces can also improve the safety of operating a system (e.g. delivering an alarm in both visual and auditory media).

The range of media available and the interaction of these media have a variety of perceptual, cognitive and other ergonomic implications for the users of multimedia applications. Multimedia can potentially impose on users a high perceptual load, structural and semantic complexity, or a large volume of information to be conveyed through the system. Manipulation of data or information presented in multimedia applications is also often part of the user's activity.

This part of ISO 14915 provides guidance on the selection, combination and integration of media. The focus is primarily on presentational aspects of multimedia (i.e. from system to user) as opposed to control and navigation issues, which are addressed in ISO 14915-2. This part of ISO 14915 starts from information requirements, which are stated in logical terms, and addresses the design issues concerning which media combinations to choose for the information requirements. This is followed by guidance on how the user’s reading/viewing sequence can be directed by design effects to ensure that the user acquires the desired information. Supplementary design guidance for different media combinations and integration are presented in informative annexes A to D.

Software ergonomics for multimedia user interfaces —

Part 3:
Media selection and combination

1 Scope

This part of ISO 14915 gives recommendations for, and guidance on, the design, selection and combination of interactive user interfaces that integrate and synchronize different media. It addresses user interfaces for applications that incorporate, integrate and synchronize different media. This includes static media such as text, graphics, images; and dynamic media such as audio, animation, video or media related to other sensory modalities. Detailed design issues within a single medium (e.g. the graphical design of an animation sequence) are only addressed as far as they imply ergonomic consequences for the user.

This part of ISO 14915 applies to

— presentational techniques for computer-based multimedia applications in general, including stand-alone and networked applications when the prime goal is to support the user's work task or provision of information,

— the design of the software user interface, and

— training and tutorial multimedia insofar as its recommendations bear on effective delivery of information.

This part of ISO 14915 does not deal with pedagogical design issues for tutorial applications and does not address hardware issues, such as input or output devices. The recommendations in this part do not specifically address applications whose primary purpose is entertainment, such as games. The focus of this part of ISO 14915 is on multimedia presentation issues; multimodal input which uses different media, such as speech, in combination with pointing for entering information is not considered in the recommendations provided.

2 Normative references

The following normative documents contain provisions which, through reference in this text, constitute provisions of this part of ISO 14915. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this part of ISO 14915 are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards.

ISO 14915-1:2002, *Software ergonomics for multimedia user interfaces — Part 1: Design principles and framework*

ISO 14915-2:—[1)] *Software ergonomics for multimedia user interfaces — Part 2: Multimedia navigation and control*

1) To be published.

3 Terms and definitions

For the purposes of this part of ISO 14915, the following terms and definitions apply.

3.1
medium, sing.
media, plur.
different specific forms of presenting information to the human user

EXAMPLES These include text, video, graphics, animation, audio.

[ISO 14915-1:2002]

3.2
multimedia
combinations of static and/or dynamic media which can be interactively controlled and simultaneously presented in an application

EXAMPLES These include combinations of text and video, or audio and animation.

[ISO 14915-1:2002]

3.3
static medium
medium in which the presentation to the user does not change over time

EXAMPLES These include text and pictures.

NOTE Adapted from ISO 14915-1:2002.

3.4
dynamic medium
medium in which the presentation to the user changes according to time

EXAMPLES These include video, music, animation, simulations.

NOTE Adapted from ISO 14915-1:2002.

3.5
content
information to be communicated by means of a multimedia application from the originator to the user according to certain communication goals

[ISO 14915-1:2002]

3.6
information type
media-neutral description of information categories that constitute the content and components

NOTE Information types can be used to specify a message to be delivered in a multimedia application. As with media types, information types embed dimensions and categories. An approach to classifying information components with information types is given in informative annex A, which provides a decision tree (Figure A.1) that focuses first on whether a component is physical or conceptual, then whether it is static (not changing) or dynamic and finally the category for the information content.

3.6.1
causal information
information describing the cause and effect of an event, including a sequence of events that describe causation

EXAMPLES Heat causing a liquid to boil. Behaviour of an algorithm that results in the desired goal.

3.6.2
conceptual information
facts, opinions or information about objects which do not have a physical existence

EXAMPLES Taxonomic classes of animals and plants. Opinions about politics.

3.6.3
continuous action information
information describing movement and other activity that is perceived to occur over a period of time

EXAMPLES Making a meal. Driving a car.

NOTE Continuous actions are normally described in the present continuous tense in English.

3.6.4
descriptive information
information which describes an object, entity, or agent

EXAMPLES Red apples, texture of stone.

NOTE This can include states and histories of objects.

3.6.5
discrete action information
information describing movement and other activity that is perceived to occur at a point in time

EXAMPLES Switching a computer on. Closing a door.

NOTE Discrete actions are a set of steps.

3.6.6
event information
information about a state change, message indicating the occurrence of an action or conveying a significant change in the world

EXAMPLES Telephone rings. E-mail message arrives, is sent.

NOTE Events may emanate from the environment as well as from objects.

3.6.7
physical information
information about phenomena which have a concrete existence; objects, agents or scenes that have a physical existence

EXAMPLES Chair, table, landscape.

3.6.8
procedural information
information about a sequence of actions organized to achieve a goal or task

EXAMPLE Instructions to assemble a bookshelf from ready-made components.

3.6.9
relationship information
information about an association between objects or agents

EXAMPLES Seat and legs are part of a chair. A product is manufactured in a factory.

3.6.10
spatial information
information about the spatial properties of the world, such as dimensions of structures, pathways, spatial distribution, location

EXAMPLES Layout of furniture in a room. Directions to the metro station.

3.6.11
state information
properties of the environment, objects or agents that remain constant during a period of time

EXAMPLES The music is being played. A person is sleeping.

3.6.12
value information
quantitative information describing properties of an object

EXAMPLE Person's height 1,80 m.

NOTE Relationships between many values may be shown by graphs and charts (see 3.6.9).

3.7
media type
categories of media that are used to present information to the user

NOTE Media types reflect the psychological properties of a medium as perceived by the user. Figure A.2 and the decision tree (Figure A.3) in informative annex A help to classify media using the following definitions of media sub-types:

3.7.1
audio medium
any medium which can be heard (received by an audio channel)

EXAMPLES Dog barking, music, traffic noise, speech.

3.7.2
language-based medium
medium with content based on natural or formal languages

EXAMPLES Alphanumeric text. Spoken language. Symbols interpreted in language-based terms, such as hieroglyphics, mathematical equations, and chemical formulae.

3.7.3
moving image medium
visual medium that is delivered at a rate that is judged by the human viewer to be a continuous image

EXAMPLES Video, film, animated diagrams, simulations.

NOTE For example, at a frame refresh rate above or near the flicker fusion frequency of 30 frames per second.

3.7.4
non-realistic medium
medium that is perceived by the user as not faithfully representing the natural world

EXAMPLES Diagrams, graphs, cartoons.

3.7.5
realistic medium
medium perceived by the user to faithfully represent the natural world

EXAMPLES Natural sounds. Photographic images. Film showing people and natural scenes.

NOTE Realistic media may be designed to create the illusion of the natural world, e.g. realistic animations of dinosaurs.

3.7.6
still-image medium
visual medium that is not presented continuously, although frames may be shown in a sequence either controlled by the user or by the system with a time delay

EXAMPLES Photographs, drawings, graphs.

3.8
agent
person or machine which carries out actions and creates events

EXAMPLES User, designer, computer program.

3.9
concurrent media
two or more media that are juxtaposed/used simultaneously during a certain period when presented

EXAMPLE A voice-over describes action in a video.

3.10
direct contact point
thematic link between two media implemented with a designed effect in both the source and destination medium

EXAMPLE A text caption is linked by an arrow to an image component that is highlighted.

3.11
indirect contact point
thematic link between two media that is implemented with a designed effect only in the source medium

EXAMPLE The text refers to a diagram with the instruction “see Figure 1”.

3.12
media combination
sequential or concurrent combination of two or more media

EXAMPLES A video is presented in a window embedded in another window containing a still image. Speech introduces a video which is then played.

3.13
sequential presentation
arrangement of two or more media that are represented one after another but do not overlap in time

EXAMPLE A video is shown, followed by a text summary.

3.14
thematic link
requirement to direct the users’ reading/viewing sequence between two media

NOTE For implementation, see contact points 3.10 and 3.11.

4 Application of this part of ISO 14915

4.1 Intended user groups

The following groups are the intended users of this part of ISO 14915:

— user interface and multimedia designers who will apply this part of ISO 14915 during the development process;

— evaluators responsible for quality assurance who will ensure that products meet the recommendations of this part of ISO 14915;

— potential buyers in selecting appropriately designed multimedia products;

— designers of multimedia development tools to be used by user interface and multimedia developers.

4.2 Applying the recommendations

The design guidelines provided in this part of ISO 14915 extend the principles described in ISO 14915-1. Multimedia user interfaces should be designed according to both the principles of ISO 14915-1 and the guidelines described in this part of ISO 14915. For certain contexts (e.g. certain tasks or user groups), the designer may be forced to follow one guideline at the expense of another in order to achieve the optimal design.

4.3 Reporting conformance to this part of ISO 14915

If a claim of product or application conformity with this part of ISO 14915 is made, the procedure used in establishing requirements for developing and/or evaluating the multimedia user interface shall be specified. The level of specification of the procedure is a matter of negotiation between the involved parties. ISO 14915 is a multi-part standard and therefore claims of conformity are related to the individual parts and not to the International Standard as a whole.

5 General guidelines for media selection and combination

5.1 General

This clause contains general guidelines for media selection and combination followed by more specific guidelines that map information types to appropriate media types for effective delivery. The principles in ISO 14915-1 provide a basis for interpreting the following guidelines. Two or more media are considered combined if their presentation is concurrent. Media are also considered to be combined or contiguous where the content is closely related or where the media are explicitly grouped in a display in adjacent windows or displayed sequentially; for example, presenting a picture followed by a text description.

Combining media can have advantages for the user. First, interfaces can be created that present information in a way similar to the real world. Depending on the context of use, this can make users' tasks easier, or more natural, especially where features of the information match the users' experiences of the real world. For example, a picture of a beach presented together with the sound of waves provides a better evocation of the subject than the picture or the sound alone. Combining media can also help accommodate user preferences for information in a particular format. For example, presenting text as well as pictures can accommodate users with a preference for either.

5.2 Supporting user tasks

Media should be selected and combined to support the user's tasks.

EXAMPLE To compare two views, an architect's drawings with corresponding photographs showing the side and front elevation of a building are placed side by side.

NOTE Some tasks benefit more from combination than others. If the task involves learning, or drawing attention to specific information, the users can benefit from media combination that presents information redundantly. If, however, the task is focused predominantly on one medium, for example visual inspection of diagrams, there can be considerably less benefit from combination. The characteristics of the user's task can also influence the sequence/concurrency of presentation; for instance, if comparison is required, two images can be presented concurrently.

5.3 Supporting communication goals

Media should be selected to achieve the communication goal in the application.

EXAMPLES In an application that is critical for safety, the communication goal is to warn the users and protect them from danger. In an aircraft emergency evacuation demonstration, speech is used for the instructions, with a diagram to show the evacuation route.

5.4 Ensuring compatibility with the users' understanding

Media should be selected to convey the content in a manner compatible with the user's existing knowledge.

EXAMPLES A radiation symbol is used to convey danger to users who have the appropriate knowledge. An architectural diagram is used to convey the structural layout of a building to architects and design engineers.

The user's ability to understand the message conveyed by a particular medium should influence selection. This is particularly important for non-realistic image media (diagrams, graphs) when interpretation is dependent on the users' knowledge and culture.

5.5 Selecting media appropriate for the users' characteristics

The characteristics of the user population should be considered when selecting media.

EXAMPLE Text is substituted by speech for blind users. Large point size text is accompanied by spoken representation of the text for older users.

NOTE Users can be categorized as visualizers or verbalizers using psychologically based questionnaires. This information can be used to aid selection of image- or language-based media.

5.6 Supporting users' preferences

If appropriate to the task, users should be provided with alternative media from which they can select a preferred medium or suppress certain media.

EXAMPLES The user chooses to display text captions on a picture rather than a speech commentary, or suppresses an audio dialogue in a noisy environment. A blind user selects speech rather than text.

NOTE Some users can prefer to interact with systems using a particular medium. The abilities of the users and properties of the users' machine are important, so users may be given the option of low- or high-resolution graphics displays.

5.7 Considering the context of use

Selection and combination of media should be appropriate in the context of use.

EXAMPLES An inappropriate combination is the presentation of auditory as well as visual display of bank-account details which could compromise the user's privacy. A training video depicting an action accompanied by the speech "This is not correct" could be missed if the speech is inaudible in a noisy environment.

NOTE Certain environments can impede accurate perception of information presented in a specific medium; for example, an auditory warning might not be heard if it is presented in a noisy environment.

5.8 Using redundancy for critical information

If important information is to be presented then the same subject matter should be presented in two or more media.

EXAMPLE Displaying an alarm-clock function visually as well as aurally. In a language-learning application, words are spoken and displayed in text.

NOTE Effective redundant combinations present similar but not identical content on different media. Redundant representation is useful for training and educational applications.

5.9 Avoiding conflicting perceptual channels

The same perceptual channel (e.g. hearing or vision) should not be used in concurrently presented dynamic media, if the user needs to extract information from both media.

EXAMPLE Playing two or more videos with unrelated content is avoided because these will interfere with each other and distract attention.

NOTE Concurrent presentation of two or more dynamic media makes it difficult for the user to perceive information from each individual source unless the information is easy to integrate. Exceptions to this guideline occur in entertainment applications, such as playing two unrelated videos in popular-music promotions.

5.10 Avoiding semantic conflicts

Presentation of conflicting information in any combination of media should be avoided.

EXAMPLE Avoid aural presentation of the word “Press the blue button” while the visual display shows a black and white image.

NOTE Users cannot comprehend or gain an integrated understanding of conflicting information from different media, especially in concurrent presentations.

5.11 Designing for simplicity

Minimal combination of media should be used to convey the information necessary for the user’s task.

EXAMPLE In a musical tutorial, the sound of music is combined with textual representation of the musical score. Adding a video of a concert performance gives little extra information and distracts the user.

NOTE As the number of media used increases, the user’s effort required to attend to and process each medium also increases, as does the effort in cross-referencing media. The trade-off between simplicity and more complex combinations will depend on the user and task (see clause 5.2 and 5.3).

5.12 Combining media for different viewpoints

Wherever appropriate to the task, different views on the same subject matter should be provided by media combination.

EXAMPLES Musical notation in a diagram gives the composer’s structural view of a symphony, speech commentary gives the musician’s view of the structure, and music on the sound channel provides the aesthetic view. Two movies are played in separate windows to show different viewpoints on the same scene, one showing a long-shot context of a football game, the other a close up of a foul between two players in the long shot.

NOTE Presenting different views by media combination helps the user to assimilate information that is related to the same topic or theme.

5.13 Choosing media combinations to elaborate information

Whenever appropriate to the task, media combinations should be selected to extend the information content.

EXAMPLE Showing a diagram of planets revolving around the sun with speech explaining forces of gravity and momentum.

NOTE Media combination is used to add information to an existing topic, whereas combining media for different viewpoints presents different aspects of the same topic.

5.14 Guarding against degradation

Technical constraints should be considered when selecting media delivery to avoid degraded quality or unacceptable response times.

EXAMPLES To avoid delay in downloading a web page, moving images are segmented into storyboard stills and displayed as a slide show. The display area of a moving image is reduced rather than slowing the frame rate. Simple images with lower bandwidth requirements are used rather than photographic quality images. Users are warned of transmission delays.

NOTE Visual media, especially moving images, are more prone to degradation if there are bandwidth or network constraints in distributed multimedia. Degradation can result in poor image quality, slower than acceptable frame rates for moving images, and poor audio quality.

5.15 Previewing media selections

If appropriate for the task, the media available for selection should be viewable by the user in a preview facility.

EXAMPLE A web link to video allows the user to view miniature samples of the video before it is downloaded.

NOTE When control over media selection is given to the user, previewing can be combined with controls to choose the way media are downloaded in high or low resolution. Guidance on controls is given in ISO 14915-2.

5.16 Using static media for important messages

A still image and text should be used for important information other than time-critical warnings.

EXAMPLE Key points in an engine assembly task are shown with still images and bullet points in text.

NOTE Very little detail is remembered from video and speech. Dynamic media can be used to alert the user and direct their attention to important messages which are conveyed in static media.

6 Media selection for information types

6.1 General

This clause gives recommendations on media selection. Information requirements can be defined logically, without reference to the physical medium that conveys them, to make the user's requirement explicit.

One way to approach media selection is as follows:

— segment the content into information components according to the task and users' requirements;

— assign information types to information components using the decision trees in Figures A.1 and A.2 in informative annex A and the definition in 3.6;

— choose media types for information types using Table 1 and the guidelines in 6.2 to 6.4.

In many cases, multiple media types may be selected for one information type. Table 1 provides examples to illustrate the possible combinations of media for a particular information type. The guidelines in 6.2 to 6.4 give the major mapping between an information type and media types with other, possible mappings in the notes. Table 1 illustrates the selection guidelines expressed in the order: preferred selection, other possible media-type selections.

Table 1 — Summary of examples for selection and combination of media for information types

Information type/ Media type	Causal	Conceptual	Continuous action	Descriptive	Discrete action	Event	Physical	Procedural	Relationship	Spatial	State	Value
Realistic audio	sound of rain and storms	**	sound of skiing	sounds of the countryside	click of ON switch	sound of the starting gun	*noise of a tornado*	**	bird song of related species	echos in a cave	sound of snoring	musical note encodes a value
Non-realistic audio	**	rising tone illustrates increasing magnetic force	continuous tone signals progress of action	morse code describes a ship	tones signal open/close door	*alarm siren*	**	**	tones associate two objects	sonar and doppler effect	continuous sound in a heart-beat monitor	Morse code for numbers
Speech	*tell someone why El Niño happens*	tell someone about your religious beliefs	tell someone what a ski turn looks like	verbal description of a person	tell someone how to turn computer on	*tell someone the race has started*	tell someone how it feels to be in a storm	speak instructions on engine assembly	tell someone that Jack and Jill are related	tell someone the pathway to and location of railway station	tell someone "Jane's asleep"	verbal report of numbers, figures
Realistic still image	*photograph of El Niño storms and ocean currents*	Statue of Liberty photograph represents "freedom"	set of photographs showing snapshots of action	*overview and detail photographs of a car*	*photograph of computer ON switch*	photograph of the start of a race	*photograph of a person's face*	*photographs showing engine assembly*	juxtapose photographs of twins	*photograph of a landscape*	photograph of a person sleeping	**
Non-realistic still image	*diagrams of ocean currents and sea temperature to explain El Niño*	*hierarchy diagram of plant taxonomy*	diagram with arrow depicting ski-turn motion	histogram of ageing population	diagram showing where and how to press ON switch	event symbol in a race-sequence diagram	**	explode parts diagram of engine with assembly numbers	*graphs, histograms, ER diagrams*	*map of the landscape*	waiting-state symbol in race-sequence diagram	*charts, graphs, scatter plots*
Text	*describe reasons for El Niño storms*	*explain taxonomy of animals*	describe ski-turn action	*describe a person's appearance*	describe how to turn computer on	report that the race has started	report of the storm's properties	*bullet-point steps in assembling engine*	*describe brother and sister relationship*	describe dimensions of a room	*report that the person is asleep*	*written number one, two*
Realistic moving image	*video of El Niño storms and ocean currents*	**	*movie of person turning while skiing*	aircraft flying	**	*movie of the start of a race*	*movie of a storm*	*video of engine-assembly sequence*	tracking shot of generations in a family	fly through landscape	video of a person sleeping	**
Non-realistic moving image	*animation of ocean temperature change and current reversal*	animated diagram of force of gravity	animated mannequin doing ski turn	**	animation showing operation of ON switch	animation of start event symbol in a diagram	**	*animation of parts diagram in assembly sequence*	animation of links on ER diagram	animate architect's drawings of a building	**	animate graphs and charts
Language-based: formal, numeric	equations, functions formalizing cause and effect	*symbols denoting concepts, e.g pi*	**	data type definitions	finite state automata	event-based notations	**	procedural logics, process algebras	functions, equations, grammars	graph theory, topological grammars	state-based languages, e.g. Z	numeric symbols

Italics denote the first preference selections.

** Indicates that mappings are unlikely.

6.2 Consider information types

Information types should be considered when selecting and combining media.

NOTE The users' characteristics and the users' tasks are also considered when selecting media; see ISO 14915-1, 5.2 and 5.5.

6.3 Consider multiple-information types

If the user's information requirement consists of multiple-information types then a combination of media should be considered.

EXAMPLE A procedure for explaining a physical task; first realistic image media will be selected, then a series of images and text.

NOTE The information types differentiate physical from conceptual, then static from dynamic information. Media combination is also covered in 5.8 and 5.12.

6.4 Selecting and combining media

6.4.1 Physical information

For physical information, a realistic still or moving image should be considered, unless user or task characteristics override this choice.

EXAMPLE A photograph is used to portray the landscape in a national park.

NOTE When physical details, such as the dimensions of a building, need to be communicated precisely, language-based media can be overlaid on an image. When a partial abstraction of physical information is desired, a non-realistic image can be used (e.g. sketch or diagram).

6.4.2 Conceptual information

Language-based media (text, speech) and/or non-realistic image media should be considered for conceptual information.

EXAMPLES Sales objectives and commentary on the market strategy are conveyed by text bullet points or spoken commentary. A flowchart is used to portray the functions of a chemical process, with speech to describe the functions in detail. Categories of animals are shown in a tree diagram.

NOTE Conceptual information with complex relationships can be shown by non-realistic images (graphs, sketches, diagrams) or by graphical images embedding text. Conceptual information can also be conveyed using realistic images and a metaphor, for instance, a rainbow photograph illustrates the colours of the spectrum.

6.4.3 Descriptive information

Language-based (text, speech) and/or realistic image media should be considered for descriptive information.

EXAMPLES Narrative text describes the properties of a chemical compound such as salt. The attributes and properties of an apple are shown by a photograph of a red apple with a caption "organically grown".

NOTE When describing objects' behaviour or movement, realistic moving image media can be used.

6.4.4 Spatial information

A realistic and/or non-realistic still image should be considered for spatial information.

EXAMPLE The location of cargo on a ship is shown by a diagram.

NOTE Detailed spatial information can be presented in a realistic image, e.g. photographs. Spatial information which involves complex pathways can be conveyed by a moving image, e.g. animate the pathway. However, location, orientation and pathway information is remembered more effectively from still images than from moving images.

6.4.5 Value information

A language-based medium (numeric text, tables) should be considered for numeric values and quantitative information.

EXAMPLE The height and mass of a person is given as 1.80 m, 75 kg.

NOTE Speech is not effective for communicating several numbers because holding many numbers in the working memory is difficult. A persistent medium is advised so that numbers can be visually inspected during a task.

6.4.6 Relationship information

A non-realistic image (e.g. chart, graph, diagram) should be considered for displaying relationships within and between sets of values or between concepts.

EXAMPLES The values for rainfall in London for each month are displayed using a histogram. The relationship between words in a thesaurus is shown as a hypertext diagram.

6.4.7 Discrete action information

For discrete action, realistic still image media should be considered.

EXAMPLE An image of the coffee machine showing a person performing the action illustrates filling a coffee percolator with water.

NOTE Use of still image media for discrete actions allows the relationship between the action, the object acted on and the agent performing the action to be inspected. Conceptual actions, e.g. mental processes, can be described using speech or text.

6.4.8 Continuous action information

For complex or continuous actions, moving image media should be considered.

EXAMPLE Turning while skiing is illustrated with a video.

NOTE A complex physical action can be better illustrated with non-realistic media (animation) so the coordination of motor actions can be inspected.

6.4.9 Event information

For giving information on significant events and issuing warnings, an audio medium, e.g. speech or sound, should be considered for alerting the user.

EXAMPLE The outbreak of a fire is conveyed by sounding an alarm.

NOTE Abstract events can be explained in language. A realistic or non-realistic still image can be used to deliver further information about the context of the event, so after the fire alarm has sounded, a red marker on a diagram of the building shows the location of the fire.

6.4.10 State information

For states, still image or language-based media should be used.

EXAMPLE The state of the weather is shown by a photograph of a sunny day.

NOTE Abstract states can be explained in language-based media or in diagrams. If a sequence of discrete states is required, an animation or a series of still images can be used as a slide show.

6.4.11 Causal information

To explain causality, still and moving image media should be considered, combined with language-based media.

EXAMPLE The cause of a flood is explained by text describing excessive rainfall with an animation of rain falling on the land, then flowing into the river, causing the river to rise and overflow its banks.

NOTE With more complex information types, it is advisable to use a media combination to form a pattern. Causal explanations of physical phenomena can be given by introducing the topic using language-based media; showing the cause and effect by a combination of still image and language-based media for commentary; integrating the message by a moving image with a voice commentary; and providing a bullet-point text summary.

6.4.12 Procedural information

A series of images with text captions should be selected for procedural information.

EXAMPLE Instructions for assembling a bookshelf from a kit are given as a set of images for each step, with text captions.

NOTE To explain procedures, a combination of media can be necessary, such as a still-image sequence with text, followed by an animation of the whole sequence. Non-physical procedures can be displayed as formatted text, e.g. bullet points or numbered steps.

7 Media integration

7.1 General

Selection of media only provides the raw material for design. The selected media should be combined and integrated in a presentation sequence so that the presentation delivers a coherent message. The following issues should be considered when combining or sequencing the presentation.

7.2 Design issues

This subclause and 7.3 give design advice on choosing combinations of media that deliver particular types of information effectively. They also address the issue of how the users' reading/viewing sequence can be influenced by the designer to ensure that important information is perceived.

Media selection also raises the issue of how a medium is presented on the user interface. This issue affects visual media in particular. For example, text can be displayed in a separate window from an image; alternatively, text can be overlaid on an image as captions. In the former, the user will tend to treat the text and image as separate entities; in the latter, the image and captions will usually be viewed as an integrated whole. This changes the user's reading/viewing sequence. Whereas media displayed in separate windows tend to be viewed sequentially, text captions will be viewed with the image on which they are overlaid. Integration can be more effective when diagrams are overlaid on a naturalistic image, although the disadvantage of making the image too complex, and hence making information extraction more difficult, should be considered.

Timing and synchronization are important issues in the design of concurrent presentations with dynamic media (speech, video). Further guidance is given in ISO 14915-2.

7.3 Media integration guidelines

7.3.1 General

This clause gives general guidelines for integrating media. Further examples are given in informative annex B to explain the potential uses of different media.

7.3.2 Advance organizers

A language medium should be considered for introducing material presented in another medium.

EXAMPLE Speech is used to introduce the subject matter of a video, then the video is shown.

NOTE The first medium introduces a topic that is expanded on by the content in the following media. When users are unfamiliar with the content of the following media, they are made aware of the content to be presented in the first (advance organizer) medium.

7.3.3 Synchronized, related media

Media that are presented concurrently and are related in content should be synchronized to match the user's perception.

EXAMPLE Lip movements on film are synchronized with speaking by an actor within 70 ms. The presentation of speech and display of text are timed for each word.

NOTE Tightly coupled synchronization is not always possible across a network.

7.3.4 Separating audio content sources

When combining two audio media, the perception of each medium should be dissimilar so as to distinguish the sources in the presentation.

EXAMPLE Tones overlaying recordings of bird song signal a change in the species of bird which is singing.

7.3.5 Avoiding interference in audio media

Presentation of two audio media should not be concurrent if the background sound is intrusive or masks another more important sound. The designer should ensure that the amplitude of two sounds do not interfere with each other.

EXAMPLE A voice-over is interleaved as a commentary on bird song so that the listener can hear the bird song without interference from the speech.

7.3.6 Limiting speech interruptions in audio and language-based media

Interruptions by non-realistic audio should be brief and placed at pauses, sentence or phrase boundaries in the speech, or be triggered by the user's specific command.

EXAMPLE Tones between voice-mail messages.

7.3.7 Integrating non-realistic images with realistic images

When realistic images are complemented with a non-realistic image, one image should be simple and the subject matter of the images should be related.

EXAMPLE A photograph of an engine is overlaid by a simple diagram of its components.

NOTE Two complex non-realistic images make cross-referencing between the images difficult. Choosing which image will be the more complex or simple may depend on the user’s task.

7.3.8 Use of captions with images

If the image is of greater significance, brief captions should be used rather than separate text.

EXAMPLE In a text description of a scene, accompanied by a photograph, a caption points out a camouflaged bird hiding in a photograph of a landscape.

NOTE If cross-referencing is required, both the image and the text can be presented concurrently, with text captions to draw the user's attention to important image components.

8 Directing users' attention

8.1 General

An important consideration of multimedia design is to link the thread of a message across several different media. This clause gives recommendations on planning the users' reading/viewing sequence, and guidelines for realizing these recommendations in presentation design. However, reading/viewing sequences can also be implemented by hypermedia dialogues and navigation controls. The essential differences are timing and user control. In a presentational design, the reading/viewing sequence and timing are set by the designer.

User attention is sequential for time-varying media, and the sequence in which a user reads is directed by the layout of text, although this is culturally dependent. For example, western languages read from left to right, Arabic languages in the opposite direction. However, the viewing order in images is unpredictable unless the design specifically selects the user's attention. The following design issues may be considered:

a) planning the overall thematic thread of the message through the presentation or dialogue;

b) drawing the user's attention to important information;

c) establishing a clear reading/viewing sequence;

d) providing clear links when the theme crosses from one medium to another.

An important consideration in multimedia design is to control the user's reading/viewing sequence so that important information is noticed. If the message theme is important and the cross-reference is critical then the design needs to draw the user's attention in both the source and destination medium (a direct contact point); otherwise, drawing the user's attention to related information in the source medium will suffice (an indirect contact point). Designing for attention is particularly important when using images. User attention to media varying with time is determined by the medium itself, i.e. we have little choice but to listen to speech or to view animations in the order in which they are presented.

Further guidance on hypermedia implementation of contact points is given in ISO 14915-2.

8.2 Direct-contact points for key thematic links

A direct-contact point should be used if the connection between information in two different media is important.

EXAMPLE Speech is used to direct the user to the object in the image while highlighting the object which is being spoken about: "Look at the map; the road to London is ... (highlight)"; or a text caption is revealed with an arrow pointing to the road.

NOTE Direct-contact points direct attention by a command or stimulus from the source medium and an anchor stimulus in the destination medium. Direct-contact points emphasize the links between media but can become intrusive if overused.

8.3 Direct-contact points for linked components

Direct-contact points should be used if components in both the source and destination medium are important and have to be perceived.

EXAMPLE "Find the spark plug illustrated in the photograph" (speech track), spark plug is highlighted (in image), "adjust the gap by rotating the body …", highlight gap (in image), caption overlay on moving image.

NOTE If the message or information items in the destination medium are less critical then attention control in the source medium alone can suffice.

8.4 Indirect-contact points

Indirect contact should be used when the connection between information in two media is necessary but perception of subcomponents in the destination medium is less important.

EXAMPLES Direct the attention to a following media component, “look at Figure 1”. Speaking about an object while displaying the image. Freeze a frame video while describing objects.

NOTE Indirect contact points are less intrusive on the user’s reading/viewing sequence so they can be used more frequently than direct contact points without becoming disruptive.

8.5 Sequence of contact points to connect a thread of topics

Multiple-contact points in the same medium should be organized in a logical order to follow the theme.

EXAMPLE In a biology tutorial, explaining parts of a cell is organized with interleaved speech segments and a diagram describing the cell’s components from top to bottom and left to right. Highlighting techniques locate each component in turn, following the order of the spoken explanation.

8.6 Guidelines for contact points between media pairs

8.6.1 General

The guidelines in 8.6.2 to 8.6.7 provide techniques for directing attention to contact points. Guidelines specific to the medium are given first, followed by guidelines for media combinations (source medium to destination medium). In some cases, the source and destination medium can be presented concurrently; however, the message theme is assumed to appear in the source medium first.

The guidelines are stated to implement direct-contact points with an attention effect in the source and destination medium. Indirect-contact points can be derived from each guideline by ignoring the attention effect in the destination medium. A summary of the design techniques for implementing contact points between different media is given in Table 2. The guidelines in 8.6.2 to 8.6.7 elaborate Table 2 but, where media combinations contain similar advice, these have been merged to reduce duplication.

Table 2 — Summary of examples of design effects for implementing contact points between media pairs

Destination medium / Source medium	Realistic audio	Non-realistic audio	Speech	Realistic still image	Non-realistic still image	Text	Realistic moving image	Non-realistic moving image
Realistic audio	No reference possible – amplitude may be used to mark segments	No reference possible- amplitude may be used to mark segments	Amplitude may be used to mark a segment of audio which is then described by speech	Hot spot on the image links to the audio tone; amplitude or replay for emphasis	Amplitude or replay for emphasis; end of audio links to image	The audio is paused, or replayed, followed by text highlight; text cue link to audio	The audio is paused and replayed; movie frame triggers link to audio, freeze frame	The audio is paused and replayed, movie freeze frame or change shot
Non-realistic audio		Tone marker refers to following audio segment	Speech commands to operate a link to audio, tone alert for speech warning	Hot spot/ button on the image links to audio, end of audio links to image	Music is used to introduce an image; hot spot/button on the image links to audio	Audio tone alerts user to a text cue	Hot spot/ button on the image links to audio; end of audio links to image	Movie frame triggers link to audio; freeze frame and button cue to link to audio
Speech			Speech refers to a following speech segment with prosodic emphasis	Speech command links to image; image hot spot cues speech	Speech command links to an image highlight	Text cue links to speech; speech keyword activates link to text	Hot spot/ button on the image links to speech; end of audio links to speech	Speech introduces cartoon; movie frame triggers link to speech; freeze frame
Realistic still image				Hot spot/ button cues may link images in both directions	Hot spot/ button cues may link images in both directions	Text cue links to image; hot spot/button on image links to text	Hot spot/ button on the movie links to still image; hot spot on still image activates link to movie	Hot spot/ button on the movie links to still image; hot spot on still image activates link to movie
Non-realistic still image					Hot spot/ button cues may link images in both directions	Text cue links to image; hot spot/button on image links to text	Movie frame triggers link to image; freeze frame and button cue to link to image; hot spot on image links to movie	Movie frame triggers link to image, freeze frame and button cue to link to image; hot spot on image links to movie
Text						Text cues link in either direction	Text cue activates link to movie; movie freeze frame/button links to text	Text cue activates link to movie; movie freeze frame/button links to text
Realistic moving image							Movie frame automatically links to second movie, freeze frame and button activates link	Movie frame automatically links to second movie, freeze frame and button activates link
Non-realistic moving image								Movie frame automatically links to second movie, freeze frame and button activates link

8.6.2 Source medium: realistic audio

8.6.2.1 General

This medium contains natural sounds. Contact points can be implemented by changing the amplitude of sound; however, changes in frequency will interfere with the quality of realism. The design effects described in 8.6.2.2 to 8.6.2.5 should be considered when implementing contact points from realistic audio to other media.

8.6.2.2 Realistic audio to non-realistic audio

No direct reference is possible; however, amplitude may be used to mark segments.

EXAMPLE Segment of bird song is played louder, followed by a tone to indicate the length of the song segment.

8.6.2.3 Realistic audio to speech

Amplitude may be used to mark a segment of audio which is then described by speech.

EXAMPLE A segment of bird song is played louder, followed by speech which describes the importance of the segment: “as you heard, the signature of this bird has a rising tone…”

8.6.2.4 Realistic audio to image (moving and still)

Amplitude may be used to mark the audio segment or the segment may be paused and replayed for emphasis, followed by displaying the image with a highlight if appropriate.

EXAMPLE The recording of a monkey’s vocalization is played twice, followed by an image of the monkey.

8.6.2.5 Realistic audio to text

The audio is paused and replayed, or an amplitude marks a segment for attention; text is then displayed, with highlighting if appropriate.

EXAMPLE The recording of noise from an aircraft is replayed with a text caption describing the type of aircraft.

8.6.3 Source medium: non-realistic audio

8.6.3.1 General

This medium includes all artificially created sounds, including music. Attention can be directed by using change in amplitude and frequency (tone), change in intervals between sounds, and patterns of sound. Music can be used for powerful effects but these differ between individuals and cultures according to the type of music. The design effects described in 8.6.3.2 to 8.6.3.5 should be considered when implementing contact points between non-realistic audio and other media.

8.6.3.2 Non-realistic audio to realistic audio

A short segment of music is used to alert the user, followed by playing the natural sound.

EXAMPLE The “Jaws” theme is played, followed by sound of an unfortunate victim being attacked by a shark.

8.6.3.3 Non-realistic audio to speech

A tone is used to alert the user before speech is played.

EXAMPLE In language teaching, a tone sounds before the instruction “please repeat after me”.

8.6.3.4 Non-realistic audio to image (still or moving)

Music is used to introduce an image.

EXAMPLE 'Ode to joy' theme from Beethoven's 9th Symphony is played, followed by an image of the European flag or a movie clip showing the European Parliament.

8.6.3.5 Non-realistic audio to text

A short tone is used, followed by highlighting a segment of text.

EXAMPLE A beep warning is used to alert the user to a spelling error in a word processor.

8.6.4 Source medium: speech

8.6.4.1 General

This is a powerful medium as it allows directed commands for attention. Explicit commands can be expressed in a variety of syntactic forms and depend on the destination medium. Emphasis can be created by use of voice amplitude, frequency change, changing the speaker, or explicit commands to the listener. The design effects described in 8.6.4.2 to 8.6.4.5 should be considered when implementing contact points from speech to other media.

8.6.4.2 Speech to image

A component in the image, or the whole image, may be referred to by name with voice-tonality emphasis, while highlighting the destination component.

EXAMPLE Speech "Look at the object (x) in the photograph/diagram", and the image component is highlighted.

NOTE Emphasis in speech can be achieved by either a voice change or by loudness or tone. A spoken reference alone, "Look at the following diagram", may be used for an indirect-contact point.

8.6.4.3 Speech to moving image

A component in the moving image or the whole sequence is referred to by name, or described combined with slow motion or freeze frame in the moving image.

EXAMPLE "As can be seen in the following clip, the first character is Hamlet", using slow motion or freeze frame for a direct-contact point; "the next video clip shows the characters in the first scene", followed by the image sequence for an indirect-contact point.

NOTE Reference to components within moving images must be checked to ensure that the objects referred to are visible. If the objects or agents are not familiar and the sequence is short then the contact point can fail. Replay controls can be provided when the contact point is difficult to establish (see also ISO 14915-2).

8.6.4.4 Speech to language-based text

Words or phrases in the text are referred to in speech, while highlighting the appropriate text segment.

EXAMPLE Speech "Inspect the third line of the current text." This is the important line (highlighted).

NOTE Speech and text are easier to comprehend if they are consistent with each other. Listening to speech and reading text concurrently is difficult; a short delay after the speech before displaying the destination reference can help comprehension.

8.6.4.5 Speech to realistic audio

A change in the speaker's voice tone or amplitude is used to draw attention to the following sound.

EXAMPLE "Listen to the following sound track"; play bird-song clip; "Listen to the rising note in the lark's song"; play another clip; "This is followed by a descending note".

NOTE Speech and sounds cannot be processed concurrently if the message theme is contained in both, as both media compete for audio processing. Speech can be played first to direct attention to the sound. Amplitude can be used to emphasize a segment of the sound. Speech descriptions can be interleaved with sound clips to refer directly to specific sounds.

8.6.5 Source medium: still image

8.6.5.1 General

The following attention-directing effects can be applied to realistic and non-realistic images. Visual attention is directed by highlighting objects or by the use of symbols, such as arrows, arcs or marker icons, or links by similar symbols. Highlighting can be achieved by the use of colour or size/shape change. The design effects described in 8.6.5.2 to 8.6.5.5 should be considered when implementing contact points from still images (realistic and non-realistic) to other media.

8.6.5.2 Still image to still image

Image highlighting links components in the source and destination images.

EXAMPLES Arrows can be used to direct attention from one photograph/diagram to another; alternatively, highlighting two resistors in two images can be used to compare them to electrical circuit designs.

NOTE The onset of highlighting or the direction of arrows will control the direction of attention from source to destination. Components in each image can be highlighted either one after the other or concurrently.

8.6.5.3 Still image to moving image

Image highlighting directs attention from the image or a component therein to segments of the video, which is then paused to emphasize the destination component if it is hard to perceive or transient.

EXAMPLE A photograph of an actor is shown followed by a movie which is paused, then an arrow points to the actor in a crowd scene.

NOTE Objects which appear in both the still and moving image can be highlighted to direct attention. Arrows can be used to direct attention from the still image to the movie window, or markers overlaid on the video to pick out referenced objects; however, arc pointing within the moving image can obscure the information content.

8.6.5.4 Still image to text

Image components and the corresponding text segments are highlighted while the image components are referred to by bold name/identifiers in the following text.

EXAMPLE Birds in a photograph and words in a text describing it are highlighted or an arc links the bird to text describing it.

NOTE An arc or arrow can be used to link the image, or image components, to the appropriate text or segments of text. Arrows and arcs can be less effective than highlights as the expected convention is to use this technique from text to image as a caption. Timing the onset will control the source and destination reference. Contact points between a source image and language-based destination medium can cause confusion if the reason for drawing attention to the image is not apparent to the user until text explains it.

8.6.5.5 Still image to speech and realistic/non-realistic audio

The image component is highlighted and referred to by name or described in the speech while the segment in the following speech or sound is emphasized by change of voice, tone or amplitude.

EXAMPLE Highlight a danger area on a map followed by speech, "The military area is nearby, as can be seen from the part of the map coloured red".

NOTE The context of the highlight is not always clear before the speech track is played. Highlighting an image component and then describing it in speech can appear unnatural to the user. In natural discourse, references point from language to an image. Concurrent presentation of a highlighted object with sound can help to link the two media (e.g. play the sound of a whale and highlight a drawing of the relevant species).

8.6.6 Source medium: text

8.6.6.1 General

Text uses the same language-based commands as speech to direct attention; however, text can also be used as captions with arrows and other graphical conventions to link to other media. Highlighting, larger or different fonts, use of colour, bold print and underlining all help to make words salient. The design effects described in 8.6.6.2 to 8.6.6.5 should be considered when implementing contact points from text to other media.

8.6.6.2 Text to still image

Text captions are linked to image components by a line or arrow while the image component is highlighted.

EXAMPLE “Inspect the object (x) in Figure 1”, followed by display of a caption (object (x) is a component of...) with a line to connect the caption to the highlighted object illustrated in the figure.

NOTE The text can refer to the object in the image or to the whole image by a name or identifier. Sequential display of text captions with direct-contact points to image components can be used as an effective means of directing the user's viewing sequence through a complex image. Displaying several text captions and image links concurrently can cause confusion.

8.6.6.3 Text to moving image

Subtitles are used to explain components in the moving image with a freeze frame to make the destination component clear, if it is hard to perceive or transient.

EXAMPLE “As can be seen in above clip, first the cover is removed then the fuser assembly is taken out...” The movie is run until the cover is removed and then paused.

NOTE A text caption can be used to explain a scene either before the movie clip, as in silent movie storyboards for a priming effect, or shown concurrently as overlays.

8.6.6.4 Text to speech

Text segments are highlighted while the speech segment is emphasized by voice change or amplitude.

EXAMPLE Display “Listen to the speaker's voice in the following extract”, play speech segment in a foreign language.

NOTE Narrative text can be used to explain aspects of speech such as dialect or intonation. Concurrent presentation can be useful for associating spoken words with their written equivalent, although sequential presentation is advised for most applications.

8.6.6.5 Text to audio (realistic and non-realistic)

The text is highlighted and refers explicitly to the subject matter of the following sound.

EXAMPLE Text display “Listen to the following recording of bird song and identify the bird”; play sound clip.

NOTE Concurrent presentation will impair information delivery because reading the text and listening to the sound will compete for attention. If information extraction from the sound is not vital then concurrent presentation can be used for aesthetic purposes, e.g. playing background music while a text is displayed.

8.6.7 Source medium: moving image

8.6.7.1 General

Directing attention from a moving image is difficult because the information content changes rapidly and effects have to be displayed for some time. As a moving image automatically focuses attention, it is rarely necessary to use explicit attention effects to direct attention from within a moving image. Instead, sequencing can be used to switch attention at the end of a movie. However, speech can be used concurrently to focus attention within a moving image. Freeze frame/pause, zoom and highlighting components that appear throughout a moving image clip can be used to draw attention from a point in a movie to another medium. The following design effects should be considered when implementing contact points from moving images (realistic and non-realistic) to other media.

8.6.7.2 Moving image to speech and realistic/non-realistic audio

Objects in the moving images (animated diagrams, cartoons) are highlighted while being referred to in the speech/sound track with a tone/prosody or amplitude emphasis.

EXAMPLE The video is played showing the heated sulfur turning brown. Speech “as can be seen in the video the first part of the chemical reaction turns sulfur dark brown”.

NOTE Similarly, sound can be linked with animated objects.

8.6.7.3 Moving image to still image/text

Objects in the moving images (animated diagrams, cartoons) are highlighted, or the movie paused while being referred to in the text/image by highlighting.

EXAMPLE A dance movie is shown, then paused. A still image of the dancer is used with text captions to explain details of the position of arms and legs for the movement.

NOTE Highlighting of objects in both images can be used but this is not advised as the user’s attention will be consumed by the moving image, so concurrently displayed information is unlikely to be comprehended effectively.

Annex A
(informative)

Decision trees for classification of types

A.1 Classifying information components and media resources

The definitions in this part of ISO 14915 are grouped into information types that define the amodal requirements for an application, and media types to describe possible representations of content. The motivation for classifying content into amodal information types is to describe the type of information that needs to be represented clearly before selecting media. The information types are applied to components in the application content. The granularity of components will depend on the task and communication goals, so for detailed instructions small components may be necessary, while for a general introduction large components are probably more appropriate. Information types describe the logical arguments that are to be conveyed, and are based on functional theories of language and ontologies in knowledge engineering. For the purpose of this part of ISO 14915, the information types from those sources have been simplified. Components are classified by 'walking through' the table using the definitions and the following questions (see Figure A1):

— Is the information contained in the component physical or conceptual?

— Is the information static or dynamic, i.e. does it relate to change or not?

— Which type in the terminal branch of the tree does the information component belong to?

The first two questions guide the user towards subsets of the types, while the third question identifies the individual type. It is important to note that one component may be classified with more than one type; for instance, instructions on how to get to the railway station may contain procedural information (the instructions <turn left, straight ahead, etc.>), and spatial or descriptive information (the station is in the corner of the square, painted blue). The information types are "tools for thought" which can either be used to classify specifications of content or be used to consider what content may be necessary. To illustrate this for the task of navigating to the railway station the content may be minimally specified as "instructions how to get there", in which case the information types function as questions in the form "what sort of information does the user need to fulfil the task/user goal?". Alternatively, the content may be specified as a scenario narrative of directions, waymarks to recognize and description of the target. In the latter case, the information types may be used to segment the scenario into components. Components in the required content will usually imply several information types. The granularity of components is a matter for the designer's choice and will depend on the level of detail demanded by the application. To illustrate the analysis:

a) Communication goal: Explain how to assemble a bookshelf from ready-made parts.

b) Information component 1:

   — parts of the bookshelf, sides, back, shelves, connecting screws;

   — mapping to information types;

   — Physical-Static-Descriptive, parts of the bookshelf are tangible, don't change and need to be described;

   — Physical-Static-Spatial, dimensions of the parts, how they are organized;

   — Physical-Static-Relationship type could also be added to describe which parts fit together.

c) Information component 2:

— instructions how to assemble parts;

— mapping to information types;

— physical-dynamic-discrete action;

— physical-dynamic-procedure;

— physical-static-state to show final assembled bookshelf.

The information types facilitate mappings to types of media resource.

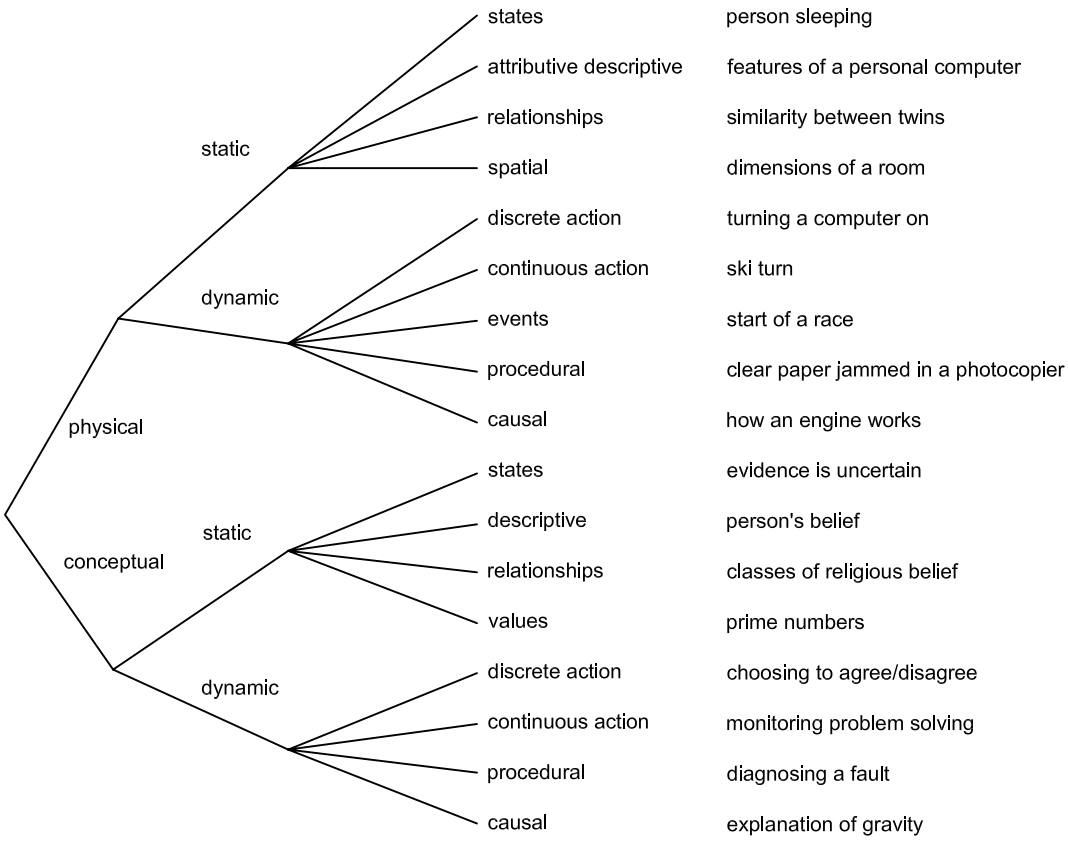


Figure A.1 — Decision tree for classifying information types

A.2 Classifying media resources

Media resources are classified to create a logical description of their perceived properties to facilitate matching to the information types. The media definitions are combined to describe any specific medium, so speech is classified as an audio, language-based medium, while a cartoon is classified as a non-realistic (designed) moving image. The definitions may be usefully considered in two dimensions of abstraction, illustrated in Figure A.2: the influence of design involved in creating the medium, from realistic to non-realistic; and rate of change from static to dynamic. The third aspect of media definitions is to consider the modality, which is divided into audio and visual modalities and than subdivided into language based or non-language based. These aspects form a faceted classification scheme rather than strictly orthogonal categories. Hence any example of a media resource can be described as a tuple of <realism, rate of change and modality>.

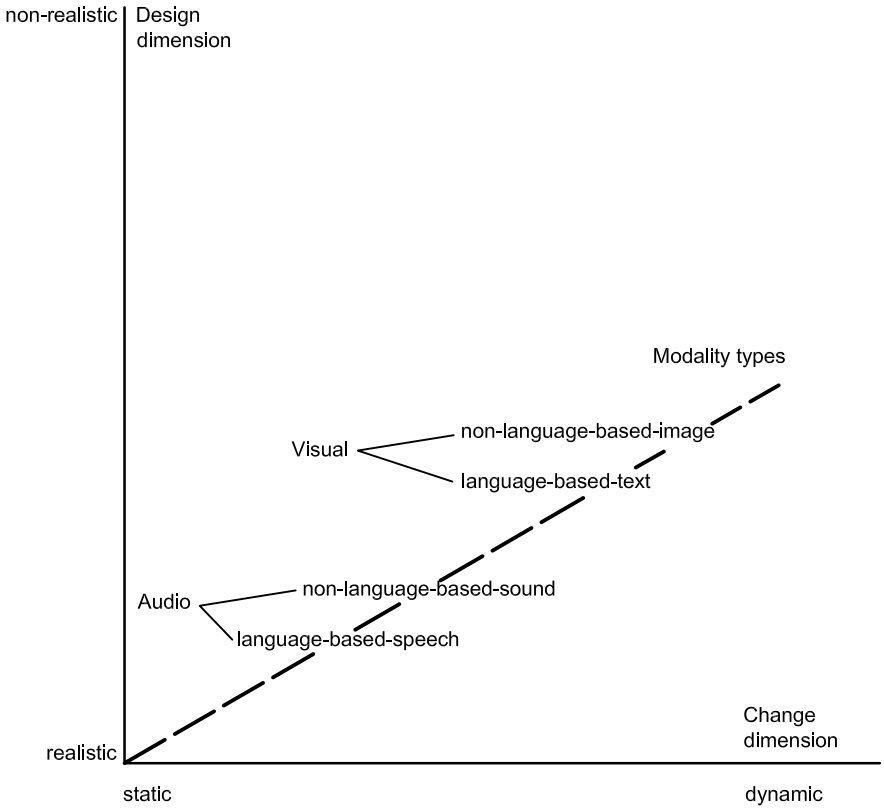


Figure A.2 — Dimensions of media-resource classification

The approach to classifying media uses a walkthrough of the decision tree (see Figure A.3) with the following questions that reflect the facets of the classification.

— Is the medium perceived to be realistic or not? Media resources captured directly from the real world will usually be realistic, e.g. photographs of landscapes, bird-song sound recordings. However, the category boundary case that illustrates the dimensionality is a realistic painting of a landscape.

— Does the medium change over time or not? The boundary case here is the rate of change, particularly in animations where some people might judge 10 frames per second to still be a video but 5 slides in 1 minute shown by a Powerpoint presentation to be a sequence of static images.

— Which modality does the resource belong to? In this case, the categories are orthogonal, although one resource may exhibit two modalities, e.g. a film with a sound track communicates in both visual and audio modalities.

The classification may be used from different viewpoints; for example, a film of a professor illustrating a diagram may be classified as a realistic, moving image medium portraying a non-realistic, still image medium. This part of ISO 14915 does not address design issues of nested media, so in this example only the moving image (film) would be considered. Classification of media resources facilitates mapping of information types to media resources; however, the process of selection may also guide the acquisition or creation of appropriate media resources. So if the selection process indicates the need for a resource that is not in the designer's media resource library, then use of this part of ISO 14915 indicates the need for acquisition/creation. Cost trade-offs will naturally be considered in this process. Finally, the classification provides a mechanism for indexing media-resource libraries.

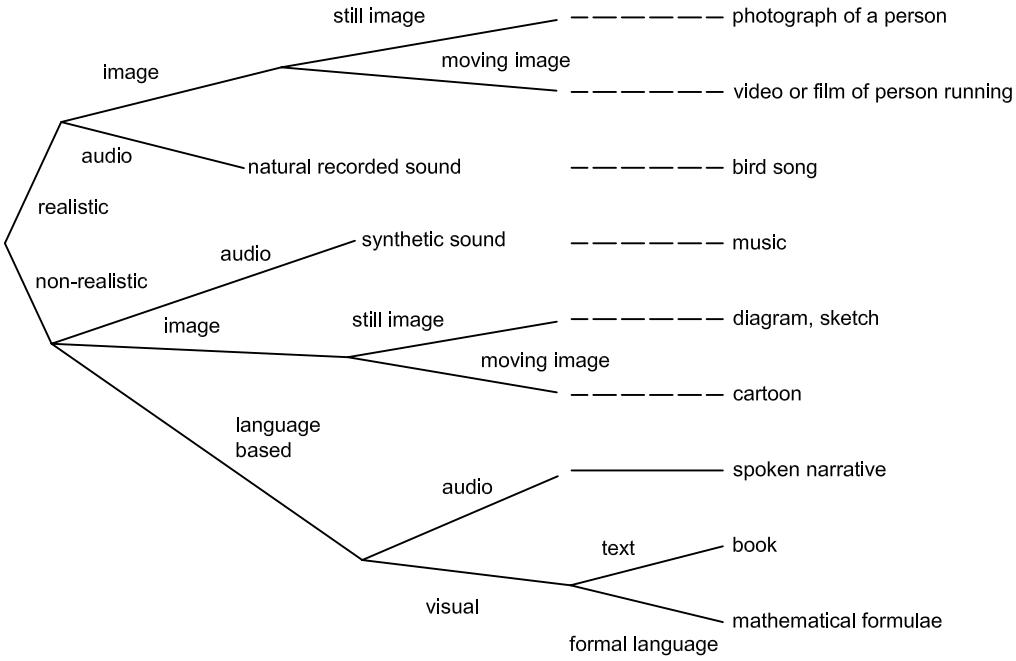


Figure A.3 — Decision tree for classifying media types

Following the example in Figure A.3, the mapping of information types to media types would be as follows:

a) Information component 1:

- parts of the bookshelf, sides, back, shelves, connecting screws;
- mapping to information types to media types;
- Physical-Static-Descriptive.....→ Realistic-Image-Still Image;
- Physical-Static-Spatial.....→ Realistic-Image-Still Image;
- Conceptual-Static-Relationship.....→ text captions, diagram.

b) Information component 2:

- instructions how to assemble parts;
- mapping to information types to media types;
- physical-dynamic-discrete action.....→ Realistic-Image-Still Image (series of snap shots for each action);
- physical-dynamic-procedure.....→ text, still image plus speech, animation to integrate still-image sequence, text summary;
- physical-static-state to show final assembled bookshelf.....→ still image, text caption.

Annex B
(informative)

Guidelines for media-combination pairs

B.1 General

Specific media combinations are given with guidance on the appropriate and inappropriate conditions of use. These recommendations should be interpreted in a specific task/application context and can be ignored if they are not compatible with that context.

B.2 Realistic audio and realistic audio

Effect: this combination will generally be aesthetic or illustrative, although stereo can be used to make one of the sounds stand out to draw attention or to provide a direct reference. Sounds with different frequency ranges (pitch) will complement each other more effectively than sounds with similar frequency ranges.

The amplitude of each source should be similar to avoid one source being difficult to perceive.

EXAMPLE Combining the sounds of an aeroplane and bird song.

Typically appropriate: where the stimulus can be missed without suffering a degradation in task performance; for aesthetic purposes; entertainment; scene-setting.

Inappropriate: to portray specific information; to provide information that later needs to be recalled.

B.3 Realistic audio and non-realistic audio

Effect: this combination overlays natural sound with specifically non-realistic auditory information. Generally, the realistic audio will be continuous, and the non-realistic audio will be discrete. When this combination is used, the non-realistic audio should be dissimilar in content in order to distinguish it from the realistic audio and the distinction between the sources should be made clear before presentation.

EXAMPLE Tones overlaying recordings of bird song to signal a change in the species of bird which is singing.

Typically appropriate: to draw attention to some aspect of the realistic audio source in a language-independent manner; to provide language-independent warning of change in status, e.g. a beep communicates a “low battery'” message on a mobile cassette recorder.

Inappropriate: to provide information that later needs to be recalled; where realistic audio is redundant.

B.4 Realistic audio and speech

Effect: this combination is used to overlay background sound with speech. The information contained in the speech should be consistent with the background sound.

EXAMPLES Voice-over; commentary.

Typically appropriate: to describe or draw attention to some aspect of the realistic audio source; to provide explicit indication of a change of status; to mark the beginning or end of the realistic audio presentation.

Inappropriate: where realistic audio is redundant; where the speech information needs to be recalled later without the realistic audio.

B.5 Realistic audio and realistic still image

Effect: natural sounds can be used to illustrate and add information to a photographic image.

EXAMPLE Whale-song sound is played with a photograph of whales.

Typically appropriate: to add information to the photograph, especially sounds made by objects or agents in the image.

Inappropriate: when sound is incompatible with, or distracts the user from, the image.

B.6 Realistic audio and non-realistic still image

Effect: designed tones and music can be used to convey meaning during interaction or to illustrate components of the image.

EXAMPLE Musical tones are used to express correct or incorrect answers when a student points to parts on an anatomy diagram in response to a question “where is the liver?”

Typically appropriate: to add audio information that describes a diagram/sketch component.

Inappropriate: when realistic audio can be difficult to hear in noisy environments, or doesn’t add any extra information.

B.7 Realistic audio and text

Effect: natural sound can be used to make a text more interesting; sounds made by objects or agents referred to in the text can be played to illustrate them. Concurrent presentation can make cross-referencing between text and audio difficult, so it is preferable to direct the user from within the text, then to play the sound.

EXAMPLE Text describes animal sounds with instructions to play the tape-recorded sounds.

Typically appropriate: when image media degrade, and when objects or agents that make sound are best described in text, e.g. abstract objects illustrated by analogy with natural sounds.

B.8 Realistic audio and realistic moving image

Effect: this combination is used to integrate or overlay audio on video. In some cases, the audio will have been recorded with the video.

EXAMPLES Movie with a sound track of natural sounds; nature film with animal sounds.

Typically appropriate: for presenting real-world/“natural” events; for presenting dynamic information; where information conveyed has both auditory and visual components; where the content needs to be engaging or entertaining.

Inappropriate: where the information is inherently static; where the user needs continuous access to all the information; where the task environment precludes audio.

NOTE If it is necessary to degrade the quality of one of the media, either during the design or when the media are played, it is generally better to degrade video than audio (e.g. reduce frame rate rather than remove sections of the audio).

B.9 Realistic audio and non-realistic moving image

Effect: sound enhances a cartoon or animated diagram.

EXAMPLES Sound track of a dog's bark with movie showing a cartoon dog; diagram of a poorly tuned motor operating with engine sounds and a back-fire; schematic diagram of a heart beating with recordings of heartbeats.

Typically appropriate: natural sounds are used to illustrate designed animations or diagrams. Useful when there is a good association between the sound and the objects/agents showing in the image.

Inappropriate: when the binding between the sound and the portrayed agents/objects is poor.

B.10 Non-realistic audio and non-realistic audio

Effect: one sound is used to refer to or augment the other.

EXAMPLES Two or more scales played in stereo for 2D orientation; warning sound with audio-pitch feedback in weighing scales for the blind.

Typically appropriate: the audio streams should complement each other and not interfere; one audio medium is short and can be overlaid on the second; to signal a point or event in a continuous audio sequence.

Inappropriate: if the audio streams interfere, or the signals are not compatible.

B.11 Non-realistic audio and speech

Effect: information is presented by both media sources.

EXAMPLE Tones between voice-mail messages.

Typically appropriate: to describe or draw attention to some aspect of the language-based audio source; to provide explicit indication of a change of status; to mark the beginning or end of the language-based audio presentation.

Inappropriate: when the audio media interfere with each other; when one audio channel masks the other channel, e.g. amplitude too loud to complement other channel.

NOTE This combination will usually be played sequentially and for short durations because interruptions by the non-realistic audio will interfere with the user's ability to comprehend the speech.

B.12 Non-realistic audio and realistic still image

Effect: music is used to enhance the aesthetic impression of the image; tones refer to the image or components therein (earcons).

EXAMPLES Warning noise played when danger zones are selected on a photographic image of a nuclear power plant; music played to enhance aesthetic appeal of holiday beach photograph.

Typically appropriate: when audio enhances description of, or can draw attention to, an image component; earcons played when part of an image is pointed to; useful when the audio medium can be easily interpreted and linked to the image; also to enhance aesthetic qualities of images.

Inappropriate: when the meaning of the audio channel cannot be easily interpreted or associated with the image.

NOTE Music can either have a detrimental effect, or no effect, in improving learning when played with images or other media.

B.13 Non-realistic audio and non-realistic still image

Effect: music is used to enhance the aesthetic impression of the image; tones refer to the image or components therein (earcons).

EXAMPLES Warning tones played when dangerous locations on a pollution map are pointed to; music complements a design sketch for a new car.

Typically appropriate: when audio enhances description of, or can draw attention to, an image component; earcons played when part of an image is pointed to; useful when the audio medium can be easily interpreted and linked to the image; also to enhance aesthetic qualities of images.

Inappropriate: when the meaning of the audio channel cannot be easily interpreted or associated with the image.

B.14 Non-realistic audio and text

Effect: music is used to enhance the aesthetic impression of a text; tones refer to the text or sentences therein (earcons).

EXAMPLES Tones for interactive selection of words, sentences and paragraphs; earcons played for feedback when answering a multiple-choice quiz.

Typically appropriate: when audio does not interfere with the text-reading sequence, short audio clips or low-amplitude tones can be useful for associating a meaning with a text, for marking items, interactive editing.

Inappropriate: when audio interferes with text reading (too loud, continuous), or audio cannot be easily associated with text components.

B.15 Non-realistic audio and realistic moving image

Effect: music is used to enhance the aesthetic impression of the movie; tones refer to the image or components therein (earcons).

EXAMPLES Music sound track on a film, tone markers for events or scene boundaries on an instructional movie (tone alerts user); tone scales used to augment action of movie, e.g. a descending scale is played as a whale dives deeper.

Typically appropriate: if sound enhances understanding of, and can be easily associated with, objects or agents in the moving image.

Inappropriate: when audio interferes with the image interpretation or is dissonant with it.

B.16 Non-realistic audio and non-realistic moving image

Effect: music is used to enhance the aesthetic impression of a cartoon or animated diagram; tones refer to diagram components (earcons) to draw attention or provide supplementary information.

EXAMPLES Music for cartoon sequence; tones or earcons on event, scene boundaries; warning tones in a schematic physics experiment point out that, when the voltage is increased, the resistor will burn out.

Typically appropriate: when the audio medium enhances the whole animation; can be associated with and augment understanding of agents or objects in the animation.

Inappropriate: when audio interferes with the image interpretation or is dissonant with it.

B.17 Speech and speech

Effect: apart from natural conversation, which can be considered as a single speech track, it is difficult to listen to two speech tracks concurrently, so it is advisable to avoid this combination.

EXAMPLES Discussion between two parties; speech against background conversation.

Typically appropriate: for emphasis (synchronous); to present dialogue (sequential).

Inappropriate: to provide instructional information, when the content is to be easily integrated.

NOTE People can separate several concurrent speech tracks, but can only extract information from one at a time (the cocktail-party effect).

B.18 Speech and realistic still image

Effect: speech augments information in the image or draws attention to components therein.

EXAMPLES Voice description of a photograph in a museum; voice-over of a photograph of the city in a tourist information kiosk.

Typically appropriate: if the image components need to be described, pointed out, or the whole image needs to be explained. A useful alternative to text that could interfere with viewing the image, or when the additional information does not need to be persistent, or for interactive use when a large amount of supplementary detail is available for a single image.

Inappropriate: when the supplementary information needs to be persistent, or is important.

B.19 Speech and non-realistic still image

Effect: speech augments information in the diagram or draws attention to components therein.

EXAMPLES Voice commentary on a diagram; speech explanation of a schematic diagram of a jet engine.

Typically appropriate: when image components need to be described, pointed out, or the whole image needs to be explained. Use when text would interfere with viewing the image, when the additional information does not need to be persistent, or for interactive use when a large amount of supplementary detail is available for a single image.

Inappropriate: when the supplementary information needs to be persistent; when diagrams are hard to interpret.

B.20 Speech and text

Effect: This combination emphasizes components of speech with text or vice versa, using redundant or complementary presentation.

EXAMPLE Text is displayed and is read by an actor.

Typically appropriate: when the combination adds information; to convey pronunciation or prosody in the written text.

Inappropriate: when the content of the media are unrelated; where the media cannot be synchronized.

B.21 Speech and realistic moving image

Effect: speech augments information in the movie and draws attention to components therein.

EXAMPLE Movie clip of person speaking.

Typically appropriate: for instructional videos; to explain the moving image (narration).

Inappropriate: when a video of a person speaking cannot be synchronized with the speech.

B.22 Speech and non-realistic moving image

Speech augments information in the cartoon and draws attention to components therein.

EXAMPLES Speech sound track on a cartoon; speech explanation of an animated diagram.

Typically appropriate: when spoken dialogue supplements understanding of agents/objects in the moving image; when action sequences, movement or views need to be explained in diagrams and animations.

Inappropriate: if the information is important and needs to be persistent, when speech explanation of action or movement is cumbersome.

B.23 Realistic still image and realistic still image

Effect: for comparison of two images; to express sequences and temporal ordering with several images. Different sizes may portray relative importance. For a combination of different images, the layout or sequence of presentation is significant.

EXAMPLES Comparison of two scenes taken as stills from a movie; comparing photographs before and after an accident.

Typically appropriate: to represent sequences in place of video; for comparison of images.

Inappropriate: where the task involves recall of specific attributes of one of the images.

B.24 Realistic still image and non-realistic still image

Effect: the non-realistic image explains features of the realistic image or draws attention to abstract information from the realistic image.

EXAMPLES Photograph of an engine accompanied by a diagram of the engine.

Typically appropriate: to draw attention to specific attributes of a realistic image in a language-independent manner; to provide language-independent information to supplement the realistic image; to explain concrete concepts associated with the realistic image.

Inappropriate: where the images are unrelated.

B.25 Realistic still image and text

Effect: the image is explained by the text or the text supplements information in the image.

EXAMPLES Text description of a scene, accompanied by a photograph; a caption points out a camouflaged bird hiding in a photograph of a landscape.

Typically appropriate: to draw attention to specific attributes of a realistic image; to provide information to supplement the realistic image; to explain abstract concepts associated with the image.

Inappropriate: where the images and text are unrelated.

B.26 Realistic still image and realistic moving image

Effect: to augment information in the movie and refer to an agent/object in the movie.

EXAMPLES Photograph overlaid as a mini-window insert on a film or video; picture of the author/director of the movie, picture of a character in the film, illustration of where the shots were taken from.

Typically appropriate: to provide additional static information to complement action, movements or scene in the movie; to provide commentary (image and speech) on the movie; to provide more detail on an agent or object in the movie by an enlarged still image.

Inappropriate: when subject matter in the photograph doesn't match the movie; if the realistic image occludes most of the movie.

B.27 Realistic still image and non-realistic moving image

Effect: to augment information in a cartoon and provide detail on a component/object in an animated diagram.

EXAMPLES Photograph of agent/object in a cartoon overlaid on the animation; photograph of the author of the animation; photograph of bird with diagram of flight movement.

Typically appropriate: to provide additional information or a different view on agents/objects appearing in the animation, to enlarge an object or scene in the movie for inspection of detail.

Inappropriate: when subject matter in the photograph doesn't match the content of the animation/cartoon.

B.28 Non-realistic still image and non-realistic still image

Effect: to compare two cartoons or animations; to show two views on the same scene of movement.

EXAMPLES Technical drawings showing two views of the same object.

Typically appropriate: to provide a simple abstraction of features from a more complex image; to present different views of the same object.

Inappropriate: where the images are unrelated and complex.

B.29 Non-realistic still image and text

Effect: the diagram illustrates the text description. Generally, the content of one medium is used to illustrate the other, so the information presented through at least one of the media will be simpler than the other.

EXAMPLES Diagram of a pump is described by a text explaining the pump’s operation while text captions point out the parts of the pump.

Typically appropriate: to illustrate simple concepts; to provide explanatory information to supplement the non-realistic image; to emphasize concepts.

Inappropriate: where the information presented through each medium contains different complex concepts; where the image and text are unrelated.

B.30 Non-realistic still image and realistic moving image

Effect: the diagram explains or summarizes abstract information in or about the movie; a diagram showing relationships between agent and objects in a scene.

EXAMPLES Diagram explaining an action sequence in the movie, such as a diagram overlay on video of a football match to identify offside zones, and possible future moves.

Typically appropriate: to provide additional, abstract, static information relevant to the structure of the movie or the agents/objects therein; to explain relationships between agents/objects in the movie.

Inappropriate: diagram overlays should not be used if motion is obscured; separate diagrams should not be used if mapping between the diagram components and moving image components are difficult to comprehend, for instance when movement makes relationships hard to trace.

B.31 Non-realistic still image and non-realistic moving image

Effect: diagram explains or summarizes information in or about the animation or objects therein.

EXAMPLES Diagram illustrating structure of scene in a cartoon clip; diagram showing properties of the planets in an animation of the solar system with movement of the planets.

Typically appropriate: to provide additional, abstract, static information relevant to the structure of the animation or the agents/objects therein; to explain relationships between agents/objects in the animation; to provide different viewpoints on action and motion.

Inappropriate: when diagram overlays will obscure motion; if mapping between the diagram components and moving image components are difficult to comprehend, for instance when movement makes relationships hard to trace.

B.32 Text and text

Effect: two texts are compared or one text supplements information in the other.

EXAMPLES Two texts are juxtaposed for comparison, for example the original and a translation, or the same text in two different fonts; text overlays as captions, footnotes, or expansion boxes in hypertext to explain the background text, for instance, a tutorial on a medieval manuscript.

Typically appropriate: to compare two texts with different properties; to provide supplementary information on the primary text.

Inappropriate: when the purpose of the comparison is not clear; when text captions obscure important components of the primary text.

B.33 Text and realistic moving image

Effect: the text summarizes, describes or supplements information in the movie.

EXAMPLES Subtitles on movies; text “speech bubbles” overlaid on movies, captions linked to agents, objects or scene in the movie.

Typically appropriate: if speech is not available; when ambient noise may make speech difficult to hear, to deliver persistent, supplementary information on the structure of the movie or on agents and objects therein.

Inappropriate: when information in the text and movie are difficult to relate, when reading and viewing action in the movie will clash.

B.34 Text and non-realistic moving image

Effect: the text summarizes, describes or supplements information in the cartoon.

EXAMPLES Subtitles; text “speech bubbles” overlaid on animations; captions linked to agents, objects or scene in the cartoon, for instance, captions appear to describe the planets in a diagram of the solar system at the user’s request.

Typically appropriate: if speech is not available; when ambient noise makes speech difficult to hear; to deliver persistent, supplementary information on the structure of the animation or on agents and objects therein.

Inappropriate: when information in the text and animation are difficult to relate; when reading and viewing action in the animation will clash.

B.35 Realistic moving image and realistic moving image

Effect: the two movies are compared or enhance each other for aesthetic appeal.

EXAMPLES Two movies are played in separate windows to show different viewpoints on the same scene; two movies are played, one showing a long-shot context, the other a close-up of action detail, for instance a long shot of a football game, the other a close-up of a foul between two players in the long shot.

Typically appropriate: to provide different viewpoints on congruent information; to give context and detail on an action, agent or object, for aesthetic purposes.

Inappropriate: if the two viewpoints don't complement each other; when the user has to attend to both images concurrently to extract useful information.

B.36 Realistic moving image and non-realistic moving image

Effect: the movies and cartoon may be compared or the animation may present an abstract view of information in the movie.

EXAMPLES A diagram illustrates motion, action or relationships between agents or objects in the movie, such as a diagram of a horse galloping with a skeleton view juxtaposed on a film of the horse; animation overlay on a film of a football game to explain players' tactics.

Typically appropriate: to provide abstract, dynamic information about the structure of a movie or the agents, objects and relationships therein; to illustrate an abstract explanation of action or movement with natural examples.

Inappropriate: if the two viewpoints don't complement each other; when the user has to attend to both images concurrently to extract useful information.

B.37 Non-realistic moving image and non-realistic moving image

Effect: the two animations may be compared or enhance each other by showing different viewpoints.

EXAMPLE Two animations show complementary views; for instance, air-flow velocity and temperature in a jet engine.

Typically appropriate: if two viewpoints are required on the action, motion or state change of agents or objects; when two viewpoints on abstract information need to be compared; when the patterns of change in the two images can be easily related and interpreted by the user.

Inappropriate: if the two viewpoints don't complement each other; when the user has to attend to both images concurrently to extract useful information.

Table B.1 — Summary of media combinations with examples

First medium	Second medium	Media combination example
Realistic audio	Realistic audio	Combining the sounds of an aeroplane and bird song
	Non-realistic audio	Music with natural environmental sounds
	Speech	Tones between voice-mail messages
	Realistic still image	Bird song with a photograph of the bird
	Non-realistic still image	Musical tones illustrate diagram parts
	Text	Environmental sounds illustrate a nature-study text
	Realistic moving image	Movie with a sound track
	Non-realistic moving image	Cartoon sound track with environmental sounds
Non-realistic audio	Non-realistic audio	Two tones in weighing scales for the blind
	Speech	Tones with speech for emphasis
	Realistic still image	Music to illustrate a photograph
	Non-realistic still image	Tones to illustrate a music-scale diagram
	Text	Tones for interactive selection of words
	Realistic moving image	Movie clip with music sound track
	Non-realistic moving image	Music and sound effects in a cartoon
Speech	Speech	Speech over a background conversation
	Realistic still image	Voice commentary on a photograph
	Non-realistic still image	Speech explanation of a circuit diagram
	Text	Text which is read and recited by an actor
	Realistic moving image	Film sound track with an actor speaking
	Non-realistic moving image	Cartoon sound track with characters speaking
Realistic still image	Realistic still image	Comparison of two scenes: summer and winter photographs
	Non-realistic still image	Photograph accompanied by a diagram of an engine
	Text	Text description of a scene, accompanied by a photograph
	Realistic moving image	Still image of the actor with the movie he is playing in
	Non-realistic moving image	Photograph of a person with an animated cartoon based on his character
Non-realistic still image	Non-realistic still image	Technical drawings showing two views of the same object
	Text	Diagram described by a text
	Realistic moving image	Flow diagram with steps illustrated in a video
	Non-realistic moving image	Storyboard diagram with animated sequence
Text	Text	Two texts for comparison: Greek and Latin
	Realistic moving image	Subtitles on a silent movie
	Non-realistic moving image	Text captions on an animated flow diagram
Realistic moving image	Realistic moving image	Two movies related by a theme: Abel Gance's triptych of Napoleon
	Non-realistic moving image	Film of a man running with an animated anatomy diagram
Non-realistic moving image	Non-realistic moving image	Two cartoons related by a theme: compare two animation styles

Annex C
(informative)

Examples of media-combination patterns

C.1 General

This annex describes commonly used combinations with guidance on the design issues raised by each pattern. Media can be combined concurrently or sequentially to achieve a communication goal.

C.2 Moving image, sound, speech

Moving image with a voice track and sound is a familiar combination from film. This combination can be used effectively for explaining visual information via the voice channel, with sounds to give additional information about the image.

EXAMPLE A video of a fire-evacuation procedure with speech instructions and the sound of a real fire incident to dramatize the presentation.

C.3 Still image, speech, text

Photographs or diagrams can be augmented with text captions or presented separately with explanatory text. Speech can be used to draw attention to specific parts of the image and reference important parts of the text.

In this combination, important information can be conveyed by text, and speech used as a supplementary commentary.

EXAMPLE In a tourist-information kiosk, a photograph of a landscape is described in speech with text captions to point out important locations of interest.

C.4 Still image, speech, sound

In this combination, the audio channels can clash, so care should be taken that speech and sound are suitably integrated. Sound can be used to convey information related to the image with a voice commentary.

EXAMPLE A photograph of a bird is accompanied by bird song and a speech clip introducing the species.

C.5 Two (or more) still images, speech, text

Images can be compared or linked together in a sequence using text captions and speech to draw the user's attention to important information.

EXAMPLE Two similar flowers are shown and the user's attention is drawn to differences by text captions and a voice commentary.

C.6 Two (or more) texts, speech

Different texts can be compared or linked together using speech to draw the user's attention to important words or phrases. This combination can be used when single words or phrases need to be found; however, only one text can be read at a time, so if the complete text has to be assimilated, this combination should not be used.

EXAMPLE A modern Greek text is displayed alongside an ancient Greek text with the same subject matter, so that differences in the written language can be evaluated.

Annex D
(informative)

Design issues and cognitive background

D.1 General

The purpose of this annex is to give the reader a brief overview of cognitive psychology and its influence on multimedia design to help interpretation of the guidelines.

D.2 Perception and comprehension

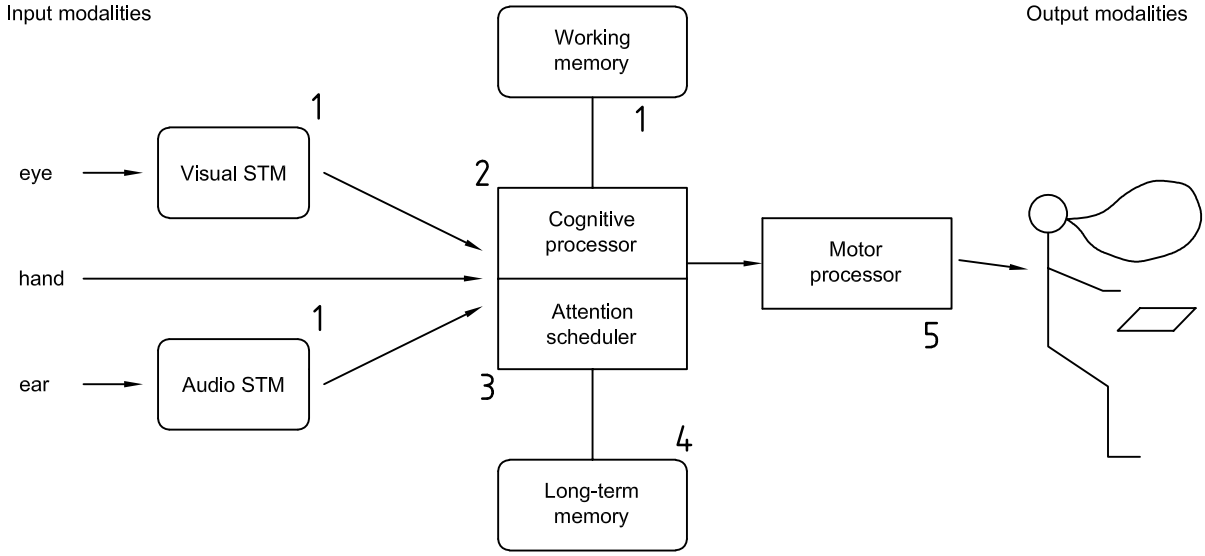
Our eyes scan images in a series of rapid jumps called *saccades*, interleaved with fixations where the eye dwells on a particular area. Fixations allow image detail to be inspected, so eyetracking gives an impression of the detail inspected in images. Generally our eyes are drawn to moving shapes, then complex, different, and colourful objects. Visual comprehension can be summarized as “what you see depends on what you look at and what you know”. Multimedia designers can influence what users look at by controlling attention with display techniques such as use of movement, highlighting and salient icons. However, designers should be aware that the information that people assimilate from an image also depends on their internal motivation, what they want to find and how well they know the domain (Treisman, 1988). A novice will not see interesting plant species in a tropical jungle, whereas a trained botanist will. Selection of visual content therefore has to take the user’s knowledge and task into account. Moreover, because the visual sense receives information continuously, it gets overwritten in the working memory (Baddeley, 1986). This means that memorization of visually transmitted information is not always effective, unless users are given time to view and comprehend images. Furthermore, users only extract very high level or “gist” (general sense) information from moving images. Visual information has to be understood by using memory. In realistic images this process is automatic; however, with non-realistic images we have to think carefully about the meaning, for example of a diagram. While extraction of information from images is rapid, it does vary according to the complexity of the image and how much we know about the domain.

Sound is a transient medium, so unless it is processed quickly, the message can be lost. Even though people are remarkably effective at comprehending spoken language and can interpret other sounds quickly, the audio medium is prone to interference because other sounds can compete with the principal message. Because sound is transient, information in speech will not be assimilated in detail, and so only the “gist” will be memorized (Gardiner and Christie, 1987).

D.3 Selective attention

We can only attend to a limited number of inputs at once. While people are remarkably good at integrating information received by different senses (e.g. watching a film and listening to the sound track), there are limits determined by the psychology of human information processing (Wickens *et al*., 1983). Our attention is selective and closely related to perception; for instance, we can overhear a conversation at a party with many people speaking (the cocktail-party effect). Furthermore, selective attention differs between individuals and can be improved by learning factors: for example, a conductor can distinguish the different instruments in an orchestra, whereas a typical listener cannot. However, all users have cognitive resource limitations which means that information delivered on different modalities (e.g. by vision and sound) has to compete for the same resource. For instance, speech and printed text both require a language-understanding resource, while video and a still image use image-interpretation resources. Cognitive models of information-processing architectures (e.g. Interacting cognitive subsystems: Barnard, 1985) can show that certain media combinations and media design will not result in effective comprehension because they compete for the same cognitive resources, thus creating a processing bottleneck. We have two main perceptual channels for receiving information: vision and hearing; information going into these channels has to be comprehended before it can be used. Information can be received in a language-based form as either speech or as written text; writing can also be viewed with an image or a movie. All such input

competes for language-understanding resources, hence making sense of speech and reading text concurrently is difficult (Barnard, 1985). Figure D.1 shows the cognitive architecture of human information processing and resource limitations that lead to multimedia usability problems.



Bottlenecks

1 Capacity overflow: information overload
2 Integration: common message?
3 Contention: conflicting channels
4 Comprehension
5 Multi-tasking input/output

Figure D.1 — Approximate model of human information processing using a "human as computer system" analogy, based on the Model Human Processor (Card *et al.*, 1983)

Capacity overflow (1 in Figure D.1) may happen when too much information is presented in a short period, swamping the user's limited working memory and cognitive processor's capability to comprehend, chunk, and then memorize or use the information. The connotation is to give users control over the pace of information delivery. Integration problems (2) arise when the message on two media is different, making integration in the working memory difficult; this leads to the thematic congruence principle. Contention problems (3) are caused by conflicting attention between dynamic media, and when two inputs compete for the same cognitive resources, e.g. speech and text require language understanding. Comprehension (4) is related to congruence; we understand the world by making sense of it with our existing long-term memory. Consequently, if multimedia material is unfamiliar we can't make sense of it. Finally, multitasking (5) makes further demands on our cognitive processing, so we will experience difficulty in attending to multimedia input while performing output tasks.

Making a theme in a multimedia presentation clearly involves directing the user's reading and viewing sequence across different media segments. Predicting the user's reading/viewing sequence is difficult. Text and speech are usually processed in sequence; however, viewing image media depends on the size and complexity of the image, the user's knowledge of the contents, as well as the user's task and motivation (Norman and Shallice, 1986). Text also enforces a reading order by the syntactic convention of language; however, what the user looks at in a still image is less predictable. Ultimately, the order of viewing objects in a still image depends on the user's task, motivation, knowledge of the domain and designed effects for salience. Attention-directing design effects can increase the probability that the user will attend to an image component, although no guarantee can be given that a component will be perceived or understood. Cross-referencing establishes a contact point between two media, and how to design such contact points depends on the properties of the source and destination media. Generally, attention control is easier in language-based media, as language provides direct commands.

D.4 Learning and memorization

Learning is the prime objective in tutorial multimedia. However, the type of learning can be either skill training, in which case carrying out an operational task efficiently and without errors is the aim, or conceptual learning, when a deeper understanding of the knowledge is required. In both cases, the main objective is to create a rich memory schema which can be accessed easily in the future. We learn more effectively by active problem solving or learning by doing. This approach is at the heart of the constructivist learning theory (Papert, 1980), which has connotations for tutorial multimedia. Interactive microworlds where users learn by interacting with simulations, or constructing and testing the simulation, give a more vivid experience that forms better memories (Rogers *et al.*, 1998). Multiple viewpoints help to develop rich schemata by presenting different aspects of the same problem, so the whole concept can be integrated from its parts. An example might be to explain the structure of an engine, then how it operates and finally a causal model of why it works. The schema integration then fits the separate viewpoints together.

D.5 Users' tasks

Recommendations for selecting and influencing the user's attention have to be interpreted according to the users' task and design goal. If information provision is the main design goal, e.g. a tourist-kiosk information system, then persistence of information and drawing attention to specific items is not necessarily as critical as in tutorial applications. Dynamic media are not generally efficient for conveying detailed information as we only remember the "gist", or outline, of what was presented (Dwyer, 1967). However, if the design goal is simply to inform the user about a high-level plan of action then a movie and speech can suffice.

Bibliography

[1] ISO 9241-10, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 10: Dialogue principles*

[2] ISO 9241-11, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability*

[3] ISO 9241-12, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 12: Presentation of information*

[4] ISO 13407, *Human-centred design processes for interactive systems*

[5] ALTY, J.L. (1991). Multimedia: what is it and how do we exploit it? In Diaper, D. & Hammond N.V. (Eds.) *People and Computers VI, Proceedings of HCI'91*. Cambridge University Press, pp. 31-41

[6] ALTY, J.L. (1997). Multimedia. In Tucker A.B. (Ed.) *Computer science and engineering handbook*. CRC Press, pp. 1551-1570

[7] ALTY, J.L. and BERGAN, M. (1992). The design of multimedia interfaces for process control. In *Proceedings of 5th IFAC IFORS IFIP IEA Conference on Man Machine Systems*, The Hague

[8] ALTY, J.L., BERGAN, M., DOLPHIN, C. and SCHEPENS, A. Multimedia and process control: some initial experimental results. *Computers & Graphics*, **17**(3), pp. 205-218

[9] ANDRÉ, E. and RIST, T. (1993). The design of illustrated documents as a planning task. In Maybury, M. (Ed.) *Intelligent multimedia interfaces*, AAAI/MIT Press, pp. 94-166

[10] ARENS, Y., HOVY, E. and VOSSERS, M. (1993). On the knowledge underlying multimedia presentations. In Maybury M. (Ed.) *Intelligent multimedia interfaces*, AAAI/MIT Press, pp. 280-306

[11] BABER, C., STANTON, N.A. and STOCKLEY, A. (1992). Can speech be used for alarm displays in process control type tasks? *Behaviour and Information Technology*, **11**(4), pp. 216-226

[12] BADDELEY, A.D. (1986). *Working memory*. Oxford University Press

[13] BAGGETT, P. (1989). Understanding visual and verbal messages. In Mandl, H. & Levin, J.R. (Eds.) *Knowledge acquisition from text and pictures*, Elsevier Science, pp. 101-124

[14] BAGGETT, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations, *Journal of Educational Psychology*, **76**(3), pp. 408-417

[15] BARNARD, P. (1985). Interacting cognitive subsystems: a psycholinguistic approach to short term memory. In Ellis, A. (Ed.) *Progress in psychology of language*, **2**, pp. 197-258, LEA

[16] BARTRAM, D.J. (1980). Comprehending spatial information: the relative efficiency of different methods of presenting information about bus routes. *Journal of Applied Psychology*, **65**, pp. 103-110

[17] BERNSEN, N.O. (1994). Foundations of multimodal representations: a taxonomy of representational modalities. *Interacting with Computers*, **6**(4), pp. 347-371

[18] BIEGER, G.R. and GLOCK, M.D. (1984). The information content of picture-text instructions. *Journal of Experimental Education*, **53**, pp. 68-76

[19] BOOHER, H.R. (1975). Relative comprehensibility of pictorial information and printed word in proceduralized instructions. *Human Factors*, **17**(3), pp. 266-277

[20] BROADBENT, D.E. (1958). *Perception and communication*. Pergamon Press

[21] CARD, S.K., MORAN, T.P. and NEWELL, A. (1983). *The psychology of human computer interaction*. Lawrence Erlbaum Associates

[22] CARPENTER, P.A. and JUST, M.A. (1992). *Understanding mechanical systems through computer animation. Report No. ONR 92-1*. U.S. Office of Naval Research

[23] DAVIS, B. and MARKS, L. (1993). *Integrative multimedia design*. Tutorial presented at INTERCHI '93, Amsterdam

[24] DINNEL, D. and GLOVER, J. (1988). Advanced organizers: encoding manipulations. *Journal of Educational Psychology*, **77**(5), pp. 514-521

[25] DWYER, F.M. (1967). The effect of varying the amount of realistic detail in visual illustrations. *Journal of Experimental Education*, 36, pp. 34-42

[26] D'YDEWALLE, G. and GIELEN, I. (1993). Attention allocation with overlapping sound, image and text. In Rayner, K. (Ed.) *Eye movements and visual cognition*. Springer Verlag, pp. 415-427

[27] FARADAY, P.M. and SUTCLIFFE, A.G. (1993). A method for multimedia interface design. In Alty, J.L. & Diaper, D. (Eds.) *People & Computers VIII*. Cambridge University Press, pp. 173-190

[28] FARADAY, P.M. and SUTCLIFFE, A.G. (1997a). Designing effective multimedia presentations. In *Proceedings of CHI 97*, ACM, pp. 272-279

[29] FARADAY, P.M. and SUTCLIFFE, A.G. (1997b). Multimedia: design for the moment. In *Proceedings of Multimedia '97*, ACM, pp. 183-193

[30] FARADAY, P.M. and SUTCLIFFE, A.G. (1998). Providing advice for multimedia designers. In *Proceedings of CHI '98*, ACM, pp. 124-131

[31] FRIEDMAN, A. (1979). Framing pictures: role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology*: General, **108**(3), pp. 316-355

[32] GARDINER, M. and CHRISTIE, B. (1987). *Applying cognitive psychology to user-interface design*. John Wiley

[33] GERNSBACHER, M.A. (1985). Surface information loss in comprehension. *Cognitive Psychology*, **17**, pp. 324-363

[34] GLYNN, S., BRITTON, B. and TILLMAN, M. (1985). Typographic cues in text: management of the reader's attention. In Jonassen, D. (Ed.) *Technology of text*. Educational Technology Publishing

[35] HANNFIN M.J. and PHILIPS, R. (1987). Perspectives in the design of interactive video: beyond tape versus disc. *Journal of Research and Development in Education*, **21**, pp. 44-60

[36] HEGARTY, M. and JUST, M.A. (1993). Constructing mental models of text and diagrams. *Journal of Memory and Language*, **32**, pp. 717-742

[37] HELLER, R.S. and MARTIN, C. (1995). A media taxonomy. *IEEE Multimedia*, Winter 1995, pp. 36-45

[38] HOCHBERG, J. (1986). Presentation of motion and space in video and cinematic displays. In Boff, K.R., Kaufman, L. & Thomaas, J.P. (Eds.) *Handbook of perception and human performance 1: sensory processes and perception*. John Wiley

[39] HOOGEVEN M. (1997). Towards a theory of the effectiveness of multimedia systems. *International Journal of Human-Computer Interaction*, **9**(2), pp. 151-168

[40] JOHNSON, P. and NEMETZ, F. (1998). Developing multimedia principles from design features. In *Designing effective and usable multimedia systems, Proceedings of IFIP 13.2 Conference*, North Holland

[41] JOHNSTON, W.A. and DARK, V.J. (1986). Selective attention. *Annual Review of Psychology*, **37**, pp. 43-75

[42] JORG, S. and HORMANN, H. (1978). The influence of general and specific labels on the recognition of labelled and unlabelled parts of pictures. *Journal of Verbal Learning and Verbal Behaviour*, **17**, pp. 445-454

[43] KOZMA, R.B. (1991). Learning with media. *Review of Educational Research*, **61**(2), pp. 179-211

[44] KITAJIMA, M. and POLSON P.G. (1998). A comprehension based model of exploration. *Human-Computer Interaction*, **12**(4), pp. 345-390

[45] KRAFT, R. (1986). The role of cutting in the evaluation and attention of film. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **12**(1), pp. 155-163

[46] LARGE, A., BEHESHTI, J. and BRELEUX, A. (1996). Effect of animation in enhancing descriptive and procedural texts in a multimedia learning environment. *Journal of American Society for Information Science*, **47**(6), pp. 437-448

[47] LEVIE, W.H. and LENTZ, R. (1982). Effects of text illustrations: a review of research. *Educational Computing and Technology Journal*, **30**(4), pp. 195-232

[48] MANDLER, J. and RITCHEY, G. (1977). Long term memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, **3**(4), pp. 386-396

[49] MARCEL, T. and BARNARD, P. (1979). Paragraphs of pictographs: the use of non-verbal instructions for equipment. In Kolers, P., Wrolstaid, M. & Bouma, H. (Eds.) *Processing of visible language*. Plenum

[50] MAY, J. and BARNARD, P. (1995). Cinematography and interface design. In Nordbyn, K., Helmersen, P.H., Gilmore, D.J. & Arnesen, S.A. (Eds.) *Proceedings of Fifth IFIP TC 13 International Conference on Human-Computer Interaction*. Chapman & Hall, pp. 26-31

[51] MAYER, R.E. and ANDERSON, R.B. (1991). Animation needs narration: an experimental test of a dual coding hypothesis. *Journal of Educational Psychology*, **83**(4), pp. 484-490

[52] MAYER, R.E. and ANDERSON, R.B. (1992). The instructive animation: helping students build connections between words and pictures in multimedia. *Journal of Education Psychology*, **84**(4), pp. 444-452

[53] MOUNTFORD, S.J. and GAVER, W. (1989). Talking and listening to computers. In Laurel, B. (Ed.) *The art of human-computer interface design*. Addison Wesley, pp. 319-334

[54] NARAYANAN, N.H. and HEGARTY, M. (1998). On designing comprehensible interactive hypermedia manuals. *International Journal of Human-Computer Studies*, **48**, pp. 267-301

[55] NAVON, D. and MILLER J. (1987). The role of outcome conflict in dual task interference. *Journal of Experimental Psychology: Human Perception and Performance*, **13**(3), pp. 435-448

[56] NICKERSON, R. (1980). *Speech understanding and reading: some differences and similarities*. In Spiro, R., Bruce, B. & Brewer, W (Eds.) Theoretical issues in reading comprehension. Lawrence Erlbaum Associates

[57] NORMAN, D.A. and SHALLICE, T. (1986). Attention to action: willed and automatic control of behaviour. In Davidson, G.E. & Schwartz, G.E. (Eds.) *Consciousness and self regulation,* v4, Plenum Press, pp. 1-18

[58] PLAMITER, S., ELKERTON, J. and BAGGETT, P. (1991). Animated instruction versus procedural instructions for learning procedural tasks. *International Journal of Man-Machine Systems*, **34**, pp. 687-701

[59] PARK, I. and HANNAFIN, J. (1994). Empirically-based guidelines for the design of interactive multimedia. *Education Technology Research and Development*, **41**(3), pp. 63-85

[60] PARK, O. and HOPKINS, R. (1993). Instructional conditions for using dynamic visual displays: a review. *Instructional Science*, **21**, pp. 427-449

[61] PATTERSON, R.D. (1982). *Guidelines for auditory warning systems on civil aircraft*. CAA Paper 82017. Civil Aviation Authority

[62] PEZDEK, K., MAKI, R., et al. (1988). Picture memory: recognizing added and deleted details. Journal of Experimental Psychology: Learning, Memory & Cognition, **14**(3), pp. 468-476

[63] POSNER, M. (1976). Visual dominance: an information processing account of its origins and significance. *Psychological Review*, **83**, pp. 157-171

[64] PURNELL, K. and SOLMAN, R. (1991). The influence of technical illustrations on students' comprehension in geography. *Reading Research Quarterly*, pp. 277-299

[65] REISER, R. and GAGNE, R. (1982). Characteristics of media selection model. *Review of Educational Research*, **52**(4), pp. 499-512

[66] RIEBER, L.P. (1990). Animation in computer based instruction. *Educational Technology Research and Development*, **38**, pp. 77-86

[67] ROGERS, Y. and SCAIFE, M. (1998). *How can interactive multimedia facilitate learning*. In Lee, J. (Ed.) Intelligence and multimodality in multimedia interfaces: research and applications. AAAI Press

[68] RUBIN, A. (1980). *A theoretical taxonomy of the differences between oral and written language*. In Spiro, R., Bruce, B. & Brewer, W. (Eds.) Theoretical issues in reading comprehension. Lawrence Erlbaum Associates

[69] SALOMON, G. (1979). *Interaction of media, cognition and learning*. Jossey Bass

[70] SCAIFE, M. and ROGERS, Y. (1996). External cognition: how do graphical representations work? *International Journal of Human-Computer Studies*, **45**, pp. 185-213

[71] STASKO, J., BADRE, A. and LEWIS, C. (1993). Do algorithm moving images assist learning? In *Proceedings CHI'93*. ACM, pp. 61-66

[72] SUTCLIFFE, A.G. and FARADAY, P.M. (1994). Designing presentation in multimedia interfaces. In *Proceedings of CHI '94*. ACM, pp. 92-98

[73] TRAVERS, R.M. (1967). *Research and theory related to audio-visual information transmission: US Office contract No OES-16-006*. Western Michigan University

[74] TREISMAN, A. (1988). Features and objects: fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, **40**A(2), pp. 201-237

[75] TREISMAN, A. and DAVIES, A. (1973). Divided attention to eye and ear. In Kornblum, S. (Ed.) *Attention and performance IV*, Academic Press

[76] VETERE, F., HOWARD, S. and LEUNG, Y. (1997). A multimedia interaction space. In Howard, S., Hammond, J. & Lingaard, G. (Eds.) *Proceedings of INTERACT 97*. Chapman and Hall, pp. 205-211

[77] VANDERVONCKT, J. and BODART, F. (1993). Encapsulating knowledge for intelligent automatic interaction objects selection. In *Proceedings of INTERCHI '93*. ACM

[78] WETZEL, C.D., RADTKE, P.H. and STERN H.W. (1994). *Instructional effectiveness of video media*. Lawrence Erlbaum Associates

[79] WICKENS, C.D., SANDRY, D. and VIDULICH, M. (1983). Compatibility and resource competition between modalities of input, output and central processing. *Human Factors*, **25**, pp. 227-248

[80] WILLIAMS, D., DUNCUMB, I. and ALTY, J.L. (1996). Matching media to goals: an approach based on expressiveness. In Sasse, M.A., Cunningham, R.J. & Winder, R.L. (Eds.) *People and Computers XI, Proceedings of HCI'96*, Springer

ICS 13.180; 35.200

Price based on 46 pages