Interactive Multimedia on the Web

Borko Furht and Oge Marques
Florida Atlantic University
Boca Raton, Florida 33431
Email: {borko,oge}@cse.fau.edu

1. INTRODUCTION

Personal computing, multimedia, and computer networks have experienced a tremendous boom during the past decade. From expensive, stand-alone, MPC-1-compatible PCs that landed in the marketplace in the early 1990s to the popularization of the Internet – particularly the World Wide Web –, +the production, archival, and distribution of multimedia has evolved from few, technologically limited devices, to a plethora of hardware and software available for these tasks. It is now possible to develop higher-quality multimedia content, in less time, and distribute it to a much wider audience (using the Web as opposed to CD-ROMs) than anyone could have imagined a decade ago.

The Internet has become ubiquitous and is quickly extending its arms to the wireless arena. The hunger for multimedia-rich content over the Web has fostered developments in many fields, from the technological infrastructure necessary to carry multimedia data, to the tools used to create, manage, and distribute multimedia content. Interactivity is the key concept in many such applications, such as games, maps, and distance learning applications.

In this chapter we present a summary of the network technologies that enable transmission of interactive multimedia information over the Web followed by a collection of tools, examples, and applications.

2. MULTIMEDIA NETWORK TECHNOLOGIES

There are five main types of communication network technologies used to carry multimedia traffic (Halsall, 2001): telephone networks, data networks (including the Internet), broadcast television networks, integrated services digital networks (ISDN), and broadband multiservice networks, such as B-ISDN. In this chapter we focus mostly on the Internet. Besides enabling a number of interpersonal communication tools such as email, the Internet also hosts a large number of interactive multimedia applications stored in Web servers (Figure 1).
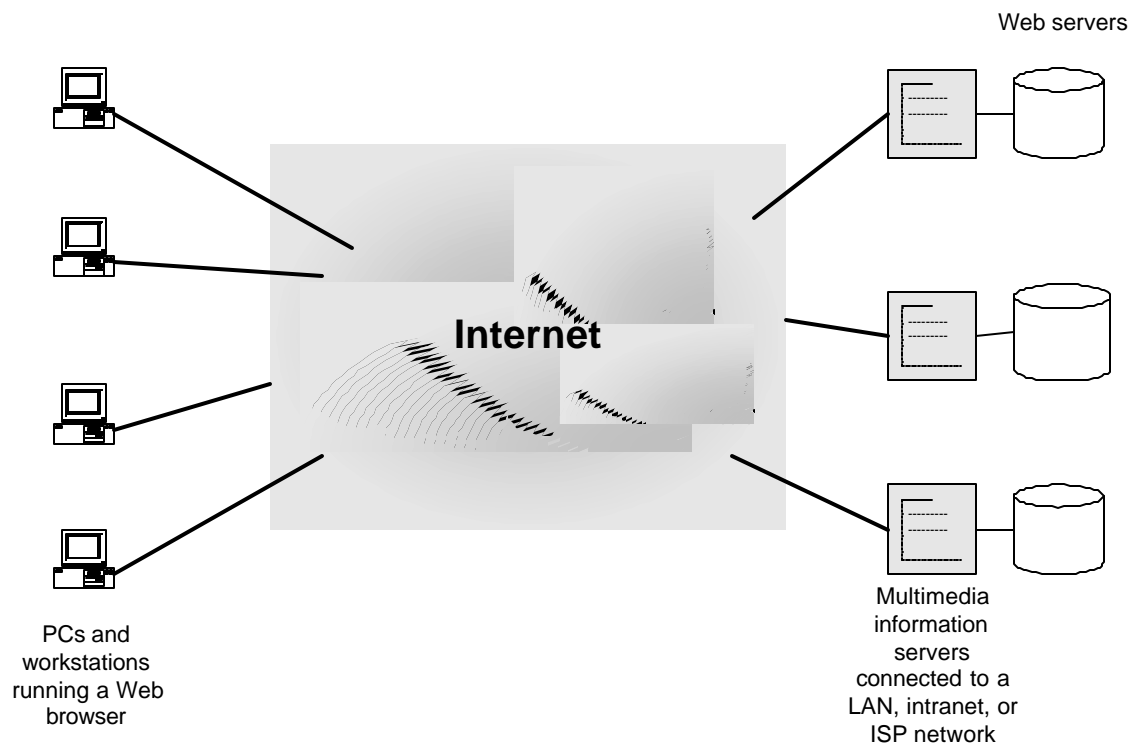

Figure 1. Interactive multimedia over the World Wide Web

2.1   Characteristics of multimedia traffic

Interactive multimedia information may be real-time and non-real-time. Real-time or *continuous* information is transferred from source to destination as it is generated and played out at the destination as it is received. This mode of operation is commonly known

as *streaming*. Audio and video streaming over the Web have become very popular over the past few years. Examples include radio and TV broadcasting over the Web. In non-real-time (or block-mode) media, the information is generated and stored in a way that is not time sensitive and downloaded by request by the interested users.

Multimedia traffic can be classified into two main types: *constant bit rate* (CBR) and *variable bit rate* (VBR). Audio streams, for example, are typically generated at a constant rate, determined by the sampling frequency and number of bits per sample used to convert the audio information into digital format. Video streams, on the other hand, exhibit variable bit rate, even though the frames that make up a video program are generated and displayed at a constant rate. This is due – among other things –to the fact that most video compression algorithms use predictive techniques that employ fewer bits on frames that resemble previous frames.

The transfer of multimedia information streams between any source and destination of can occur in a number of different ways:

- *Simplex*: information flows in one direction only.
- *Half-duplex*: information flows in both directions but alternately.
- *Full-duplex*: information flows in both directions simultaneously.
- *Unicast*: information flows from the source to a single destination.
- *Multicast*: information flows from the source to multiple destinations, comprising a subset of the nodes connected to the network.

- *Broadcast*: information flows from the source to all destinations connected to the network.

In the case of half- or full-duplex communications, the flow of information can be *symmetric* – where each party sends a comparable amount of information to the other – or *asymmetric*, where the flow of information is much higher in one direction than in the opposite direction.

The Internet uses packet-switching techniques to transfer information between two end computers using intermediate nodes – routers – along the way. This mode of operation is also sometimes referred to as store-and-forward, indicating that there is a built-in delay in the process, every time a packet arrives at a router and awaits for the router's decision (upon looking at its routing tables) as to where the packet must be forwarded next. The sum of the store-and-forward delays in each router contributes to the overall end-to-end delay between source and destination. The average of this delay is known as *mean packet transfer delay* and the variance is usually known as *delay variation* or *jitter*. The inevitability of these delays is one of the technical challenges behind enabling interactive multimedia applications on the Web.

Besides the delay limitations, the Internet has another characteristic that makes for less-than-ideal conditions for carrying multimedia traffic: it offers a best-effort service to its end users, which means that the TCP/IP-based transfer of packets is subject to errors (damaged packets, lost packets, timeouts, etc.), and the typical way to handle such errors is to discard packets and (optionally) request/wait for retransmission. Since much of

multimedia traffic may be sensitive, retransmissions are usually not desired, because they add up to traffic inside the network and also because by the time the retransmitted packet arrives, it is no longer needed. A common way of handling with lost/retransmitted packets while providing the user with the illusion of continuous flow of information is the use of buffers at the receiving side.

Designers of networked multimedia applications usually specify the minimum network requirements that must be met in order to enable the application to work. These requirements are collectively known as network *Quality of Service (QoS)* parameters. Examples of QoS parameters are: maximum packet size, average bit rate, average packet error rate (PER), mean packet transfer delay, worst-case jitter, and transmission delay. In addition to the network QoS parameters, the application itself has QoS parameters associated with it, such as: the required bit rate or mean packet transfer rate, the maximum start-up delay, the maximum end-to-end delay, the maximum jitter, and the maximum round-trip delay. To simplify the process of determining whether a network meets the QoS requirements of a particular application, several *service classes* have been defined, each of which contains a specific set of QoS parameters. A network may support two or more service classes, in which case different classes are usually assigned different priorities.

2.2 Standards for multimedia communications

Multimedia information is typically transmitted in compressed form over the Web. Many standards have recently been developed for image, audio, and video compression and transmission. The most important standards are:

- MPEG-1: the first widely deployed standard proposed by the Motion Pictures Expert Group (MPEG). It aims at audiovisual encoding at a bit rate of 1.5 Mbps, which allows for VHS-like video with stereo audio to be encoded and stored in a CD-ROM.

- MPEG-2: enables the transmission of audio and video over broadband networks. It aims at encoding digital TV (and HDTV) signals and has been widely used in the consumer electronics field, from delivery of TV programming via personal satellite (dish) systems to personal digital video recorders, such as TiVo.

- MPEG-4: follows an object-based representation approach to audiovisual scenes, whereas each scene is composed of objects, both natural as well as synthetic. MPEG-4-based systems are not yet fully deployed at a consumer level at the time of this writing. It is expected, though, that the object-based nature of the standard will significantly enhance interactive multimedia applications.

- MPEG-7: standard that describes multimedia content, so that users can search, browse, and retrieve multimedia contents more effectively than using current text-based methods.

- MPEG-21: the latest effort of the group focuses on providing a multimedia framework that should enable transparent use of multimedia resources across a wide range of networks and devices.

- JPEG and JPEG 2000: The JPEG (Joint Photographic Experts Group) compression standard for still images was developed in the late 1980s and has become tremendously popular ever since. JPEG 2000 is a new standard that uses

Wavelet subband coding techniques and allows for numerous compression options that has recently been approved and is starting to enjoy some popularity among Internet imaging users.
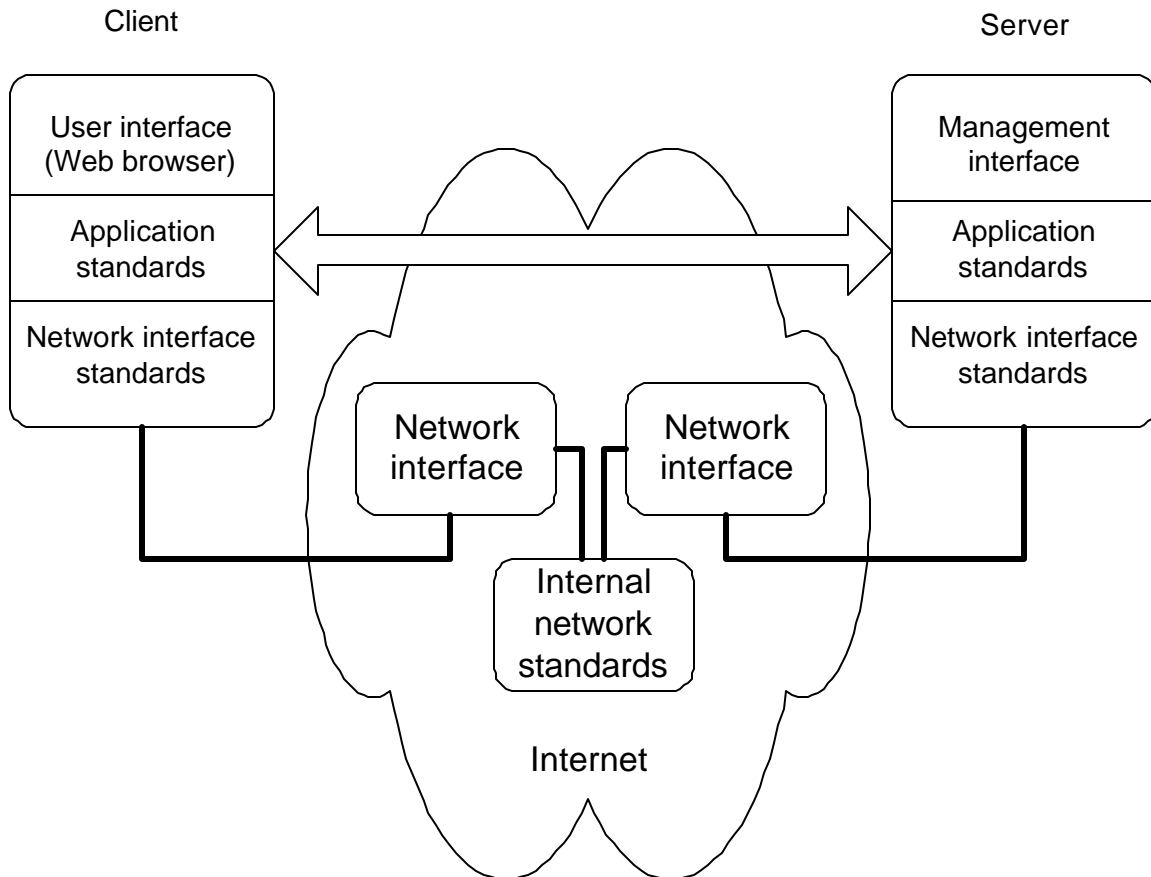

Figure 2 – Standards requirements for interactive multimedia applications over the Internet.

In the context of interactive multimedia applications on the Web (Figure 2), standards are required both at application and the networking level. Examples of application-level standards include the JPEG and MPEG compression standards and the hypertext transfer protocol (HTTP) implemented by all Web browsers to allow HTML pages to be displayed at the client's side. Network-level standards include the TCP/IP protocol suite.

3. CREATING INTERACTIVE MULTIMEDIA WEB PAGES

Contemporary tools for designing Web pages (such as Macromedia Dreamweaver and Microsoft FrontPage) allow including multimedia objects, such as videos, images, animation, and audio, into Web pages. In order to play sound, animation, and video, the user system may require plug-ins, which are small programs that work in cooperation with a Web browser. Some plug-ins have become very popular, so Netscape and Microsoft browsers have built them into their latest versions. Popular multimedia plug-ins include RealNetwork's Real Audio and Real Video, Apple's Quick Time, and Macromedia Shockwave and Flash.

Besides plug-ins, which allow multimedia objects to be included into Web pages, more sophisticated Web pages also include Java applets and VRML (Virtual Reality Modeling Language) 3-D images. <u>Java applets</u> are programs that are embedded in Web pages (in HTML code) and are used to create multimedia Web pages that include animation, sound, images, and videos. The applet runs locally on the client machine, which is running a Web browser. <u>Virtual Reality Modeling Language</u> is a generic text-based language used to construct 3-D images and 3-D text for Web pages. A VRML document is an ASCII file, much like an HTML document. In VRML, the Web designer specifies objects – called *nodes* – that define shapes and their attributes. These nodes create 3-D scene graphs, which a VRML browser renders. Examples of nodes include cube, sphere, cone, cylinder, rotation, scan, etc.

Although including all these components makes Web pages multimedia rich, they are still not interactive. The complete process of creating interactive multimedia Web pages, and related tools and languages, is shown in Figure 3. This process is described next.

SERVER   CLIENT

Original Web pages
Include:
multimedia objects
Java applets
Flash and Shockwave
    animations
VRML images
CGI forms

Web browser with plug-ins

Original Web pages displayed on client

HTTP

CGI forms
- Perl
- JavaScript
- VBScript

Stored information

Active Server Pages (ASP)

HTTP

Information from client

Information from Internet

Dynamically created Web pages
-VBScript, JavaScript

HTTP

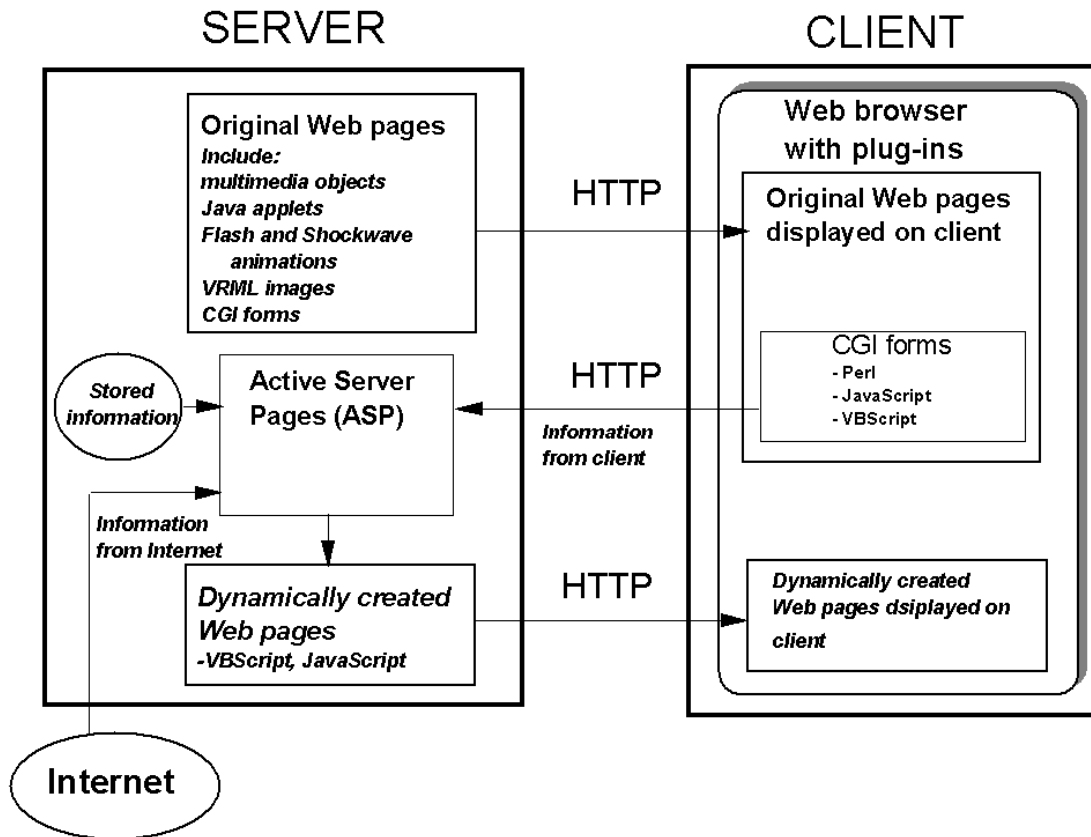*Dynamically created Web pages dsiplayed on client*

Internet

Figure 3 – A process and tools in creating interactive Web pages.

The original (home) Web pages, hosted at a server, include multimedia objects (such as music, sound, images, animations, and video clips), Java applets, and VRML images. To provide interactivity, a Web designer should also apply client-side and/or server-side scripting.

<u>Client-side scripting</u> includes CGI (Common Gateway Interface) scripts that provide interactivity by processing and responding to user inputs. When a user submits a CGI form, a program that resides on the server system receives and processes the form's content. This program then creates an HTML document dynamically, which the server displays back to the user on the client machine. CGI scripts are usually created using Perl, JavaScript, or VBScript.

<u>Perl</u> is an interpreted scripting language for creating CGI forms, which is well suited for text manipulation. Its features include flexibility, compact, form, security, and support for file and database operations.

<u>JavaScript</u> is another scripting language that is used to create Web pages with CGI forms. However, the browser, and not a program on the server, executes JavaScript adding interactivity and dynamic content to Web pages.

<u>VBScript</u> is a scripting language from Microsoft that is based on Visual Basic, which is used to create interactive multimedia Web pages. Like JavaScript, VBScript statements are embedded into an HTML code, which a browser executes. VBScript provides access to OLE (Object Linking and Embedding) objects, called ActiveX controls.

<u>ActiveX</u> is an OLE object, which is a class library that contains a complete set of functions for manipulating an object. The code for an OLE object already exist, so programmers can include the objects in applications quickly and easily. ActiveX controls

provide programmers a way to extend browser capabilities and create a wide range of Web site capabilities. Some examples of ActiveX controls include: button (control to display various frame sequences), chart (enables the programmer to draw various types of charts with different styles), and gradient (shades the area with a range of colors).

Server-side scripting provides an effective way to dynamically build documents in response to client (or user) requests. As shown in Figure 2.1, server-side scripting uses information sent by clients, as well as information stored on the server, and information from the Internet, to dynamically create Web pages (in XTML or XML), which are then sent back to the client.

Active Server Pages (ASP) is a popular server-side technology from Microsoft, which is processed by an ActiveX component (server-side ActiveX control) called a scripting engine. An ASP file has an extension .asp, and contains XHTML tags and scripting code. The VBScript is most widely used for ASP scripting, but other languages, such as JavaScript, can be also used.

In summary, a number of tools, languages, and components, which are introduced in this section, are used to create contemporary and effective interactive multimedia Web pages.

4. TOOLS, SERVICES, AND APPLICATIONS

Today, there is a wide range of interactive multimedia services and applications on the Web. In this section, we present the basic services and applications as well as related tools. The list of these services and applications is given in Table 1.

Table 1. Interactive multimedia services and applications on the Web

APPLICATION	DESCRIPTION
Multimedia portals	Users have a single, customizable access to multimedia services, including video, audio, animation, interactive banners, etc. Multimedia portals can be classified as entertainment, educational, and information portals.
Interactive e-learning and distance learning	Interactive e-learning applications include a wide variety of media files to interactively explain complex concepts. Distance learning allows customers to subscribe to on-line courses taught at remote sites.
Media sharing	Media sharing provide collaborative music and video production by a group of artists from dispersed locations.
Interactive multiplayer games	Multiple players participate in games over the Internet
Web 3D virtual environments	The content is presented in three dimensions, and allows users to fly through the virtual space.
Participatory publishing	The user can contribute and influence the Web content along with authors.
Interactive Web maps	Interactive Web maps are time-enabled maps that typically display archeological and historical data.
Integration of the Web with traditional media	These applications include radio and television broadcasting on the Web, video-on-demand, and interactive TV applications.
Interactive multimedia retrieval	Large multimedia databases can be accessed, search, and retrieved through the Internet.

4.1 Multimedia Portals

Market environment is changing rapidly and therefore new functionality is important to gain competitive advantage. Time-to-market is critical and businesses need to integrate systems, automate processes, and provide each other with access to key functionality. Recently, Web portals evolved as a single, integrated point of access of information,

applications, and people. Web portals are Web sites that integrate new content with existing content, server-side applications, and Web-based services. There are two basic Web portals: (i) Public portals, such as Yahoo, that bring together information from various sources, applications, and people, and offer personalized Web sites, and (ii) Enterprise, or Business portals that give employees access to organization-specific information and applications.

Multimedia portals are special Web portals, which offer multimedia services, such as streaming video, audio, interactive banners, animations, etc (Dimitrova, 2000). Multimedia portals can be classified in the following groups: (1) Entertainment portals, (2) Educational portals, and (3) Information portals.

<u>Entertainment portals</u> typically provide customizable music, television, and movie news. For example, a portal from Warner Brothers, shown in Figure 4, provides features such as email center, customizable access to music, DVD videos, television, and movies, entertainment news, a chat room, and interactive advertiser banners.
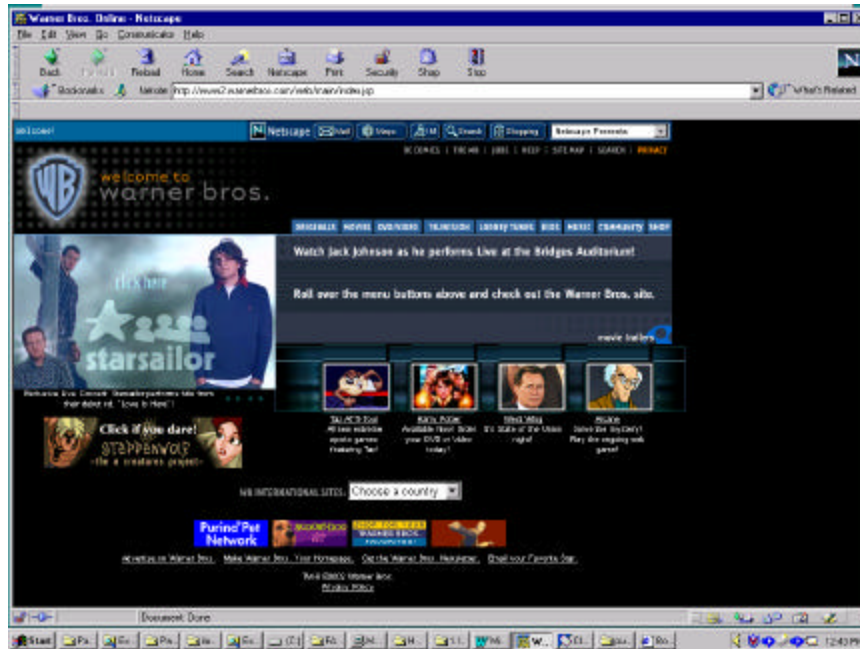
Figure 4. An example of the entertainment portal from Warner Brothers
(www.entertaindom.com)

Educational portals include on-line learning applications that provide access to a diverse

set of content. An example of a tool for developing educational portals is Aspen learning

Content Management Server from click2learn (www.click2learn.com), which allows a

team of instructional designers, media developers, subject experts, and content reviewers

to work together effectively through project management, workflow and collaboration.

Information portals … ?

4.2 Interactive e-Learning Tools and Distance Learning Applications

The Web coupled with multimedia content has a great potential of changing the

education. Contemporary e-learning tools allow the development of on-line interactive

learning applications. An example of the tool is ToolBook Instructor and Assistant

(www.click2learn.com) that supports a wide variety of media files, including sound,

animation, streaming media, and still images. Some interactive learning applications

using ToolBook include innovative technical and maintenance training, electronic service

manuals, software training courses, and others. Figure 5 shows a screen from the interactive Boeing 777 course, showing how skilled aircraft mechanics share techniques for installing safety hardware in confined spaces.
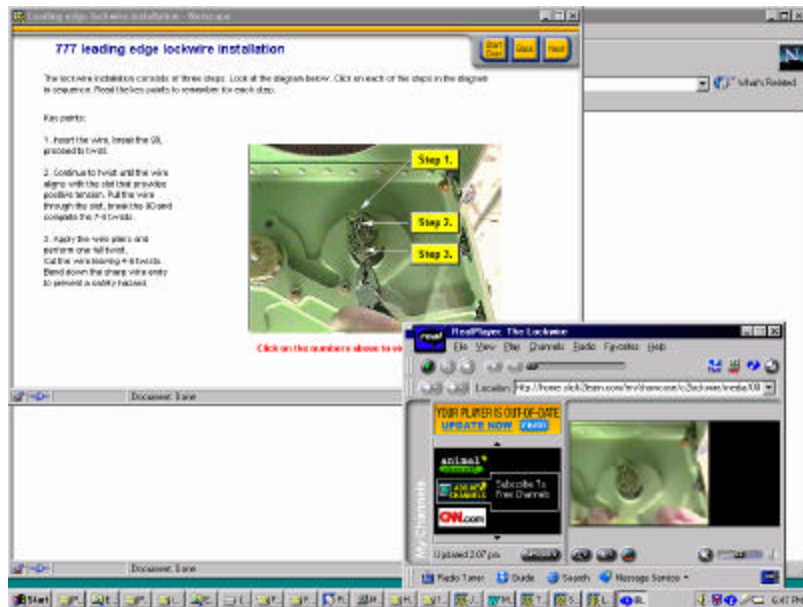

Figure 5. Boeing 777 interactive maintenance course created
using the ToolBook e-learning tool.

Distance learning allows remote participants to receive classes over the Internet. Advanced distance learning systems include a real-time interactive virtual classroom, which allows a remote participant not only to receive a live class feed, but also to interact in a live class using audio and video over the Internet (Deshpande, 2001).

4.3 Media Sharing Applications

Media sharing includes the design of tools that support collaborative music or video production by a group of artists from geographically dispersed areas. A Sharable Media project team at MIT Media Lab, led by Glorianna Devenport (www.ic.media.mit.edu),

has developed set of tools that allow online groups of filmmakers to collaborate on the same film project by contributing and sharing material, and comment to one another's creations. An example is PlusShorts application (Davenport, ???), which is implemented as a Java applet, and allows distributed groups of users to contribute and collaborate in the creation of shared movie sequences.

RocketNetwork (www.rocketnetwork.com) has developed a tool for online musicians that provides simultaneous access to audio files and allows collaboration of musicians and producers. This system, referred to as "global production network" allows updating files dynamically from anywhere in the world, and applications include films, TV and radio sound track production, and sound track for broadcast media. The system is based on a RocketNetwork's Central Server that coordinates sharing, dynamically updates, and distributes audio and MIDI files.
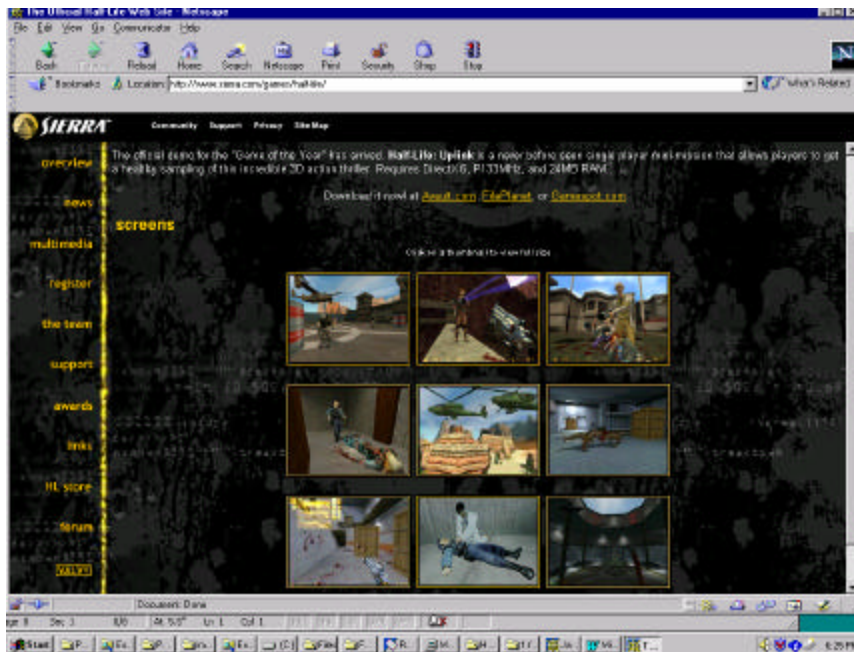

Figure 6 – Half-life from Sierra.com - illustration of a multiplayer Web game.

Multiplayer Web Games

The challenge of Web-based multiplayer games is in the quality of the connection and the ability of the designer to implement the plot in a cyberspace way. Real-time games also require that players visit a chat room and register and that all clients must be synchronized. Figure 6 illustrates an example from the game Half-Life, created by Sierra.com.

Web 3D Virtual Environments

Support for VRML in Web browsers allows developers to present their content in three dimensions, rather than just as documents. The user can "fly" through the virtual space and experience real car, house, or town. The browser requires a plug-in a Web browser that provides complex VRML rendering. Figure 7 illustrates the Cortona VRML plug in, by ParallelGraphics for a 3D model of a sport car.
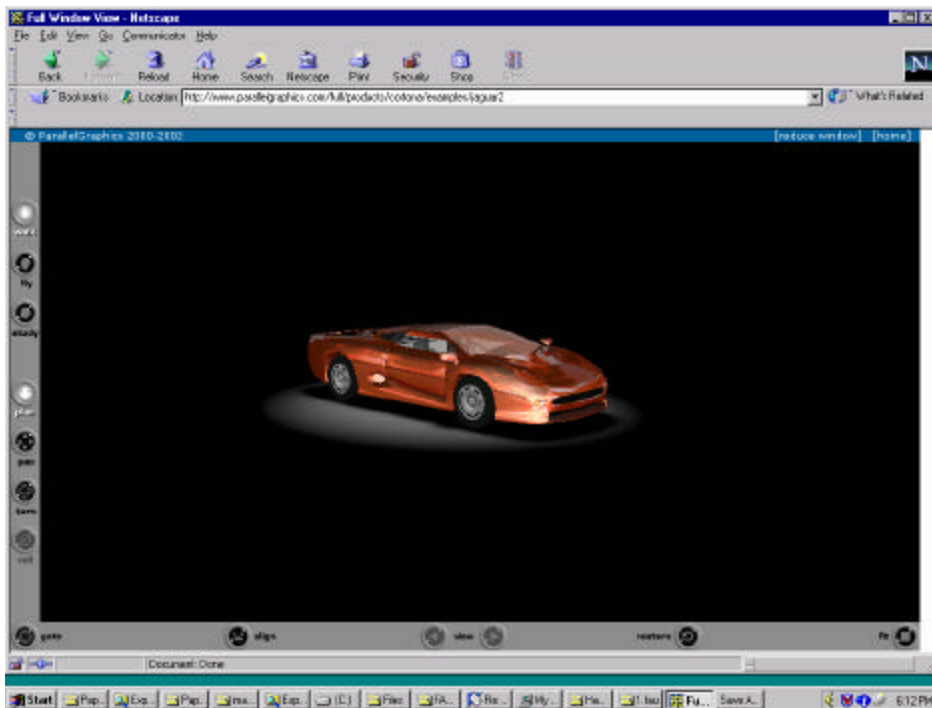
Figure 7 – VRML model of classic sport cars. The navigation buttons allow changing the user's view, opening the door, and the car movement (www.parallelgraphics.com)

The same company ParallelGraphics offers Outline3D interior design application that enables users to quickly customize a virtual environment specific to their design needs. Potential applications include:

- The 3D room for potential buyers to view the finished interior design of an office or a home.

- Real-estate application for real estate agents to show customers the look of their future home

Participatory Publishing

Participatory publishing refers to Web pages where readers can contribute and influence the content along with authors.

An example of participatory publishing is WebTour, an application developed by Siemens Research Center in Princeton, in which viewers can make verbal comments and dynamic annotations, or drawings to illustrate some points. The system is capable to record, playback, store, search, and distribute personalized dynamic multimedia annotation on Web documents in the form of guided Web tours.

Potential applications of participatory publishing tools including WebTour include:

E-commerce. When a customer enters the Web site of a retail store, a Web clerk would greet the customer and guide him/her through various Web pages explaining key products and adding an advertisement when necessary.

Corporate Intranet. Different users, such as employees and/or customers, can make annotations on the same Web document, and share ideas.

Distance learning. A teacher can make annotations on Web pages explaining complex problems. The annotations could then be retrieved by students and played back.

WebTour is implemented using open Web technologies described earlier – standard browsers, Java, Dynamic HTML and JavaScript, and Active Server Pages.

Interactive Web Maps

Interactive Web maps are time-enabled maps that are created from distributed databases. Interactive maps typically display archaeological and historical data, such as historical maps and images, the growth of cities, the spread of dynasties and empires, or the distribution of archeological sites. However, there is a wide-range of other applications that could include time-stamped data. Interactive Web maps could also link map objects to images, other Web pages, multimedia, or a database. In addition, interactive features include zooming and panning, temporal filtering, data-driven hotlinks, and other functionality.

An example of a tool for creating interactive Web maps is TimeMap from University of Sydney (www.archaeology.usyd.edu.au/research/time_map). Figure 8 illustrates an interactive map developed by the TimeMap team for the International Dunhuang Project and the British Library. By selecting an area and clicking the mouse brings the viewer to another interactive screen (Figure 9). This interactive map allows the viewer to select and view archeological sites, monumental remains, routes, etc.
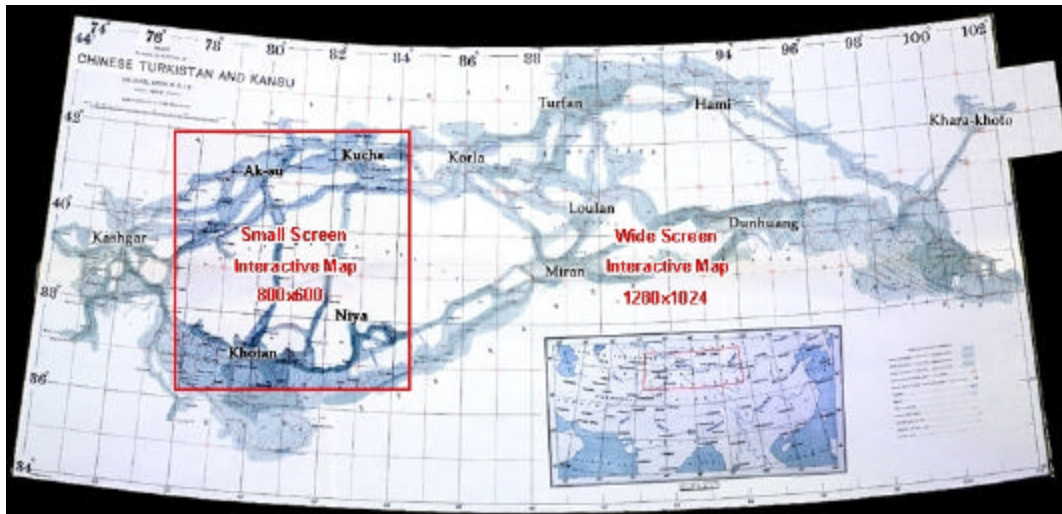
Figure 8 – An example of an interactive Web map. The user can click and select an area that brings the next interactive screen
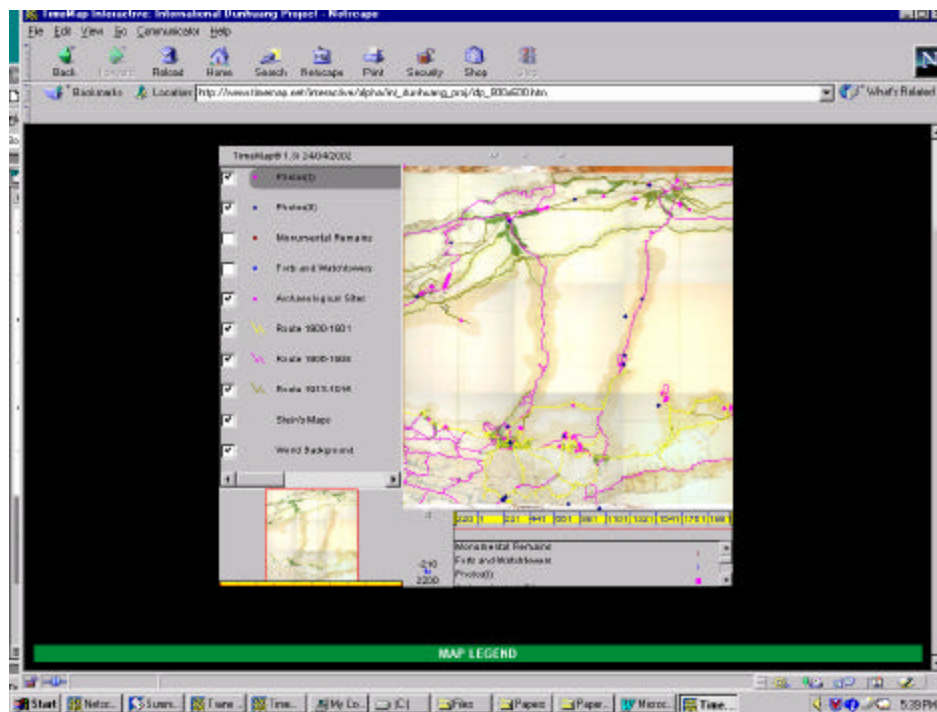[www.archaeology.usyd.edu.au/research/time_map)

Figure 9 – The user can select to view archeological sites, monumental remains, routes, etc.

Another tool, JMapView from Hexa Tech, written in Java, allows creating any type of interactive maps or Web geographical information systems. An example is shown in Figure
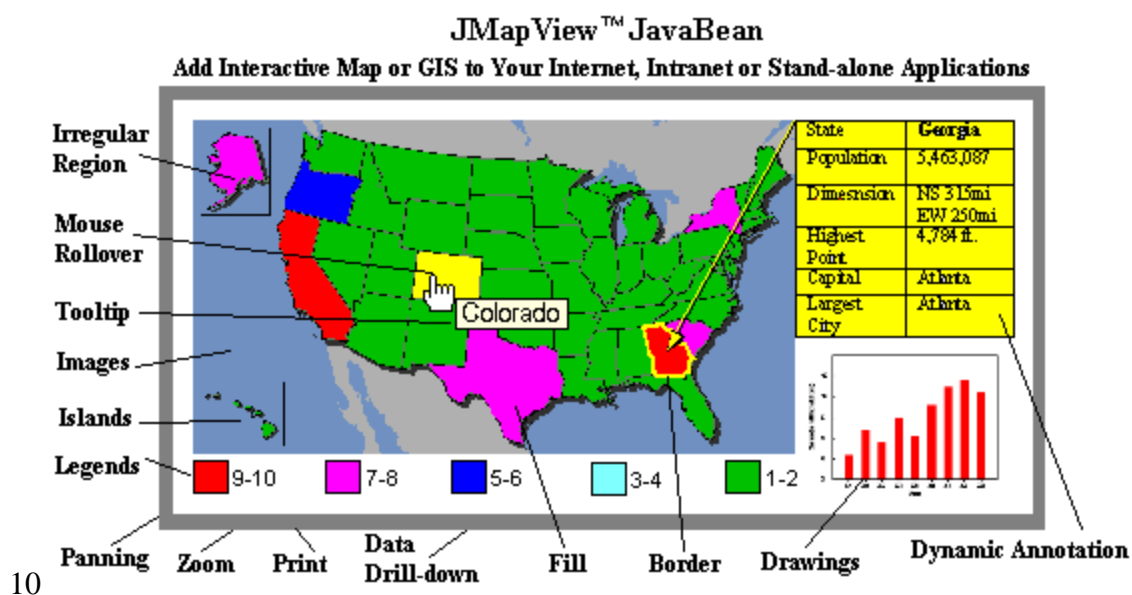

10

Figure 10Interactive map created using JMapView (www.hexatech.com)

Multimedia Search over the Internet

Today, there are more that 275 million multimedia files on the Web. On average, Web surfers interact with sites containing multimedia content up to 25 times longer than sites with static content. (reference???)

FAST Multimedia Search (www.fastseearch.com) is currently one of the largest integrated multimedia search product on the Internet, allowing users to receive results for image, audio, and video files on a single result page. Users simply search by keyword, and as the result, all retrieved multimedia files are accessible from the result page. This

tool has been used by Lycos to create their multimedia search engine (www.lycos.com), which is illustrated in Figure 11

FAST Multimedia Search (www.fastseearch.com) is currently one of the largest integrated multimedia search product on the Internet, allowing users to receive results for image, audio, and video files on a single result page. Users simply search by keyword, and as the result, all retrieved multimedia files are accessible from the result page. This tool has been used by Lycos to create their multimedia search engine (www.lycos.com), which is illustrated in Figure 10.
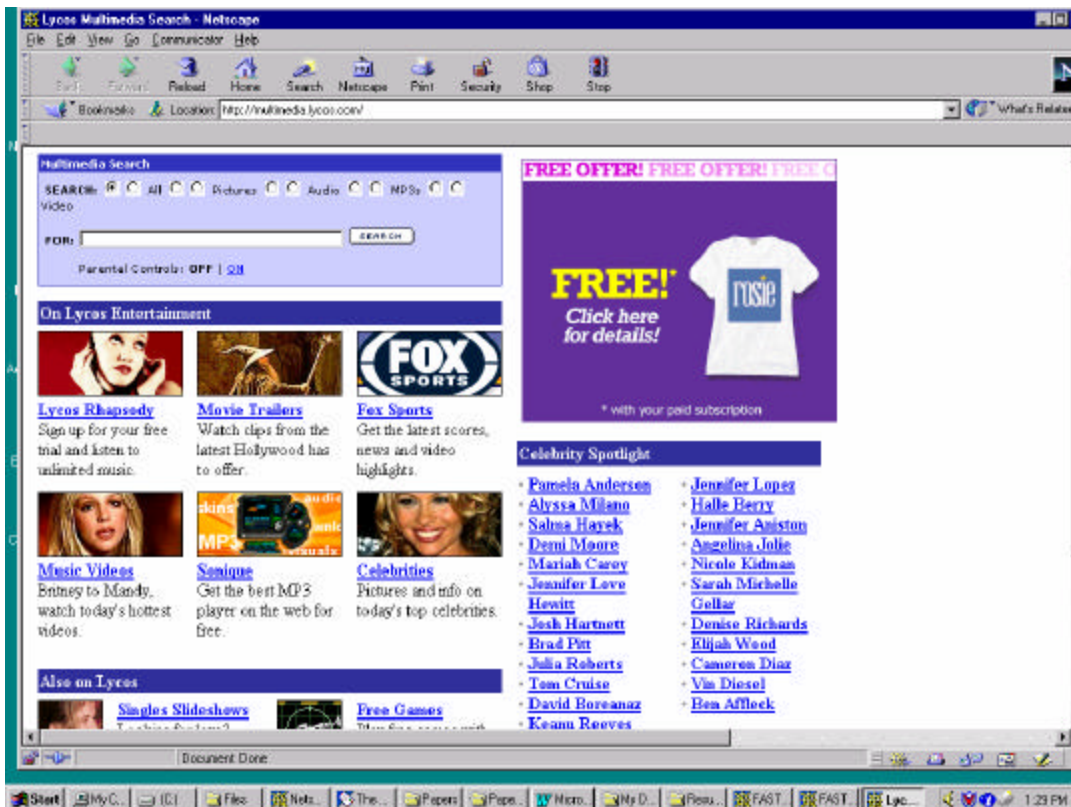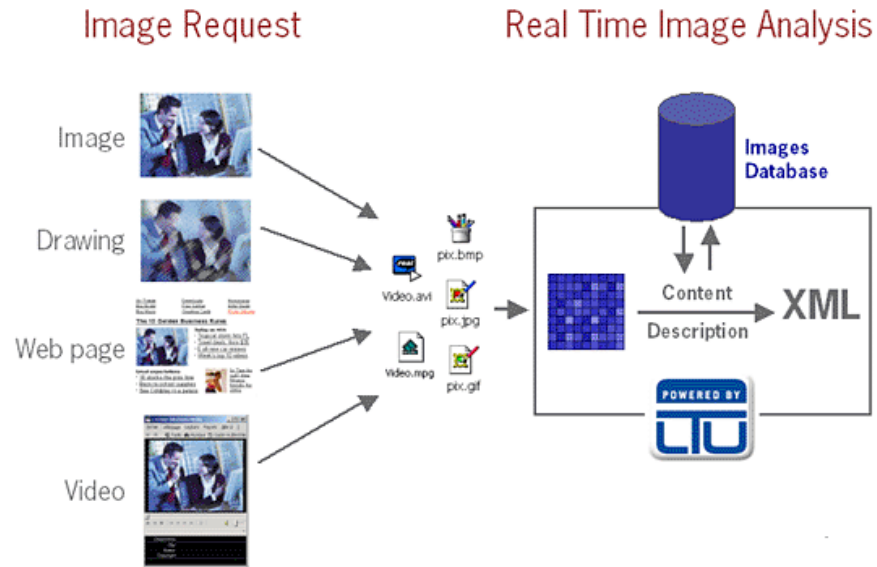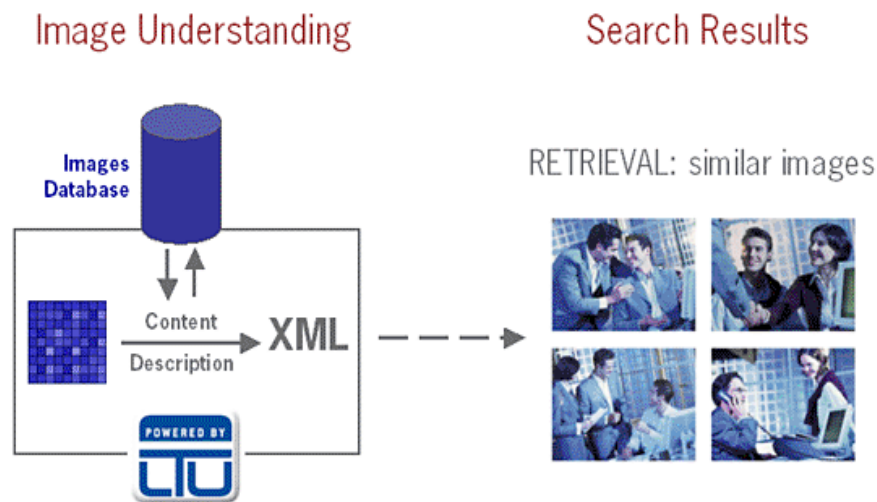

Figure 11 Lycos offers the largest integrated multimedia search engine on the Web that includes 60 million images, sounds, and videos (www.multimedia.lycos.com)

Another, more advanced search of the multimedia repository is content-based retrieval, where the user provides information about the actual contents of the image, audio, or video data, rather than just keywords (Marques, 2002). A content-based search engine translates this information to query the multimedia database and retrieve the candidates that are likely to satisfy the user's request.

An example of a commercial content-based image retrieval system is a product-suite from LTU Technologies [www. ltutech.com], which includes five products: Image-Filter, Image-Indexer, Image-Seeker, Image-Shooper, and Image-Watcher. The core of the system is a high-level perceptual image analyzer that is capable of indexing, recognizing, and describing images according to their visual features. Figure 12 illustrates the operation of the system.

(a)

(b)

Figure 12. (a) An image is presented to the system and real-time image analysis is performed by extracting features of the image, (b) The system than matches these features with the features of images in the database and retrieves similar images [www.ltutech.com].

Integration of the Web with Traditional Media

There is a growing trend in integrating the Web with traditional media such as radio and television. As the streamed multimedia protocols for audio and video has matured, a number of radio and television stations begin offering live and near-live audio and TV content over the Web. Examples include yahoo! Broadcast radio and television Web sites (www.broadcast.com/radio and www.broadcast.com/televsion).

In addition, some programs are capable to support active user feedback, or even participation, so we can expect a number of new video-on-demand and interactive TV applications.

Summary

The synergy between new multimedia technologies and the Web have brought new delivery mechanisms and new types of applications. We are still in the first phase of this marriage, where we are faced with technical challenges such as how to efficiently transmit multimedia data over the Internet and how to best use our new tools. However, the focus begin already moving to some new areas, such as new metaphors for publishing, how mass media can better reach and involve their audiences, how to users can better access the Web's wealth of multimedia information, and others (Wynblatt, 2000).

References

N. Dimitrova, H. Yu, F. Galliano, R. Koenen, A. Zakhor, C. Bouman, "Entry into the Content Forest: The Role of Multimedia Portals," IEEE Multimedia, Vol. 7, No. 3, July-September 2000, pp. 14-20.

A. Kelliher, J. J. Seo, P. Pan, C. Lin, and G. Davenport, "Visual Interfaces for Shareable Media," ISEA 2000 International Symposium on Electronic Art, December 2000.

M. Wynblatt, D. Benson, and A. Hsu, "Multimedia Applications on the Internet," a chapter in the "Handbook of Internet Computing," editor-in-chief and author B. Furht, CRC Press, Boca Raton, FL, 2000.

S. Deshpande and J.-N. Hwang, "A Real-Time Interactive Virtual Classroom Multimedia Distance Learning System," IEEE Transactions on Multimedia, Vol. 3, No. 4, December 2001, pp. 432-444.

O. Marques and B. Furht, "Content-Based Video and Image Retrieval," Kluwer Academic Publishers, Norwell, MA, 2002.

Appendix: Web Resources

Multiplayer games
- www.sierrastudio.com/games/half-life
- www.microsoft.com/games
- www.uo.com
- www.cryo-networks.com/uk/cryonetworks.htm
- www. idsoftware.com
- www.epicgames.com

Web multimedia portals
- www. eUniverse.com
- www. tollywoodlive.com
- www.oz.net/blam
- www.click2learn.com

Interactive Web maps
- www.eparka.com
- www.hexatech.com
- www.archaeology.usyd.edu/au/research/time_map

Web multimedia search engines
- www.fastsearch.com
- www.multimedia.lycos.com

Content-based search engines
- www.ccrl.com/amore
- www.qbic.almaden.ibm.com
- www.ctr.colombia.edu/WebSEEK
- www.convera.com
- www.virage.com
- www.cobion.com
- www.lutech.com

Radio and television broadcasting over the Web
- www. broadcast.com/television
- www. broadcast.com/radio