Chapter 17

Multimedia Applications

17.1 Introduction

The availability of multimedia hardware and software components has driven the enhancement of existing applications towards being more user-friendly (known as re-engineering) It has also initiated the continuous development of new multimedia applications. Applications are crucial for the whole domain of multimedia computing and communications because they are the only reason why anybody would invest in this area. However, so far no serious attempt has been made to classify these applications. Because of the burgeoning number of multimedia applications, with this chapter we aim to provide a structured view on the field of multimedia applications.

Data regarding projects, products and other issues discussed in this chapter were collected from a variety of sources: several computer magazines, sets of product catalogues, project descriptions in the context of the U.S. Federal *High Performance Computing and Communications (HPCC)* program, the European *ESPRIT* program, RACE, ACTS, DELTA and other programs, worldwide-supported initiatives such as the *World Wide Web Initiative*, market surveys and extensive customer feedback.

17.1.1 Programs

Several programs for the development of multimedia applications have been established during the last few years, some well-known from the U.S. and Europe are outlined below.

- *USA*

  The High Performance Computing and Communication *(HPCC)* program accelerates the development of scalable, high-performance computers, advanced high-speed computer communications networks and advanced software – all critical components of a new *National Information Infrastructure (NII)* initiative [HPC94]. The HPCC program evolved in the early 1980's out of recognition by American scientists, engineers and leaders in government and industry. One of the most significant program components in the HPCC program is the *Information Infrastructure Technology and Applications (IITA)* program. The IITA's research and development efforts are directed towards *National Challenge* problems such as *civil infrastructure, digital libraries, education and life-long learning, energy management, the environment, health care, manufacturing processes and products, national security* and *public access to government information.* IITA technologies will support advanced applications such as:

  - *Tele-medicine*

    An individual's medical records (including X-ray and CAT scan images) will be sent to a consulting physician located miles away.

  - *Remote Education and Training*

    The access and study of books, films, music, photographs and works of art in the Library of Congress and in the nation's great libraries, galleries and museums will be available on a regular basis to teachers and students anywhere in the country.

  - *Tele-operation*

    The flexible incorporation of improved design and manufacturing, which may be performed in a distributed manner, will produce safer and more energy-efficient cars, airplanes and homes.

– *Information Access*

  Universal access to government data and information products by industry and the public will be supported.

HPCC program management consists of several working groups coordinating activities in specific areas. The application group, led by NASA (the National Aeronautics and Space Administration), coordinates activities related to national challenge applications, software tools needed for application development and software development at high performance computing centers. The education group, led by the National Institute of Health and Department of Health and Human Services, coordinates HPCC education and training activities. The communication group, led by NSF (National Science Foundation), coordinates network integration activities. The research group, led by ARPA (the Advanced Research Projects Agency, Department of Defense), focuses on basic research, technology trends and alternative approaches to address the technological limits of information technology.

- *Europe*

  ESPRIT (European Strategic Program for Research in Information Technology) is a well-known scientific program of the European Community. The primary goal is to support development of technology and science similar to the HPCC program. The smaller RACE (Research in Advanced Communication in Europe) program is similar to ESPRIT, but focuses on communication issues. In the second phase, the RACE II program focused on the residential and small business user market to use multimedia communication applications (tele-applications) [RAC93]. *ACTS* (Advanced Communication Technology)is the follow-up to the RACE program. Emphasis is on the customer's access connection, which is the most cost-sensitive issue of the whole network and the one most closely related to the service demand. The RACE projects, for example, cover applications such as:

  – *Tele-interaction*

    Tele-services will be used in information systems (e.g., information kiosks) and entertainment (e.g., telegames) as part of the information age.

  – *Tele-shopping*

Shopping through remote electronic catalogues will support faster and more convenient sale and advertisement of products.

- *Thematic channels with interactive TV and electronic newspaper*

  The development of current TV technology towards interactive TV and the use of thematic channels will make it possible to create new programs for education and entertainment. Thematic channels might provide access to electronic newspapers and other information.

- *Tele-working*

  Further development of interactive tele-services will provide an environment for the reliable setup of home offices, as well as industrial collaboration and remote education. Already today, services such as videotelephony and video-conferencing are part of collaborations among remotely located laboratories and colleagues.

For these applications, new tools and system components are being implemented for incorporation into commercial products.

Among the national programs, the German Telekom project *BERKOM* (BERliner KOMmunikationssystem) is one of the most prominent, having run over five years and incorporating the multimedia work of the most active researchers in the field.

17.1.2 Structure

There are many views on how multimedia applications should be classified. For example, a market-oriented view and pragmatic view may divide the current multimedia applications into *kiosk applications, educational applications* and applications in the area of *cooperative work.* Another view would be a communication-oriented view, dividing multimedia applications into *interactive* or *distribution-oriented applications.* A third possibility is some view derived from the hypertext/hypermedia area.

Our classification evolves mainly from the need to describe and present a coherent view on this important area, discussed at numerous commercial and scientific events;

it looks at *multimedia processing* from the computer user perspective. Hence, we distinguish among tools and applications which support the user in *media preparation, media composition, media integration and media communication.* Furthermore, the user is exposed through multimedia applications to *media consumption* and *media entertainment.* This has become known, colloquially, as the "media food chain."

17.2 Media Preparation

Media preparation is performed by multimedia I/O hardware and its supporting software. Therefore, hardware and software are the basic components for introducing media into the digital world of a computer. Appropriate hardware is the prerequisite for working with multimedia applications. The software creates the environment to work actively with the multimedia applications. It allows the computer user to use and interactively work with the multimedia hardware. We discussed system software issues in Chapters 8, 9, 11 and 12 and application software issues in Chapters 13, 14 and 16. We shall also present some application-specific software issues in this chapter when we discuss different kinds of user interaction with media. For the purpose of better understanding, we concentrate in this section on some specialized multimedia devices.

17.2.1 Means

New hardware technology is needed for multimedia applications and their interactive experience. Chapters 3, 4 and 5 discuss the basic principles of media and their hardware support in more detail. Here we want to expand briefly on other devices also available for media preparation.

Audio Support

Some *audio* support with *multiple-channel digital sound tracks* is already available. For example, a six-channel digital sound track (front-left, center, front-right, surround-left, surround-right and subwoofer) has been developed. In the area of virtual reality entertainment, sound interaction occurs via a *helmet*. The same degree of attention was paid to the design and development of digital stereo sound.

Video Support

Video boards and *digitizers* aim toward a high-resolution picture presentation. The ultimate goal is high resolution and a film rate of 60 frames per second (HDTV) or faster, "a la Showscan," which provides an extremely clear picture. An important capability of the video hardware is to provide a constant frame rate with a minimum of jitter. This property is more important than a faster rate or even increased color or pixel resolution because a large amount of jitter between frames causes the perception of jerky motion video, which is more disturbing than a slower frame rate with minimal jitter. For example, the video compression technique MPEG creates frame-differencing. Because the differences in content between frames are not really controllable, a constant delivery rate of frames to the viewer is not always maintained, thereby causing a perception of jerky MPEG-compressed motion video.

Graphical displays provide high resolution for graphical, image and motion video applications. An important component of the raster system display (see Figure 4.2 for the architecture of a raster display) is the *video controller*, which constantly refreshes the display. For applications where *mixing of video* is required, the video controller provides this function. Two images, one defined in the frame buffer and the other defined by a video signal coming from a television camera, recorder or other source, can be merged to form a composite image. Examples of this kind of merging are regularly seen on television news, sports and weather shows.

Currently, several basic kinds of displays are used in virtual reality applications [Tri87, FDFH92, Cla94]:

- *Head-Mounted Displays (HMD)*

  An HMD includes one or two displays. Special lenses allow the user to focus on the display as if they were further away.

- *Surround Displays*

Surround displays surround the user, meaning the user is situated in a room with walls serving as displays. To provide stereoscopy and sensitivity to head motions, a stereoscopic display system and head position tracker are used.

- *Digital Holography*

  Holography is a method for displaying 3D images without using special head-gear or tracking the viewer's location. Traditional holograms are produced by exposing photographic film simultaneously to laser light scattered from the object to be recorded, and to a reference beam from the same laser. The interference patterns recorded on the film encode the object's appearance from a range of viewpoints. The hologram is viewed by illuminating it with laser light from the opposite direction.

Scanner Devices

Image scanners and *photo CD devices* support input and output of images and photographs. Although *data tablets* can be used to manually digitize existing line drawings, this is a slow process, unsuitable for more than a few simple drawings. Image scanners provide an efficient solution. A television camera, in conjunction with a digital frame grabber, is an inexpensive way to obtain moderate resolution (1000×1000 pixels, with multiple intensity levels) raster images of black-and-white or color photographs. Slow-scan Charge-Coupled-Device (CCD) television cameras can produce an image of 2000×2000 pixels in about 30 seconds [FDFH92].

For high-quality publication work, a *photo scanner* is used. The photograph is mounted on a rotating drum. A light beam is directed at the photo, and the amount of light reflected is measured by a photocell. For a negative, transmitted light is measured by a photocell inside the drum. As the drum rotates, the light source slowly moves from one end to the other, thus doing a raster scan of the entire photograph. The highest resolution scanners use laser light sources, and have resolutions greater than 2000 pixels per inch.

Another type of scanner uses a long thin strip of CCDs, called a *CCD array.* A drawing is digitized by passing it under the CCD array. A single pass, taking one or two minutes, is sufficient to digitize a large drawing. Resolution of the CCD array is 200 to 1000 pixels per inch, which is less than the photo scanner technique.

Recognition Devices

Recognizers are built to recognize different media. An example is the *object-oriented character recognition engine AQUIRE* [KW93b]. AQUIRE is used in a pen-based computer environment. From a programmer's point of view replacing a keyboard with a pen requires a pen-based user interface that provides a complete control mechanism for a pen as a central input device. To support the same user interface methods, where the user is accustomed to a keyboard as the input device, a highly sophisticated character recognition engine must be embedded in the pen-based applications. This recognizer can be used to transform drawings (e.g., cup drawing) into their semantics (e.g., cup meaning).

Other recognizers may perform, for instance, *image recognition* to convert images into text, or *speech recognition* to convert audio into text.

Tracking Devices

Trackers report information about position, orientation, acceleration, pressure or joint angles of tracked objects [MAB92]. There are several technologies which have been deployed:

- *Electromagnetic trackers* produced by, for example, Polhemus and Ascension, determine the receiver's position and orientation using coils, pulsed with an electrical signal, in the transmitter and receiver for producing and sensing a magnetic field.

- *Ultrasonic trackers* use ultrasonic pulses in the transmitter and receiver to determine the orientation of the receiver.

- *Optical tracking systems* use video cameras to track objects. Camera-based approaches may track hands and bodies without requiring that users wear special apparatus [Ino93].

Other tracking systems are used, such as coarser-grained, position-only tracking, finger-joint-angles tracking with various sensory gloves or eye tracking technologies. Eye tracking is employed to focus a camera onto the object envisaged by the photograph itself.

Motion-based Devices

Motion-bases are typically hydraulic systems that manipulate the viewer along several axes of motion (up to six degrees-of-freedom of lateral or rotational). The movement of the platform, vehicle or chair is programmed to mimic the real-world motion that would correspond to the visual image. A motion-base is best at simulating acceleration. For instance, quick movement of a motion-base simulates bumps and impacts. The challenge is to implement a synchronized motion-base with its visual reality.

17.2.2 Remarks on the Current Status

An important issue in multimedia performance is the proper selection of media based on the media hardware availability. This means that depending on what the multimedia application should achieve, the computing system with its devices and particular media quality is selected. It is crucial to point out that this approach is still very dependent on hardware, and causes problems with portability of multimedia applications.

Another important problem is that the currently available multimedia computer hardware is still not fast enough to meet the goals of applications, such as *virtual reality entertainment* and *motion-based simulators*.

17.3 Media Composition

Media composition involves *editing* single media, i.e., changing its objects, such as characters, audio sentences, video frames and attributes such as the font of a character, recording speed of an audio sentence or color of an image.

17.3.1 Text and Graphics Editors

Text Editors

Text editors provide writing and modifying facilities to compose text in a document. There are either separate text editors (e.g., GNU emacs text editor in combination with the LaTeX document preparation tool on workstations, WordPerfect on PCs) or text is embedded in graphical tools such as drawing programs – *xfig, MacDraw, CorelDRAW (Corel Corp.), etc..* When editing text, one must deal with the issues of font selection, text style and text effects:

- *Fonts*

  The exact description of each character of text is determined by its font. Font files contain these descriptions, either in bitmap or vector form. Vector fonts are mathematical descriptions of the characters, which can be rendered in a wide range of sizes. Bitmap fonts are stored as bitmaps in predefined sizes.

- *Text Styles*

  Text can be presented in different styles, such as italicized, emboldened, underlined, etc. There are many possible text styles. Therefore, a writer of a document should make a careful choice for its uniform use.

- *Text Effects*

  More advanced text processing systems provide text effects such as shadowing, extrusion, textured fills, text-on-curve, etc. Such possibilities are offered in *CorelDRAW.*

Text editors are also beginning to be enhanced through other media, such as graphic objects. This is the same trend we see in the expansion of graphical tools with text manipulation capabilities. An example of an advanced word processor with graphical capabilities is *Microsoft Word.* This tool provides, in addition to text capabilities, a new toolbar and ribbon that can be customized for the creation of tables, envelopes, bullets and more. Furthermore, a built-in drawing program allows one to work with graphics without leaving the *Word* application.

Graphics Editors

Graphics editors use facilities at the user interface for editing structural representations of graphical objects (structure-level editing) and for modifying higher-level operations on graphical objects (object-level editing). These two levels of editing are possible because the graphical system stores object primitives and their structural representations, which can be manipulated. A simple example of a graphical editor is *xfig* – an X Windows drawing program running on UNIX machines. This drawing application, also called a *layout editor or graphical illustrator* [FDFH92], supports editing of structural representations (the box to the right in Figure 17.1), as well as editing through the use of graphical icons (the two columns of graphical icons to the left in Figure 17.1). Graphical objects can be drawn, moved, copied, etc. via the user interface. Similar interactive programs that allow users to assemble complex 3D

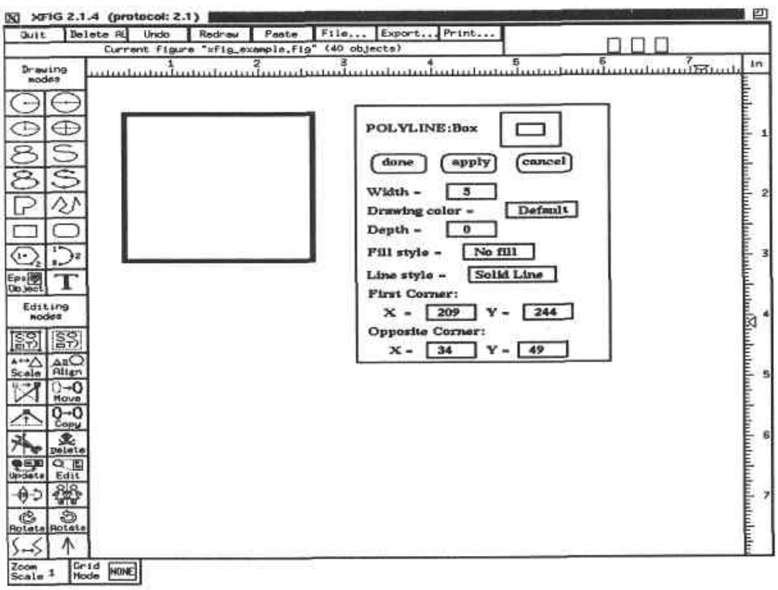

Figure 17.1: *Drawing application (xfig) with graphical and structural editing capabilities.*

objects from simpler objects are called *geometric editors* or *construction programs.*

Object-level editing, screen refreshing and scaling in graphical editors require storage and re-specification of primitives by the application or by the graphics package. If the application stores the primitives, it can perform the re-specification; thus, these object-level operations are more complex. For example, in a graphics application with motion or update dynamics, the editor can modify an object database, which involves modifications of viewing, modeling transformations and changes in or replacement of objects.

17.3.2 Image Editors

Image editors are suitable for applications when neither the application nor the underlying software package keeps a record of the primitives (as is typical in most painting programs).

Scaling (one of the functionalities of an image editor) cannot be implemented by re-specifying the primitives with scaled endpoint coordinates. All that can be done is to scale/edit the contents of the image frame (also called canvas) using *read-pixel* and *write-pixel* operations. For example, a simple and fast way to scale up a bitmap/pixmap (make it larger) is via *pixel replication,* as shown in Figure 17.2. With pixel replication, the image becomes larger, but also coarser, since no new information is provided beyond that contained in the original pixel-level representation (compare Figure 17.2(a) and Figure 17.2(b)). Moreover, pixel replication can increase an image's size only by an integer factor.

Hence, a second scaling technique is used – *sampling and filtering* [FDFH92]. *Sampling* is the process of selecting a finite set of values (pixels) from a continuous signal. Once the pixels (samples) have been selected, they must be displayed, using *reconstruction,* to recreate the original continuous signal from the samples. The continuous signal itself includes noise (high frequencies in the original signal), hence to get a clear image, a *filtering* process (removal of the high frequencies) must be used either before sampling (pre-filtering) or after reconstruction of the image (post-filtering) to remove the noise.

One technique of pre-filtering is *aliasing.* Here, the high frequency components are converted into lower frequency components. The visual artifact of this is that the observed image is *staircasing*, i.e., some edges are not smooth. This visual artifact

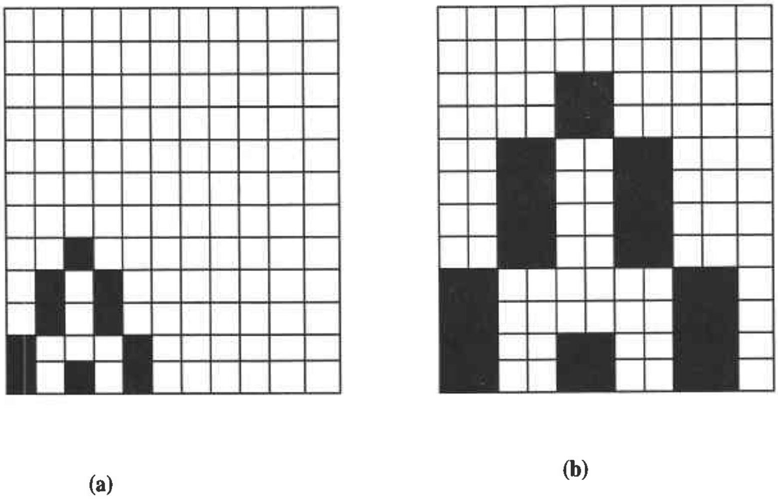

Figure 17.2: *Effect of pixel replication (a) original image at screen resolution. (b) zoomed (double size) image at the same screen resolution.*

must be removed. One possibility for removing *staircasing* is to increase the screen resolution. Another possibility to remove *staircasing* is to use different methods of *anti-aliasing* (see [FDFH92]).

Image editors provide other functionalities such as *increase of resolution, change of intensity, modification of RGB (Red, Green, Blue) colors, colormap editing*, etc. These functionalities can be found in an image editor such as *xv*, developed for UNIX workstations by John Bradley of the University of Pennsylvania in 1993.

An example of a graphics/image editor is *Adobe's Photoshop*™. This tool allows one to draw, edit and paste objects on several layers. Experimenting with different combinations of graphics, text and special effects without altering the original background image is possible. Furthermore, new filters let the user create 3D lighting effects, and remove dust and scratches from scanned images.

Another example of a photograph editor is Aldus *PhotoStyler*™. This tool edits photographic images created by different types of scanners and includes several editing options, such as editing of resolution, scaling of the photographic images, and providing access to various image processing techniques to apply them to the image after it has been scanned for producing a proper photographic image.

17.3.3 Animation Editors

Animation editing is based on graphical editors with respect to 2D or 3D spatial graphic objects. The additional component in animation is time, which can also be edited (4D editing). The functionalities of such editors include *cutting frames from an animation clip, adding new frames to an animation clip*, etc.

The most advanced animation tools already provide the animator with the capability to draw only the key frame. The intermediate frames are then drawn by the computer animation program. This process is called *tweening*. Further, some animation tools include *morphing* (polymorphic tweening), which is a transformation from one shape to another. This transformation can be seen in the cartoon *The Lion King* from Disney studios (1994). With this technique, many special effects can be created.

A general problem in temporal editing of animations is *temporal aliasing* (*staircasing* in time). The temporal aliasing problem in animation can be partially solved by increasing temporal resolution (i.e., increase the frame rate). In another approach, *temporal anti-aliasing* takes multiple samples of a signal and computes their weighted average. In this case, however, the multiple samples must smooth the time continuity and not the space continuity (as was the case in image editors), so the intensity at a point in the image for several sequential times is computed and these are weighted to get a value at a particular frame. Many other approaches have been developed to address temporal aliasing: supersampling, box filtering in the time domain, etc. [FDFH92].

Note that animation tools may include several integrated media editors. For example, Gold Disc, Inc. offers *Animation Works Interactive (AWI)* [Lut94] where *Movie Editor* is used for the assembly of complete animations, *Cel Editor* for building cels (this is an old name for an animation object drawn on a sheet of celluloid) and actors, and *Background Editor* (painting tool) for building the background.

17.3.4 Sound Editors

Sound tools support a number of operations that let the user access, modify and play sound data. The operations fall into four categories:

- *Locating and Storing Sounds*

  Location and storage of sounds can be done in four ways: (1) *record* a sound using an A/D audio device (analog-to-digital converter), (2) *read* sound data from a sound file, (3) *retrieve* a sound from a pasteboard, and (4) *create* sound data algorithmically.

- *Recording and Playback*

  The record operation continuously records sound from a microphone input until it is stopped or paused. Recorded sound is *m-law encoded*, which corresponds to CCITT G.711 and is the standard for voice data used by telephone companies in the U.S., Canada and Japan. *A-law and μ-law encoding* is also part of G.711, and it is the standard encoding for telephony elsewhere in the world. The data are sampled at a rate of 8000 samples per second with 12-bit precision, but if the digitized sound is being compressed, then 8-bit precision per sample is achieved.

  The playback operation plays a sound using a D/A audio device (digital-to-analog converter) speaker output.

- *Editing*

  The editing operation allows one to copy/paste, cut, delete, insert or replace sampled sound data. One problem ought to be pointed out here: audio data are normally contiguous in memory. However, when a sound object is edited, its data can become *fragmented* or discontiguous. Fragmented sounds are played less efficiently. Hence, it is important to have an operation which *compacts* the samples into a contiguous object. Note that compacting a large sound object that has been considerably fragmented can take quite some time.

Digital music tools may include, for example, MIDI editors for music data stored as a *midi-file* – a music representation format that supports the conventions of the standard MIDI audio. MIDI audio is a series of commands for controlling a music synthesizer to produce, for example, orchestral music.

Generally, music editors include functionalities such as modification of loudness, amplitude, tone control, retrieval of a note from a part, removal of a note from a part, addition and removal of groups of notes, etc.

With music editors, special effects such as hall echos can be created. Furthermore, if *Digital Signal Processors (DSP)* are available, music can be synthesized on the DSP and new music can be created with a particular music tool (e.g., MusicKit [Tec89]).

17.3.5 Video Editors

Video editors are based on image editors for editing individual frames, but as in the case of animation editing, temporal considerations are important. Therefore, time resolution (time aliasing) is solved if frames are deleted, added or replaced. Editing functionalities of video editors may combine several cuts into one sequence, adjust audio separately from video and add video transition effects. An example of such a motion video editor is *VidEdit,* which works with Microsoft *Video for Windows* [Lut94].

Some advanced motion video editors (e.g., *D/Vision Pro* from TouchVisionSystems, a PC-based professional video editor running under DOS), can open several windows with different source videos. The editor can roll through each video source at variable speeds and select *edit-cut points.* This kind of editing causes fragmentation of the video. In the case of a conventional videotape, the edited sequence of video frames must be recorded to a new tape to view the new video clip. This kind of editing is called *linear* editing. An advanced video editor (e.g., D/Vision Pro) provides an *Edit Decision List (EDL)* from which the final video can be reconstructed, i.e., the edited video does not have to be recorded to a new tape because it can be played continuously using the EDL. This is called *non-linear* editing. Such tools may have further editing capabilities, e.g., adding dynamic transitions such as wipes, dissolves or fades between cuts.

Advanced motion video editors include several editors for editing video, sound and music in an integrated fashion. Examples of such tools come with Macintosh's

QuickTime.

17.4 Media Integration

Media integration specifies relationships between various media elements to represent and manipulate a multimedia object (e.g., document). Integration is still very much dependent on technology, i.e., platform-specific and format-specific, although there are attempts to provide tools which will integrate media on any platform with any format. An example of media integration in a multimedia application can be found in an authoring tool (Section 17.4.3). In this application, some levels of integration are platform-independent, but others, such as authoring languages, are not.

17.4.1 Multimedia Editors

Multimedia editors support the ability to manipulate multimedia documents that include structured text, multi-font text, bitmap images, graphics, video, digitized voice and other modifiable objects. Most editors use the *What You See Is What You Get* (WYSISYG) editing approach. An example of an early multimedia editor is the BBN's *Diamond Multimedia Editor* [CFLT87].

Several design issues need to be considered when implementing editors for multimedia documents:

- *Document Structure*

  The editor's functionality depends on the structure of the multimedia documents. To enable the exchange of documents and to be prepared for CSCW (Computer-Supported Cooperative Work), the document structure should be compatible with international standards, such as SGML (HTML) or ODA.

- *Media Editor Integration*

  Each medium has its own structure, therefore a multimedia editor actually consists of a collection of powerful and complex editors for individual media provided in an integrated fashion. There are several different issues regarding integration:

1. *Display Surface*

   Different media should be viewed in an integrated fashion, although for editing purposes either an integrated surface or separate window can be used.

2. *Processes*

   A design choice must be made with respect to the implementation of individual editors. They can be implemented as separate processes and the multimedia editor serves then as a parent process for management, or all the editors are implemented within a single monolithic process.

3. *User Interface*

   Media editors should be consistent with respect to the appearance of menus, dialogues, and terminology for commands and prompts.

4. *Data Levels*

   It should be possible to transfer data from one media type to another.

- *Multiple Buffers and Multiple Panes*

  It may be necessary to simultaneously manage multiple documents during one editing session. This can be done by providing *multiple buffers within the editor*. However, with a windowed environment (where multiple instances of an editor can run), several documents can be processed too.

  The capability of viewing multiple parts of the same document can also be provided. *Multiple panes* allow a single window to be split into multiple views. Panes are useful as lightweight devices for allocating screen space without suffering the overhead of going through the window manager. Their existence is often short-lived.

- *Large Documents*

  The multimedia editor must be able to handle large documents, perhaps stored in a distributed fashion. It might even happen that we need to work on *partial* documents because the editor does not have access to the whole document, only to parts of it.

- *External Representation*

  The individual media should be stored in their standardized formats.

The editor framework provides facilities for managing multiple buffers and panes and dispatching events to the individual media editors. It provides functions which operate on entire objects and documents, as well as a set of services which are used by the media editors. Each media editor makes available an array of generic functions, often addressed as the editor framework, as well as defined whatever media-specific editing operations are appropriate.

17.4.2 Hypermedia/Hypertext Editors

Hypermedia/hypertext documents consist of multimedia and non-linear links among the information. The documents are stored in multimedia databases in a structured representation (e.g., HTML database for HTML documents). Hence, the editing process means *accessing* document structures through links (associations) and *editing* objects according to their characteristics (text editors, graphics editors and others are executed).

Hypermedia/hypertext documents might be created and modified through hypermedia/hypertext tools such as:

- Apple's *Hypercard* runs on a Macintosh. It follows the card model (the page size of the document is the same as the size of a card), but has scrolling fields too (for more explanation, see Chapter 13). The documents incorporate text and graphics. Other media are accessible through extensions. The system provides a powerful scripting language.

- *DynaText* is a hypertext system based on SGML and a large number of graphics standards. It comes from Electronic Book Technology. DynaText is now available for SUN workstations.

- *NoteCard* (Xerox PARC) uses the card metaphor, meaning that it structures information in card-size chunks at a hypertext node.

- *Hyperbole* is a flexible, extensible *Personal Information Manager* (PIM) tool, which runs on top of the GNU Emacs text editor, available on any UNIX system. It is written in the programming language Lisp. It is the first step towards a distributed multimedia architecture which will create a *Personalized Information Environment (PIR)*. Hyperbole brings techniques of associative information management offered by hypertext systems to the domain of PIMs. Hyperbole allows the user to use any comfortable tool to generate information. At any point, the information can be easily integrated and further adapted for use within Hyperbole. Rather than the structured approach taken by most PIMs, Hyperbole offers an open framework under which many styles of information and task management can be performed [Wei91a].

- *Guide* is a hypertext editor running on Macintoshs and PCs using Windows. The tool uses an SGML-compatible language called HML. Documents may contain text and graphics.

The editors of hypermedia documents should not only include the editing facilities of the individual media, but also networking capabilities for accessing distributed hypermedia documents. Hence, a *tele-service* that follows logical links should be included. There already exist such hypermedia systems, for example, on the *World Wide Web (WWW)* and *HyperBase.*

17.4.3 Authoring Tools

Consider an application which coordinates a *multimedia presentation.* This application needs to provide a dynamic behavior and support several users' actions to integrate media to a required multimedia presentation. To implement an application with such dynamic support requirements, several processes must be programmed. This kind of application can be either written in a programming language, or implemented using an *authoring system.*

Hence, an authoring system is a set of software tools for creating multimedia applications embedded in an authoring environment. A person who creates applications for multimedia integration, for example, presentation, is called an *author.* The processes together are called *authoring* [Lut94]. There are also other components which belong to the authoring environment, such as multimedia *hardware*, *firmware* (software that is permanently built into the hardware) and an *assembly tool* (an authoring tool that arranges multimedia objects into a presentation or an application, dealing with their relationships in space and time).

When a multimedia application is produced via an authoring system, the author goes through several stages. Note that often at the various steps, a user's feedback is needed which might imply some additional work at any previous step:

- *Concept*

  This step identifies the *application audience*, the *application type* (presentation, interaction, etc.), the *application purpose* (inform, entertain, teach, etc.) and the general subject matter. At this stage, the authoring system cannot help.

- *Design*

  The style and content of the application must be specified. The object should include and generate enough detail so that the following stages of content collection and assembly can be carried out by the authoring system without further interruptions. However, the authoring system should still be tolerant of some kind of revisions. At this stage, the design parameters are entered into the authoring system. The authoring system can take over the task of documenting the design and keeping the information for the next steps of outlining, storyboarding, flow charting, slide sorting, and scripting.

  The other task in the design stage is to decide which data files will be needed in the application, such as audio, video and image files. A list of the material should be generated. The authoring system is only rarely involved in this task (a few authoring systems include entries for dummy file names).

- *Content Collection*

  The content material is collected and entered into the authoring system. In general, this includes taking pictures, making a video clip and producing an audio sequence. When the existing content is available either from internal or external sources, no creation tools are needed. It may be necessary to use a conversion tool to convert external source formats into formats with which the authoring system works. If the author creates the content himself/herself, creation tools are needed, such as word processing, paint and drawing software, image capture hardware and software, audio capture hardware and software and video animation hardware/software. Some authoring systems have some of these features, but with limited capabilities compared to the stand-alone tools.

- *Assembly*

  The entire application is put together in the assembly stage. Presentation packages, for example, do their assembly while the author is entering the content for the various screens. Once the screens are defined and placed in order, the presentation is ready to run. The limitation of this approach is that the author has little or no chance to interact with the authoring system. When the application includes a lot of interaction and very complex or dynamic screens, most authoring tools require details of the work, sometimes even programming. Most high-end authoring software packages, especially those which have a full authoring language, can be operated in a modular mode. In this case, a programmer can create customized modules that fit the specific needs of the particular application.

- *Testing*

  The created application must be tested. More sophisticated authoring systems provide advanced features such as single-stepping or tracing the program flow.

As is clear from the above steps, authoring tools are still platform-dependent, although they are closer to the goal of independence than editors or hypermedia tools. It is worthwhile to mention Kaleida Labs, a joint venture between Apple Computer and IBM for multimedia technologies. One of their major projects is *ScriptX*, a universal language that will allow multiple digital platforms to play the same digital file without modification. Because different platforms have different features and capabilities, the *ScriptX* run-time environment includes *dynamic adaptation* that allows an application to query the environment of the current platform and decide in real-time how it can best present itself. *ScriptX* is fully object-oriented with the capability for the user to combine objects at run-time, similar to MHEG, described in Chapter 13. Authoring tools might use this language to become more platform-independent.

Several authoring products are currently available which help to develop applications, such as:

- Information delivery applications can be developed using the authoring tools *Mediascript OS/2 Pro* (Network Technology Corp.), *IconAuthor* (Unisys), *ToolBook* (Asymetrix / ADI), *Authorware Professional*, IBM's *InfoDesigner 2* and others.
- Professional presentations can be developed using presentation authoring tools such as *PowerPoint* (Microsoft, Inc.), *FreeLance Graphics* and *Harward Graphics*. All the tools provide many features for enhancing the presentation by adding professional styles, images, graphics, audio, video, animation or charts.
- QuickTime movies and interactive projects can be created by movie authoring tools such as *MovieWorks* (from Interactive Solutions, Inc.). MovieWorks has several advanced capabilities: first, MovieWorks allows the user to create and edit objects in text, sound and paint; second, the user determines the project's look and feel; third, the user uses the *Composer* component to integrate the objects into a scene, and to add animations and special effects (scaling and transitions); and fourth, scenes can be linked together, either sequentially or interactively, to create a project.

17.5 Media Communication

Media communication denotes applications which exchange different media over a network via *tele-services* (e.g., video conferencing, cooperative work, mailing, etc.) to multimedia application end users.

The advantage of tele-services in multimedia applications is that the end users can be located in different places, and (1) still interact closely in a quite natural way or (2) operate on remote data and resources in the same way as with local data and resources. The disadvantage (currently) is that the delivery time of the tele-services is longer than the processing time of local multimedia applications. For example, the retrieval of a video clip takes longer when the information must be retrieved from a remote video server than if it is located on the video disc of a local computer. Therefore, the tradeoff between location and time needs to be kept in mind.

In the following section (17.5.1), we briefly describe *tele-service* mechanisms. In Sections 17.5.2 through 17.5.6 we present some implementation architectures for tele-services, and in Section 17.5.7 discuss some tele-applications.

17.5.1 Tele-Services

Tele-services are services provided by communication systems which are based on and make use of audio and video data. With current networks and the further development of high-speed networks, technology will enable distributed multimedia applications which need tele-services.

In this section we concentrate only on basic communication paradigms, such as tele-interaction and retrieval of information, without going into the details of communication systems. The details and basic principles of communication systems with respect to multimedia are described in Chapters 10 and 11.

Interactive Services

Interactive services include an exchange of control data between remote sites to influence the presentation of continuous media data. Communication between the sender and receiver can be performed either *synchronously*, which means that data arrive with a well-defined end-to-end delay, or *asynchronously*, which means that data arrive at any time. For example, a video conferencing application uses synchronous communication when remote conference participants are viewing a speaker. Mailing systems use asynchronous communication.

With respect to their task, interactive services are roughly divided into *conversational services* (e.g., used in video conferencing), *messaging services* (e.g., used in mailing systems), *retrieval services* (e.g., used in document retrieval systems), *tele-action services* (e.g., used in banking systems) and *tele-operation services* (e.g., used in tele-robotics systems). We will briefly describe the communication behavior of each service below:

- *Conversational Service*

  A conversational service supports *conversation* between remotely located end users. The service uses a two-way communication connection between the sender and receiver. The bi-directional delivery of multimedia is done in synchronous mode. Another feature of a conversational service is the time of the data delivery. The data must be delivered quickly (i.e., end-to-end delay must be minimal) in both directions so that conversation among the users can flow smoothly and in real-time without disturbing the human perception of a dialogue.

- *Messaging Service*

  A messaging service provides an exchange of messages between a sender and receiver where the end users are human users. The exchange of messages in both directions is done asynchronously, such that the time of delivery can be pre-determined. Since the end users can send messages whenever they like, fast delivery time is not required (i.e., end-to-end delay does not have to be minimal). The content of a message, a mail, can include all kinds of media data.

- *Retrieval Service*

  A retrieval service provides an exchange of messages between a sender and receiver, where the sender, also called the *client*, is a human user and the receiver, also called the *server*, is a computer with database provision. The client requests information from the server, where the information is stored; the server retrieves the information and sends it back.

  There is two-way communication, but the communication has the following characteristics. From the client to the server, the communication is asynchronous because the request from the client does not comply to timing constraints. The communication from the servers to the client may require a synchronous or asynchronous mode of operation depending on the information retrieved. If the retrieved information includes continuous media, the delivery must be synchronous. For example, if a user requests a movie from a remote video server, the video and audio delivery must occur according to stringent timing constraints. If the retrieved information includes a text file,

General Categories	**Applications**
Transaction Processing	credit cards, lottery, automatic teller machines, medical insurance claims
Alarm and Surveillance	burglary, fire, smoke, medical, disabled persons, environmental surveillance
Business Automation	information access, data processing
Utility Resource Management	automatic meter reading, time-of-day rate information
Control and Command	appliances, thermostats, lights, vending machines, industrial equipment monitoring, hospital equipment
Interactive Video Support	home shopping

Table 17.1: *Tele-action services [Sco94].*

then the delivery can be asynchronous. There is a strict requirement on the reliability of the transmitted data, while the requirement on the delivery time is relaxed. Although, to provide an acceptable user service, the delivery time (response time) from the server to the client should be minimal.

- *Tele-action Service*

  Tele-action means to act at a distance [Sco94]. Such actions include reading or writing some information to a remote location, or possibly both. Instead of sending a person to write or read information on a device, it is done remotely via a communication network. This form of data collection is useful when coupled with a computer system that can use the data to initiate an action (e.g., generate a bill). The tele-action services can be further classified with respect to their relation to the industry as shown in Table 17.1.

  - *Transaction Processing*

Transaction processing services perform business transactions. They remotely check databases for available funds (possibly transferring the funds) and print receipts. Some applications which use these services are listed in Table 17.1.

– *Alarm and Surveillance*

Alarm and surveillance is a service that monitors factors affecting individual or public safety.

– *Business Automation*

Business automation services provide information access to medical or legal databases, real state listings, etc. Furthermore, they can be used for management of databases for individual distributors and suppliers.

– *Utility Resource Management*

Utility resource management services provide better management and a distribution of resources, such as gas, water and electricity. These services help, for example, a utility company track and control consumption, as well as monitor safety factors.

– *Control and Command*

Control and command services, often called *tele-metry* services (telemetry means to measure at a distance), are used to remotely read data and monitor the status of remote control equipment.

– *Interactive Video Support*

A growing number of industries are trying to develop services that offer smart control of a television set. In particular, there is interest in services that provide interactive control of video services.

- *Tele-operation Service*

Tele-operation services have bi-directional communication. They allow the user to perform a task at a distance, typically by manipulating a master controller that causes a slave effector to move remotely mimicking the master controller's movement.

Distribution Services

Distribution services are services for the distribution of information to different remote sites. They are one-way communication from the broadcasting source to the remote destinations. For example, TV broadcasting or radio broadcasting use distribution services. There are two kinds of subservices: *distribution services without individual user presentation control* and *distribution services with individual user presentation control.*

The development of these services continues to move towards interactive services [Ran94]. The reason is that with the new technology of *video on demand* over cable TV networks, several types of control have been and will be given to the viewer:

- *Pay-per-view*

  Cable subscribers can order movies and programs using today's *set-top decoder box*. However, these set-top boxes do not give the viewers any freedom other than the choice off whether to view the program or not. The viewer has no control over the movie shown.

- *Near Video-On-Demand*

  This service is achieved by having many channels broadcast the same program, but with a definite temporal variation or delay between the channels. With this approach, the viewer can simulate forward and reverse functions by changing channels appropriately. One suggestion for the time delay is to start the same video clip every ten minutes. This gives the user some control over the time at which (s)he can view the movie. In the case of a live program, a *deferred airing* has been suggested to give the user the possibility: (1) to decide if (s)he wants to watch the program, and (2) to see the live program later in case the user missed the first specified broadcasting time. *Deferred airing* is a concept in which live programs are broadcast after a specific time delay.

- *True Video-On-Demand*

  This service provides the functions of a VCR and hence gives the user complete freedom to temporally alter the viewing. Further steps are *Interactive TV* and

Cyber Vision, where the user is involved in the content and can at least make choices of how the movie will proceed.

The first two types of services do not require many changes in current cable TV networks. The third service will require a switching system to be installed to support bi-directional signaling.

17.5.2 Implementation of Conversational Services

Conversational services are implemented as tools like multimedia conferencing, video-telephony or computer-supported cooperative work. These tools are then used, for example, in a group of applications called *tele-working.*

Video Conferencing

Tele-conferencing systems allow the user to achieve most of the efficiency and productivity of traditional meetings with one main difference: the user can stay at his/her desk as can the remote conference participants. A multimedia conferencing system enables people to work together across geographically distant locations without the need to meet at one site. They communicate among each other in multi-party or face-to-face mode using motion video, audio and textual information in each direction. The audio and video quality heavily depends on the platform. Therefore, a big factor in the success of a tele-conferencing system is to achieve high media quality over any platform and interconnectivity among various platforms and vendors. A possible setup of a video conferencing system is shown in Figure 17.3.

Video conferencing is used either in an office environment, where the video is displayed on a PC or workstation screen, or in a conference room, where the video is displayed on a *video wall* (large TV screen). For the office environment, desktop video conferencing systems have been developed. The name suggests that the PCs or workstations in offices are placed at users' desks so that they can easily communicate any time. For a conference room environment, large TV screens in conference rooms are used for meetings of groups located at different geographical places. We will discuss some of the conference room solutions in Section 17.5.7.

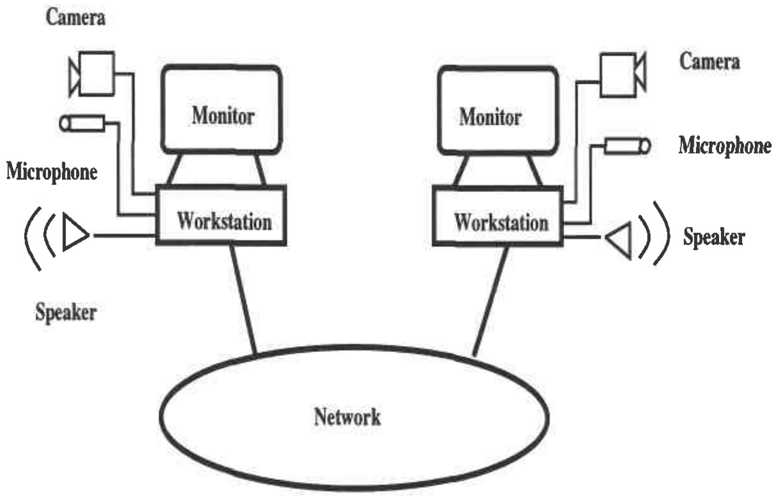

Figure 17.3: *Video conferencing system.*

Desktop video conferencing systems often include a dedicated shared white-board application (i.e., drawing and writing software support for multiple users) [MR94]. *Application sharing* denotes techniques which replicate the user interface of the particular software (e.g., the user's favorite text processor) so that the software can be used simultaneously by all participants of a conference. The concurrency in the activities underlies the mechanisms of "floor passing" (also called "chalk passing") to determine which one of the users may actually interact with the software at a given time (see Chapter 11).

Some examples of conferencing tools are *vat* for audio conferencing and *nv* for video conferencing running on SUN workstations, and BERKOM's *MMC (Multimedia Conferencing)* on a network of UNIX-based machines such as IBM's RISC System/6000, SUN, HP and further workstations [KW93a].

Video-phone with Conversational Service

A video-phone is basically a sophisticated telephone with a screen for the presentation of the caller(s). A video-phone with conversational services supports *video-telephony* applications. Video-telephony is used for telephone meetings between two or more persons in which image transfer may be voice-controlled, such that the speaker is seen by the others. A possible architecture is shown in Figure 17.4.

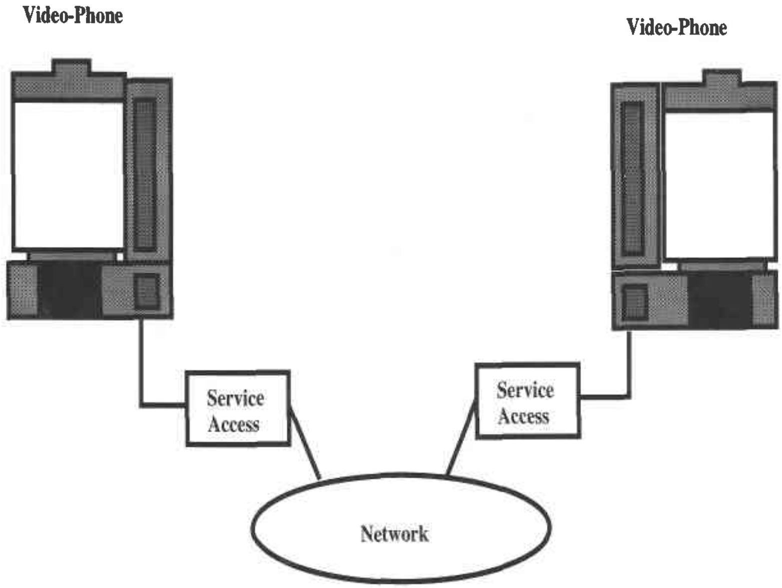

Figure 17.4: *Video-telephony system.*

Computer Supported Cooperative Work (CSCW)

The current infrastructure of networked workstations and PCs makes it easier for people to cooperate. The cooperative work done in this environment is called *Computer Supported Cooperative Work* (CSCW). CSCW systems allow several people at different locations to work on the same data, most often a document.

CSCW systems are divided into *asynchronous CSCW* and *synchronous CSCW* systems. *Asynchronous cooperative work* specifies processing activities which do not happen at the same time. A typical example is a health insurance claim for coverage of surgery expenses processed separately by persons responsible for different functions in the insurance company. *Synchronous cooperative work* happens at the same time. The fast development of multimedia has established video conferencing as an integral part of CSCW systems.

Systems that support group collaboration are called *groupware.* The essence of groupware is the creation of a shared workspace among collaborators and it is often used as a synonym for CSCW. Groupware may consist of video conferencing, together with shared computer-based applications (e.g., shared editors, white-boards). If groupware is supported in real-time, it belongs to the area of *synchronous tele-collaboration.*

Commercial CSCW applications still include few audio-visual components and are built as only video-telephony, video-conferencing or desktop-conferencing applications.

17.5.3 Implementation of Messaging Services

Messaging services are used in electronic mailing systems. A number of extensions to the functionalities of electronic mail have been implemented to allow the exchange of multimedia messages. Some examples of electronic mail prototypes are DARPA's experimental *Multimedia Mail System, a Distributed Inter-office Mail System* or *Diamond Mail* from BBN.

We will present two approaches implementing multimedia mail systems: *MIME (Multipurpose Internet Mail Extension)* (Internet Standard) and *the Multimedia Mail Tele-service* based on CCITT recommendation X.400 (88) [Sch94b].

MIME

MIME is an extension of Internet Mail defined by the Internet Engineering Task Force (IETF) working group [BF93]. It offers a simple standardized way to represent and encode a wide variety of media types. MIME messages can include seven types of media – text, images (image/gif, image/jpeg), audio, video (video/mpeg), message for encapsulated messages, multiparts for including multiple types of data in a single message and application data (application/PostScript).

Various strategies for MIME implementation are possible. A flexible approach is to use *metamail.* It is a simple program which is called when the mailing program gets non-textual information. Metamail does not understand any MIME data type, but instead knows how to read a set of configuration files, called *mailcap files.* The mailcap program, handling the mailcap file, recognizes the data type and calls the particular program for viewing the message. For example, if the message includes an image, mailcap may call the image viewing program *xv.*

Multimedia Mail Tele-service

Multimedia Mail Tele-service, developed within the BERKOM project [RK92], is based on the X.400/88 CCITT recommendation. This service supports message types such as ASCII text, PostScript, teletext and other *body part* types of the X.400 standard. The BERKOM profile defines externally-defined body parts for audio, video and image as new information types.

The mailing system is implemented as follows: the X.400 *Message Transfer System* (MTS) delivers messages submitted from either a *Multimedia Mail User Agent* (MM-Mail UA) or a *Global Store Server (GSS)* to one or more recipient UAs or MSs and returns notification to the originator, if requested. The transfer system is based on the *store-and-forward* principle. The user agent includes composers, editors, viewers or converters for multimedia messages. The storage server stores multimedia components and makes any data, especially high-volume data, accessible worldwide. It can be considered as a public or private service for the temporary deposition of bulk data in a global network. Major components of this multimedia mail system are shown in Figure 17.5. Unlike text messages, the size of multimedia messages may range from a few Kbytes to many Mbytes. Messages of many Mbytes might be too large for MTS, as well as for the storage capacities at the recipient's site. A possible solution is to include *references* to large message components within the message rather than include the contents. These references are like pointers to a remote store that can be accessed by originators and recipients using specialized protocols for data transfer.

17.5.4 Implementation of Retrieval Services

Multimedia retrieval systems allow users to access large multimedia storage and databases, located at servers, and to view this information. The servers are shared

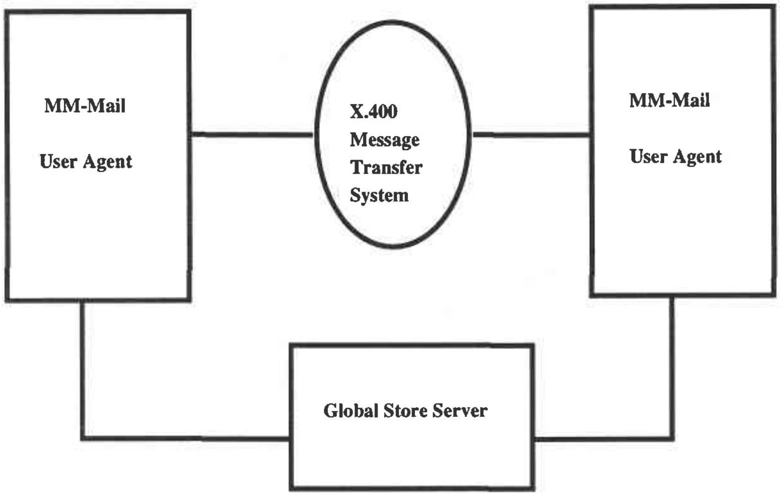

Figure 17.5: *Multimedia mail system.*

among the clients. We will describe a multimedia retrieval system with the focus on video retrieval. Multimedia retrieval services are used in applications such as the World Wide Web and Internet Gopher, which we briefly describe.

Video Server

The main types of information stored in video servers are movies and other video information. It is more economical to store videos on central servers because of the considerable amount of valuable disk space they consume. To retrieve video data from the central server over today's computer networks, a software, which carefully manages resources in networks, end-systems and routers, is necessary. Client workstations and PCs, which are used to retrieve and display data, are connected to the video server using, for example, local area networks.

The communication protocol to retrieve video from a video server is as follows: the user issues a request, which may include the name of the video; the server sends the required video over the network and the video is then presented in a client window. A similar protocol applies to audio and other media. All multimedia data should be retrieved from a server in real-time. A possible video retrieval system setup is shown in Figure 17.6.

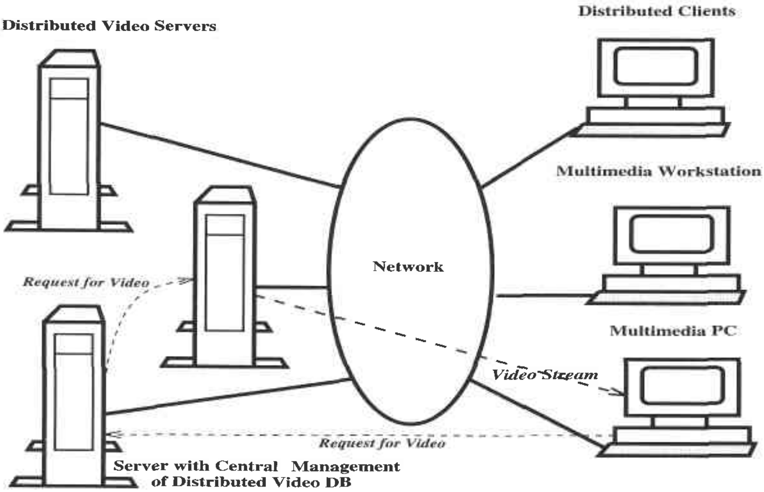

Figure 17.6: *Multimedia retrieval system.*

Examples of video server prototypes are:

- The *Ultimedia Server/6000*, which allows users to store, record, share and retrieve multimedia data using IBM Heidelberg's HeiTS communication technology.

- The V^3 *video server* [RM93], which allows a user to interactively store, retrieve, manipulate and present analog and short digital video clips. The picture and sound data of a video clip are stored in a database in digital form or on a laser as analog streams. The conventional properties of database systems, such as queries, multiple-user access and recovery, are supported. Digital and analog data are transported via digital and analog networks, respectively. Video clips can be accessed by methods that initialize, present and control them. For example, the presentation of a clip can be started by the operation *play* and interrupted by the operation *stop.* Due to the storage of analog video clips on laser discs and their WORM (Write Once Read Many) characteristics, picture and sound data cannot be edited. The video data of a digital clip, modeled as a sequence of frames, can be manipulated by inserting, cutting and moving parts. The server system is developed on top of the VODAK database management system. VODAK is a prototype of an object-oriented and distributed management system developed by GMD-IPSI, Darmstadt, Germany.

Retrieval services are widely used, for example, for the preparation of *news broadcast*, storage and retrieval of multimedia data about accidents for *insurance purposes*, *employee information systems* and *digital libraries*.

World Wide Web

World Wide Web (WWW) is a system that allows *clients* to move through many different remote *servers* throughout the world and retrieve hypermedia documents.

It has a body of software (clients, servers, gateways, library, tools) and a set of protocols and conventions such as: a *WWW Server* is a program, like *ftp server*, that responds to an incoming caller request and provides a service to the caller. An example of a WWW server is the CERN (the European Laboratory for Particle Physics) server or the NCSA (the National Center for Supercomputing Applications) server. *Gateways* are servers that provide data extracted from other systems. For example, the *VMS Help* gateway allows any VMS help file to be made available to WWW clients. The *Common WWW Code Library* is the basis for most WWW browsers because it contains network access and format handling. *WWW and HTML (HyperText Markup Language)* tools are parts of the available WWW software whose purpose is to manage WWW servers, generate hypertext, view retrieved information, etc. For example, the user can view a hypermedia document (a HTML document) with the *NCSA Mosaic*©tool for X Windows running on SUN workstations.

Internet Gopher

The Internet Gopher is a client/server-based worldwide information delivery service from the University of Minnesota. The user can search and retrieve various information, using the *Internet Gopher protocol*, such as newspapers, newsletters, weather forecasts, phone-books, libraries, information about universities, etc. This information can include media types, such as text, sound and/or image. Furthermore, gopher objects can be files (e.g., UNIX uuencoded files, BinHexed Macintosh files), images (in GIF format), MIME information (multimedia mailing format) or HTML documents (hypertext format).

17.5.5 Implementation of Tele-action Services

We will briefly outline two possible architectures for the implementation of tele-action services. One architecture demonstrates the implementation of utility resource management and the other architecture demonstrates the implementation of an alarm and surveillance system.

Message Switch and Store for Utility Resource Management

A possible (and quite common) setup for the implementation of utility resource management is shown in Figure 17.7. The system is based on a *Message Switch and*

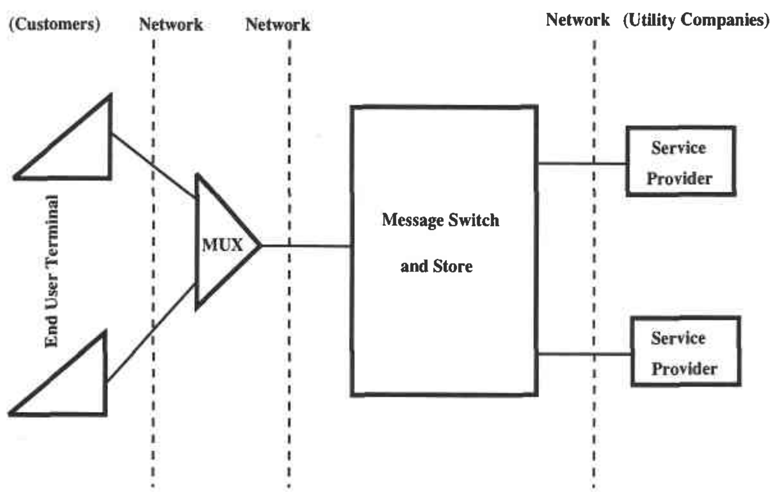

Figure 17.7: *Message switch and store for utility resource management.*

Store device that provides much of the functionality of the message delivery service.

It is a key component because it collects messages from the multiplexer and sends them to the service provider, which is the utility company. The multiplexers are typically located in the local exchanges and are remotely controlled by the *message switch and store* device. Permanent monitoring and polling of the end user terminals (e.g., tariff meter, pulse meter) is performed by the multiplexer.

Remote Camera Control for an Alarm and Surveillance Service

One possible implementation of an *alarm and surveillance* service is a *remote camera control system* and it is used in areas such as: *production monitoring*, *computer-integrated manufacturing* or *monitoring of security areas.* The architecture of such a system is shown in Figure 17.8. The camera is located at a remote site in need of

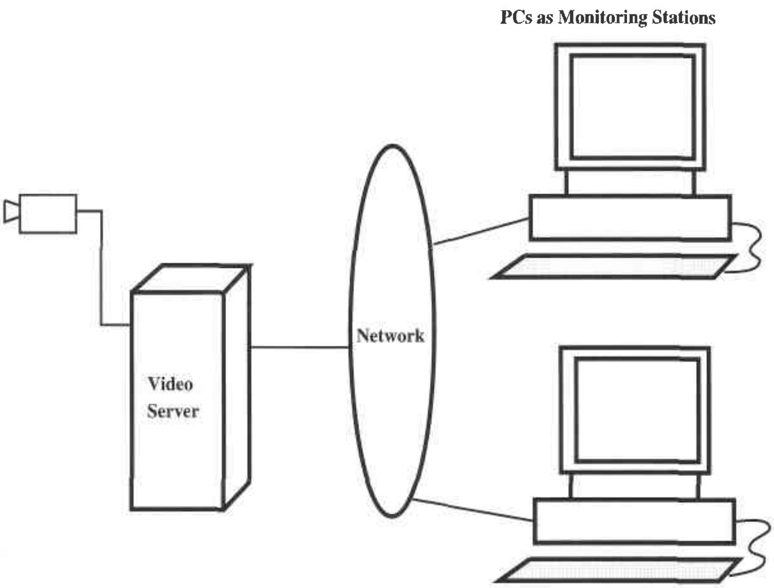

Figure 17.8: *Remote camera control system.*

monitoring. It is connected to a video server which digitizes the motion video and sends it over a communication network to a user (observer). The user views the motion video on his/her PC or workstation [WSS94].

17.5.6 Implementation of Tele-operation Services

Tele-operation used, for example, in a tele-robotics application presents challenges quite distinct from tele-conferencing. It allows a remote operator to exert force or to impart motion to a slave manipulator. The operator experiences the force and resulting motion of the slave manipulator, known as *kinesthetic feedback.* An operator is also provided with visual feedback and possible audio feedback as well. An example of such a system is shown in Figure 17.9 [NS93].

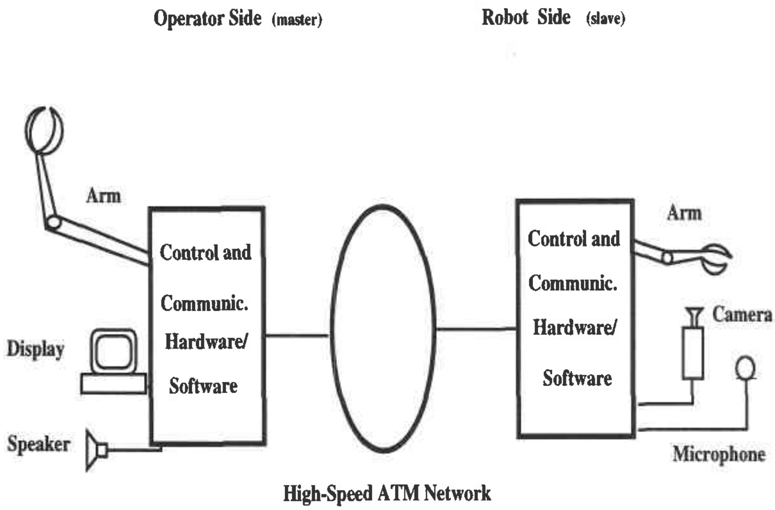

Figure 17.9: *Tele-robotics system.*

17.5.7 Applications of Tele-services

The services discussed above are used in our society in many ways. One group of applications that use these services are *tele-working* applications. People can stay at their work or home and connect with other people for different purposes. Tele-working includes tele-activities such as *remote education*, *offices in the home*, *tele-collaboration* and *tele-medicine.*

Remote Education

There are two remote education scenarios:

- The first scenario is a group of people sitting together in one location with the necessary equipment needed for their remote education. The remotely located teacher gives a lecture to the group. The visual equipment in the classroom (i.e., room where the group is gathered) is most often a *video wall.* A video wall is a large TV screen connected to a computer for video reception and transmission. The audio can be transmitted either through a telephone or together with the video signal.

  This type of education can use video-telephony, video-conferencing or just live (cheap) video can be transmitted without any conference management. If no conference control is provided, the teacher and students control the communication as in a normal class situation.

- The second scenario is individual tutoring, where there is no need for the students to be located together (in space and often also in time). Both the teacher and student have a PC with a monitor, keyboard and other devices. A video-phone can also be added to allow interactive conversation. The teacher may distribute and receive papers to and from students using electronic mail.

Examples of remote educational systems are:

- *BETEL* and *BETEUS* – In the *Broadband Exchange for Trans-European Links (BETEL)* project, high-speed links (ATM-based, cross connected) were provided by France Telecom, Swiss PTT and Alcatel to demonstrate multimedia remote tutoring services (and further advanced services). The EPFL (Lausanne) and Eurocom (Sophia-Antipolis) jointly provide a tele-teaching application [BDG+94]. As a successor of BETEL, the *Broadband Exchange for Trans-European USage (BETEUS)* project supports the interconnection of six locations in Europe through the ATM pilot network. In the summer of 1995, this network will be used for tele-teaching among four sites [BDG+94]. Further research at EPFL, under the guidance of J.P. Hubaux, will be devoted to an open architecture for advanced multimedia services over ATM with emphasis on network management issues.

- *CO-LEARN* – a multimedia system for distributed teaching and learning with participants at different sites (PCs running MS Windows or UNIX workstations) connected by an ISDN network [Hau93]. This system offers four scenarios: tele-teaching, real-time tele-assistance, real-time multimedia conferencing and an asynchronous forum for exchange of multimedia and learning material.

- *Tele-mentoring System* – a system for interactive distance learning using the Asynchronous Transfer Mode (ATM) network of the AURORA Gigabit Testbed [SDF+94]. The experimental trial for distance learning used tele-conferencing hardware (Video-Window from Bellcore), which converts NTSC television and audio signals to and from ATM cells. This hardware connected Bellcore's and University of Pennsylvania's video walls with other apparatus to create a realistic two-way interaction.

Tele-office

Currently, employees can work at home and still access various information sources and communicate with their colleagues, management and others via telecommunication networks. A home office might consist of a telephone, PC, printer (the minimal configuration), video-phone, fax (the advanced configuration) and telecommunication services such as conversational services, retrieval services and/or messaging services.

Tele-collaboration

Tele-collaboration has become an important part of our working environment. To make this service successful, a *media space* [BBI93] – a collaborative work environment – is created. A *media space* may consist of a network supporting the transmission of audio, video and other data, fixed connections (local or remote) to employees' offices, a crossbar switch that links cameras, monitors and microphones in each office and computers that allow switched access to persons in each office. The possible architecture of a media space is shown in Figure 17.10.

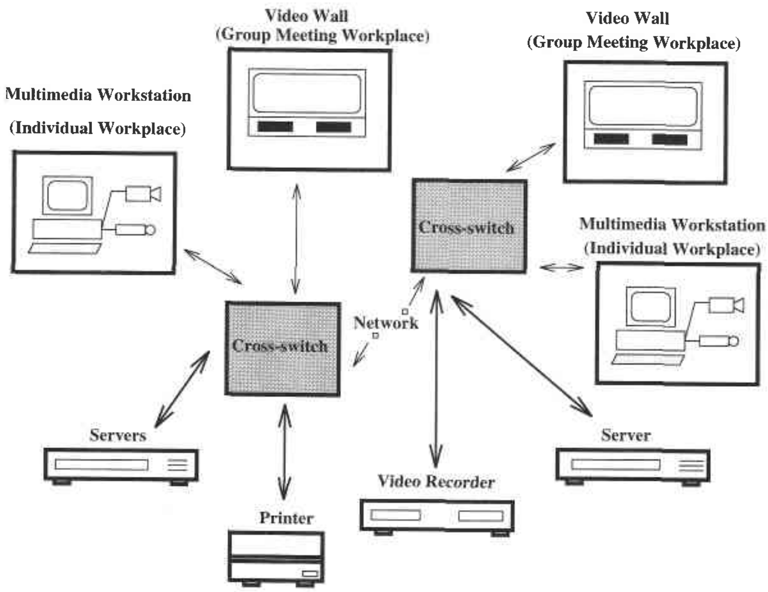

Figure 17.10: *Media space as the example for tele-collaboration.*

In addition to the hardware, a media space may support conversational service capabilities for one-on-one conversations (e.g., video-telephony, e-mail), many-to-one conversations (e.g,, bulletin board newsgroups, e-mail, tele-conferencing), many-to-many conversations for group meetings (e.g., video-conferencing, groupware), shared applications (e.g., white-board) and recording and retrieving video records.

There are several research projects that have implemented a media space into their working environment and have included trials to evaluate the effectiveness of the media space:

- *Media Space at Xerox PARC*

  This media space uses cameras, monitors, microphones and computing to connect employees' offices. Media space in this project is geographically split between Palo Alto, California and Portland, Oregon. It is important to note that during the trial, the media space was constantly in use, functioning like an extension of physical space. It was not something that was turned off or on during the day, but was continually available. The office cameras were most often open and in one of three states: *open and on* (the user is visible); *open but focused on some nonuser location* (friendly but not personally visible); and *closed* (lens cap was on and/or the camera was off). Microphones were often off, but could be switched on quickly as needed. The media space was mostly used for project and lab meetings.

- *VideoWindow at Bellcore*

  The VideoWindow system provides a large screen display with live audio between two public areas on different floors of a research lab building. During the trial, the system was on 24 hours a day for three months to support informal interactions among researchers and staff. People typically came to the area for some specific task (e.g., to pick up mail) and could engage in conversation with others who happened to coincidentally appear either in the same space or in the space provided by the VideoWindow.

- *Cruiser at Bellcore*

  Media space in the Cruiser system [FKRR93] connects research offices for informal interactions. It uses a desktop video-telephony system. Cruiser is based on the model of walking down a hallway and popping one's head into a doorway. During a cruise, either the observed or the observer have the option of changing the cruise activity to a two-way conversation activity.

There are also other research projects, such as CAVETAT/Telepresense from the University of Toronto, RAVE from Rank Xerox EuroPARC, Kasmer from Xerox PARC, and TeleCollaboration from US West Advanced Technologies.

Tele-medicine

Tele-medicine services address elderly, sick or disabled people who cannot leave their homes. Over the telecommunication network, these people can consult their doctor (tele-diagnosis) and get medical information and other administrative health information. Tele-diagnosis uses a conversational service implemented, for example, through video-telephony. Access to medical and health care information can be achieved through retrieval services.

Tele-surgery allows one to consult a specialist on demand for crucial and difficult operations performed at local hospitals. The easiest way to implement this application is to provide remote camera control with conferencing capabilities.

17.6 Media Consumption

Media consumption is the act of *viewing, listening or feeling multimedia information.* Viewing and listening are the most common ways users consume media. Feeling multimedia information can be experienced in motion-based entertainment parks, for example, through virtual reality. This is still a very new area in computer science and there are few results. Therefore, we will discuss it only briefly in Section 17.7 and Chapter 18.

The major emphasis of this section is on viewing multimedia information (presentation). Presentation of multimedia information is often done through authoring tools, as well as by other tools. One major requirement of media consumption, which needs to be considered by each public multimedia application, is the *simplicity of presentation.* Especially, when user interfaces with new media are introduced, it is important to convince users to consume them because people like to do things the same or in ways similar to how they used to do them in the past. Therefore, (1) familiar human user interfaces must be created, (2) the users need to be educated, and (3) the users need to be carefully navigated through the new application.

We will analyze some design issues with respect to viewing multimedia documents and then discuss some applications where proper media consumption is important (e.g., books, electronic newspapers, kiosks).

17.6.1 Viewing Multimedia Documents

Multimedia documents can be viewed ("consumed") in two modes: by *browsing* and by *detailed media consumption.*

Browsing means that the user goes quickly through the document to get an overview of what the document includes. For example, the user may just read the titles of articles in a newspaper, table of contents of a book or a brief abstract of a scientific article. To browse through text, graphics or image information, the viewing tool provides a *menu* (e.g., table of contents), *sliding bars* on the side of the window where the document is displayed in a *highlighted phrase* (colored or underlined), *graphical icons* (e.g., arrow icon, audio icon), *search functions* or *small images.* Highlighted phrases, graphical icons and small images are hyperlinks to other information. For browsing through video information, the tool provides functions such as *move forward, move backward, search for a certain scene* or *play a short video clip.* For audio information, the tool provides analogous functions such as *move forward, move backward, play a short audio clip* or *search for a certain song/musical passage.*

Detailed media consumption means a detailed reading, viewing or listening of the multimedia entity. This mode requires functions such as *display a document, quit, play video/audio, stop video/audio.*

Many viewing tools provide additional functions which help the user view multimedia information:

- *Navigation* includes prespecified steps to view a document. The *navigate* menu may, for example, specify how to get to the beginning of a document, bibliography or other specific part of the document.
- *Annotation* allows the user to add personal annotations to any document during the viewing process. Annotations in the *Mosaic* tool are inlined as hypertext links at the end of the document. The annotation can be *personal, public* or *workgroup*-oriented.

Section 17.5.4 already discussed the implementation of retrieval services. We now extend the discussion with respect to consumption issues by the two following examples.

Music Consumption

Music consumption is widespread among PC users. Current music tools offer various functionalities for working with MIDI and WAVE audio data.

Music tools provide opportunities to: (1) learn the basic theory necessary to understand music (MIBAC's Music Lessons), (2) create a powerful recording studio (OpCode Systems' Musicshop), (3) manipulate sounds and CD audio tracks (OpCode Systems' Audioshop), (4) compose, play, record, edit and print out music on a PC (Passport Designs' MusicTime), (5) mix and edit MIDI instruments (Master Tracks Pro 5), and (6) learn to play musical instruments (e.g., The Pianist, The Jazz Guitarist).

NCSA Mosaic

NCSA Mosaic is an Internet-based *global hypermedia* information browser and World Wide Web client developed at the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign.

A single-click with the left mouse button on a hyperlink (i.e., a highlighted phrase or icon) causes Mosaic to retrieve the document associated with that link and display it in the *Document View* window. A single-click with the middle mouse button also causes Mosaic to follow the link and open a new Document View window on top of the existing one. From that point, either Document View window can be used for further navigating and viewing of hypermedia documents.

17.6.2 Books, Proceedings and Newspapers

Books, proceedings and newspapers can be interactive multimedia documents which may be electronically distributed to the home. The user may either print the data or navigate through the information on some computer. Instead of simply broadcasting the same newspaper to all readers, the user accesses the electronic versions of newspapers, magazines, book, etc. The access can be customized according to individual profiles. This approach saves paper, has a potential for personal selection and provides fast delivery. It means that news can be written into the server as soon as it arrives in the news studio and the reader gets the most recent information. Despite the advantages, experiments have shown that customization is not always wanted and readers often tend to prefer paper over electronic versions.

17.6.3 Kiosks

Recent technological advances have made possible the high-quality delivery of video and audio integrated into the desktop computing environment. This capability, combined with the increasingly common use of digital information acquisition and storage, provides an opportunity to create public information services known as *multimedia kiosk systems.*

Kiosk systems are often located in public areas, accessible to visitors or customers. The kiosk are controlled by a computer that allows the user to interactively control the information or service (s)he wants to obtain. Since kiosk customers might be unskilled users, the user interface must be simple and easy to handle.

A further requirement for a kiosk system is that response time must be short. At present, this can be achieved only if the kiosk system is local or connected to its server by a local area network. The kiosk applications use the retrieval services (query mode) or some tele-action services for communication with the servers.

The equipment constituting a multimedia kiosk system might take a variety of forms. The hardware requirements include a processor, storage device, display, speakers and a touch screen. An advanced configuration could also include a video disk player, high-speed network connection, high-resolution screen, keyboard, printer and camera, plus any other hardware needed for the specific application, such as a device to accept money and dispense change.

Applications of information-providing services within the realm of multimedia kiosk systems include *airport or train station kiosks* with maps of terminals, arrival/departure times and gate numbers; *museum showcase kiosks* with preview information of forthcoming exhibits and schedules of forthcoming attractions; *bank assistant kiosks* with information on banking products and worksheets for planning savings accounts; *cinema information kiosks* with information on times and places of movies, selected clips from movies and movie trivia; *retail store kiosks* with information on product highlights, special pricing, and store layout; and *real-estate catalog kiosks* with information on real estate categorized by price, location, information about schools, libraries, stores in the neighborhood, pictures and videos of objects.

Some examples of kiosk systems as information manipulators are: a *ticket counter* that provides reservation and purchase of plane/concert/etc. tickets, and interactive seating according to a floor plan; a *bank teller* that supports sale of life-insurance, transfer of funds, and tracking of investments; an *education* system that supports an on-line student interaction with lessons and immediate feedback/review (learning tools); a *cooperative work* system that supports team development of document draft, dynamic work assignment and status reports.

17.6.4 Tele-shopping

Multimedia tele-shopping enables users to shop from their homes. For example, a household installs a PC and a tele-service (retrieval service) to set up a connection to a database or multimedia catalogue. The service allows the user to search for different products from the catalogue. The products may be presented either with video and sound or as a text accompanying still images. Analogous to kiosk systems, the user interface and manipulation of information during the viewing process must be easy to work with because of the broad range of users. A product may also be ordered and paid for electronically (i.e., tele-action service).

Examples of tele-shopping applications are: home ordering/shopping of goods, ticket reservation (theater, cinema, concerts, shows, travels, etc.) and advertising with multimedia.

17.7 Media Entertainment

Virtual Reality entertainment (VR), Location-Based Entertainment (LBE), motion-based simulators, large-screen film and games (based on interactive audiovisual support) are applications that use multimedia for entertainment and bring a different and more involved entertainment experience than what is available with a standard TV or movie theater.

17.7.1 Virtual Reality

The term *Virtual Reality* (VR) promises far more than our technology can currently deliver. It has been variously used to describe user interfaces ranging from synthesized physical environments presented on Head-Mounted Displays (HMDs), to ordinary graphics displayed on conventional CRTs, to text-based multi-user games.

Computer-based VR systems are three-dimensional, interactive as opposed to passive, and use one or more devices in an attempt to provide the user with a sense of presence, be it visual, auditory or tactile[HPC94, Ear93]. Among these devices are head-tracked displays and stereo displays (both visual and audio), hand trackers and haptic displays (devices which provide force feedback).

The first VR systems appeared before computers were used for VR. Morton Heilig developed a machine called the Sensorama [Hei62], which involved all of the senses except taste in a virtual motorcycle ride through Manhattan. Early flight simulators also created virtual environments without the aid of computers. They used movies or created live video by shooting model boards with TV cameras [Sch83].

Currently, the hardware platform of virtual environments consists of color stereo HMDs, haptic displays, spatial sound, data gloves and 3D graphics [Bro88, WF90]. The software architectures for virtual environments have been developed to support a single hardware platform or a small number of tightly coupled platforms. As a result, systems were originally modeled after traditional interactive programs. Current systems still ignore issues that would arise if such applications were used on a larger scale, as will be required for the real world. The first virtual environment applications were simple event-loop-based programs. There are several problems with this approach because the following requirements need to be satisfied:

1. VR displays should respond to changes in tracked objects, especially the user's head, at least ten times per second for the virtual environment to be convincing

[FD94]. A solution to this requirement is to distribute the VR system over multiple processes, decoupling the simulation steps from the redisplay loop.

2. VR systems should not have tightly coupled distributed processes because this approach does not scale towards new hardware and system software solutions. A solution is to use structured applications (modular approach), i.e., applications structured as a large set of asynchronous, event-driven processes. Each process is independent of the others and communication is performed via a well-defined message protocol. Hence, as the VR technology advances, individual modules (processes) can be independently modified.

3. VR systems should scale up gracefully. Solutions for this requirement can be achieved using adaptive algorithms, dynamic environments and adaptive protocols (e.g., the DIS protocol used in SIMNET [FD94]).

4. VR systems should have immersive, fully synthesized, photo-realistic graphical displays. The solution for this requirement is still far away because current technology still does not provide such displays. Partial solutions to this requirement might be in a graphical display with (a) rendering of fully scenes, (b) rendering of selected objects, or (c) rendering images from the viewpoint of a given user.

There are many VR systems based on various types of implementation approaches. One implementation approach is demonstrated in *MR Toolkit* [SGLS93], where a VR system decouples the simulation steps from the redisplay loop. Since the simulation determines what is displayed and the user's head position determines from which angle it is displayed, MR Toolkit distributes the VR system over multiple processes. Another implementation approach took toolkits such as *dVS* (Division) [Gri91], *VR-DECK* [CJKL93] or *DIVE* [CH93]. They implemented VR systems as a large set of asynchronous, event-driven processes. This approach allowed the system to be more easily reconfigurable at run-time and more fault-tolerant. A third approach is taken in the *WAVES* system [Kaz93], where a large-scale distribution of virtual environments over communication media of varying bandwidth is supported.

17.7.2 Interactive Video

Interactive video research addresses various problems in the area of *interactive TV* and *Video-On-Demand.* Interactive TV research concentrates on cable and public television, whereas Video-On-Demand concentrates computer-oriented television. Since both areas merge, in the future we will see the results of both areas in one *interactive video* service. We described the individual steps in Section 17.5.1 on distributed services.

Interactive TV

Interactive TV specifies that the TV viewer can become a more active participant than is the case today. There are several types of interactivity. The simplest is when the viewers can "produce" the programs they are watching. For instance, the user might select one out of several camera angles from a televised sporting event, or ask for supplementary information about the teams or players. Another example could be an educational program where one out of several educational levels could be selected and/or extra tutorials could be requested.

Interactive TV is an application that may require different types of technological solutions because the interactive programs would be too specialized to be transmitted on ordinary channels. This means that one has to subscribe either to special cable TV channels, or to a telecommunication service. Both cases require a decoder for receiving the TV signal and equipment for communicating with the TV/producer studio.

Video-On-Demand

Video-On-Demand (VOD) services represent a class of applications where video information is accessed from one or more video servers.

More generally, VOD systems include many more components that are necessary for the provision of a complete service, such as video server(s), administration and maintenance systems, networking services, backbone networks for linking geographically distributed video servers and set-top units for receiving, demodulating, decoding and converting video for television playback. Elements of a VOD system are shown in Figure 17.11 [CCP+94].

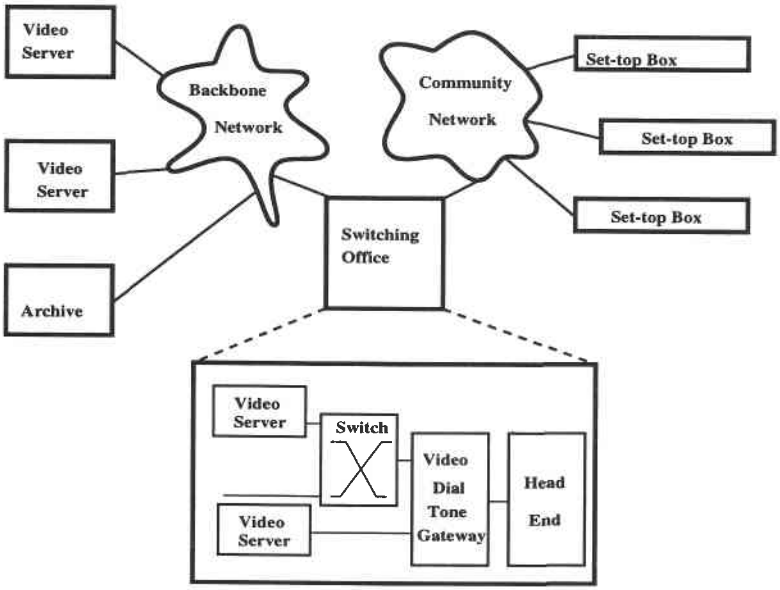

Figure 17.11: *Video-On-Demand system.*

VOD services need *retrieval* tele-services. Furthermore, the video service is an asymmetrically switched service in which the customer chooses among a wide selection of video material and receives, on-demand, a real-time response. The service is asymmetric in the sense that the downstream (to the customer) channel is much higher bandwidth than the upstream channel.

The best-known application of VOD is the *video library* which uses *Interactive VOD.* Interactive VOD allows a user to gain access to a movie (i.e., digitized video sequence stored on a storage medium such as hard disk) via point-to-point connection. This connection allows the user individual and instantaneous control of the storage medium in terms of *start, fast-forward, pause* and *rewind* actions.

There are two basic types of interactive VOD service [DVV94]:

- *Interactive VOD with instantaneous access*, whereby the user can instantly retrieve and individually control program information from a library instantly, with instant control response. The service is provided as follows: the customer selects a movie out of a large set of movies; the transmission starts within a few seconds; the user can stop and continue the transmission instantaneously; the functions *fast forward* and *rewind* are performed instantaneously, and; the user gets uninterrupted response. To provide this service, a video load buffer must be created at the start of the program so that responses to different functions can be performed immediately.

- *Interactive VOD with delayed access*, whereby users retrieve and individually control program information from a library, but there is a waiting time depending on the available bandwidth resources in the network, and/or popularity index of the requested program. In this case, the user needs to wait a few minutes before the movie starts, while (s)he still has full pause control capability. For this case, the video load buffer is only created when the function pause is performed and not at the start of the program. Therefore, this service consumes less video buffer resources and does not require fast load.

17.7.3 Interactive Audio

Similar to interactive video services, *CD-on-Demand* is likely to be established. The audio server will store music libraries, and the listeners will be able to retrieve their requested song from such a library. An example of such a system is the *LyricTime* research prototype from Bellcore [LHB92]. *LyricTime* is a personalized music system that allows the listeners to select songs (using a filter function) from a music server. The songs are played at a listener's workstation using its built-in audio capability. At the same time, a still image from the album cover is presented on the display. The listener is free to stop and start playing at any time, step forward and backward through the list of selected songs, change the volume, tell the filter function what mood (s)he is in and provide evaluative feedback on the current song.

Another interactive audio application might be the availability of *thematic audio channels* , for example, to car drivers. Over these channels, specific information (news) on road conditioning can be requested from the driver, although future development in this area is headed towards *Intelligent Vehicle Highway Systems*, where an instrumented car enters an instrumented line and the driver becomes a passenger. Already, route guidance systems based on digital maps, such as the U.S. system marketed by Oldsmobile [Col94], can pilot drivers through optimal routes with timely voice instructions.

17.7.4 Games

The modern computer game is an audiovisual engine capable of keeping an internal model of some dynamic system. Unlike a real application, a game often deliberately hides certain pieces of information about its internal state, working against the user to create a greater challenge. Games are based on *interactivity* between the user and the computer [Joi94].

Games can be divided according to storage location, environment sophistication and number of players. With respect to the storage location, electronic games can be stored on a *local* computer or a *remote* computer (tele-games). With respect to the environment, games can be placed in an *interactive environment with audio-visual components* and/or *an interactive environment with advanced technology components,* such as movies and VR groupware, etc.

With respect to the number of players, there are *one-on-one games* and *terminal-sharing games* with two players, where each player has a different input device and they play at the same time. We also experience games which make use of communication networks and allow players to interact remotely (tele-games).

Games in an Interactive Environment

One interactive game technique is referred to as *branching.* Using this technique in a game, the player experiences short, linear story segments. At the end of each segment, there are a small number of choices, each leading to a new linear segment, etc. The games use CD-ROM technology to store massive amounts of artwork, music, sound effects and animation on a single disc.

Most interactive games do not use the branching technique at all. They are based on *clever algorithms.* An example of an algorithm-based game with a rich interactivity is the highly acclaimed game *Sim City.* This product simulates the growth and development of an urban metropolis, with the player in the role of a city mayor. At each point in the game, the player can perform various operations on the landscape, such as zoning land, demolishing buildings and layering down roads, water mains and electric power lines. Since the game map is large, and since the player can perform these operations almost everywhere on the map, the number of possible moves at any given moment in time is clearly immense. Here, algorithms from the game theory, and other application domains, such as the networks' routing algorithms, genetic algorithms, neural nets and cellular automata are used.

Tele-games

The main use of tele-games is for video games and VR games. A video game menu might be connected via a network to a centralized video game machine, from which a game is chosen. A scenario is then sent from the central machine to a PC monitor in the home, and the home participant seemingly has control over one or more objects within the scenario. As the game starts, only the new coordinates of the game object and moving targets are sent from the machine to the PC at home, while the actions performed by the home participant are sent to the central machine. Instead of playing with a centralized machine, one could also play games against another home participant, or several participants, over the network.

VR games might be similar to the video game described above, but will give the user the illusion of actually being in the game by using a helmet with an HMD and headphones, so that (s)he will be surrounded by a synthetically generated environment. The user might also be wearing a data glove with motion sensors, or using haptic displays, which enhance the user's sense of natural participation even further.

Networked graphical games (like DOOM) are proliferating rapidly, using a range of technologies. Much of the current development is in 3D *Distributed Interactive Simulation (DIS).* The basis for today's DIS is SIMNET, a large-scale tactile training environment developed by DARPA (Defense Advanced Research Projects Agency) in the early 1980's. In this environment, vehicle simulators, each incorporating real-time simulation and high resolution image generation systems, interact over local and wide area networks. However, there are still considerable obstacles to achieving the goals of DIS. In the meantime, shared 3D database gaming will continue to evolve on-line (e.g., Gemie's Air Warrior) and begin to appear on LAN, in themed arcades, family entertainment centers and festival retail locations [Dod94].

An example of a WAN (Wide Area Network) game environment is the MUD (Multi-User Dungeon) game environment. Other game environments are created over commercial networks, such as ImagiNation and MPGNet, which allow users to have a fully graphical, multi-player experience in their home for an hourly fee.

17.8 Trends

We observe several trends in multimedia applications:

- Applications are going from reengineering of existing applications to establishing new application domains. The new applications may require reengineering of user-interfaces, new integration techniques, etc. The problem is that the multimedia application designers and developers do not always know what will be the future applications.

- Multimedia applications are moving from a single PC user environment to either a multi-user environment or to a personalized user environment using PowerBooks and other personalized tools.

- Multimedia applications are designed less and less for local environments only and more and more for distributed environments.

- The solutions of current applications are often platform-specific and system-dependent. The trend is going toward open solutions, so that applications are portable across various platforms.

- Media consumption is going from a passive mode of user-computer interaction to an active mode of interaction, although not every user is comfortable with this change.

- Media communication services are going from unidirectional to bidirectional information flow. Interactive-TV is the best example.

- Technical improvements and changes in multimedia applications improve productivity through better collaboration opportunities, visualization of different manufacturing processes, etc.

Last but not least, we enthusiasts in multimedia research and development must always keep in mind that even the best and most advanced system features are only visible to the user by exciting and productive applications. Going back to the roots when Alexander Graham Bell invented the telephone, the major application – also advertised at that time – was listening to a concert (and not to act as the synchronous communication vehicle as it is today). Therefore, we surveyed and categorized this plethora of applications without giving our personal impression of the "best" multimedia application domain.