Networked Multimedia

François Fluckiger,
CERN, Geneva, Switzerland

What is Multimedia?

Multimedia is a rapidly evolving field with growing applications involving economic sectors as diverse as consumer electronics and entertainment, cable television and broadcasting, education and professional training, publishing, telecommunications and the computer industry, as well as telecommunications operators. This uniquely broad spectrum of involved sectors, the variety of those involved, and the fast advances of the enabling technology makes multimedia one of the most fashionable themes of the decade, as well as a rather confusing concept with regard to its actual meaning and scope. Let us define the term and explain why it seems so popular.

Digital multimedia is the field concerned with the computer-controlled integration of text, graphics, still and moving images, animation, sounds, and any other medium where every type of information can be represented, stored, transmitted and processed digitally.

Thus, multimedia may be used as a name ("multimedia is a new field") but is more generally employed as an adjective ("multimedia product", "multimedia technology"). Text, graphics, still images, moving images, computer animation and sounds are six forms of information with different properties. Graphics are described by objects and therefore retain the semantic content, whereas images are represented as bitmaps which can not be revised. Moving images, animation and sound are time-dependent media, unlike text, graphics and still images which are called discrete media.

One characteristic of many multimedia systems is that they may stimulate several senses. Such systems are called *multisensory* systems. The most notable advance in Human Computer Interaction (HCI) has been the introduction of audio media, either speech or non-speech sound. The tactile sense has been historically and is still nearly exclusively involved in capturing the user input (keyboard and mouse), and new input devices are emerging (gloves in virtual reality, as we shall see in Chapter 15). The benefits of multisensory systems are manifold. First, the interaction with the computer system may appear more "natural" and "friendly" to the user. Second, useful redundancy of information may be possible. For example, user attention may be better caught when visual and audio alarms are mixed. Third, multisensory complementary information presented concurrently will improve the memorization of knowledge. A typical illustration is the mix of complementary visual and audio information in lecturing. Fourth, emotional information is easier to convey. An example is videophony where the video image of the correspondent is added to voice telephony. Last but not least, multisensory systems may be of invaluable benefit to users with special needs.

Another characteristic of many multimedia systems is that they allow *interactivity*, that is, the user may have some form of control on the presentation of the information. Depending on the degree of interactivity, the user may select the time when the presentation starts, the order in which the various information items are presented, and the speed at which they are displayed.

In multimedia systems, all information is represented in a digital form, including time-dependent media types such as sound and motion video. The key advantage of digital representation lies in the universality of representation. Since any medium, be it a text, an image or a sound is coded in a unique form which ultimately results in a sequence of bits, all kinds of information can be handled in the same way and by the same type of equipment. Furthermore, transformations of digital information are error-free, while analog transformations introduce distortions and noise. However, the digitization process introduces a distortion of the information. Reducing this distortion may be achieved by increasing the sampling rate and the number of bits to code each sample. Images, sound and motion video require large amount of digital storage capacity.

Why is multimedia so popular?

There is no single reason for the increasing popularity of multimedia, which on the contrary results from a combination of several factors.

First, multimedia results first from the emergence of complementary form of technology which are making possible what was unthinkable 10 years ago. One may say that multimedia has been initially more pushed by technology than pulled by the market. That is, multimedia was not initially generated by a genuine and strong user demand.

One existing problem in which multimedia systems may help is the risk of information overkill. Organizations are exposed to ever larger volumes and more numerous types of information. How can we ensure that the information is sorted and reaches those who need it, and only those, without excessive delay? And if this is possible, how can we do it at reasonable cost, in terms of manpower and storage? Since multimedia is computer-controlled, it may also help in mastering massive information more efficiently and cheaply. Image document management where paper documentation is digitized and progressively eliminated is an example. Savings are possible in manpower (sorting, indexing, filing and retrieving documents), in storage cost and space, and in time to retrieve and display information.

Another relevant area is knowledge acquisition, where multimedia or even multisensory forms may improve learning processes. Multimedia Computer Based Training and Computer Based Education, where the learning is basically self-guided, are two multimedia applications areas with the best market prospects.

Another field where a problem exists is that of collaborative work. Modern economies tend to increase requirements for remote teams, collaborators, or business players to interwork as efficiently as possible. Multimedia may offer alternatives to the traditional solution: the combination of telephone calls, facsimiles and travel. An example is light-weight audio-videoconferencing, delivered straight to the office.

In computing, three types of technological advances have particularly enabled the emergence of multimedia. First, very large scale integration (VLSI) electronics: the power of central processing units keeps doubling every two years; the density of transistors in memory chips reached 64 million, so that a single chip can store 16 million bytes. This permits the handling of still and moving images —both very demanding in processing power and immediate storage capacity— by desktop personal computers or workstations. Second, the technology and cost of storage capacity, which is essential for continuous media, keep improving fast. Non-removable magnetic disks cost as little as $1 per megabyte. CD-ROMs can store up to 72 minutes of stereophonic sound of nearly VCR-quality compressed motion video. They cost as little as $2 per disk when mass-produced. Write-Once-Read-Many disks (WORMs), also removable, are the technology of choice for storage of semi-static massive multimedia information. Third, a number of parallel advances in software and user interface concepts are particularly relevant to multimedia. This includes graphical user interfaces (GUIs), object-oriented concepts, and the client-server paradigm.

Networking is the other field of dramatic progress. Transmission techniques —not only those concerning the popular fiber optics, but also the more conventional copper cables— give affordable cost transmission rates in the range of 100 to 1000 million bits per second. New services offered by public operators are also instrumental. First, the narrow-band integrated-service-digital-network (N-ISDN) can bring to every organization or home, the potential to establish digital calls on which data, speech-quality sound, and low frame rate low resolution video can be mixed. The next step is the advent of broad-band integrated-service-digital-network (B-ISDN) services, supported by the famous Asynchronous Transfer Mode (ATM) technology. Applications such as audio-videoconferencing at broadcast TV-quality, or video-on-demand are thus made possible.

Finally, algorithms to compress sound, images and motion video have progressed considerably. With recent off-the shelf compression products, the storage volumes of colour images can be reduced by a factor of 30. Compression ratios with motion video are even more dramatic. Nearly VCR-quality can consume 130 times less storage or bit rate than uncompressed studio-quality digital TV. For normal TV broadcast-quality the ratio is in the order of 40.

Information superhighway is a generic name for programs aiming in several countries at building a high-speed communications infrastructure where eventually every home is connected by an optical fiber. The initial idea comes from the USA. The expected integrated services include conventional telephony and telefax, videophony and electronic messaging, access to entertainment servers for services such as movie-on-demand or computer-games, education, access to cultural and general information, and tele-transactions such as shopping from home. In parallel, other companies are investing in wireless transmission, generally satellite based, to offer similar services.

Most of the potential superhighway markets are in fact unknown, and many technological issues have to be resolved. A key factor for the success of applications such a multimedia information or video -on-demand will be the nature and the quality of the information content.

Overview of multimedia applications

There are many possible ways of classifying multimedia applications. One existing taxonomy distinguishes between applications dealing with communications between people (people-to-people), and applications aiming at improving access of people to systems or servers (people-to-systems). People-to-people multimedia applications fall into two broad categories: synchronous, sometime called real-time applications such as audio-videoconferencing, and asynchronous applications such as multimedia electronic mail. People-to-systems applications may be categorized according to the mode of access to the information server. Interactive access applications are those where the end-user at least triggers the access to information, such as in video-on-demand services. Conversely, in distribution applications, the server determines the time and form of transmission. This is analogous to TV broadcast, which may either be universal, or be restricted to closed -e.g. paying- groups of recipients.

Let us start by discussing a few synchronous people-to-people applications. One example is packet-mode telephony. This is an inter-personal conversation over ordinary packet networks. You call through your regular desktop computer, talk and listen to it, and no physical telephone is involved. Products based on conventional audio encoding require 64 Kbps per call, but the rate may be as low as 16 Kbps with efficient compression technique. Any LAN has sufficient capacity to carry at least a dozen of simultaneous calls with no loss. But WANs require in the order of T1 speed to support multiple calls, and maintain the audio quality, while not degrading the ordinary data service for which they have usually been installed. Because of the lack of dedicated products and standards, little use is made of packet telephony which is normally restricted to on-site calls within close corporate environments.

Packet videophony is another inter-personal synchronous application. It supports audio and visual conversation via a regular desktop computer over the ordinary packet networks it is attached to. Video adds some 80 to 200 Kbps to the audio channel involved in packet telephony, and most LANs can afford half a dozen of such simultaneous video conversations, whereas WANs require megabit lines to support a few parallel sessions. A number or freeware of commercial implementations, designed in fact for multiparty videoconferencing, are available off-the shelf, most operating over the IP protocol. In contrast to packet telephony, users tend to employ packet videophony more for long haul than on-site conversation, probably because the service is completely new and does not suffer from comparison with the telephone. In theory, ISDN is the natural service to support audio-video services between two individuals. In practice, its general use requires ISDN access to be available in every office where the service is delivered, and the desktop computer to be equipped with an ISDN interface in addition to its regular LAN interface.

Let us continue our exploration of asynchronous people-to-people applications. In the context of the support by computers of cooperative work, a family of tools are available to allow the remote sharing of computer display windows between participating individuals involved in common tasks. Thus, a least one portion of computer display is viewed by all participating individuals. Shared workspace applications aim at creating at a distance the sharing of ephemeral information such as a whiteboard sketch and also the actual more permanent documents which may be the object of the joint work.

Another synchronous application is audio-video distribution, that is, the propagation of an information to a possibly large subset of the potential destinations (multicast) or to all potential recipients (broadcast). In professional environments, this is used for the dissemination of seminars, meetings or lectures. Multicasting or broadcasting requires a network capable of replicating information flows, so as to offload the source itself from generating multiple identical streams. Conventional packet LANs and a few packet WANs technologies (IP, ST-2) are multicasting- or broadcasting-capable. In local environments, the main advantage of digital audio-video distribution lies is the possibility to deliver the stream to every office over the local-area network. Though the bit rate of the combined audio and video streams may be as low as 200 Kbps —this is for a very low frame rate— care must be taken if the LAN is a normally loaded production Ethernet. A choice of public domain and commercial implementations exist to equip most personal computers and workstations with the software and possibly the hardware to decode and display incoming packet audio-video streams. Large-scale distribution over wide-area data networks is more tricky. The technology of choice for corporate professional audio-video broadcast where the location of the recipient systems is stable, is satellite transmission. Over the Internet, open multicast groups are used for broadcasting conferences or public events. Multicast flows are carried over an experimental overlay network called MBone. Usually, the broadcast of an event results in several streams, each being propagated within a multicast group: the video and the audio streams are separated.

Let us conclude our review of synchronous application with videoconferencing, Videoconferencing involves groups of people, not single individuals only, and is different from video distribution as it necessarily implies bidirectional communications. Videoconferencing is supported by two types of technology but the line between the two has progressively blurred. Circuit-mode videoconference systems, also called conventional video-codecs, or simply codecs, rely on dedicated circuits offering a guaranteed bit rate, such leased line or public circuit-switched services. Systems from different vendors will generally interwork as most support the ITU H.320 standard, though many manufacturers find their proprietary compression schemes more efficient. Packet mode videoconferencing results from an evolution of packet videophony. The rationale is identical: exploiting the potential of computers on the desktop and that of the networks they are connected to, to support videoconference services. Note that packet videoconference systems from different computer manufacturers will generally not interoperate, but public domain software will.

Let us now turn to asynchronous people-to-people applications. Desktop voice-mail, desktop video-mail and desktop multimedia mail are examples of asynchronous applications. A frequent use of text and voice mail is the annotation of a text message with a short oral preamble, for example to add a comment to a sensitive message. Usually, multimedia messages are composed of several parts, each of a different information type. An important concept in multimedia mail is the way in which certain parts can be included. Parts of messages can be explicitly or implicitly included. When explicitly included, they are transferred at the same time as the rest of the message. When implicitly included, the message contains only a reference, a pointer to a remote copy the part. A number of multimedia mail implementations are available, either commercially or in the public domain. They generally implement the Internet MIME standard.

Let us conclude our review with applications where people access information servers. We shall concentrate on the description of an hypermedia application. Hypermedia is an extension of the hypertext concept where information is not only text but any type of media. So, what is hypertext? This text with links. Information is not sequential, like in a book, but formed of documents linked together. Documents contains fragments of texts called anchors which point to other documents. World-Wide Web (WWW) is the name of a cooperative project initiated at CERN to design and develop a series of concepts, communication protocols and systems according to the hypermedia concept. The project has now evolved into a world-wide cooperative initiative involving more than 300 hundred contributors. The World-Wide Web designates the project itself, a series of protocols and conventions, and the resulting space of digitized and interlinked multimedia information, called the hyperspace. The World-Wide Web hyperspace is supported by a network of multimedia information servers. An essential characteristic is that there is no central authority for accepting new information is the overall web of information. Anybody can create a document and insert it in the web, and anyone can make hyperlink references to any other document. Another essential feature is that the principle of geography-independence of the various items of the web has been implemented through a simple mechanism for referencing and locating unambiguously individual documents. The exploration of the web follows the hypertext principles. Users generally "click" on anchors to activate a link. The document, which may be a piece of text, an image or a sound or motion video sequence is fetched. It may come from another country or continent without the user knowing it. The universal information web contains things like the Vatican Library of Renaissance Culture or several virtual museums that anyone in the world can explore from his desktop computer. The World-Wide Web has created a universal web of information distributed transparently around the world, accessible by anyone having an Internet access, from anywhere in the world, covering all types of knowledge and culture built up over ages by mankind.

Invitation to further reading

[1] is a reference textbook on networked multimedia, the applications, their requirements on the underlying network, and a discussion on how current and future networks will satisfy them. [2] is a comprehensive practical guide, with special emphasis on multimedia authoring.

Bibliography

[1] **Fluckiger, F.** *Understanding Networked Multimedia*, Prentice Hall, Hemel Hempstead, UK (available March 1995)

[2] **Vaughan, T.** *Multimedia, Making it Work,* Osborne McGraw-Hill, Berkeley, California (1994)