IPR No.: IPR2016-00500
Patent No. 7,864,163

EXHIBIT 1003

12

MULTIMEDIA USER INTERFACE DESIGN

Alistair Sutcliffe
University of Manchester Institute of Science and Technology

INTRODUCTION

Design of multimedia interfaces currently leaves a lot to be desired. As with many emerging technologies, it is the fascination with new devices, functions, and forms of interaction that has motivated design rather than ease of use, or even utility of practical applications. Poor usability limits the effectiveness of multimedia products that might look good, but do not deliver effective use (Scaife, Rogers, Aldrich, & Davies, 1997). The multimedia market has progressed beyond the initial hype, and customers are looking for well-designed, effective, and mature products.

The distinguishing characteristics of multimedia are information-intensive applications that have a complex design space for presenting information to people. Design, therefore, has to start by modeling information requirements. This chapter describes a design process that starts with an information analysis, then progresses to deal with issues of media selection and integration. The background to the method and its evolution with experience can be found in several publications (Faraday & Sutcliffe, 1996, 1997b, 1998b; Sutcliffe & Faraday, 1994). A more detailed description is given in Sutcliffe (2002). The time-to-market pressure gives little incentive for design; so at first reading, a systematic approach may seem to be counter to the commercial drivers of development. However, I would argue that if multimedia design does not adopt a usability engineering approach, it will fail to deliver effective and usable products.

Multimedia applications have significant markets in education and training, although dialogue in many systems is restricted to drill-and-quiz interaction and simple navigation. This approach, however, is oversimplified: For training and education, interactive simulations, and microworlds are more effective (Rogers & Scaife, 1998). Multimedia has been used extensively in task-based applications in process control and safety critical systems (Alty, 1991; Hollan, Hutchins, & Weitzman, 1984); however, most transaction processing applications are currently treated as standard interfaces rather than multimedia-based designs. With the advent of the web and e-commerce, this view may change.

Design issues for multimedia user interfaces expand conventional definitions of usability (e.g., ISO 9241 part 11) into five components:

- *Operational usability* is the conventional sense of usability that concerns design of graphical user interface features such as menus, icons, metaphors, and navigation in hypermedia.
- *Information delivery* is a prime concern for multimedia or any information-intensive application, and raises issues of media selection, integration, and design for attention.
- *Learning*: Training and education are both important markets for multimedia, and hence learnability of the product and its content are key quality attributes. However, design of educational technology is a complex subject in its own right, and multimedia is only one part of the design problem (see chapter 42, Quintana et al., which deals with educational software design).
- *Utility*: In some applications, this will be the functionality that supports the user's task; in others, information delivery and learning will represent the value perceived by the user.
- *Aesthetic appeal*: The attractiveness of multimedia is now a key factor, especially for Web sites. Multimedia interfaces have to attract users and motivate them, as well as being easy to use and learn.

Multimedia design involves several specialisms that are technical subjects in their own right. For instance, design of text is the science (or art) of calligraphy that has developed new fonts over many years; visualization design encompasses the creation of images, either drawn or captured as photographs. Design of moving images, cartoons, video, and film are further specializations, as are musical composition and design of sound effects. Multimedia design lies on an interesting cultural boundary between the creative artistic community and science-based engineering. One implication of this cultural collision (or rather, one hopes, synthesis) is that space precludes "within media" design (i.e., guidelines for design of one particular medium) being dealt with in depth in this chapter. Successful multimedia design often requires teams of specialists who contribute from their own skill sets (Kristof & Satran, 1995; Mullet & Sano, 1995).

DEFINITIONS AND TERMINOLOGY

Multimedia essentially extends the graphical user interface paradigm by providing a richer means of representing information for the user by use of image, video, sound, and speech. Some views of what constitutes multimedia can be found in Bernsen (1994), who proposed a taxonomy of analogue versus discrete media, which he calls modalities, as well as visual, audio, and tactile dimensions. Heller and Martin (1995) take a more conventional view of classifying image, text, video, and graphics for educational purposes. The following definitions broadly follow those in the ISO standard 14915 on Multimedia User Interface Design (ISO, 1998). The starting point is to ask about the difference between what is perceived by someone and what is stored on a machine.

Communication concepts in multimedia can be separated into:

- *Message*: The content of communication between a sender and receiver.
- *Medium (plural media)*: The means by which that content is delivered. Note that this is how the message is represented rather than the technology for storing or delivering a message. There is a distinction between perceived media and physical media, such as CD-ROM, hard disk, etc.
- *Modality*: The sense by which a message is sent or received by people or machines. This refers to the senses of vision, hearing, touch, smell, and taste.

A message is conveyed by a medium and received through a modality. A modality is the sensory channel that we use to send and receive messages to and from the world, essentially our senses. Two principal modalities are used in human–computer communication:

- *Vision*: All information received through our eyes, including text and image-based media.
- *Hearing*: All information received through our ears, as sound, music, and speech.

In the future, as multimedia converges with virtual reality, we will use other modalities more frequently: ***haptic*** (sense of touch), ***kinaesthetic*** (sense of body posture and balance), ***gustation*** (taste), and *olfaction* (smell). These issues are dealt with in chapter 14, Multimodal Interfaces (Oviatt), and chapter 31, Virtual Environments (Stanney).

Defining a medium is not simple because it depends on how it was captured in the first place, how it was designed, and how it has been stored. For example, a photograph can be taken on film, developed, and then scanned into a computer as a digitized image. The same image may have been captured directly by a digital camera and sent to a computer as an e-mail file. At the physical level, media may be stored by different techniques.

Physical media storage has usability implications for the quality of image and response time in networked multimedia. A screen image with 640 × 480 VGA resolution using 24 bits per pixel for good color coding gives 921,600 bytes; so, at 30 frames/s, 1 s needs around 25 megabytes of memory or disk space. Compression algorithms (e.g., MPEG [Moving Pictures Expert Group]) reduce this by a factor of 10. Even so, storing more than a few minutes of moving image consumes megabytes. The usability trade-off is between the size of the display footprint (i.e., window size), the resolution measured in dots per inch, and the frame rate. The ideal might be full screen high resolution (600 dpi) at 30 frames/s; with current technology, a 10-cm window at 300 dpi and 15 frames/s is more realistic. Physical image media constraints become more important on networks, when bandwidth will limit the desired display quality. Sound, in comparison, is less of a problem. Storage demands depend on the fidelity required for replay. Full stereo with a complete range of harmonic frequencies only consumes 100 kilobytes for 5 mins, so there are few technology constraints on delivery of high-quality audio.

COGNITIVE BACKGROUND

The purpose of this section is to give a brief overview of cognitive psychology as it affects multimedia design. More details can be found in section I, Humans in Human–Computer Interaction.

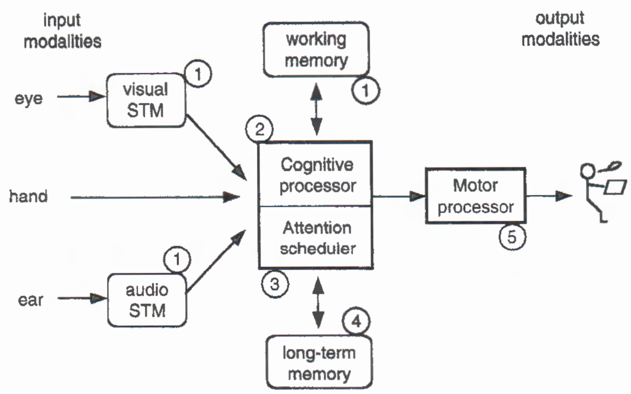
Perception and Comprehension

Our eyes scan images in a series of rapid jumps called saccades interleaved with fixations in which the eye dwells on a particular area. Fixations allow image detail to be inspected, so eye tracking gives some impression of the detail inspected in images. Generally, our eyes are drawn to moving shapes, then complex, different, and colorful 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 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. Because the visual sense receives information continuously, it gets overwritten in 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 nonrealistic images, we have to think carefully about the meaning, for example to interpret a diagram. Although 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 & Christie, 1987).

Selective Attention

We can only attend to a limited number of inputs at once. Although 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, Sandry, & Vidulich, 1983). Our attention is selective and closely related to perception; for instance, we can overhear a conversation in a room 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, whereas 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



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

FIGURE 12.1. Approximate model of human information processing using a human as computer system analogy, based on the Model Human Processor (Card et al., 1983). For more on cognitive models, see chapter 2 (Proctor and Vu) and chapter 5 (Byrne). STM = short-term memory.

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 either as speech or as written text viewed in an image. All such input competes for language understanding resources, hence making sense of speech and reading text concurrently is difficult (Barnard, 1985). Figure 12.1 shows the cognitive architecture of human information processing and resource limitations that lead to multimedia usability problems.

Capacity overflow (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 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 cannot make sense of it. Finally, multitasking (5) makes further demands on our cognitive processing, so we will experience difficulty in attending to multimedia input when performing output tasks.

Making a theme in a multimedia presentation clear involves directing the user's reading and viewing sequence across different media segments. Video and speech are processed in sequence, and text enforces a serial reading order by the syntactic convention of language; however, viewing image media is less predictable, because it depends on the size and complexity of the image, the user's knowledge of the contents, task and motivation (Norman & Shallice, 1986), and designed effects for salience. Attention-directing 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.

Learning and Memorization

Learning is the prime objective in tutorial multimedia. However, the type of learning can be either skill training, in which case conducting an operational task efficiently and without errors is the aim, or a deeper understanding of the knowledge may be required. In both cases, the objective is to create a rich memory schema that 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 constructivist learning theory (Papert, 1980), which has connotations for tutorial multimedia. Interactive microworlds in which users learn by interacting with simulations, or constructing and testing the simulation, give a more vivid experience that forms better memories (Rogers & Scaife, 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 display a causal model of why it works. Schema integration during memorization fits the separate viewpoints together.

The implications from psychology are summarized in the form of multimedia design principles that amplify and extend those proposed for general UI design (e.g., ISO 9241 part 10 [ISO, 1997]). The principles are high-level concepts that are useful for general guidance, but they have to be interpreted in a context to give more specific advice.

- *Thematic congruence*: Messages presented in different media should be linked together to form a coherent whole. This helps comprehension as the different parts of the message make sense by fitting together. Congruence is partly a matter of designing the content so it follows a logical theme (e.g., the script or story line makes sense and does not assume too much about the user's domain knowledge) and partly a matter of attentional design to help the user follow the message thread across different media.
- *Manageable information loading*: Messages presented in multimedia should be delivered at a pace that is either under the user's control or at a rate that allows for effective assimilation of information without causing fatigue. The rate of information delivery depends on the quantity and complexity of information in the message, the effectiveness of the design in helping the user extract the message from the media, and the user's domain knowledge and motivation. Some ways of reducing information overload are to avoid excessive use of concurrent dynamic media and give the user time to assimilate complex messages.
- *Ensure compatibility with the user's understanding*: Media should be selected that convey the content in a manner

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**