Adaptive Synchronization of Semantically Compressed Instructional Videos for Collaborative Distance Learning

Dan Phung, Columbia University, USA
Giuseppe Valetto, IBM T.J. Watson Research Center, USA
Gail E. Kaiser, Columbia University, USA
Tiecheng Liu, University of South Carolina, USA
John R. Kender, Columbia University, USA

ABSTRACT

The increasing popularity of online courses has highlighted the need for collaborative learning tools for student groups. In this article, we present an e-Learning architecture and adaptation model called AI2TV (Adaptive Interactive Internet Team Video), which allows groups of students to collaboratively view instructional videos in synchrony. Video player actions, like play, pause and stop, can be initiated by any group member and the results of those actions are synchronized with all the other students. These features allow students to review a lecture video in tandem, facilitating the learning process. AI2TV upholds the invariant that each student will receive semantically equivalent content at all times. Experimental trials show that AI2TV successfully synchronizes instructional videos for distributed students while concurrently optimizing the video quality, even under conditions of fluctuating bandwidth, by adaptively adjusting the quality level for each student while still maintaining the semantic content of the instructional videos.

Keywords: distance learning; distributed systems; collaborative learning; computer-assisted education; multimedia systems; video display terminal;

INTRODUCTION

Distance learning programs such as the Columbia Video Network (www.cvn.columbia.edu) have evolved from Fedexing lecture video tapes to their off-campus students to streaming videos over the Web. The lectures might be delivered “live,” but are more frequently post-processed and packaged for students to watch (and re-watch) at their convenience. This introduces the possibility of forming “study groups” among students who can view the lecture videos together and pause, rewind, or fast-forward the video to discussion points, thus approximating the pedagogically valuable discussions that occur during on-campus
lectures. However, conventional Internet video technology does not yet support collaborative video viewing by multiple geographically dispersed users. It is particularly challenging to support What I See Is What You See (WISIWYS) when some users are relatively disadvantaged with respect to bandwidth (e.g., dial-up modems) and local resources (e.g., old graphics cards, small disks).

AFTV is an e-learning architecture supporting virtual student groups. To that end, we have developed the technology for “semantically adapting” standard MPEG videos into sequences of still images. This technology automatically selects the most semantically meaningful frames to show for each time epoch and can generate different sequences of JPEG images for a range of different compression (bandwidth) levels. It was designed with typical lecture videos in mind: for instance, it recognizes that it is more important to see the blackboard content after the instructor has finished writing, rather than showing the instructor’s back as he/she writes it on the board.

Other technical challenges are synchronizing and adapting the downloading and display of the image sequences among the distributed students, including support for shared video player actions. We have developed an approach that achieves this using three mechanisms working in tandem: First, the software clocks of the video clients for each student are synchronized using NTP (Network Time Protocol), hence they use the same time reference with respect to the image sequences, where each image is associated with its start and end times relative to the beginning of the sequence. Second, the video clients communicate with each other over a distributed publish-subscribe event bus, which propagates video actions taken by any user to the rest of the group, as well as other events occurring on the video clients. Finally, since we are particularly concerned about disenfranchised user communities that have relatively low bandwidth, the final contribution of AFTV concerns enabling the optimization of the video quality according to the bandwidth constraints of each user, while enforcing group synchronization through a distributed feedback control loop that dynamically adapts each video client.

This paper presents the architecture and dynamic adaptation model of AFTV, describes how it tackles the challenges of quality optimization and synchronization in collaborative video viewing, and provides an evaluation of the effectiveness of our approach, with empirical results obtained using real lecture videos from Columbia’s Video Network.

**MOTIVATION AND BACKGROUND**

Correspondence courses have been available for over a century, for example, the University of Wyoming began offering extension courses in 1892 (Miller, Ditzler, & Lamb, 2003) correspondence courses have traditionally been designed for individual students with a self-motivated learning style, studying primarily from text materials.

A National Science Foundation report (NSF, 2002) discusses how technology, from radio to television; to audio and video cassettes; and to audio and video conferencing has affected distance education. The report states that the recent use of Internet technologies, especially the Web, has “allowed both synchronous and asynchronous communication among students and between faculty and students” (p. 1) and has
Faculty Perceptions and Participation in Distance Education