T-Cube: A Multimedia Authoring System for eLearning[*]

Mingchao Ma
Volker Schillings
Tongbo Chen
Christoph Meinel
Dept. Of Theoretical Computer Science and New Application
FB IV – Informatik
Universität Trier, Germany
{ma, schillings, chen, meinel}@ti.uni-trier.de

Abstract: eLearning has become an alternative approach to the traditional college and university of education. Experiments have shown that producing of the teaching contents for eLearning is a time-consuming and costly job. This paper introduces a programming-free and easy-to-use multimedia authoring system, T-Cube, designed and has already been used at the University of Trier for eLearning. As a rich-media authoring system, T-cube constructs and presents multimedia contents to students with multi-rate for either offline (CD/DVD/download) or online (in real time or on demand) usage. The rich-media teaching contents, including video, audio and screenshot, are recorded and encoded at the classroom and simultaneously published on Internet. Broadband users can access the lecture in real time as if they were sitting in the classroom. After being compressed and encoded, the teaching contents are also available for narrowband users on demand. Further, the contents on CD/DVD are also available on request. This paper focuses on describing the features, practical experiments and assessments of T-cube system.

Introduction

Modern information and communication technologies are transforming the way people learn. E-learning not only provides new possibilities for personalized learning at home or workplace, but also reduces the requirements of the expensive classroom training and well balances the traditional and innovative forms of knowledge transferring.

Today many universities and institutes provide the students with web-based online courses (MIT OpenCourseWare; Stanford Online). However, less efforts and expenses are put to develop the effective e-learning materials. Some commercial multimedia authoring tools, Authorware, Flash and Multimedia Fusion, are adopted to create multimedia documents and integrate them into web-based teaching and learning environment. (Authorware; Multimedia Fusion; ToolBook). Some other experimental research tools (Brent Martin, 2002; Brian Bailey, 1998; Denis Helic, 2002; Kuo-Yu Liu, 2002; M. Jourdan, 1998; R. Baecker, 1996; Suzanne Little, 2002) are used to support interactively authoring. Despite all these advances, creating multimedia presentations is still primarily a manual, labor-intensive process. To solve this problem, a simple and efficient system, named T-cube, is designed to capture and deliver the multimedia contents for both online and offline usages. The rich-media contents including video, audio and screenshots of lecturer's computer are recorded and encoded in the classroom on one hand, and they are simultaneously published on the Internet on the other hand. Broadband users can access the lecture online in real time as if they were sitting in the classroom. After post-processing, the rich-media contents are also reachable for narrowband users on demand. The attractiveness of the proposed T-cube system is that it supports multiple rates connection, so that not only broadband customer (using DSL) but also narrowband users (using ISDN or modem). In other words, as long as an end user can access Internet and has the software of Real player in hand, he can access the T-cube at anytime anywhere.

Instead of simply inventing "yet another wheel" allowing one to capture the lecture at a classroom, our aim is to use off-the-shelf components and (de facto) standard techniques (SMIL, MPEG4, RealMedia etc.) to implement an easy-to-use and programming-free authoring system for eLearning. To develop such authoring system, a number of requirements were identified. Firstly, the system should not expect the authors have any programming skills. Secondly, the multimedia contents can be reached not only by

---

broadband users but also by narrowband users. Thirdly, users (students) can use (de facto) standard software (IE, RealPlayer, QuickTime etc.) to replay the teaching contents. Lastly, instructors should not be limited to any proprietary software for preparing their lectures/presentations.

The T-cube system has already been used by the department of computer science at the University of Trier since the summer semester of 2002. Up to now, the length of recorded multimedia content is more than 500 hours. At the winter semester of 2002, the lecture "Internet security" (at University of Trier, Germany) is simultaneously transmitted to Beijing University of Technology, China. The practical usage proves that our solution is very successful. All recorded teaching content are also available on the website for later usage(*Tele - TASK*).

The rest of this paper is organized as follows. We first describe the system structure and management. Then discuss how to reduce bandwidth to support narrowband users. The experiments and assessment are presented subsequently, followed by related works. Finally we conclude our paper and discuss the future work.

System structure and management

Structure: Figure 1 presents the system structure and components. Its main part is 'classroom' section. The contents of a lecture are divided into two parts. The first part, video and audio, are captured by video camera and microphone respectively. After being transferred via IEEE 1394 interface, they are delivered to the encoding server. The second part, screen contents and handwriting, are transmitted via LAN to the encoding server as well. Then at the encoding server, video, audio and the content of screen are correlated in the same timeline and encoded to RM format streaming. After the streaming being transferred via Internet/Intranet to Real streaming server, the end users can access it either by HTTP protocol or RTSP (Real Time Streaming Protocol), using Real player (Volker Schillings, 2002).
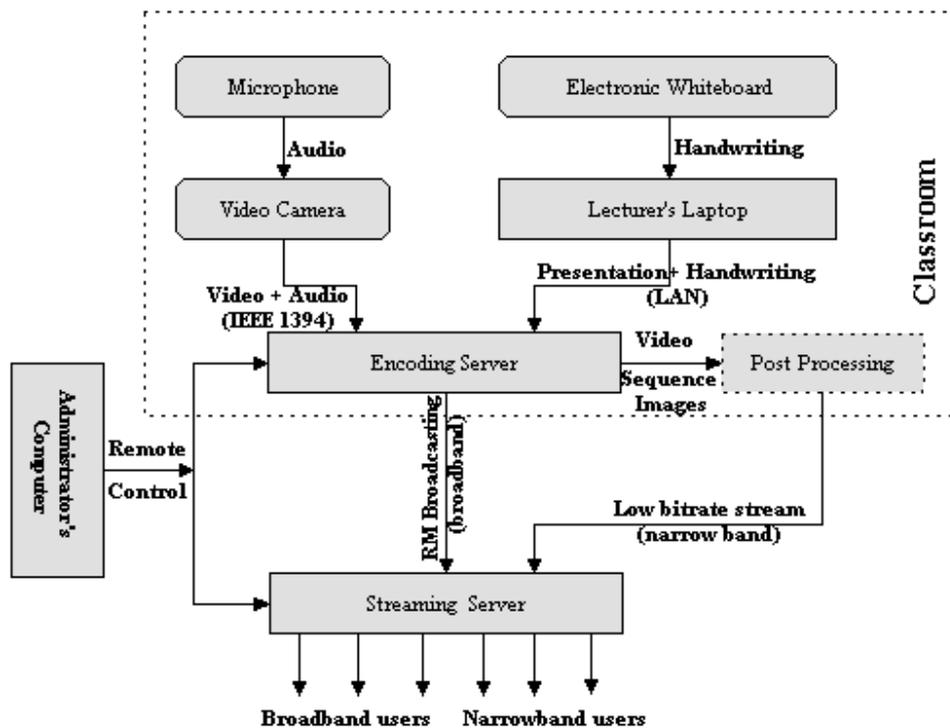

Figure 1 System Structure

Application components: The system includes three application components: encoding, post-processing and administrating. Encoding program is running on the encoding Server, and it has two

responsibilities. Firstly, it compresses and encodes the input data (video and screenshots). Its output (RealVideo, Support for MPEG4 is developing), are transmitted to Real streaming server. Secondly, in order to support the narrowband users, encoding program also saves the uncompressed video streaming (AVI) and screenshot sequences (BMP) on hard disk separately, and correlates them in the same timeline. Different from encoding program, post-processing program can run either on the encoding server or on the other computers. The details about how to support narrowband user will be discussed later on. Administrator program remotely controls and monitors the encoding server and streaming server. It can start, stop, pause the capturing process.
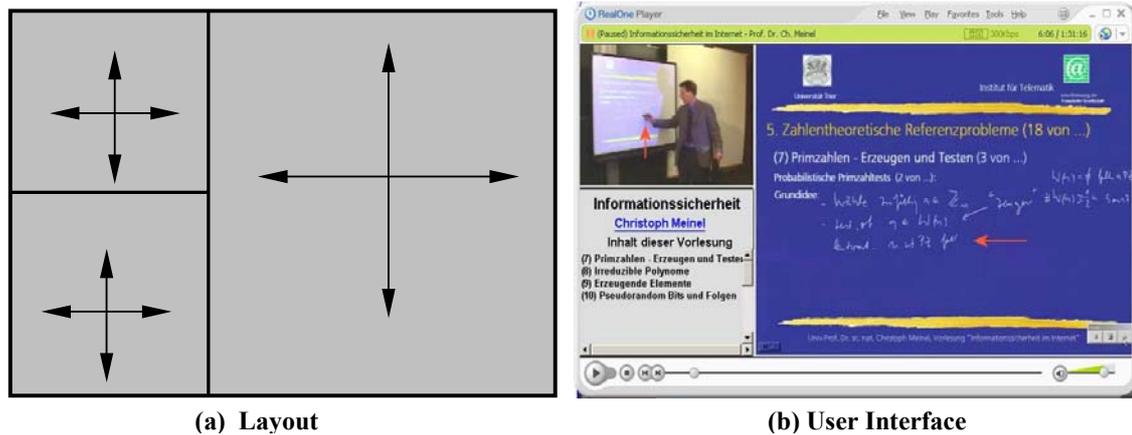

| (a) Layout | (b) User Interface |

Figure 2 Views Layout and User Interface

User Interface: Figure 2 (a) is the layout of the three views. Part I is the video (including audio) of the lecturer. Part II is the screenshot and the handwriting of lecturer. Part III is the logo and index of part II. By the index, user can navigate fore- and backward across the streaming. Figure 2 (b) is the practical user interface. Please note the handwriting of the lecturer is also displayed.

Management: The T-Cube system is a self-service system. It supports Plug & Play and is very convenient to use. What the lecture needs to do is to connect his laptop and the video camera to the T-Cube system, and need not do any reconfiguration on the laptop at all. The system provides web-based management interface, therefore the user can very conveniently manage, configure and monitor the system remotely using the standard browser. The encoded streaming is automatically burned on CD/DVD after the lecture.

Supports for narrowband users

In order to meet the requirements of narrowband customers using ISDN, reducing the bandwidth is necessary. The easiest way is to use Real Codec, which supports as low as 28.8 kbps rate, to compress and encode the streaming. Some application tools adopt this solution. Unfortunately, the quality of the streaming is poor. As we previously discussed, there are two types of streaming: video streaming (see Figure 2, part I) and the sequence of screenshots (see Figure 2, part II). We treat them as different media, and compress and encode them by different technologies. Traditional video compressing technology is only adopted to encode the video streaming. Since the window of video is 220 X 168 in pixel, it is not difficult to reduce it to about 30 kbps bandwidth, and guarantee the quality of video and audio as well. Therefore, the key job is to reduce the bandwidth of Part II.

Challenges: Streaming the multimedia contents and make them reachable by modem or ISDN is a challenge. Since effective bandwidth of standard 64K ISDN is about 56 kb (the extra bandwidth is for noise, data loss and packet overhead), for our application, the video and audio will consume about 30~32 kbps bandwidth, therefore only about 25kb bandwidth is left to Part II. Furthermore, 56K modem has lower bandwidth than 64K ISDN (only 35~40kb effective bandwidth). Due to the feature of teaching, the audio

and teaching contents are more important than the video, one feasible solution is to transmit the audio and contents to the modem users only, by which the bandwidth is dramatically reduced.

Compressing and Encoding images: In order to obtain the screenshot of lecturer's computer and the annotation of the lecturer, we use a free software VNC (*VNC - Virtual Network Computing*), which is running on the lecturer's computer. The screenshots of the lecturer's computer are captured in the rate of 1 frame per two seconds according to the characteristic of slideshow. Usually the slideshow doesn't change very quickly and thus the capturing rate is high enough to reflect the changes of the screen. All these screenshots (BMP format) together with the timestamp are saved. An uncompressed BMP file is about 543 KB. One and half hour's lecture generates about 2700 files with a total size of about 1.5G bytes. The bandwidth requirement of streaming uncompressed images is roughly about 200 kbps, which is too high to support the ISDN user. Further reduction of bandwidth requirement is required.
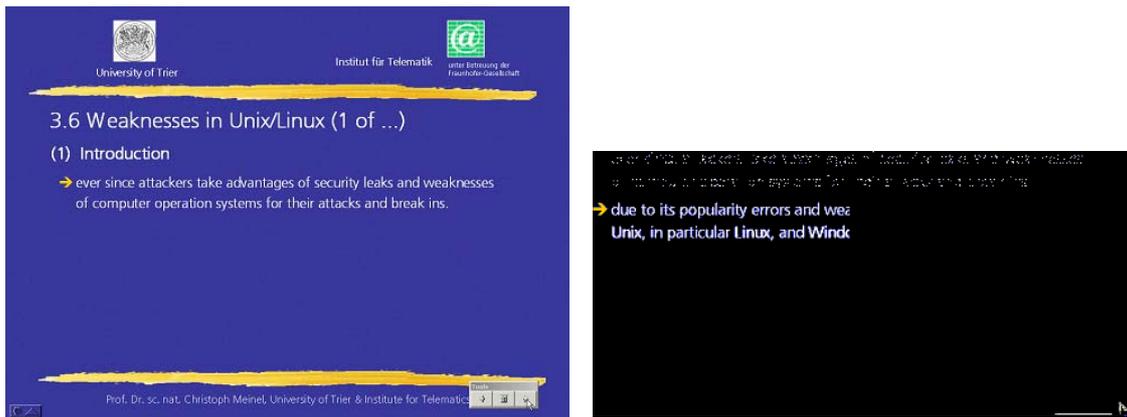
(a) Key Frame (K-th)

(b) Transparent Frame (K+1-th)
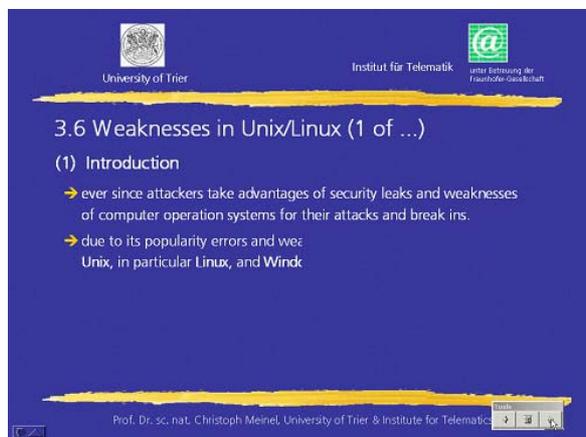
(c) Reconstructing = (a+b)

Figure 3 Compressing and Encoding

Inspired by the video-encoding technology, we address the above issues in following steps. First, we compare the sequence of images to detect all duplicate images and delete the redundant images so that the number of the images is dramatically reduced by 85%. Next, we convert BMP files to PNG files to reduce the size of image to 1/50 of the size of uncompressed image (*PNG Programming Resources*, ; *Portable Network Graphics Specification 1.0*, 1996). Finally, we make some PNG files transparent as shown in Figure 3(b). Those unchanged PNGs files are defined as Key Frame of the streaming as show in Figure 3(a). The basic idea is as follows. Let the ith frame (Figure 3(a)) be the key frame. Then, compare the next frame (i.e., the (i+1)st frame) with the ith frame to find all unchanged pixels, which are made

transparent. Only the changed areas are left as shown in Figure 3(b). At last, the (i+1)st frame (Figure 3(c)) is recovered by combining the two frames (a+b). By keeping only the changed pixels between two adjacent frames, we further reduce the size of image and save the bandwidth. After the process of compressing and transparence, the images of one and half hours' lecture are about 800K in size only. Streaming 800K images within one and a half hours need roughly (in average) 2 kbps bandwidth. Considering the wave of the size of images and the condition of the network, we allocate the left 25 kbps bandwidth for the Part II. Therefore, the video & audio (30 kbps bandwidth) and the streaming images (25 kbps bandwidth) are reachable for the standard 64K ISDN.

Integration: Now we have a video stream (with audio) and a sequence of compressed images. We use SMIL and RealPix to combine the different medias in the same timeline, therefore the users (students) can replay these rich-media with Real Player.

RealPix, a tool from RealSystem, can stream images across an intranet or the World Wide Web(*RealNetworks Company*). Designed to integrate with SMIL, the RealPix language makes images to appear and transit in the RealPlayer window. Being still images that utilize transitions coded into the player, RealPix provides a low-bandwidth way but creating a dynamic slide show style presentation on the Web without extra programming. By combining RealPix with SMIL, we create multimedia presentations that playback over standard ISDN and modems. Our post-processing program can automatically generate RealPix file.

Synchronized Multimedia Integration Language (SMIL) is a W3C Recommendation designed for choreographing web-based multimedia presentations which combine audio, video, text and graphics in real-time (*Synchronized Multimedia Integration Language*). It is easy to learn and deploy on Web sites. A simple XML-based mark-up language, similar to HTML, enables an author to describe the temporal behavior of a multimedia presentation, associating hyperlinks with media objects and describing the layout of the presentation on a screen. SMIL is created specifically to solve the problems of coordinating the display of a variety of media (multimedia) on Web sites. By using a single time line for all of the media on a page, their display can be properly time coordinated and synchronized.

Our system integrates three types of media: Video, Text and Images. The three types of media are synchronized by SMIL and displayed as a whole streaming to the end user. Because we stream a sequence of images instead of video steaming, we can dramatically reduce the bandwidth of part II while still guarantee its quality.

Experiments and Assessments

T-Cube system has already been used in various scenarios including distance education and online symposium/conference. We also test the possibility of producing video-based Software Demo and eTutoring. Because different scenarios have different requirements, we discuss them separately.

Distance education: T-cube system has been used for distance education since 2002. At first beginning, the online lectures are only for the students within the campus. From the winter semester of 2002, the lecture "Internet Security: Weaknesses and Targets" has been given at University of Trier, Germany and simultaneously transmitted to Beijing University of Technology, China. The practical usage proofs that our system is successful. The feedbacks from students show that the quality of the online lecture is good. It is also very useful for students for review, homework and exam preparation.

Online Symposium/Conference: We used our system to record two German Symposiums and one International Symposium. Although the same system is used, we still face some different requirements. One difference from distance education is preparation for the presentation. Because attendees come from different areas or countries, it is infeasible to ask all attendees to prepare their presentations with some proprietary software. Although all attendees in the symposiums use MS PowerPoint, some of them insert some web page into their slides or run their demo program to enhance the impression of audience during their presentation. Since our system is able to capture the screenshot of the presentation, there is no problem to record these contents. Another difference is that most of attendees take their own laptops, thus our system must support Plug & Play. Usually these laptops have their own network configuration, in order to avoid changing their default configuration, we developed a small software to masquerade all IP address so that the laptop can have the network connection without changing its default setting.

Software/Product Demo and eTutoring: One interesting application of our system is to work as computer screen recording application to produce the Software/Product Demonstration or eTutoring. Some

commercial software (e.g. TechSmith SnagIt, Camtasia, IBM ScreenCam) have the same function by capturing screenshots. Traditionally captured screenshots are encoded as video streaming. However, we achieve it in a different way. Although we still capture the screenshots, we encode them into sequence of images instead of video streaming (see preceding Compressing and Encoding Images section). By this way, we can achieve the same purpose but dramatically reduce the size of the file.

Related Works

Automatically recording a live lecture/presentation and producing dynamic rich-media presentations for online or offline usage is not new. Several other systems and commercial software follow the similar idea (Abowd, 1999; *BMRC Lecture Browser*; *Cornell Lecture Browser*; Gil Cruz, 1994; Gregory D. Abowd, 1996; *Microsoft Producer*; R. Mueller, 2000; *Real Presenter*; Scott Minneman, 1995). Due to the limited network bandwidth available on the web, early systems only transmit audio of speaker and slides to the users(Alan F. Smeaton, 1997; *MANIC*; Mia Stern, 1997). There are also a growing number of multimedia authoring tools which combine slide presentations with video/audio of the speaker. Although the above mentioned systems and our T-Cube system may look similar superficially, there are several foundational differences. In order to better understand the distinction and to evaluate various approaches, we identified four separate phases of the execution of a capture and access system.

Preparation: Most of systems require the lecturers to use their proprietary authoring software for preparing the materials. Some of them provide a web-based interface (e.g. BMRC Lecture Browser), others (Lienhard J, 1998) are stand-alone applications (e.g. declass ZenPad, MS Producer, RealPresenter). These authoring software can import PowerPoint slides, images, WebPages, a slip of Video/audio etc. But some software (e.g. Lecture Browser, MS Producer, RealPresenter) can not work without PowerPoint software, because they depend on the Add-on component (installed on the lecturer's computer) special for PowerPoint. In principle, our system does not require the lecturers to use some proprietary tool for preparation. Because we capture the screenshot of the computer, any contents displaying on the screen are recorded. This feature is very important for the online conference/Symposium scenarios.

Presentation: Generally, all recent systems record a small, low-quality video of the lecturer using a digital camera. However, how to record the content of presentation and the annotation of the lecturer is quite different. BMLB and MS producer transfer the slides of PowerPoint into HTML. The advantage is the need of bandwidth is very small, but these system can not record the annotation of the lecturer. Besides, they can not work without PowerPoint. RealPresenter has the same problem. The ACE system at Kyushu University (Suganuma, 2002) uses another digital video camera to "photo" the blackboard. This method is easy and does not require other special hardware (e.g. electronic whiteboard). But its quality of capturing is not good and video streaming consumes too much bandwidth. Many recent systems (e.g. eClass, AOF etc.) record the contents and annotation with the help of electronic whiteboard. AOF and eClass developed their own recording tools (ZenPad, ZenViewer, AOFwb, aofShell etc.) to achieve this aim. These software record the electronic whiteboard activity (e.g. slides up/down, strokes of annotation, timestamp etc.) for later integration. Thus, recorded lecture can only be available after post-processing (integration). Although AOF supports so call "telepresentation", it assumes that all the high-volume data used in a telepresentation, such as images, movie clips, application programs, etc., have been transmitted to and installed on the connected remote hosts before the telepresentation actually starts. Distinguished from these systems, T-cube system records the slides and annotation by capturing the computer screen, and is able to broadcast the recorded multimedia contents in real time for the broadband users.

post-processing: Once class/presentation is complete, there are some efforts necessary to provide integrating the captured electronic whiteboard activity and the audio/video stream of the class. Hence some post-production works are necessary to produce a Web-based interface, and makes it easy for the students to access. Most of systems provide some tools to automate this work. But since our system integrates all streaming in real time, the left work for post-processing program is to compress the captured content to support the narrowband user (see preceding section).

Access/Replay: eClass, Berkeley Lecturer Brower and Cornell Lecturer Brower provide a web-based interface for the students. Students use the web browser (IE/Netscape) and a video player tool (Real Play, MS Media Player) to replay the recorded content. AOF provides its proprietary player (aofSync) to replay its specific media file. Currently we encode the lecture/presentation into RealVideo format, students

can use RealPlayer (and web browser) to access the recorded lecture. Meanwhile, the research to support MPEG4 has already been in progress.

Conclusion and Future Works

This paper presents a new easy-to-use multimedia content creating system. It produces a rich-media presentation from any seminar, talk or class easily and efficiently. It allows the speaker to go into the classroom, press a button, and give a presentation without caring about any technical details of the system. The contents of the presentation, including the annotation of the speaker, are automatically broadcasted on Internet and available for users in real time as well. After post-processing, the rich-media contents can also reach the narrowband users on demand. Through the proposed T-Cube system, virtual classroom is created for students. The contents of presentation, including video and audio, can reach across distances economically and effectively.

On the other hand, during the practical application of the T-Cube system, we also find some limitations and problems. One is the storage space. Because we have to save the uncompressed video (AVI) and images (BMP) for post-processing, these files consume too much disk space. One and a half hour lecture requires about 20 GB. We may solve this problem by compressing the AVI and BMP on the fly, but it requires a very powerful computer. Another problem is the information retrieval. Conversion of text to the image leads to the difficulty in indexing information. Our solution is to retrieve the text information before it is converted to image. A prototype application for information retrieval is under study. Currently we use SMIL 1.0 and RealPix to integrate multi medias (text, image and video). Upgrading the system to SMIL 2.0, which provides greater interactivity and embedded metadata, may result in further improvements.

Last but not least, capturing a live classroom is an essential building block for distance education instead of a complete solution. In the near future, this system will be included in a robust Learning Management Systems (LMSs), and will be capable to support Web-based multimedia and interactivity. The students on the computers can not only join the online course, but also interact with the lecturer and other online students. Multiply streaming formats, including MPEG4, will also be supported. With the help of RDF(Resource Description Framework) and knowledge base, multimedia content indexing and retrieval will be provided as well.

Acknowledgements

References

Abowd, Gregory D. (1999). Classroom 2000: An experiment with the instrumentation of a living educational environment. *IBM System Journal, 38*(4), 508-530.

Alan F. Smeaton, Francis Crimmins. (1997), *Virtual Lectures for Undergraduate Teaching: Delivery Using Realaudio and the WWW*, World Conference on Educational Multimedia and Hypermedia/World Conference on Educational Telecommunications (MEDIA/ED-TELECOM'97), Calgary, Canada.

Authorware. Macromedia Company. http://www.macromedia.com/software/authorware/

BMRC Lecture Browser. The Berkeler Mulimedia Research Center, University of California at Berkeley. http://bmrc.berkeley.edu/projects/lb/

Brent Martin, Antonija Mitrovic. (2002), *Authoring Web-Based Tutoring Systems with WETAS*, World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education (E-Learn 2002), Montreal, Canada.

Brian Bailey, Joseph A. Konstan, Robert Cooley, Moses Dejong. (1998), *Nsync: a toolkit for building interactive multimedia presentations*, The sixth ACM international conference on Multimedia, Bristol, United Kingdom.

Cornell Lecture Browser. Multimedia Research Group, University of Cornell. http://www.cs.cornell.edu/zeno/

Denis Helic, Hermann Maurer, Nick Scerbakov. (2002), *Implementing Complex Web-Based Training Strategies with Virtual Classrooms*, World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education (E-Learn 2002), Montreal, Canada.

Gil Cruz, Ralph Hill. (1994), *Capturing and playing multimedia events with STREAMS*, The second ACM international conference on Multimedia, San Francisco, California, USA.

Gregory D. Abowd, Christopher G. Atkeson, Ami Feinstein, Cindy Hmelo, Rob Kooper, Sue Long, Nitin Sawhney, Mikiya Tani. (1996), *Teaching and learning as multimedia authoring: the classroom 2000 project*, The fourth ACM international conference on Multimedia, Boston, Massachusetts, USA.

Kuo-Yu Liu, Natalius Huang, Bo-Hung Wu, Wei-Ta Chu, Herng-Yow Chen. (2002), *The WSML System: Web-based Synchronization Multimedia Lecture System*, ACM Multimedia.

Lienhard J, Maass G. (1998), *AOFwb: A new alternative for the MBone whiteboard wb*, Conference on Educational Multimedia and Hypermedia (ED-MEDIA'98), Freiburg, Germany.

M. Jourdan, C. Roisin, L. Tardif. (1998), *Multiviews Interfaces for Multimedia Authoring Environments*, The International Conference on Multimedia Modeling, Lausanne, Switzerland.

MANIC. University of Massachusetts. http://manic.cs.umass.edu/

Mia Stern, Jesse Steinberg, Hu Imm Lee, Jitendra Padhye, James F. Kurose. (1997), *MANIC: Multimedia Asynchronous Networked Individualized Courseware*, World Conference on Educational Multimedia and Hypermedia/World Conference on Educational Telecommunications (MEDIA/ED-TELECOM'97), Calgary, Canada.

Microsoft Producer. Microsoft Company.
http://www.microsoft.com/windows/windowsmedia/technologies/producer.aspx

MIT OpenCourseWare. Massachusetts Institute of Technology. http://ocw.mit.edu/index.html

Multimedia Fusion. ClickTeam Company. http://www.clickteam.com/English/aboutus.php

PNG Programming Resources. http://www.libpng.org/

Portable Network Graphics Specification 1.0 (1996). W3C. http://www.w3.org/TR/REC-png

R. Baecker, A.J. Rosenthal, N. Friedlander, E. Smith, A. Cohen. (1996), *A Multimedia System for Authoring Motion Pictures*, The fourth ACM international conference on Multimedia, Boston, MA.

R. Mueller, T. Ottmann. (2000). The "Authoring on the Fly" system for automated recording and replay of (tele)presentations. *Multimedia Systems, 8*(3), 158-176.

Real Presenter. RealNetworks Company. http://www.realnetworks.com/products/presenterone/

RealNetworks Company. http://www.realnetworks.com/

Scott Minneman, Steve Harrison, Bill Janssen, Thomas Moran,Gordon Kurtenbach,Ian Smith. (1995), *A Confederation of Tools for Capturing and Accessing Collaborative Activity*, The third ACM international conference on Multimedia, San Francisco, California, USA.

Stanford Online. Stanford Center for Professional Development, Stanford University.
http://scpd.stanford.edu/scpd/students/onlineclass.htm

Suganuma, Akira. (2002), *Development of an Automatic Camera Control System for Videoing a Normal Classroom to Realize a Distant Lecture*, World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education (E-Learn 2002), Montreal, Canada.

Suzanne Little, Jane Hunter. (2002), *A Tool for Creating, Editing and Tracking Virtual SMIL Presentations*, World Conference on Educational Multimedia, Hypermedia & Telecommunications, Denver, Colorado USA.

Synchronized Multimedia Integration Language. W3C. http://www.w3.org/AudioVideo/

Tele - TASK. Department Of Theoretical Computer Science and New Application, University of Trier. http://www.tele-task.de/

ToolBook. click2learn company. http://home.click2learn.com/en/toolbook/index.asp

VNC - Virtual Network Computing. AT&T Laboratories Cambridge. http://www.uk.research.att.com/vnc/

Volker Schillings, Christoph Meinel. (2002), *tele - TASK Teleteaching Anyware Solution Kit*, ACM SIGUCCS'02, Providence, Rhode Island, USA.