Portable Presentation Player: Mobile Viewing of User-Controllable Movies of Slide Presentations

Yukiyo Uehori, Tohru Fuse, Laurent Denoue*, Patrick Chiu*

Fuji Xerox,
3-3-1, Roppongi Minato-ku, Tokyo 106-0032, Japan
*FX Palo Alto Laboratory, USA,
3400 Hillview Ave., Bldg. 4, Palo Alto, CA, 94304, USA
{yukiyo.uehori, tohru.fuse}@fujixerox.co.jp
{denoue, chiu}@fxpal.com

Abstract. Recorded presentations are difficult to watch on a mobile phone because of the small screen, and even more challenging when the user is traveling or commuting. This demo shows an application designed for viewing presentations in a mobile situation, and describes the design process that involved on-site observation and informal user testing at our lab. The system generates a user-controllable movie by capturing a slide presentation, extracting active regions of interest using cues from the presenter, and creating pan-and-zoom effects to direct the active regions within a small screen. During playback, the user can simply watch the movie in automatic mode using a minimal amount of effort to operate the application. When more flexible control is needed, the user can switch into manual mode to temporarily focus on specific regions of interest.

Keywords: Mobile applications, mobile phone, presentation viewing.

1   Introduction

Recorded presentations, lectures and tutorials such as screencasts are hard to watch on a mobile device with a small screen. A typical projector or wall display shows presentations at 1024x768 pixels, while a mobile phone screen has about 240x160 pixels. Even with smart phones (about 240x320 pixels), the actual physical size of the screens is likely to remain small because people prefer portable and small devices.

Another challenge is that in a mobile situation, users are frequently distracted by events around them. For example, during commuting or traveling on trains and airplanes, distractions like announcements, interruptions, and brief conversations can disrupt the viewing experience.

To deal with the first problem of limited screen size, our approach is to detect active regions of interest in the captured slides using cues from the presenter such as mouse cursor movement or hand gestures over the slide, which can be zoomed into during playback. To support viewing in a mobile situation, interaction can be minimized by generating a movie from the recorded presentation with pan-and-zoom

effects built in, augmented by manual controls to provide more flexible content navigation.

In this demo, we show *Portable Presentation Player* on a mobile phone (Nokia N95 8G). In the sections that follow, we describe the system in more detail, and report on our on-site observations and informal user testing at our lab.

2   System Details

Our system is designed to handle the captured slide presentation as synchronized media streams: (1) slide images, (2) audio, (3) user interaction events, (4) video of room activity. The video stream is optional because it needs much more infrastructure. The user interaction events can be obtained from mouse cursor movements or gestures in front of the slides detected from (4). The regions of interest can be extracted from PowerPoint API or by image segmentation algorithms. The active regions are identified as regions with user activity over them, along with a time interval. Another way to help identify the active regions is to run speech-to-text algorithms on the audio stream and do matching on the slide text. The active regions determine a time synchronized path through the slide images.

In previous related work, Multimedia Thumbnails [2], movies with pan-and-zoom were generated for documents (e.g. PDF files of papers) by extracting the document layout format and synthesizing audio of detected keywords. Because we are addressing a different domain that has the benefit of captured cues pointing to the active regions as well as the natural audio from the presenter, it can produce more robust results. Another related work is for photos and pictures [3], in which automatic paths for zooming and scrolling were generated by image analysis based on attention models.

Once we have the media contents (slide images, audio) and the active region path, an interactive Flash movie is generated. During playback on a mobile device, the user presses the Play button to view the presentation in automatic mode: the system plays the slide images and audio while zooming into the regions of interest along the active region path. At any given time during playback, the user can switch to manual mode by pressing the Mode Change button: the system allows the user to take control and manually navigate the focus to a specific region of interest independently of the timeline by using the arrow keys. When the user exits the manual mode, the automatic playback resumes to where it would have been at that time, transitioning visually by zooming in/out or panning. In both modes, the presenter cues (e.g. mouse cursor movements) are overlaid on the slide with an arrow-shaped icon.

In our current prototype, we have implemented a system that handles content (slide images, cursor events, audio) captured from a laptop used by the presenter; this component is a portable version of [1]. The authoring component is a Flash application that generates movies from the content and the metadata. The movie metadata is generated on a server PC, and for playback device we use a Nokia N95 8G mobile phone. During playback on the mobile phone (Fig. 1), the user can view the presentation as a movie in automatic mode, or manually jump around the regions using the keypad in manual mode.
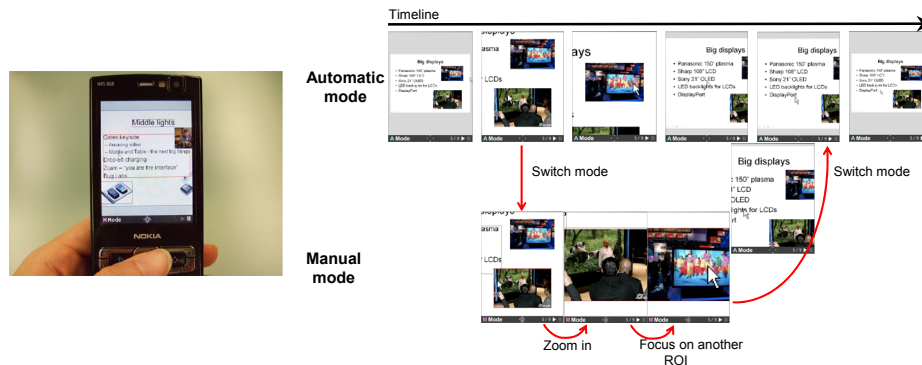
Fig 1. Left: Movie generated from captured slide presentation being viewed on a mobile phone. **Right**: During playback, user can take control and manually navigate the focus to a specific region of interest independently of the timeline.

3 On-Site Observation

We did on-site observation to acquire insights as part of our iterative design process. We tried out the application by ourselves in various mobile environments such as inside and outside of the office, and during commuting or traveling on trains.

We found that in a situation outside the office when the user's attention was distracted by the surrounding events (e.g. stopping of trains, announcements), the automatic mode with the pan-and-zoom movie was useful. Content understanding of a presentation was most effectively performed by listening to the audio, and a user could look at the screen for short periods with the automatic zoomed-in views when some interesting keywords were heard.

While a user could concentrate to some degree, the manual mode was frequently used to refocus or to regain attention. In both modes, we noticed difficulties when a phone was operated with one hand as it was not easy to use several keys required for some tasks. For example, with manual zooming, two keys (1 and 3) were used to zoom in and out. An improvement we added is a single-key quick zooming feature using the Enter button, which cycles through three zoom levels (whole page, fit selected region, optimize based on font height).

Other interesting observations include the following:

- Repeat playback and resume functions are important because many distractions occur unexpectedly in settings outside the office.
- Pan-and-zoom movie can be useful for a quick glance without having to do any other operations while in busy situations.
- Visual transitioning by pan-and-zoom animation is helpful when switching from manual mode to automatic mode, or going from the user's selected region on a small screen to the presentation's active region of interest.
- Quick bookmaking/annotation feature can be useful for reviewing a presentation in case a user cannot watch or listen in detail at a certain time.

4   User Testing

We conducted informal user testing at our lab to evaluate the prototype's usability. There were 6 participants. Each performed 13 tasks with the Portable Presentation Player on activities that include: content viewing, reading and comprehension from audio in automatic mode, reading and describing pictures on a slide in manual mode, mode transitioning.   User operations were logged by the application for later analysis. After performing the tasks, users filled out questionnaires.   The content was from a real presentation captured by our system described in Section 2.

The results show that the users were able to perform the tasks with very little problem: the number of key press errors was 2 out of the total of 327 (0.6%), and the task completion percentage was 100%.

For the information comprehension, there was a task question that 3 users could not give the correct answer. This particular question was difficult because it required the user to carefully pick out the information from the audio and slides.

From the questionnaires (based on a Likert scale from 1 [Disagree] to 5 [Agree]), users indicated that the automatic mode feature is easy to use (3.8 average score), the manual mode is easy to use (4.0), the manual mode is necessary (4.7), landscape mode is necessary (4.2), and seamless transition between auto and manual mode is necessary (4.5).   The users also indicated that it was easy to read the text (3.8), view the figures (4.2), and understand the audio (3.7).

Users were also asked for comments on preferences and difficulties.   Four users commented that the manual mode is what they liked about the application (which agrees with the high score of 4.7 above on a similar issue).   This suggests that there may be a design tradeoff between providing "hands free" operation with an automatic mode to support mobility, versus providing flexible playback control with a manual mode to satisfy user preference.

5   Conclusion

This demo showed an application designed for viewing captured slide presentations with audio in a mobile situation. From on-site observations and informal user testing in our lab, we developed a design that provides effective automatic and manual playback of a movie generated from the presentation content for mobile phones.

References

1. Denoue, L., Hilbert, D.M., Adcock, J., Billsus, D., Cooper, M. ProjectorBox: Seamless presentation capture for classrooms. *Proceedings of E-Learn '05*, pp. 1986-1991.
2. Erol, B., Berkner, K., Joshi, S. Multimedia thumbnails for documents. *Proceedings of ACM Multimedia '06*, pp. 231-240.
3. Liu, H., Xie, X., Ma, W.-Y., Zhang, H.-J. Automatic browsing of large pictures on mobile devices. *Proceedings of ACM Multimedia '03*, pp. 148-155.