Video-lectures over Internet: the impact on education

Marco Ronchetti
Dipartimento di Ingegneria e Scienza dell'Informazione
Università degli Studi di Trento, Italy

ABSTRACT

Recent trends suggest that multimedia (in particular audio and video) will increase their share of Internet traffic, and that users are becoming more and more acquainted with viewing multimedia content on their computers and mobile devices. Scholarly institutions are experimenting since several years with the distribution of video-lectures, which are generally not meant as a replacement of traditional lectures, but rather as a different kind of support of the educational process. After recalling the pioneering ideas, in this chapter we discuss the pedagogical soundness of the idea of using videos over Internet for teaching and learning. Later we review the various directions that research has taken over the last 10 years to support and enhance this modality.

KEYWORDS

Asynchronous Education, Distance Education, Distance Learning, Educational Multimedia, Educational Technology, Education Research. E-Learning, Internet-Based Instruction, Internet-Based Technology, Multimedia, Multimedia Streaming, Video Streaming, Video Retrieval, Video Archives, Web-Delivered Education

INTRODUCTION

Over the last decade, Internet popularity has been steadily growing, up to a point where the time share spent on the two media is almost equal: according to an IBM survey of consumers "66% reported viewing 1-4 hours of TV per day, versus 60% who reported the same levels of personal Internet usage"[1]. A new trend has been emerging over the last year: watching videos on Internet, plus watching time-shifted TV (which can also happen through the Internet) amounts now at 6,7% of the total time devoted to watching videos (Nielsen 2009), and the figure is steadily growing. It is very likely that these two media, which where were once quite separated, will merge in future at least to a large extent. This media hybridization is already evident: short (1 to 5 minutes) video clips on the Web are becoming more and more common – and many TV

[1] http://www.marketingcharts.com/television/ibm-consumer-study-internet-rivals-declining-tv-as-primary-media-source-1340/

channels give the possibility to access (part of) their traditional TV content through the Web. Producing and distributing (video) content through the web is getting easier and easier, even though there are still a number of open issues.

This fact boosts several research streams, ranging from assuring quality of service to indexing and extracting knowledge from videos, to usability concerns, to the many new applications that can be invented by using this new media combination.

Among the possible applications, those in learning are certainly very interesting. Actually the field of e-learning has been a precursor in using the videos over the Internet. There is in fact little doubt that education in all its forms, but especially in the field of life-long learning where just-in-time services might be important, will continue taking advantage from this technology. There is in fact an evident trend in the evolution of our society that strongly changes the paradigms of the previous millennia. In the past, learning could be mostly concentrated in the early stages of life. Once one learned a job, the acquired knowledge was sufficient for the rest of the life. In today's society this is no longer true. It is quite obvious in the technological arena, where knowledge obsolescence is quite rapid: in the computer science field, for instance, the known paradigms can radically change within a very short time span. This trend however is general, and all sectors are affected by it. Hence the ability to update one owns knowledge becomes vital to survive in the job market. The need of continuous education and long-life learning permeate our society, and it will be always more so. Training on the job, just in time training and education update need to find their way into our life. Informal education, and the ability to autonomously find and use suitable educational resources will play an important role in this scenario. In this setting, the availability of multimedia-carried educational content is an exceptionally powerful asset.

Today we see a flowering of education-related initiatives ranging from producing podcasts (i.e. audio-streams) and webcasts (i.e. synchronized slide shows + audio) to recording "live" lectures, to producing "how-to" You Tube videos. In recent years several custom systems aimed at supporting the process of acquiring and distributing videolectures have been developed: some were commercialized, some were put in the public domain, and some were used locally as prototypes. Each of these systems has a slightly different twist, and some implement new clever ideas to respond to the peculiar needs that may arise in different situations. A review of the requirements for such systems has been compiled by Ronchetti (2008).

In this chapter we shall first briefly review the pioneering ideas, then we shall discuss the pedagogical soundness of the idea of using videos over Internet for teaching and learning. Then we will review the various directions that research has taken over the last 10 years. We shall not attempt to provide an exhaustive review, but rather we shall try to highlight the ideas and the trends.

BACKGROUND

The possibility of using digital video for distance education was envisioned already 15 years ago when F. Tobagi (1995) built in Stanford a prototypical architecture for distributing digital video lectures. Although using (analogical) video for recording lectures had been in use for more than a decade, the digital approach was obviously superior in terms of ease of distribution of the didactic material, asynchronous and multiple simultaneous accesses. Moreover it was promising in terms of possibility of cross-referencing the learning resources.

The first report of actual deployment of an application of video-streaming to teaching we are aware of dates back to 1998 (Hayes). At that time, the main goal was to substitute a VHS based system for delivering lectures to a geographically remote place (from USA to France) with a digital alternative. At the beginning it was limited to transmitting an audio stream with synchronized power point images, but shortly thereafter it evolved into a video transmission that included both the teacher and the slides with a technique called chroma key[2].

One year later. Hilt, V. & Kuhmünch (1999) described the results obtained in the framework of the German Teleteaching project that was envisioned in 1995. They portrayed three scenarios: the “Remote Lecture Room Scenario", in which lectures were transmitted live via the Internet to lecture rooms at the partner universities, the “Remote Interactive Seminar" (much more interactive since a discussion between the participating actors usually took place after each talk) and the “Interactive Home-Learning", in which students running Linux or Windows 95/98 on their PC watched lectures from home via an ISDN Internet connection. Although they focused mostly on synchronous events, some attention was given to the asynchronous scenario.

As a variation, Chen et al. (1999) proposed a framework to record lectures in which the teacher presents learning material in the form of HTML pages. The web navigation events were recorded and associated to the video, so that they could be replayed when the students asynchronously watched the lecture.

In these years the main challenge was to overcome the problems caused by Internet bandwidth, as discussed by Fong & Hui (2001). Related to this issue, there was the choice of the software for performing video-streaming, the main competitors being RealVideo, Apple QuickTime and Windows Media Video as discussed e.g. by McCrohon et al. (2001). Since none of these solutions was really cross-platform (supporting at least Linux, Windows and Apple OS), many opted for the storage in multiple formats. Later also Adobe Flash Video entered into the arena, offering universal support. The situation is still in evolution, as Microsoft recently responded with its Silverlight platform in the attempt to make some previously Windows-only products portable also to other platforms. Obviously the quality of service of the streaming has an impact on users’ perception of the video content, and hence on their satisfaction and on the overall system usability, as discussed e.g. in Chen et al. (2006).

After a short time in which pioneers opened the way, showing a possible but uncertain future, big players are coming onto the scene today, such as in the cases of the Massachusetts Institute of Technology (http://watch.mit.edu/) for video-lectures and seminars and the University of California at Berkeley which, according to U.S. Government news[3] of January 2008, is the first University with a plan to offer full courses through You Tube. Also, recently Apple started a dedicated section in their iTunes platform that is called “iTunes U”[4] and is dedicated to academic videolectures. The history of creation of the OpenCourseWare initiative at MIT is reported by

---

2 http://it.wikipedia.org/wiki/Chroma_key

3 http://www.america.gov/st/educ-english/2008/January/20080122181508lCJsamohT0.1036035.html

4 http://www.apple.com/education/mobile-learning/

Abelson (2008). The MIT answer[5] to the issue of whether students will balk at paying about $26,000 per year in tuition when they can get all their materials on line was: "Our central value is people and the human experience of faculty working with students in classrooms and laboratories, and students learning from each other, and the kind of intensive environment we create in our residential university". This statement supports our view that the pedagogical experience cannot reduced to the (video)lecture, which introduces our core topic: what is the pedagogical value of the video-lectures?

VIDEOLECTURES FROM A PEDAGOGICAL PERSPECTIVE

Pedagogical effectiveness of video-lectures

Even though recently e-learning seems to be passing the trough of disillusionment –i.e. the big depression that follows greatly exaggerated expectations in the hype cycle[6] - (Friedland et al. 2009), it is generally believed to be an effective instruction methodology. Some experiments have demonstrated that, at least under certain situations, e-learning can be at least as effective as conventional classroom learning (see e.g. Zhang et al. 2004, Dutton et al. 2001). But what about the particular form of e-learning given by video-lectures?

Lauer et al. (2004) evaluated the effectiveness of some aspects of the user interface of video-lecture based systems. They found that students considered as important elements:

- Sound quality
- Good readability of the accompanying slides, annotation and clips
- The possibility to have also a local copy of the learning material (as opposed to being able to only view in streaming mode)
- The possibility to navigate the lecture both by slide title and by a time-bar
- The availability of text search (keyword or full text)

Similar results were obtained by Ronchetti (2003a), who also investigated which were the advantages perceived by the students. Some advantages came from the possibility of not being present in class during (some) lecture:

- possibility to recover lectures lost due to forced absence (illness, work or other time-frame incompatibility);
- ability to better organize their time, deciding not to be present at some lecture (elective absence);

Other (less obvious) advantages emerged from investigations with the students who regularly frequented the course:

- possibility to review some critical point (cases of poor understanding of a section due to concentration drop, excessive speed in an explanation or intrinsic difficulty);
- possibility to review (portion of) lectures as a confirmation that their understanding is correct;
- ability to check the correctness of notes taken during a lecture.

---

[5] Goldberg. C. Auditing Classes at MIT, on the Web and Free. New York Times (2001)

[6] http://en.wikipedia.org/wiki/Hype_cycle

The interesting point here is that most students did not consider video-lectures as an alternative to class, but rather as an interesting new tool to complement their traditional learning. This is probably also the explanation of why they did not consider the offer of a synchronous mode to be a value.

In all these cases, the possibility to quickly navigate the lecture was essential (and this is one of the aspects that makes digital recording deeply different from more traditional VHS-based videos, the other being the ease of access via download). The relevance of direct access to a particular point of a lecture for reviewing only a few minutes is also supported by the analysis performed by Zupancic & Horz (2002). According to Soong et al. (2006) "recordings enable them to access parts of lectures which they do not understand". Chiu et al (2006) report that students are more willing to use the taped class lectures for reviewing purpose if the taped video is from their own class (rather than from another class). The work by Zhang et al. (2006) shows that students using what they call "interactive video" (i.e. video-lectures with the possibility of random access) achieved significantly better learning performance than those in other settings, and also showed a higher level of learner satisfaction.

Ronchetti (2003a) also reports that students who did not pass the exam at the first call made (later) an intense use of the video-lectures – meaning that they help exactly those students who need more support. Zupancic & Horz (2002) report a strong increase of system usage right before the exams, which also suggests a form of just-in-time support.

Bennett & Maniar (2008) discussed the efficacy of video-lectures. They tried to offer a balanced view, presenting positive and negative aspects. Their review is not conclusive, as they cannot support their personal opinions with a scientific evaluation, so that the matter remains unanswered. Moreover, in their review they mix two different aspects: effectiveness of *videos as teaching tool* and effectiveness of *recorded video-lectures*. This confusion is prototypical: in literature, very often, different aspects get mixed. Another frequent mistake in literature is the extension of results obtained in a given domain to nearby area, where their holding may be questionable.

Berner & Adams (2004) conducted a randomized controlled trial to assess the importance of video+slides as opposed to audio+slides. According to their results, adding video to an audio presentation does not lead to greater satisfaction or greater learning.

The German instruction psychologist Glowalla (2004) expresses a different opinion: according to him, learners show a better concentration in front of a video than in a classroom, while the audio + slide version favours less the concentration, and is perceived as more boring. Data obtained by Reisslein et al. (2004) confirm that students felt that the web-carried video helped them to stay focused during the instruction. McCrohon et al. (2001) report that 83% of the students preferred video streaming to audio streaming. Fey (2002) argues that video does not increase understanding or retention, but is favoured by most people due to emotional reasons.

Moreno & Mayer (2000) state a general principle that would bring to the opposite conclusion: "Students learn better when the instructional material does not require them to split their attention between multiple sources of mutually referring information". Their work is sometime used as an argument against the standard video-lecture + slide format, as reported e.g. by Krüger & Plaza (2005). However, one should be careful in interpreting Moreno & Mayer's results. In fact, their experimental work is mostly concerned with generic multimedia rather than with the specific case of video-lectures. Krüger & Plaza extrapolate their results, considering that in video-lectures there are two areas of the screen, which are competing for the viewer's attention (the video and the slide). Hence they conclude that this leads to the “split-attention effect”. We believe that this syllogism is misleading. Our argument is that the video-lecture mimics so closely the real-world experience (with the lecturer, and an often large projected slide that is not necessarily near to the speaker) that students might find this “spilt” a natural feature, and hence not suffer of the “split-attention effect”. To ascertain if Krüger & Plaza are right, or rather if our view is correct, it would be needed to repeat Moreno & Mayer’s experiments on this particular setting rather than just extend their results by similarity.

Friedland & Rojas (2008) took the “split-attention effect” very seriously, up to the point of building a system that merges the image of the lecturer taken from the video stream with the board or slide image, so that the two sources of information get reunited.

Hermann et al. (2007) attempted to extract information on users’ preferences among the two formats (audio+slide vs. video+slide) from the access log files of a system for delivering audio- and video-lectures, but their evidence is not conclusive.

We share the opinion that audio and video are almost equivalent as far as content is concerned (and moreover it is amazing to observe that in certain cases the slides are more helpful as a guide for the speaker than as a support for the listener, even if we do not question their utility for the learner in general). The video is however probably slightly better than audio for a few reasons:

- Viewing the speaker (at least for a portion of the lecture) gives a sense of familiarity that helps getting emotionally more involved;
- Watching the video helps keeping concentration
- in the cases in which the blackboard is used to complement the content of the projected slides the video gives a clear advantage.

Recent work tends to be mostly positive in evaluating the success of video-lectures, as shown by the following anthology:

- “92% of students who access the video-streamed lectures (…) agreed that this was a useful learning resource”. (McCrohon et al. 2001)
- “students find an added value in having a multimedia version of the traditional lecture, especially if provided through a tool that has a well-thought user interface” (Ronchetti 2003b)
- video-lectures-based “distance education was at least as good as traditional classroom instruction” (Reisslein et al. 2005)
- for a good majority of students, “a distance learning course without on-line lecture would compromise learning” Chung (2005)
- “students have benefited from accessing video recorded lectures” (Soong et al. 2006)
- video-lectures “are indeed adequate alternatives to live lectures for engineering students” Maness (2006)
- “the use of video (in the absence of the teacher) in teaching primary school pupils is as effective as when the teacher uses the real objects in teaching Agricultural and Environmental Sciences” Isiaka (2007)

Looking further in the past instead, early experiments report a clear inferiority of remote lectures when compared to in-classroom lectures (see e.g. Hilt et al. 1999): “The Home Learning students as well as those in the remote classroom rated perception in a normal face-to-face lecture higher". The reasons mentioned were the unfamiliarity of the technological equipment, the poor audio quality and a merely sufficient visual quality of the slides in the whiteboard. Additionally students in remote situations reported difficulties concentrating for a full, uninterrupted 90-minute lecture and did not feel themselves perceived by the lecturer. Obviously in these years most technical issues have been fixed, the bandwidth has increased, and probably today's student are more used to videos over the net. Also, the experiment to which Hilt refers to is synchronous – and later we shall discus the difference between asynchronous video-lecture and lecture over videoconference.

Not all the problems have however disappeared. On the negative side, in a distance-learning context that was heavily based on video-lectures, students reported insufficient interaction with the instructor and insufficient interaction with fellow students. Also, the students quite frequently reported the required self-discipline as the least liked aspect of distance education (Reisslein et al. 2005). This basically means that video-lectures are an essential ingredient, but they do not suffice. This point of view is reinforced by the quite innovative work presented by Brown & Liedholm (2004). They start from the consideration that students exhibit different cognitive styles, but unlike other lines of research focusing on system personalization, they rather offer students a spectrum of possibilities (textbook, video-lectures, practice quizzes, problems, text slides, lecture slides) so that each student can choose the learning strategy that better suits her/him. They monitor and observe students behaviour during the learning process, and finally they draw conclusions on the relative usefulness of every single resource type. Their results show that video-lectures were very valuable for 77% of the students, with more than half of the students considering them one of the two most valuable tools.

A very interesting paper by Fritze & Nordkvelle (2003) analyzes the difference between a live lecture, a video-lecture and a lecture in videoconference. It should be noted that by video-lecture they mean a lecture given in front of a camera, without the presence of a real class: we shall call this a "synthetic" lecture, because in most cases in literature "video-lectures" are actual recordings of live lectures (i.e. lectures given in front of a real audience, and not created on purpose for distance learning). They studied live lectures, synthetic lectures and lectures in videoconference by analyzing the language employed, the examples used, the rhetoric figures, the presence of humour. Their results show that the synthetic lectures lose many of the immediacy qualities of the live lectures: "there are less humorous situations, the examples are de-personalized and presented with less succulence, contain very few questions and with higher intensity (….) video-lectures contain shorter and more focused examples (…) The lecturers compensate for the lack of immediacy by employing a more inclusive linguistic style." In summary, synthetic lectures are more serious, stringent and linear, with the de-personalized and distanced used of examples and humour. According to the authors, synthetic lectures are closely associated to a scientific frame of reference: to science as a social system. Live lectures instead present a "more dialogic context (…) with more humour, personal examples, questioning and gestures. (…) such techniques (are used) in order to reduce complexities, and supply more cues to understand the subject matter. The lecturer's building of trust differs similarly: the serious video-lecture builds confidence in the lecturer primarily as a person speaking on behalf of the science", whilst in the live-lecture the speaker builds personal trust primarily as an *educator*.

In our view, these findings are not against the use of video-lectures, but rather suggest that video-lectures should be the recording of live events, and not “canned” presentations. By the way, the cost of preparing a synthetic lecture is vastly superior to the cost of recoding a live event (that happens anyway).

Another interesting result by Fritze & Nordkvelle concerns lectures in videoconference. Although they should in principle share the qualities of a live lecture, this is (at present) not the case due to the mediation of the technological infrastructure. The difficulty of the teacher to actually “see well” the audience induces in her/him a limitation of gestures, walking, situated humor, questions that in the end contribute to a minor immediacy than obtained in live-lectures, making lectures in videoconference more similar to synthetic lectures.

Using the video to change the didactic paradigm

A word of caution comes from Demetriadis & Pombortsis (2007). They bring evidence that “e-lectures can be safely used as students’ introductory learning material to increase flexibility of learning, but only within a pedagogically limited perspective of learning as knowledge acquisition (as opposed to construction)”. Their statement echoes what many pedagogues say, when they criticize the (traditional) lecture model because during a lecture students are passive. There is no space for autonomous construction of knowledge, and often even the (potentially possible) degree of interaction is close to zero: especially in large classrooms, the number of students’ questions and observation is low. Teachers often respond that they’re busy covering the program, and that they have no time for fostering interaction and stimulate discussions. As a side effect, certain learning styles are frustrated. Lage et al. (2000) envisioned a solution that they called “the inverted classroom”. After presenting an interesting review of different taxonomies of learning styles, they suggest that events that traditionally take place inside the classroom should be placed outside the classroom and vice versa. In practice, they suggest that time in the classroom should be spent only for discussion, cooperative work in small groups etc. The traditional classroom work (i.e. the traditional lecture) should be placed outside the classroom – thanks to multimedia technology. They report experiments in which the perception of both students and instructors were positive.

Independently, Foertsch et al (2002) came to the same idea. They also started from the need of moving towards a learn-by-doing paradigm, and encountered the hurdle that “before students can be effective team members or problem-solvers, they need to have a basic understanding of the problem domain, some background knowledge about how problems can be solved, and instruction on how to use the tools at their disposal”. So they tried to use “distance technology” to actually reduce the distance between students and professors. The vision was that if all of a professor’s lectures, syllabi, and assignments are digitized and put online, professors could spend less of their time lecturing and more time assisting the students. This approach was experimented with a course and reported in their work. As a result, “the replacement of live lectures with online lectures and Team Labs significantly enhanced the usefulness, convenience, and value of the course for the majority of students”.

Day & Foley (2006) and Ronchetti (2009) came autonomously to the same thought, and reported similar results. The idea in itself is actually not new – in the past some teacher gave as homework the duty to read a chapter of a textbook – and then the work in class was focused on the discussion of material that students had previously read. The availability of recorded lectures however gives a fully new twist to this, as during the video-lecture (recorded with the same lecturer. e.g. during the previous academic year) students are exposed to the teaching style of the person with whom they later discuss: the teacher has the opportunity of presently *exactly* his view – and not some neutral textbook material – to the students before engaging in a conversation with them. In the end, this approach actually doubles the time students spend with the teacher (half of this time being virtualized by videos), allows for a more interactive learning style, and, according to the limited evidence collected by the four mentioned papers, ends up in a much more satisfactory experience for teacher and learners.

ECONOMICAL CONSIDERATIONS ABOUT VIDEO-LECTURES.

The previous sections demonstrated that there is a wide agreement on the advantages of adopting the video-lecture technology: the next step is to perform a cost/benefits analysis to decide if and how to employ it. Recently Rowe & Casalaina (2006) provided an estimate of the cost of recording one day of conference. According to them the cost is US$ 3000 – not including travel and living expenses – (and hence a hourly cost of at least US$ 300). We think this figure is exaggerated, and we'll try to provide our own estimate.

Getting the needed infrastructure is easy, as some commercial vendors have appeared – such as Tegrity[7] or AutoAuditorum[8]. On the other hand, open source or free software solutions are also available (e.g. ePresence[9], openEYA[10], LODE[11]). These solutions produce lectures that may or may not have some client platform limitation, and may or may not be fully automated. A detailed comparison of these systems is beyond the scope of the present chapter, but all of them are suitable to the task. Moreover, every institution can have different requirements (Ronchetti 2008), but at least one of the available free platform – with their different approaches – is likely to satisfy its basic needs.

For starting our estimate we note that – from what emerges from the literature review above – there is a rather universal consensus that this technology is pedagogically more sound and technologically more mature if one focuses on asynchronous fruition, and on recording live lectures – as opposed to creating synthetic lectures.

The cost analysis becomes then very simple. Costs concern the infrastructure (acquisition station and distribution station) plus running costs. The first one can be very low, as an acquisition station, with some of the free software solutions, costs less than 2000 Euro: it only requires a laptop, a digital camera and a radio-microphone (which is an essential element, as guaranteeing good audio quality is a crucial aspect). The distribution station – with a terabyte hard disk costing by now as little as 100 Euro – costs as much as a web server: 1000 Euros can be more than enough. Of course, if one foresees many simultaneous users, a streaming server may be a better choice – in which case another 1000 Euros might be needed for the software, even though there

[7] http://www.tegrity.com
[8] http://www.autoauditorium.com
[9] http://sourceforge.net/projects/epresence
[10] http://www.openeya.org
[11] http://latemar.science.unitn.it/LODE

exists open source solutions, such as VLC/VLS[12] or the Darwin Streaming server[13]. In all cases, the infrastructural cost should not be a major hurdle.

Running costs may be an issue. In the best case, a fully automated system will need a periodic supervision, as most services do – but this would require a very limited annual workload. Most systems can perform all the needed post-processing – up to the creation and maintenance of a dedicated web site in an automatic manner. The acquisition phase can be automatic, in which case the cost would be close to zero, or manual. In the second case the costs can be significant, but much lower than indicated by Rowe. A dedicated person could probably collect at least 1000 hours per year, and would not need a sophisticated technical specialization, as using most systems is rather straightforward. Our estimate of the acquisition cost is hence at most of the order of 25 Euro per hour, and it can be cut (at least) by two by asking a student to do simple operations during a lecture (the cognitive overhead can be minimal, so that the student can actually follow the lecture while performing those operations).

On the other hand, a non-fully automated solution brings some advantages that may be worth the cost. During acquisition there are essentially three actions that might need assistance: starting/pausing/stopping the recording, manipulating the camera (slewing, zooming, panning), synchronizing the slides. Among these, camera manipulation can sensibly increase the quality of video-lectures – especially if the teacher uses the blackboard. Not surprisingly, research has been addressing the problem of creating a fully automatic acquisition system using a wide variety of approaches, as we shall discuss in the next section.

RESEARCH ISSUES

Detecting slide transitions

The first system able to detect slide transitions we are aware of was proposed by Mukhopadhyay (1999). During the recording phase, a synchronization tone was added to the audio track. During post-processing all the available images (e.g. slides) were matched with the portion of the video showing the projection screen, and the best fit was chosen.

During the last decade several papers dealt with computer vision-based, statistical techniques to attack the same problem, trying to offer better solution. The typical pre-processing involves segmenting the lecturer from the background, localising the projection area, finding the slide transitions and matching the observed slides in the videos to their electronic versions. Reviews of these attempt can be found in the papers by Gigonzac et al (2007) and De Lucia et al. (2008). Generally such approach have the drawback that the camera should be fixed (to facilitate the localization of the screen), and most importantly that the projection screen should to be part of the video image (which in our view severely limits the usefulness of the video itself).

Chen & Heng (2003) used an automated speech recognizer and matched words from the speech with words from the slide to identify slide transitions.

Wang et al. (2007) analysed teacher gestures to predict and detect slide transitions

The problem of detecting slide transitions can be avoided by directly capturing the VGA output (as e.g. in Rowe & Casalaina, 2006). This has the advantage of allowing capturing not only slides, but also just anything that happens on the presenters screen (videos, demonstrations,

[12] http://www.videolan.org/vlc/streaming.html
[13] http://developer.apple.com/opensource/server/streaming/index.html

live tutorials etc.). On the other hand, it has the disadvantage of losing precious semantic information: timing of the slide transitions (and slide titles) offer an important navigation feature, and the content of the slides can be parsed to allow at least a limited possibility of searching the video through textual queries. Tanaka et al. (2004) detect slide change through an http proxy – but this approach only works for HTML-based presentations. Another possibility is having the speakers laptop signalling PowerPoint slide transitions (as e.g. in Baecker 2003), but this requires the speakers using Windows and accepting to mount a macro on their machines.

Virtual cameraman

Capturing visual details – such as following the teacher to better capture his/her expressions and body language, or zooming on the blackboard when needed is one of the advantages offered by a (costly) human operator. Some researchers have attempted to implement a "virtual cameraman". Such a virtual actor should be able to follow the speaker as s/he moves, and also to follow a "savvy" style. as performed by an art director, e.g. changing camera, focus on the semantically important details etc. Tracking a human object has been the focus of a research field that is too wide to review it here. Typical techniques are sensor-based tracking (in which the target is required to wear a suitable, e.g. magnetic, device), sound-based tracking (based on stereo perception or commercially available microphone array-based sound localization), video-based tracking (such as skin-colour based, motion-based, shape-based). A pioneering example of lecture room automation was Bellcore's AutoAuditorium (Bianchi 1998). Other systems are reported e.g. by Liu et al. (2001) and Wallik et al. (2004).

Wang et al. (2007) used gesture and posture recognition to simulate suitable camera motion and to perform video cutting effects.

An alternative approach is to use just one camera with panoramic video capturing, and then to extract the portion of image of interest (Sun 2005). A rather similar idea is implemented in the EYA system (Canessa et al. 2008). They used a wide-angle photo camera to record high resolution pictures every 10 seconds. The client is a little different from the usual video+slide format: it shows the video, a large thumbnail of the current picture, and when the user moves the mouse over the thumbnail, a high resolution subset of the image is shown where other systems put the slide. In this way, the user can focus on the detail he wants (be it the blackboard, the projected screen or other). The advantage is that fully traditional lectures (based on chalkboard or even on viewgraphs can be fully supported, even though also here the semantics carried by the slide change detection is lost.

Capturing the blackboard

Since in many lectures the blackboard still plays an important role, it is important to be able to effectively capture what happens there. We have already mentioned various ways to do so: using high-resolution pictures or having a virtual or real cameraman. Recent years have seen in some countries a relatively wide diffusion of interactive whiteboards: touch sensitive devices connected with a computer and a projector, that allow capturing the screen and manually interacting with the objects present on the screen (menus, windows, icons etc.). Also in this case, the research field is too wide to be discussed here, but we shall briefly mention a few cases in which such or similar devices have been used with the idea of recording live lectures for asynchronous use.

Early work (since 1996 till 2001) on the capture of several experiences in the classroom, including whiteboard traces, was performed in the framework of the Classroom2000/eClass project (see eClass 2001).
Hilt et al. (2001) report recording and replaying interactive media stream (video and whiteboard). Their demonstration was done with tele-cooperation in mind.
Joukov & Chiueh (2003) developed a digital desk to support and capture classroom activity.
The E-chalk project (Friedland & Rojas 2006) extracts handwriting from a traditional blackboard via image analysis. Various whiteboard vendors provide solution for recording audio+smartboard track. An evaluation of such tools from a pedagogical point of view would be useful.

Mobile devices

The delivery of files via RSS feeds has gained a considerable attention over the last three years. Many papers have dealt with production, distribution and assessment of educational podcasts. We do not review them here as they support only the audio channel. Other variations known as *vodcast* (video podcast) or *screencast* (audio + slide podcast) get closer to the video-lecture paradigm, and we find them more interesting. At least two examples in literature show the feasibility of the screencast approach: Hürst, Weite & Jung (2007) investigated using an Apple i-Pod to deliver lectures, and Ronchetti & Stevovic (2008) showed the feasibility of the screencast approach even on a mobile phone. Earlier Liu & Choudary (2004) used palmtop devices for testing content-aware scalable transmission of instructional videos over wireless networks. Their focus was on optimizing the delivery of videos by distinguishing between content and non-content video frames. Friedland & Rojas (2006) used mobile phones and i-pods to show the videos of their E-chalk videos.
The recent evolution of the technology, pioneered by the Apple i-Phone, allows for larger screens and already delivers you-tube videos. We expect to see some developments in the near future. The already mentioned iTunes U initiative is a first attempt in this direction, even though it remains to be understood how valuable such option may be from a usability and/or pedagogical point of view.

Annotation and other forms of interaction

Ideas about the annotation of generic web documents have been around for a while – see e.g. the Annotea project (Kahan et al. 2001). Also annotation of streaming video has been a subject of interest that spawned a research branch, and that was also applied to video-lectures (see e.g. Bargeron et al 1999, Correia and Cabral 2005). Anchoring discussions to specific resources rather than collecting them in bulletin boards or forums sound like a reasonable thing to do, since it would provide a context for the discussion. There have been proposals in this sense (Abowd et al 1998, Haga 2002, Lauer et al. 2005). Another idea, a precursor of the Web 2.0 fashion, was to have students taking lecture notes, attaching them to the video or to a slidecast and sharing them (e.g. Truong 1999, Kam et al 2005). Although all these approaches seem to be interesting and useful, and technology seems to be ripe, none of these seems to have really encountered a wide acceptance. Probably these approaches could be now revisited in the light of the social networking approach brought to us from the Web 2.0.

Another issue partly correlated is the fact that video on web is not a mere replacement of TV: being bidirectional, the Internet channel promises to make all this quite different from the traditional uni-directional TV approach. Already today some systems allow using bi-directional flows, with synchronous seminars (webinars or video webcasts) giving the remote listener the possibility to interact with the speaker, sometimes though the intermediation of a human proxy (e.g Baecker et al. 2003). Also on this front, which is very active in the research on computer supported cooperative work, not much has been harvested in the e-learning scenario.

Semantic indexing, multimodal access, and other topics

Indexing and allowing search in a distributed collection of video-lecture is a key element in an unstructured scenario like the one of self-study to support life-long learning. WLAP is a notable project (Bousdira et al. 2001) that used the notion of "learning object" (a IEEE standard[14]) to represent lectures, and that envisioned a distributed architecture where multiple archives could be integrated. Indexing terms is however not enough: what would be really needed is semantic modelling of the video-lectures, i.e. the possibility of understanding which topics are dealt with. A good semantic model would be helpful in aiding search and retrieval of the content - also in the form of query/answer (Q/A) - summarising lectures, splitting the lectures into smaller, homogeneous chunks, indexing them. Such tasks can get advantage from the information that can be extracted from the slides, notes and other associated material, and from transcripts of the speech. The advantages would be to be able to find the most interesting lectures for a given topic in a large collection, and to be able to quickly identify the most interesting portions of each candidate lecture.

A natural approach is to try to apply text techniques (such as content extraction, summarization etc.) from multimedia passing through the intermediate step of applying Automated Speech Recognition (ASR) techniques to the audio tracks. Wald (2005) suggested that text from ASR could be used to create captions (e.g. for deaf learners), to assist those who, for cognitive, physical or sensory reasons, find note taking difficult and to allow search.

Using an ASR to extract text from a video-lecture in not trivial, as a good quality of the sound is not always guaranteed. Moreover lecture language resembles conversational language (Glass et al. 2007), and often contains domain-specific, rare words, or even words in a different language. Being able to supplement the ASR with a suitable language model is therefore important. Choudhari et al. (2007) propose starting from textbooks to identify the most relevant terms that should compose the vocabulary.

In literature there are examples of segmenting lectures based on ASR text (Lin et al. 2003, Fujii et al 2006) or on text and slides (Repp et al. 2008a). Repp et al. (2008b) also built a Q/A system based on the semantic annotation generated for the lectures. Hürst et al. (2006) and Fogarolli et al. (2007) used a multimodal approach, indexing both ASR-extracted text and slides to allow searching libraries of lectures. Fogarolli et al. (2009) also annotated the lectures by using Wikipedia so as to automatically extract from the ASR text the most important topics that each lecture dealt with.

Supporting efficient and user-friendly navigation with interactive content overviews may need more research on the interface to be used. Mertens et al (2006) explored this facet.

Other topics include multimedia shrinking and summarization, video abstraction, lecture segmentation, gesture analysis. A review of these approaches applied to the videolecture domain is reported elsewhere (Ronchetti 2010).

[14] http://en.wikipedia.org/wiki/Learning_object

CONCLUSION

There is overwhelming evidence that video-lectures effectively support learning. Literature demonstrates that in most cases this is achieved in a better way by recording live lectures rather than creating *ad-hoc*, synthetic content. Also, it is now clear that video-lectures should not be considered as a mere replacement of classroom lectures. Besides the obvious support of students who cannot physically attend lectures, video lectures are in fact used by traditional students as an additional tool that allows them to quickly find information or to review their understanding. Moreover, video-lectures can be creatively used to change the didactic paradigm. Even more importantly, they start creating a body of knowledge that can be extremely useful for informal self-study, hence supporting the life-long learning perspective

Although some basic problems have been solved and tools for producing and distributing video-lectures are available in various forms, a standard has not yet emerged (and might be useful in future to allow interoperability of tools that will mine and extract information from the lectures).

Several research directions have fruitfully produced results in various areas such as improving the automatic capture of lectures and events connected with them, even though there is ample space for improving the results obtained so far.

The market and technology of portable small devices is still evolving, and for sure we shall in future see more attempts to bring (video)lectures on them: much has still to be done in terms of usability for these devices, and also to ascertain the actual usefulness of bringing these resources on small devices.

The marriage of Web 2.0 and video-based education is still not complete, and probably we shall see future development in connecting virtual (learning) communities and video-lectures. Social annotations may be one of the ingredients.

Finally, we believe that better support to search and intelligent navigation of multimedia content, probably based to more advanced semantic extraction, will be the main front on which further research will improve usability and usefulness of the video-lectures. This last fact is especially necessary and useful for providing a better support to life-long learners.

REFERENCES

Abelson (2008) The creation of OpenCourseWare at MIT. *Journal of Science Education and Technology* Vol. 17, No. 2, pp. 164-174

Abowd G.D., Atkeson C.G., Brotherton J., Enqvist T., Gulley P. & LeMon, J. (1998) Investigating the capture, integration and access problem of ubiquitous computing in an educational setting, *Proc. of the SIGCHI Conf. on Human factors in computing systems*, p.440-447

Baecker, R.M., Moore, G., Keating, D., & Zijdemans, A., (2003) Reinventing the Lecture: Webcasting made Interactive, In *Proc. of HCI Int. Conf. 2003*, Crete (Greece), June 22-27 2003.

Bargeron, Gupta A., Grudin J, & Sanocki E. (1999) Annotations for streaming video on the web. *CHI '99 Extended Abstracts on Human factors in computing systems* pp. 278-279

Bennett E. & Maniar N.. (2008) Are videoed lectures an effective teaching tool?. Retrieved May 15, 2009 from http://stream.port.ac.uk/papers

Berner E.S. & Adams B. (2004) Added value of video compared to audio lectures for distance learning. *International Journal of Medical Informatics* (73), pp.189-193

Bianchi, M. (1998) AutoAuditorium: a fully automatic, multi-camera system to televise auditorium presentations, *Proc. of Joint DARPA/NIST Smart Spaces Technology Workshop*, July 1998.

Bousdira, N., Myers, E., Neal, H.A., Severance C., Storr M. & Vitaglione G. (2001) WLAP: The Web Lecture Archive Project. CERN-OPEN 2001-066

Brown B.W. & Liedholm C.E.L. (2004) Student Preferences in Using Online Learning Resources. *Social Science Computer Review* vol. 22 (4) pp. 479-492

Canessa E., Fonda C. & Zennaro M, (2008) Academic Webcasting using the Automated EyA Recording System *INTED-International Technology, Education and Development*, Valencia/Spain, March 2008.

Chen H.-Y., Chen G.-Y. & Hong J.S. (1999) Design of a Web-based synchronized multimedia lecture system for distance education. *IEEE Int. Conf. on Multimedia Computing and Systems*. vol. 2 pp. 887 – 891

Chen Y. & Heng W.J. (2003) Automatic synchronization of speech transcript and slides in presentation. *Proc. of the Int. Symp. on Circuits and Systems, ISCAS '03*. vol. 2 pp. II-568-571

Chen, S.Y., Ghinea G. & Macredie R.D: (2006) A cognitive approach to user perception of multimedia quality: An empirical investigation *International Journal of Human-Computer Studies* Vol. 64, pp. 1200-1213

Chiu, C., Lee, G., & Yang, J. (2006). A comparative study of post-class lecture video viewing. *Proc. of the 5th IASTED Int. Conf. on Web-Based Education* ACTA Press. pp. 126-130

Choudary C., Liu T. & Huang C. (2007) Semantic Retrieval of Instructional Videos. Ninth IEEE International Symposium on Multimedia Workshops (2007)

Chung Q. B. (2005) Sage on the Stage in the Digital Age: The Role of Online Lecture in Distance Learning, *The Electronic Journal of e-Learning* Volume 3 Issue 1, pp 1-14

Correia N. & Cabral D. (2005) VideoStore: A system to store, annotate and share video based content. Recent Research Developments in Learning Technologies. FORMATEX pp. 1299-1303

Day J.& Foley J.. (2006) Evaluating web lectures: a case study from HCI. *CHI 2006 – Experience Report – User Centered Design for Learning and Education* pp.195-200

De Lucia, A. Francese, R., Passero, I. & Tortora, G. (2008) Migrating legacy video lectures to multimedia learning objects. *Software: Practice and Experience* Vol.38, pp.1499–1530

Demetriadis, S., & Pombortsis, A. (2007). e-Lectures for Flexible Learning: a Study on their Learning Efficiency. *Educational Technology & Society*, 10 (2), 147-157.

Dutton, J. Dutton, M. & Perry, J. (2001) Do online students perform as well as lecture students? *Journal of Engineering Education*, Vol. 90, no. 1, pp. 131–142

eClass (2001) for various references on the eClass project, see http://www.cc.gatech.edu/fce/eclass/pubs/index.html (last checked June 22, 2009)

Fey A. (2002) Audio vs. Video: Hilft Sehen beim Lernen? Vergleich zwischen einer audio-visuellen und auditiven virtuellen Vorlesungen. *Lernforschung*, 30. Jhg (4):331–338 (in German)

Foertsch J.. Moses G., Strickwerda J. & Litzkov M (2002) Reversing the lecture/homework paradigm using eTEACH web-based streaming video software. *Journal of Engineering Education* vol. 91 (3) pp. 267-274

Fong & Hui. (2001) Low-bandwidth Internet streaming of multimedia lectures. *Engineering Science and Education Journal* vol. 10 (6) pp. 212 - 218

Fogarolli A., Riccardi G., Ronchetti M. (2007) "Searching information in a collection of video-lectures". *Proc. of World Conf. on Educational Multimedia, Hypermedia and Telecommunications ED-MEDIA 2007*, p. 1450-1459.

Fogarolli A. & Ronchetti M. (2009) Extracting Semantics from Multimedia Content. *Special Issue of Scalable Computing: Practice and Experience*, pp. 1895-1767,

Friedland, G. & Rojas, R. (2006). Human-centered Webcasting of Interactive-Whiteboard Proc. of the Eight IEEE Int. Symposium on Multimedia,

Friedland, G. & Rojas, R. (2008). Anthropocentric video segmentation for lecture webcasts. *J. Image Video Process*. 2008, pp. 1-10.

Friedland, G. Hürst, W. Knipping, L. & Mühlhäser M. (2009) Can we escape the trough of disillusionment? A Perspective on E-Learning Technology Research from the ACM Workshop on Educational Multimedia and Multimedia Education, *ACM eLearn magazine*, February2009.

Fritze Y. & Nordkvelle Y. (2003) Comparing Lectures – Effects of the Technological Context of the Studio. In: *Education and Information Technologies*, Kluwer Academic Publisher, Netherland,, pp. 327-343

Fujii, A, Itou, K, & Ishikawa, T. (2007) LODEM: A system for on-demand video lectures. *Speech Communication* 48, pp.516–531

Gigonzac, G., Pitie, F. & Kokaram.,A. (2007) Electronic slide matching and enhancement of a lecture video. *4th European Conf. on Visual Media Production, IETCVMP 2007*, pp. 1 – 7

Glass J., Hazen T.J., Cyphers S. & Malioutov I. (2007) Recent progress in the MIT spoken lecture processing project. Proc. Interspeech 2007, pp.2553-2556

Glowalla U. (2004) Utility and Usability von E-Learning am Beispiel von Lecture-on-demand Anwendungen. In *Entwerfen und Gestalten*, 2004 (in German)

Haga H., (2002) Combining video and bulletin board systems in distance education systems. *The Internet and Higher Education*, 5, pp.119-129

Hayes, M.H. (1998) Some approaches to Internet distance learning with streaming media, In *Second IEEE Workshop on Multimedia Signal Processing*, Redondo Beach, CA, USA

Hermann, C. Welte, M., Latocha, J., Wolk, C. & Huerst W. (2007) Eine logfilebasierte Evaluation des Einsatzes von Vorlesungsaufzeichnungen, *GI-Edition Lecture Notes in Informatics, DeLFI*, pp. 141-150 (in German)

Hilt, V. & Kuhmünch, C. (1999). New Tools for Synchronous and Asynchronous Teaching and Learning in the Internet. In B. Collis & R. Oliver (Eds.), *Proc. of World Conf. on Educational Multimedia, Hypermedia and Telecommunications ED-MEDIA 1999* pp. 975-980

Hilt V., Mauve M., Kuhmünch C. & Effelsberg W. (2001) Interactive media on demand: generic recording and replay of interactive media streams *Proc. of the ninth ACM Int. Conf. on Multimedia* pp. 593 - 594

Hürst W. and Deutschmann N.. (2006) Searching in recorded lectures. *Proc. of World Conf. on Educational Multimedia, Hypermedia and Telecommunications ED-MEDIA 2006* pp. 2859-2866

Hürst W., Weite M. & Jung S. (2007) An evaluation of the mobile usage of e-lecture podcasts. *Proc. of the 4 Int. Conf. on Mobile Technology, Applications and Systems* pp. 16-23

Isiaka B. (2007) Effectiveness of video as an instructional medium in teaching rural children agricultural and environmental sciences. *International Journal of Education and Development using Information and Communication Technology* vol. 3 pp. 105-114

Joukov J. & Chiueh T. (2003). Lectern II: a multimedia lecture capturing and editing system. *Proc. of Int. Conf. on Multimedia and Expo, 2003. ICME '03*. vol. 2 pp. II - 681-4 vol.2

Kahan, J., Koivunen, M-R., Prud'Hommeaux, E. & Swick R.R. (2001), "Annotea: An Open RDF Infrastructure for Shared Web Annotations", *Proc. of the WWW10 International Conference*, HK

Kam M., Wang J., Iles A., Tse E., Chiu J., Glaser D., Tarshish O. & Canny J. (2005) Livenotes: a system for cooperative and augmented note-taking in lectures. *CHI '05: Proceedings of the SIGCHI Conf. on Human factors in computing systems* pp.531-540

Krüger & Plaza (2005). Vortragsaufzeichnungen–Ein Querschnitt über die pädagogischen Forschungsergebnisse. Workshop Proceedings DeLFI (in German)

Lage, M.J., Platt, G.J. & Treglia M. (2000) Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment *The Journal of Economic Education*, Vol. 31, No. 1, pp. 30-43

Lauer T., Müller R. & Trahasch S. (2004) Learning with lecture recordings: key issues for end-users. *Proc. of the IEEE Int. Conf. on Advanced Learning Technologies 2004.* pp. 741 – 743

Lauer T., Trahasch S. & Zupancic B. (2005) Anchored Discussions of Multimedia Lecture Recordings. *Proceedings 35th Annual Conference - Frontiers in Education - FIE'05* pp. 12-17

Lin M., Nunamaker, J., Chau M,. & ChenH. (2003) Segmentation of lecture videos based on text: a method combining multiple linguistic features. *Proc. of the 37th Annual Hawaii Int. Conf. on System Sciences*, 2004. P (2003) pp. 1-9

Liu Q, Rui Y. Gupta A. & Cadiz JJ (2001) Automating camera management for lecture room environments. *CHI '01: Proceedings of the SIGCHI Conf. on Human factors in computing systems*, *CHI Letters* 3 (1) pp. 442-449

Liu T. & Choudary C. (2004) Content-aware streaming of lecture videos over wireless networks. *Proc. of IEEE Sixth Int. Symposium on Multimedia Software Engineering*, 2004. pp. 458 – 465

Maness J.M. (2006) An evaluation of library instruction delivered to engineering students using streaming video. *Issues in Science and Technology Librarianship*

McCrohon, M., Lo,K. Dang & J., Johnston, C. (2001) Video Streaming of Lectures Via the Internet An Experience. *ASCILITE 2001 - The 18th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education*, Melbourne pp. 397-405

Mertens R., Ketterl M., Vornberger O. (2006) Interactive Content Overviews for Lecture Recordings. *Eighth IEEE International Symposium on Multimedia* pp. 933-937

Moreno, R. & Mayer, R.-E. (2000) A Learner-Centred Approach to Multimedia Explanations: Deriving Instructional Design Principles from Cognitive Theory, Interactive Multimedia *Electronic Journal of Computer-Enhanced Learning* (2)

Mukhopadhyay S. & Smith B. (1999) Passive capture and structuring of lectures. *MULTIMEDIA '99: Proceedings of the seventh ACM Int. Conf. on Multimedia* Part 1 pp.477-477

Nielsen (2009) A2/M2 Three Screen Report – 1st Quarter 2009. Retrieved June 9, 2009, from http://it.nielsen.com/site/documents/A2M2_3Screens_1Q09_FINAL.pdf

Reisslein J. Seeling P. & Reisslein M. (2005) Video in distance education: ITFS vs. web-streaming: Evaluation of student attitudes. *The Internet and Higher Education* (8) pp.25-44

Repp S. & Meinel C. (2008a) Segmentation of Lecture Videos Based on Spontaneous Speech Recognition. Tenth IEEE Int. Symp. on Multimedia ISM 2008 pp. 692 – 697

Repp S., Linckels S., & Meinel C. (2008b) Question answering from lecture videos based on an automatic semantic annotation. *Proc. of the 13th annual conf. on Innovation and technology in computer science education ITiCSE '08* pp.17-21

Ronchetti M., (2003a) Has the time come for using video-based lectures over the Internet? A Test-case report *CATE - Web Based Education Conference 2003*, Rhodes (Greece)

Ronchetti, M. (2003b). Using the Web for diffusing multimedia lectures: a case study. *Proc. of World Conf. on Educational Multimedia, Hypermedia and Telecommunications ED-MEDIA 2003* pp. 337-340

Ronchetti M. (2008) Requirements for videolectures: which system is the best for you? *World Conf. on Educational Multimedia, Hypermedia and Telecommunications EDMEDIA 2008*, pp. 2192-2199 Available on-line at http://latemar.science.unitn.it/segue_userFiles/LODE/RequirementsForVideolectures.pdf

Ronchetti M. & Stevovic J. (2008) Extending the podcasting approach: lectures on the phone, *Proc. of World Conf. on Educational Multimedia, Hypermedia and Telecommunications EDMEDIA 2008*, Chesapeake:AACE, p. 5690-5696

Ronchetti M. (2009) Using video lectures to make teaching more interactive, *Proc. of the Int. Conf. on Interactive Computer Aided Learning ICL-2009*

Ronchetti M. (2010) "Perspectives of the application of video streaming to education", accepted for publication in *Streaming Media Architectures, Techniques, and Applications: Recent Advances* Ce Zhu, Yue-nan Li, Xia-mu Niu eds,

Rowe & V.Casalaina. Capturing Conference Presentations (2006) *Multimedia, IEEE* vol. 13 (4) pp. 76-84

Soong S.K.A,, Chan L.K. & Cheers C. (2006) Impact of video recorded lectures among students. *Proceedings of the 23rd annual ascilite conf.: Who's learning? Whose technology?* Pp.789-792

Sun X., Foote L, Kimber D. & Manjunath B.S., (2005) Region of interest extraction and virtual camera control based on panoramic video capturing. *IEEE Transactions on Multimedia*, vol. 7 (5) pp. 981 - 990

Tanaka, Y., Itamiya, T., Hagino, T. & Chiyokura H. (2003) HTTP-proxy-assisted automatic video indexing for e-learning. *International Symposium on Applications and the Internet, SAINT 2004* Workshops pp. 502 – 507

Tobagi. F. (1995) "Distance learning with digital video". *Multimedia, IEEE* vol. 2 (1) pp. 90 – 93

Truong K & Abowd G. (1999) StuPad: integrating student notes with class lectures. *CHI '99: CHI '99 extended abstracts on Human factors in computing systems* pp. 208 – 209

Wald. M. 'SpeechText': Enhancing Learning and Teaching by Using Automatic Speech Recognition to Create Accessible, Synchronized Multimedia (2005) *World Conf. on Educational Multimedia, Hypermedia and Telecommunications EDMEDIA-2005*

Wallick M.N., Rui Y. & He L. (2004) A portable solution for automatic lecture room camera management. *IEEE Int. Conf. on Multimedia and Expo, 2004. ICME '04.* vol. 2 pp. 987 - 990

Wang F., Ngo C-W., Pong T-C. (2007) Lecture Video Enhancement and Editing by Integrating Posture, Gesture, and Text. *IEEE Transactions on Multimedia* vol. 9 (2) pp. 397 – 409

Zhang D., Zhao J.L., Zhu L. & Nunamaker, J.F.Jr. (2004) Can e-learning replace classroom learning?. *Communications of the ACM* vol. 47 (5) pp-75-79

Zhang D., Zhu L., Briggs L.O. & Nunamaker, J.F.Jr. (2006) Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management* (43) pp. 15-27

Zupancic, B. & Horz, H. 2002. Lecture recording and its use in a traditional university course. *Proc. of the 7th Annual Conf. on Innovation and Technology in Computer Science Education* ITiCSE '02. ACM, New York, NY, pp.24-28

KEY TERM DEFINITIONS

Automated speech recognition: a computational technique that converts spoken words to text.

Continuous education: an all-encompassing term within a broad spectrum of post-secondary learning activities and programs.

Digital library: a library in which collections are stored in digital formats (as opposed to print, microform, or other media) and accessible by computers

Distance education: a field of education that focuses on the pedagogy and andragogy, technology, and instructional systems design that aim to deliver education to students who are not physically "on site".

e-learning: a term that encompasses all forms of Technology-Enhanced Learning, i.e. support of any pedagogical approach that utilizes technology.

Gesture: a form of non-verbal communication in which visible bodily actions communicate conventionalized particular messages, either in place of speech or together and in parallel with spoken words.

Multimodality, multimodal (interaction): a form of man-machine interaction using multiple modes of input/output.

Semantic web: an evolving development of the World Wide Web in which the meaning (semantics) of information and services on the web is defined, making it possible for the web to "understand" and satisfy the requests of people and machines to use the web content.

Text segmentation: the identification of lexical units in writing systems.

Video abstracting: a research area that deals with gaining perspectives of a video document without watching it entirely

Video conference: a set of interactive telecommunication technologies which allow two or more locations to interact via two-way video and audio transmissions simultaneously.

Video-lecture: lecture recorded in a video and delivered through a variety of media.

Video segmentation: the identification of boundaries among regions that differ for content or aspect in a video.