Automation Techniques for Broadcasting and Recording Lectures and Seminars

Robert Mertens
Zentrum virtUOS
University of Osnabrück
D-49069 Osnabrück
Robert.Mertens@uni-osnabrueck.de

Rüdiger Rolf
Zentrum virtUOS
University of Osnabrück
D-49069 Osnabrück
Ruediger.Rolf@uni-osnabrueck.de

Abstract

E-Learning plays an increasingly important role in modern university teaching. One fast and easy way to produce high-quality content for selected E-Learning scenarios is broadcasting and recording lectures and seminars. Due to cost-efficiency in courses shared by two or more universities and the possibility to communicate with experts from far-away locations, broadcasting technologies such as video-/audio-conferencing provide an attractive addition to live courses. For conventional lectures, recordings have proven to be valuable as they make it possible to repeat important parts of the lecture for students, who were unable to attend or for those who did not grasp specific topics during the lecture.

For most lecturers however, these new technologies are hard to use since they require constant attention and specialised technical knowledge. It is thus crucially important to reduce technology-interaction during recording and broadcasting in order to facilitate the use of these promising technologies. In the first part of the talk, we will identify a number of criteria for automating and thus easing the use of broadcasting (using video conference technology to share a lecture between two universities) and recording a lecture (to provide the recorded material to students at the home and other universities) while maintaining and in some cases even improving the quality of the filmed material.

When broadcasting a lecture it is important to move the camera's focus to where the attention of an interested student is (to the professor, who is speaking, or to the student, who is asking a question) and to select the best audio and video inputs for this position. The automation software must make it easy for lecturers or their students to indicate what should be broadcasted.

In recording a lecture it is essential to enrich the video with a synchronized representation of the material presented by the lecturer, i.e. PowerPoint slides, simulation programs, videos. The event of changing a slide must automatically be linked to the exact position in the video. Thus the video is synchronised with the sequence of slides the students follow on their screen. In the second part of the talk we will analyse to what degree these criteria are fulfilled by existing systems and we will give a brief overview of state-of-the-art technology in this field.

Introduction

People are used to learning things from television. They get their information from the news, educational or other programs. Thus, it is no question that people could learn scientific topics from recordings or live-broadcasts of lectures and seminars, too. But unfortunately most universities cannot afford the cost for technicians and videographers that would be needed. Technicians are needed anyway to maintain the equipment, but a videographer, who films the lectures, poses an extra cost. It would be good if the technician who is present at the location could care about several events at the same time and would only be needed if there was a problem in one of the activities. This can only be achieved if the systems for recording or broadcasting are easy to use and have most functions automated.

Automation Techniques

There are several ways to automate a lecture room. It is essential that it is easy to turn on the system. It would be best if the system could be started by pressing a button and would operate on its own afterwards. But unfortunately this is not how most current systems work. Many devices have to be started separately, because they are not needed in every session and consume too much resources. There might be devices in the future which fit better into an automated system. At the moment existing systems are examined to find out how they can be automated.
Many universities use media-control-devices (e.g., from Crestron)[1]. The automation of the broadcasting or recording devices can be integrated into the panels used by the media-control. Another possibility is to use a small software-program on the PC that is used for the presentations that starts all other devices.

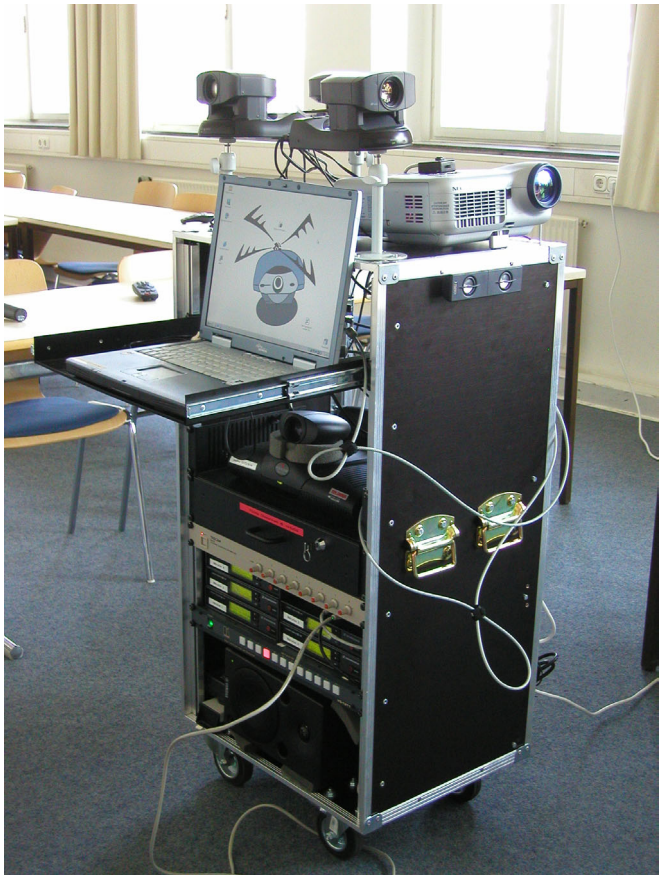
There are some tasks that the lecturer can prepare before the session. In Osnabrück we have developed a system, where the lecturers can set up the videos that they want to show, the telephone-conference participants they want to call, the internet-pages they want to show and the polls they want to make before the lecture over an internet interface. They can start all these things by pressing a button on an infrared-remote. They define the order in which they want to show the different media before the lecture over the internet interface. However, the sequence of the media can also be defined during the lecture via the infrared-remote.

We use infrared-remote for controlling the cameras and microphones as well (Knaden and Rolf, 2003). It is very important for the automation of lecture-rooms that the cameras can be moved remotely. If a person would be needed to operate a camera, he would be bound to this task and could not do anything else. He could also only work at one lecture at a time. Thus, automation is very important. We use infrared-remotes to move pan-tilt-cameras to predefined positions and switch the microphone input to the ceiling microphone next to these positions. Every person in the room has one remote (in some seminars 2 or 3 people have to share a remote). With this remote the cameras can be switched to the predefined camera-position for this remote. The lecturer is always able to switch back to the last predefined position.

The lecturer can switch to any predefined position in the room, and he can also adjust the cameras during the lecture. The system can be used remotely over the internet, too. Then, instead of a remote control a software-program controls the cameras. In this scenario the

[1] For more information on Crestron see: http://www.crestron.com

lecturer does not have to sit in the same room with the audience, but he can still address someone in particular in the audience.
We have found a way to minimize the size of our system. It has now become a mobile broadcasting unit. We call it “Flying Classroom”. It has nearly all the functions mentioned above and fits into a 19” rack. On the top there are three cameras and a beamer. To be able to move it wherever we want there are four wheels under the rack.



Picture 1 The Flying Classroom

There is another much more elegant way to control the cameras automatically. Microsoft Research (Rui et al. 2003) has developed a system that moves the cameras with computer-vision and microphone array techniques. The computer-vision approach is integrated through the use of additional static wide angle cameras that direct the pan-tilt-cameras to the positions where something is moving. Heuristics were created to minimize the movement of the cameras, as movement of the cameras is very disturbing for the viewer. The microphone-array-technique works with two or more microphones. Sound spreads with a specific speed and it travels different distances to the microphones. From the time-difference at which the sound reaches the microphones the position of the sound-source can be calculated. This technique is used to move the camera for the audience to that person who asks a question.
The described system uses several cameras with virtual camera-men and a virtual director who decides which camera will go on-air.

Broadcasting

Normally lectures are filmed to be replayed later. But there are some reasons why live-broadcasting a lecture or a seminar is worth the effort.

One scenario might be that there is a famous professor visiting a university for a talk and this university wants to share this event with partner universities. The students who watch the talk in a videoconference have the chance to ask questions and maybe even start a discussion with the lecturer.
Another scenario would be that a university does not have a lecturer for a specific subject and decides to participate in a broadcasted lecture from another university. It is essential that the students can interact with the lecturer during the session. This is the case, because the students need to be able to ask questions. Only then the lecturer can interact with the students and change his talk or maybe even change the schedule of the whole course.
These are only two prototypical examples for scenarios, where live-broadcasting would be useful. Another example would be a seminar in which students from different locations and with different skills and knowledge work together on a specific topic and share their knowledge.

There are different technologies for live broadcasting. Very common is videoconferencing with H.232 devices, like MS Netmeeting, Polycom, PictureTel, etc. The H.232 standard does not only allow point-to-point connections, with special MCU-devices multi-point conferences are possible, too. The video and audio quality is good at a medium network-bandwidth (from 64kbit/s up to 2Mbit/s, typical 512kbit/s). The time-delay for the viewers is typically under 1s.
For a better quality MPEG2-streaming is possible with special devices. Then the audio and video is nearly DVD-quality but it uses much internet-bandwidth (over 5Mbit/s) and the time-delay is about 1s. With this quality only point-to-point conferences are possible.
Both techniques need an additional internet-connection for broadcasting the presentation slides with MS Netmeeting or VNC.
There are service providers who offer tools that can broadcast the presentation slides and video over one connection (e.g., Centra or WebEx). Most of these tools need a separate audio connection over telephone. The video has a very slow frame-rate, a poor resolution and a rather high noise level. The bandwidth of these tools is moderate so that a modem or DSL connection should be fast enough. The time-delay can be several seconds. The greatest advantage of these tools is that they can easily establish multi-point connections.

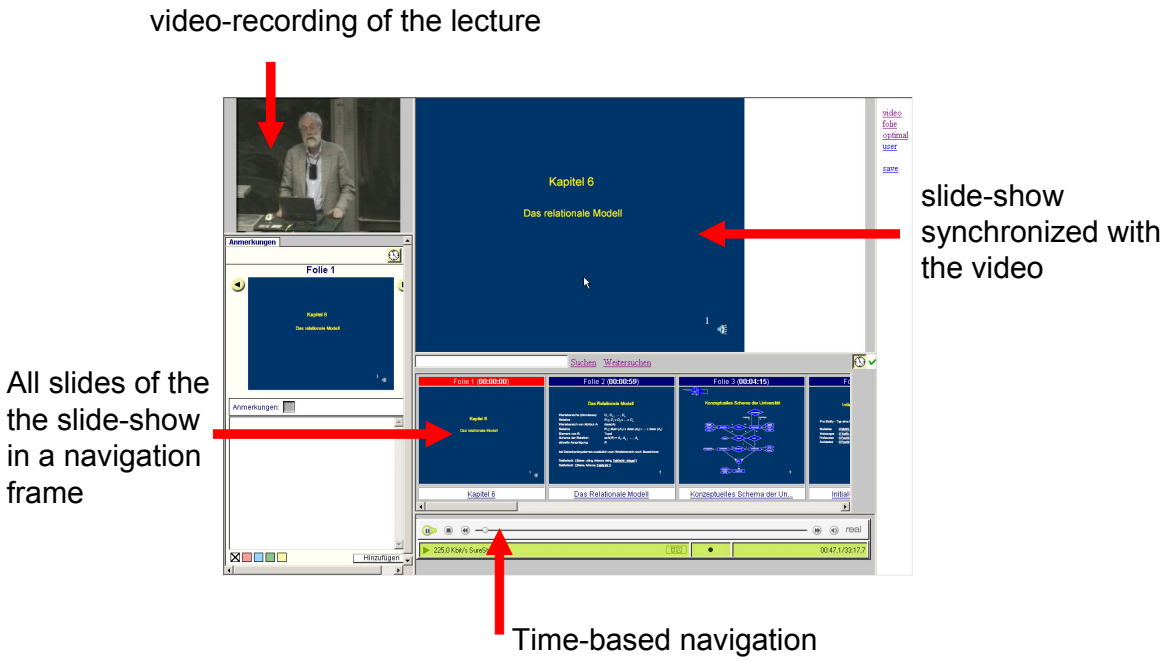
The problem with live-broadcasting is that all the different locations have to be coordinated. The rooms on all places have to be booked, the technicians and the students on all locations have to be at a certain place and so on. Another problem is that lecturers have to be convinced that they want to broadcast their lecture and other universities have to be convinced that they want to import this lecture. An infrastructure has to be created that allows an easy exchange of lectures between different universities.

Lecture recording tools

Apart from the video-related problems described above, lecture recording brings a whole set of new problems to the task. Recent overviews have treated these problems with both foci on the rather technical problems and on the information retrieval side of the matter.
Below we will cover both sides in detail, starting out from a representative paper for each focus.
In this section we will only treat PowerPoint-based lecture recording tools due to the facts that these are the easiest to use, that they are rather wide-spread, have very little hardware requirements and offer a comparably well structured presentation of the lecture to the learners. In the following section we will give a short introduction to the general look-and-feel of these tools' viewer interfaces before we proceed to discussing the topic in detail.

What does a recorded lecture look like?

To the learner, most PowerPoint-based lecture recording tools offer a user interface similar the one shown on the screenshot below. They show a video-recording of the lecture, a slide-show that is synchronized with the video as well as navigational elements. The navigation found on most viewers consists of a time-based navigation like that of the RealPlayer and of a slide-based navigation that can be represented as a miniature view of the whole slide-show or a as table of contents.



Technical requirements and their problems

(Lauer & Ottmann 2002) list a number of technical features such as symbolic representation of contents[2], dynamic capture of hand-written annotations and screen recording as well as two navigation features as the most relevant criteria for the evaluation of these tools. The navigation features advocated for by the authors are a rather coarse-grained thumbnail-overview that allows for slide-by-slide navigation and visible scrolling.

Visible scrolling is a mechanism that instantaneously reflects any changes made to the navigation-slider in the film- and in the slide-window. One major disadvantage of this approach is that it requires special, proprietary media players and/or a local copy of the film shown (Lauer & Ottmann 2002).
Another problem is constituted by the fact that visible scrolling only works if thematic changes are reflected in the visual part (slides and film) of the recording. In the filmed part of

[2] Usable for both full-text search and post-editing

a recorded lecture this is rarely the case, in the slide-part this requires the exact reproduction of the order and timing in which objects leave and enter the PowerPoint-slides. Unfortunately some types of animations are not supported by systems like Lecturnity, which convert slides to a proprietary format before the presentation (Lauer & Ottmann 2002). In this way, meaningful information gets lost on the presentation's way to the web.
This problem might in part be solved by the dynamic capture of annotations, which is implemented in Lecturnity and AOF (Lauer & Ottmann 2002). This, however, requires the lecturer to add handwriting to the slides during the presentation, which in turn is only possible with special hardware. Another problem at this point is that lecturers simply might not want to hand-write into their carefully crafted presentations.
An alternative solution is an automatic pre-processing, that decomposes a slide into several slides, one for each effect in the presentation. The challenge in this is to make both presentations look the same, so that the audience do not see a difference, neither during the lecture nor in the recordings. To remove multiple representations of one slide in the navigation-elements, a simple post-processing is needed. We have recently implemented this pre-processing and we are currently testing it in cooperation with the University of Clausthal.

Screen recording is needed whenever the lecturer switches to another program during the lecture. This is especially important for lecturers who need to demonstrate computer programs or simulations during the lecture. Automatic generation of a structured overview problem and navigation within the screen recording are still unsolved problems, even though approaches like DHL's listener-technology (Mühlhäuser & Trompler 2002) sound promising.

Another feature which is treated in more detail by (Einhorn et al. 2003) is the incorporation of metadata in the recorded document. Among other information, metadata can contain keywords, the language in which the course is taught or a course description, which makes the recorded documents accessible for content management.
Which kind of metadata is required differs from application-scenario to application-scenario and should best be evaluated by the users themselves. Currently there is a number of systems available that support the use of metadata. These systems include AOF and Lecturnity.

Information retrieval

(Brusilovsky 2000) has a special focus on retrieval-oriented features. He addresses line-level synchronisation in conjunction with direct addressability of each line as one of the most important retrievability features. There is currently no PowerPoint-based system available that combines these two features. Added a full-text search mechanism these two features[3] could directly lead the user to any desired piece of information in a recording.
Full-text search, however, is not very common in PowerPoint-based lecture recording tools. The only two systems supporting full-text search are AOF and Lecturnity, which both need special players in order to provide this functionality. An ideal solution should allow full-text search directly in the standard web-interface.

Another feature mentioned by (Brusilovsky 2000) is the supplemental delivery of annotations, links and references. Links and references can help the learner to get to related materials and annotations can clarify difficult or ambiguous parts of the lecture provided the existence of an interface that allows to add and alter these annotations after the recording. An interesting

[3] Line-level synchronisation and direct addressability.

approach in this could be the use of a wiki-like (Leuf & Cunningham 2001) mechanism that allows all participants of a course to change the online-annotations to a recorded lecture at any time. This, however, requires the course-material to be reasonably difficult, in order to motivate the students to replay parts of it repeatedly.

A feature not addressed by (Brusilovsky 2000) is web-accessibility and searchability of the documents with conventional search engines. Major search-engines like Google have added PowerPoint's file format to their list of searchable web content rather early in the history of the WWW (Fletcher 2001). This clearly shows the demand for presentations to be searchable and accessible via the web. Unfortunately there is currently no system that generates documents indexeable by web-search-engines. Systems like AOF however provide special search engines for their content. The problem with these search engines is that they can not be used for other formats which will most probably reduce the user acceptance.

The ideal lecture recording tool

After having taken a look at the current state of research in the field of lecture recording tools, we can now list the features identified above to give a short and concise overview of the relevant criteria for lecture recording tools. As a conclusion from the above paragraphs, the features of an ideal lecture recording tool can be summarised as follows:

- Advanced navigation
- Animations
- Capture of hand-written annotations
- Full-text-search
- Line based addressability
- Line based synchronisation
- Metadata
- Post-editability
- Screen Recording
- Searchability by conventional search-engines
- Supplemental delivery of annotations, links and references at any time

This list might not be complete and it might even contain some features that are not needed in all application contexts[4], but it still provides a good orientation for the selection or the development of future lecture recording tools.
As to current tools, the one that comes closest to these requirements seems to be Lecturnity, even though a number of competing but not yet competitive approaches are being developed both by commercial and academic institutions.

[4] One of these features might be the capture of hand-written annotations which is by far not used by all lecturers.

Conclusion

The technologies for broadcasting and recording became affordable for universities and they have proven in practical use that there is no other real technical problem. But there is still a technician needed to operate the devices, but affords to automate the devices show first prosperities and there will soon be a completely automated lecture room, where technicians only have to do maintenance work. Much more lectures will be recorded or broadcasted.

Our analysis of both the literature and our own experiences has identified a number of criteria for the evaluation of lecture recording systems, which are unfortunately not yet fulfilled by any system available on the market. We have created a concise list of these criteria which may serve as an orientation for the design and evaluation of future systems in this field.

Live-broadcasting of lectures needs elaborated planning. An infrastructure for the marketing of such lectures has to be built up and didactic concepts for broadcasted lectures have to be found.

Literature

Brusilovsky, P. (2000). *Web lectures: Electronic presentations in Web-based instruction*. Syllabus 13 (5), 18-23.

Einhorn, R., Olbrich, S., Nejdl, W. (2003*). A Metadata Model for Capturing Presentations*. Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT 2003), 9-11 July 2003, Athens, Greece.

Fletcher, W. (2001). *Concordancing the Web with KWiCFinder*. In Proceedings of the Third North American Symposium on Corpus Linguistics and Language Teaching. Boston, MA

Knaden, A. & Rolf, R. (2003). *Automatisierte Kamerasteuerung zur Übertragung von Seminaren und Vorlesungen*. In: V. Dötsch, G. Schade & K. Hering (Eds.), e-Learning and beyond. Proceedings of the Workshop on e-Learning 2003, HTWK Leipzig, 14.-15. July. 87-97.

Lauer, T. & Ottmann, T. (2002). *Means and Methods in Automatic Courseware Production: Experience and Technical Challenges*. World Conference on E-Learning in Corp., Govt., Health & Higher Ed. 2002(1). 553-560.

Leuf, B. & Cunningham, W. (2001). *The Wiki Way*. AddisonWesley Longman.

Mühlhäuser, M. & Trompler, C. (2002). *Digital Lecture Halls Keep Teachers in the Mood and Learners in the Loop*. Association for the Advancement of Computing in Education (AACE) (Ed.). Proc. E-Learn 2002. 714-721.

Rui, Y., Gupta, A. & Grundin, J. (2003). *Videography for Telepresentations*. In: CHI letters, Volume no.5, Issue No. 1, CHI 2003, April 5-10, 2003, Ft. Lauderdale, Florida, USA. 457-464.