ClassX - An Open Source Interactive Lecture Streaming System

Sherif Halawa, Derek Pang, Ngai-Man Cheung, and Bernd Girod
Stanford University
Packard Bldg, Serra Mall
Stanford, CA 94305, USA
{halawa, dcypang, ncheung, bgirod}@stanford.edu

ABSTRACT

The ClassX open source project is a free experimental interactive video streaming platform designed for educators, researchers and software developers. With minimal infrastructure set-up, ClassX offers educational communities a cost-effective solution for online lecture delivery. Our goal is to encourage contributions from other researchers, developers and educators in building an open, cost-effective and state-of-the-art online education video viewing system for the general public.

Categories and Subject Descriptors

H.5.1 [**Multimedia Information Systems**]: Video

General Terms

Algorithm, Design

Keywords

Education, interactive region-of-interest video streaming, automatic lecturer tracking, electronic slide matching

1. INTRODUCTION

The advent of video streaming technology in the past decade has reshaped our educational systems significantly. More educators are making their lectures available online for students in their institutions or around the world to view. Nevertheless, the current status of lecture capture remains an obstacle that hinders many educators from recording and publishing their lectures. This is due to the special equipment and considerations that must be taken to ensure that important scene regions such as boards or projector screens are captured at sufficient quality.

Conventional lecture capture systems often rely on expensive recording and production equipment and require a camera operator. Some systems use multiple cameras to capture different views of the classroom. A human operator then selects one of the feeds at a time. Other systems use a single HD camera to capture the whole scene and manual or automatic tracking to crop out what is thought to be the student's region-of-interest (RoI). Prior work based on such systems includes [1, 8, 11].

The conventional video production model encounters a number of challenges when applied to lecture capturing. Professional recording and production equipment might be unaffordable. Camera operators tend to be unskilled, often students working a temporary part-time job. Nevertheless, wages for an extra person in the classroom might be prohibitive. Moreover, the camera operator might not have a deep understanding of the material presented and thus frame the shot incorrectly. A review of conventionally produced lecture videos reveals that this is a rather common problem. For example, the video might show a close-up of the instructor, when the attentive student wants to view the material on the blackboard that the instructor is referring to. Frequently, the camera operator lags behind a shift in focus by several seconds and might even appear "out of phase" if the focus shifts back and forth rapidly, say, between two wallboards. Note that it is not possible to correct for this in the conventional postproduction stage. Contents that is not recorded is lost.

ClassX efficiently addresses the abovementioned problems. A high-quality video is captured using a single consumer-grade HD camcorder. This makes it easy and cheap to record in any classroom. The camcorder is statically mounted to capture the front of the classroom in its entirety. This eliminates the need for a camera operator and the possibility of losing important scene content. The video is afterwards processed and transcoded to a special format which allows the student to choose his/her RoI. This RoI may be selected interactively in "interactive mode", or may be estimated by the system in the "automatic tracking mode". ClassX also synchronizes the video automatically with electronic slides, thus providing powerful video annotation with no human intervention.

ClassX was the idea and initiative of a group of graduate students at Stanford almost two years ago. The software has evolved through the contributions of numerous students since then, and we have learned tremendously from the practical deployment of the system from the very start. In March 2011, ClassX was released as an open source project under the GNU GPL in order to make it free for educators to use, and in order to attract researchers' and developers' inter-

MM'11, November 28–December 1, 2011, Scottsdale, Arizona, USA.

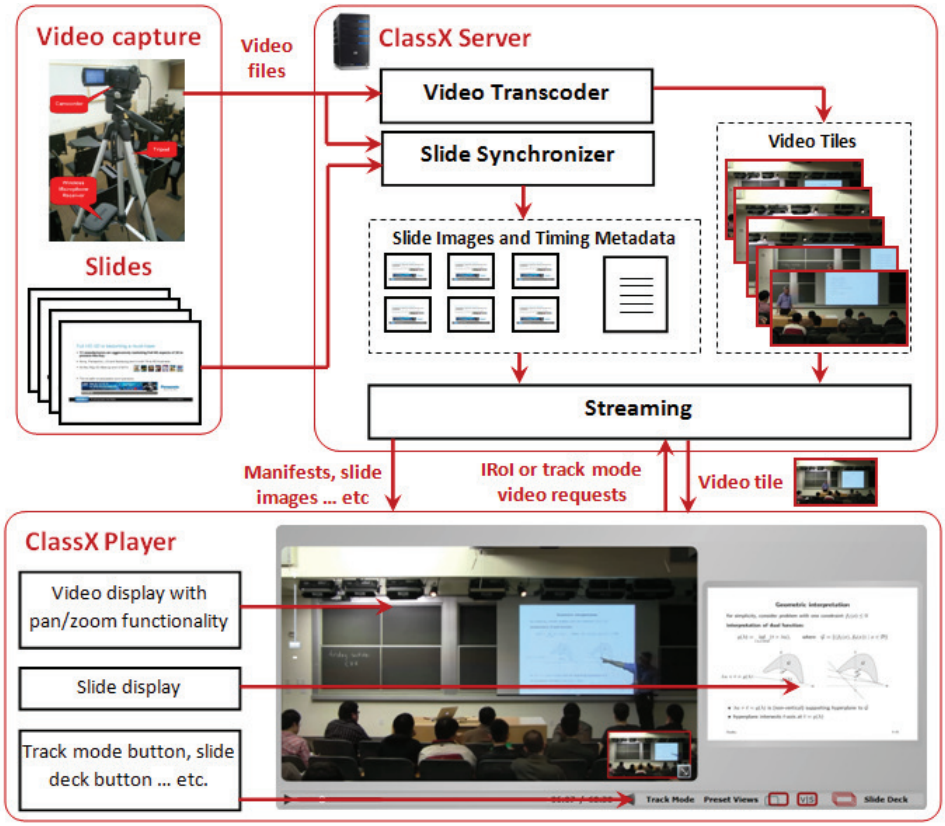

Figure 1: Overview of the ClassX System

est, who may then contribute to the project by adding new features that are useful to educators and students.

The rest of this paper is organized as follows. Section 2 provides a system overview of ClassX. Section 3 provides pointers to the locations of the resources needed to run or test ClassX. Section 4 provides usage statistics based on the deployment of ClassX as a lecture streaming system at Stanford University and as an open source project.

2. SYSTEM OVERVIEW

The ClassX system [4] is illustrated in Figure 1. It consists of the following components:

- A video transcoder
- A slide synchronizer
- Players to play the content to the user
- A web system that integrates the previous components and provides easy publishing and access interfaces

We will discuss each component in more details.

2.1 Video Transcoder

ClassX uses interactive region-of-interest (IRoI) to serve HD quality video at lower bit rates [5–7]. This is achieved as follows. When an input video is supplied to the ClassX video transcoder, it resamples the video, producing one or more representations of the video at different resolution levels. One or more tiles are then generated per resolution level, each of which is a differently cropped version of the video at that resolution level. The tiling scheme is designed such that any region of interest can be covered using one of the generated tiles. Figure 2 illustrates a scheme where an input video of resolution 1920×1080 pixels is transcoded into five tiles in two resolution levels. When the user zooms out, the ClassX server sends a tile of lower resolution. When the user zooms in on a specific RoI, the server sends a tile corresponding to that RoI generated from a higher resolution version of the video. This technique guarantees high video quality and saves 40% to 60% of the bandwidth compared to streaming the whole HD video.

In addition to the tiles, the transcoder also generates the tracking tile which automatically follows the lecturer. When the user activates tracking mode, there is no need to manually pan/zoom in the video.

2.2 Slide Synchronizer

When a slide deck is supplied, ClassX runs an automatic slide synchronizer that applies image matching techniques to determine when each slide is shown in the video and generate metadata that enables the ClassX player to automatically display the correct slide alongside the video. The synchronizer starts by locating keyframes (frames where slide transitions have occured). These keyframes are then input to a pairwise image matching stage using local features as described in [10]. The results of the image matching are used to generate a manifest listing the slides and their showtimes

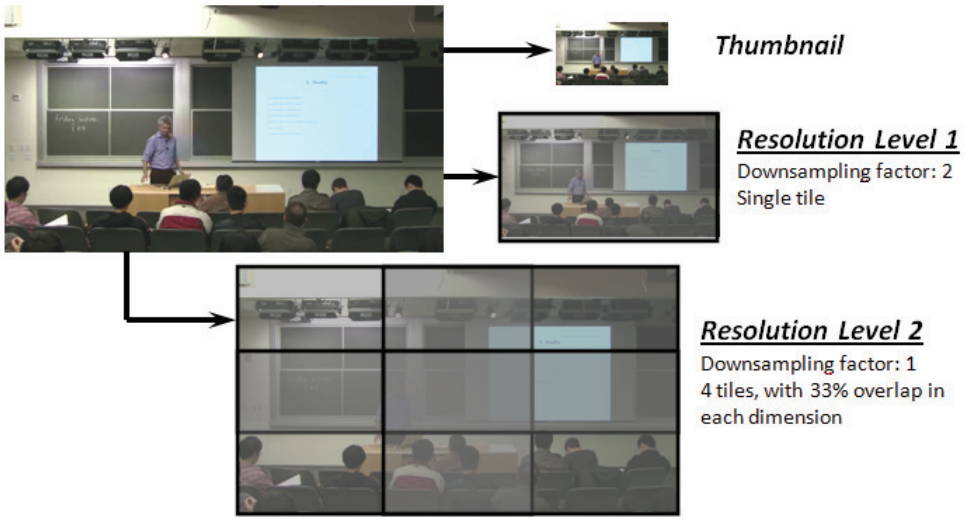

Figure 2: Tiling scheme with two resolution levels and five tiles

in the video. The ClassX player loads this manifest and uses it to show the correct slide in the slide display.

2.3 ClassX Players

The ClassX player, depicted in Figure 1, is the application which presents ClassX content to the user. It contains a video display through which the user views and interacts with the video. There is a track mode button that lets the user switch automatic lecturer tracking mode on or off.

When slides are synchronized with the video, the player also shows a slide display area and a button which can open or close a slide deck bar. This bar lists thumbnails of all available slides and lets the user seek to the point in the video when a selected slide is first shown.

Two versions of the ClassX player, running under the Microsoft Silverlight and Adobe Flash runtimes, have been released. These are currently optimized for desktop application only. A third version, built for Android, may run on any Android device with a multi-touch screen, including mobile phones and tablets [9]. This version will be released soon.

2.4 Web System

ClassX is provided as an integrated web system that is easy to install on any Unix-like server. The web system provides a publisher interface for easy publishing of ClassX presentations, as well as a user interface that the students may use to browse and view the published sessions.

3. INSTALLATION AND TEST

The ClassX open source page [2] provides links to useful resources for educators who want to use ClassX and developers who want to experiment with or contribute to the project. The web page at [3] contains a zipped file that contains all required packages and a list of instructions on how to build and test ClassX.

4. USAGE STATISTICS

ClassX was first deployed in Fall 2009 at Stanford university, and has been used since then to serve over 35 courses and events. The recordings covered a wide gamut of presentation styles, ranging from slide presentations to "chalk-and-talk" lectures to pre-recordings without an audience.. Different event types were also covered, ranging from group talks in small rooms to public events held in large auditoriums.

Users have been providing ratings for some of the aspects of ClassX, including video quality, streaming quality, and ease of use. The results are shown in figures 3, 4, and 5.

As of end of August 2011, the number of distinct IP addresses that have viewed one or more sessions on ClassX was approximately 6,390 and the total number of session views was approximately 14,830. The open source project page has been visited 447 times. The code has been downloaded approximately 108 times.

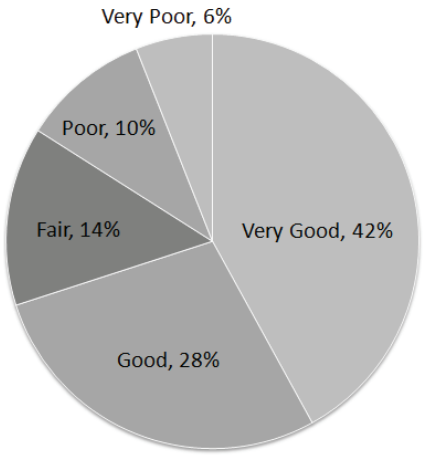

Figure 3: User feedback for video quality

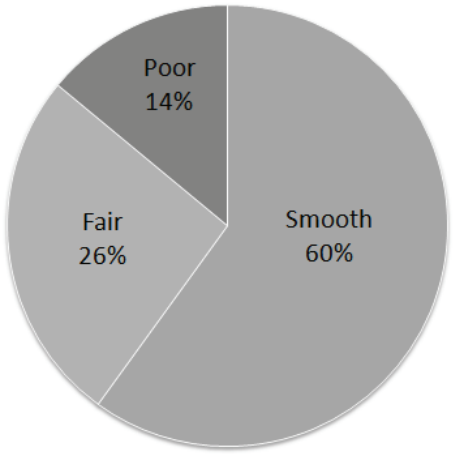

Figure 4: User feedback for streaming quality

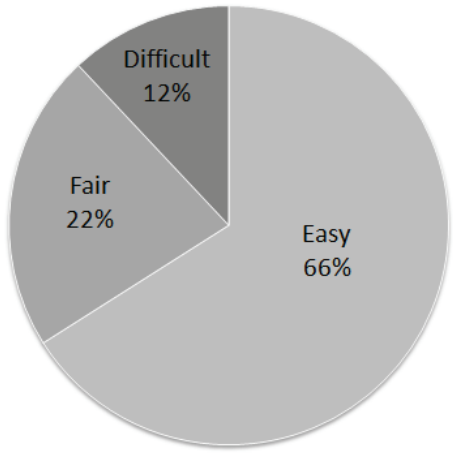

Figure 5: User feedback for system ease of use

5. CONCLUSIONS

We have presented the Class-X open source project. ClassX offers a low-cost solution for lecture capture, and has been used to cover many courses and events at Stanford University. ClassX also offers an interactive user experience through its IRoI video streaming and slide synchronization features. The viewer ratings of ClassX indicate their satisfaction with the video and streaming quality, and that they are finding the system in general easy to use.

Our user activity logs indicate that users often have different RoIs, and that the IRoI, automatic tracking, and slide synchronization features are actively being used by students. Publishers have also declared that they find the recording and publishing process easy. The system is built fully using open source tools to make it very easy to deploy and extend.

We have received many requests and suggestions to add content extraction features, such as board script and speech recognition, analytics, and accessibility features such as closed captioning and slide text-to-speech to the system. This was one of the motives behind the open source release. We aim to get more researchers and developers involved with the project to add these and other features that are useful to educators and students.

6. REFERENCES

[1] M. Bianchi. Automatic video production of lectures using an intelligent and aware environment. *Proc. 3rd Intl. Conf. on Mobile and ubiquitous multimedia (MUM'04), Maryland, USA*, pages 117–123, 2004.

[2] ClassX open source project homepage. `http://classx.sourceforge.net`.

[3] ClassX installer page. `http://classx.stanford.edu/Installer.html`.

[4] A. Mavlankar, P. Agrawal, D. Pang, S. Halawa, N. Cheung, and B. Girod. An interactive region-of-interest video streaming system for online lecture viewing. *Proc. of 18th International Packet Video Workshop*, pages 64 – 71, December 2010.

[5] A. Mavlankar, P. Baccichetand, D. Varodayan, and B. Girod. Optimal slice size for streaming regions of high resolution video with virtual pan/tilt/zoom functionality. *Proc. of 15th European Signal Processing Conference (EUSIPCO'07), Ponzan, Poland*, September 2007.

[6] A. Mavlankar and B. Girod. Background extraction and long-term memory motion-compensated prediction for spatial-random-access-enabled video coding. *in Proc. Picture Coding Symposium (PCS'09), USA*, May 2009.

[7] A. Mavlankar, J. Noh, P. Baccichet, and B. Girod. Peer-to-peer multicast live video streaming with interactive virtual pan/tilt/zoom functionality. *Proc. of IEEE Intl. Conf. on Image Processing (ICIP'08), USA*, pages 2296–2299, October 2008.

[8] S. Mukhopadhyay and B. Smith. Passive capture and structuring of lectures. *Proc. 7th ACM Intl. Conf. on Multimedia (Part 1) (MULTIMEDIA'99), FL, USA*, pages 477–487, 1999.

[9] D. Pang, S. Halawa, N.-M. Cheung, and B. Girod. ClassX mobile: An interactive video streaming system for mobile devices. Accepted for publication in Proc. ACM Intl. Conf. on Multimedia, AZ, USA, Nov. 2011.

[10] S. Tsai, D. Chen, J. Singh, and B. Girod. Rate-efficient, real-time CD cover recognition on a camera-phone. *Proc. ACM Intl. Conf. on Multimedia, Vancouver, BC, Canada*, Oct. 2008.

[11] C. Zhang, Y. Rui, J. Crawford, and L. He. An automated end-to-end lecture capture and broadcasting system. *ACM Trans. Multimedia Compt. Commun. Appl.*, 4(1):1–23, 2008.