Automated Lecture Services

Cha Zhang and Yong Rui

Microsoft Research One Microsoft Way, Redmond, WA 98052, USA
chazhang@microsoft.com, yongrui@microsoft.com

Summary. Increasingly popular, lectures are given before a live audience, while simultaneously being viewed remotely and recorded for subsequent on-demand viewing over the Internet. Traditionally, it is very expensive to offer such services due to the high labor cost involved. In this chapter, we survey existing approaches for providing automated lecture services. In particular, we examine two major challenges in providing such services, namely, how to capture, analyze and render the lecture content automatically, and how to provide live/on-demand lecture viewing/browsing experience with an automated end-to-end system. The chapter is concluded by a list of future research directions, hoping to inspire even more work on this interesting and highly useful topic.

14.1 Introduction

Live/on-demand Internet broadcasting of lectures in the workplace, at conferences and in educational settings has attracted a lot of interest recently. For instance, many corporations make seminars and training sessions available online for employees who cannot attend a live presentation [1, 2]. Conferences such as SIGGRAPH and NOSSDAV recorded their presentations and made them available for on-demand replay. The growth of e-learning systems in universities is even more significant. According to the annual online education report from the Sloan Consortium in 2006 [3], in the United States, more than 96% of the very largest institutions (more than 15,000 total enrollments) have some online offerings. Nearly 3.2 million students were taking at least one online course during the fall 2005 term, a substantial increase over the 2.3 million reported in the previous year.

Although online viewing provides a convenient way for people to watch lectures at a more convenient time and location, the cost of providing such services can be huge. Typically, there are three types of costs involved in such a system:

1. Installation cost, which includes hardware (e.g., computer servers, microphones, cameras) and software cost.
2. *Recurring infrastructure cost*, which includes bandwidth charges, equipment wear and tear cost, etc.

3. *Recurring labor cost*, which includes pre-lecture activities (e.g., setting up the equipment), during-lecture activities (e.g., controlling cameras to track presenters and audience and switching between cameras) and post-lecture activities (e.g., posting the lecture to a web site).

Installation cost is a one-time investment. Improvements in technology have led to a rapid decline in this expense over the past decade. The recurring infrastructure cost is not significant either, considering that many lectures are broadcasted within an intranet environment. The labor cost, however, occurs every lecture and has not decreased over time. As a result, there have been a lot of research efforts recently to build automated lecture services to reduce the labor cost. To name a few well-known systems, the *AutoAuditorium* system [4, 5] by Michael Bianchi was probably the first commercialized system (and so far still one of the best systems) that can capture a lecture fully automatically with a set of intelligent microphones and cameras. The *Classroom 2000* project [6] at Georgia Institute of Technology built a classroom that supports teachers and students in traditional university lecture environments, with an electronic whiteboard that records the lecturer’s notes on pre-loaded presentation materials. The *Cornell Lecture Browser* [7] was an early attempt to automate the post-production of lecture contents by synchronizing audio and video captured during the lecture with slides acquired from the speaker. The *Berkeley Internet Broadcasting System (BIBS)* [8] has been adopted as an integral part of the university’s course delivery infrastructure, and webcasts approximately 15 classes each semester. The system was designed to require as few people as possible to operate. The Microsoft *iCam* system [9–11] conducted study on how human directors select the best view out of multiple choices and applied that for automated view switching during content production. It is an automated end-to-end system that requires minimum human interaction and supports both live/on-demand lecture browsing. The system has been used on a daily basis for over 5 years. The *ePresence* system from University of Toronto [12] is an open source system for lecture capture, archive and broadcasting. It is scalable, interactive and able to support presenters and engage remote audiences with rich media. These systems demonstrate that with advanced audio/image/video processing techniques and careful engineering, the labor cost of lecture services can be reduced dramatically, which makes it possible to run such services on regular basis.

In this chapter, we will discuss various techniques for providing automated lecture services. There are two major challenges involved, namely, to capture, analyze and render the lecture content automatically; and to provide live/on-demand lecture viewing/browsing experience with an automated end-to-end system. These two challenges are not isolated, since a successful system has to address both to avoid high labor costs. Nevertheless, in order to give readers a unified view of different techniques across different systems, we will