

BU ProPane Project Background

BU ProPane Team:

Griffin Dunn

Colin Madigan

Phillip Stahlfeld

September 12, 2012

The purpose of this document is to introduce the BU ProPANE project to a third party with no prior knowledge concerning the project. The document begins with a high-level discussion of the problems that this project is aiming to solve. Next, the document provides relevant background information regarding current practices and relevant terminology. The document concludes with a technical research section that contains ideas that will prove useful during the implementation phase of the project.

Problem Statement

The goal of this project is to create a system that captures all of the information written on a board during a class in a readily accessible manner. The two driving forces behind solving this problem are: autonomous collection of notes for students with disabilities and providing a means for professors to compare their notes with the actual information presented during a lecture.

Current practices at Bucknell rely on the aid of a student volunteer to provide photocopies of his or her notes to students who are incapable of taking notes themselves during the class. This is not an optimal solution because: the student volunteer must be in class every day, turnaround time for the notes is too long, and the quality of the student volunteer notes cannot be guaranteed. In addition, several professors use different colored writing to represent different forms of information. If the student volunteer did not copy the notes in color, then any students using the notes would not have access to the same information.

Typically, professors have a lecture planned out before class starts and have notes prepared to follow their plan. However, when students ask unexpected questions or do not fully grasp the material, it is easy to deviate from this plan. For example, civil engineers at Bucknell University plan out their lectures by dividing a piece of paper into sixths and filling in each block with what they want to present (See Figure 1). Problems arise when information not included on these pieces of paper are written on the board (due to student questions etc.). When the professor is preparing to give the lecture the following year, he or she will have to remember what the previous year's students struggled with. Using a board capture system, professors could compare previous lectures with previous lecture notes to generate a more effective lecture plan.

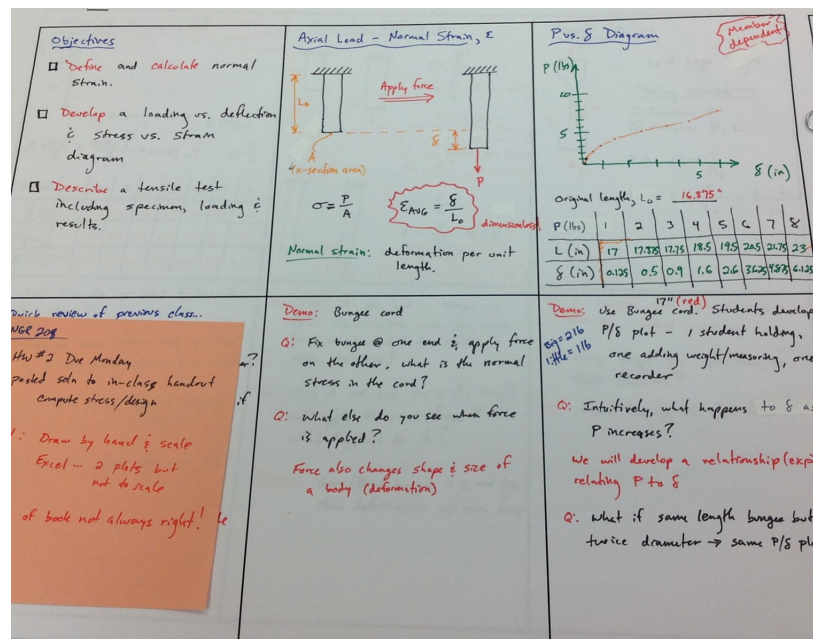


Figure 1: Image of civil engineering lecture notes. Orange note is an example of how lecture notes are recorded and altered. Image courtesy of Doug Gabauer 2012.

Background Information

How are things currently done?

Smart phone apps
Professional scribes

Useful Terminology

WE WILL GET THIS WITH MORE RESEARCH

Research

Competing Technologies

There are three general categories: phone apps, scanners, and electronic whiteboards.

Phone Apps-

Phone apps scan an image of a whiteboard and filter out the unnecessary information (See Figure 2). Several examples of phone apps are listed below.

- WhiteBoard Capture Pro
- Qipit White
- Genius Scan
- JotNot Scanner Pro

Scanners-

Scanners are pieces of hardware that attach to white boards that will track movements on the board (some with the usage of an electronic pen). These products will not be fully suited to this project because they require installation and calibration, which violates the idea of portability. Several examples of scanners include:

- MimioCapture
- eBeam System 3
- Interlink
- FreeBeam

Related Technologies

SMART boards

Research Data

Microsoft Whiteboard Capture System

Source: Microsoft Research Technical Report MSR-TR-2002-89
<ftp://ftp.research.microsoft.com/pub/tr/tr-2002-89.pdf>
All images used in this section are from the source document.

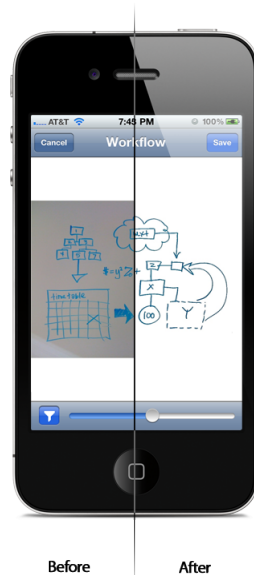


Figure 2: Advertisement demonstrating the image filtering capabilities of the WhiteBoard Capture Pro iPhone app. App eliminates discoloration of board due to glare. Taken from <http://www.beetlebugsoftware.com>.

Summary

The Whiteboard Capture System came out of Microsoft Research in 2002. It was research for a bigger Distributed Meetings project. It can record all information placed on a whiteboard during a session, as well as the audio in the room, and it provides a platform for users to review key information from the session later from their own computers.

Key Features

The Whiteboard Capture System (WCS) requires only a camera, microphone, and computer to operate with any existing whiteboard. It records all information written on the board during a meeting, and all audio in the room during the meeting. The system then analyzes the data and detects “key frames”, or pictures of the board which the system determines to contain the most important information. The key frames are time-stamped so they can be associated with the audio that was recorded at the same point in time. Users can access the key frames securely via a web-based browsing system. Audio can be played by clicking on a key frame, and the option exists for the presentation to unfold for the viewer like a movie, where information will appear on their web-based whiteboard as the audio recording talks about it, as shown below in Figure 1.

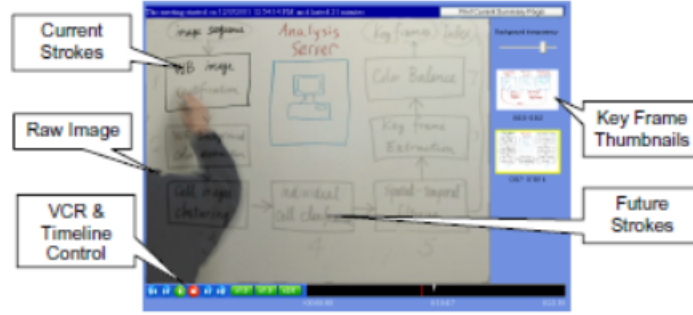


Figure 3: Browsing system representation of meeting playback

Technical information

The WCS uses a Canon PowerShot G2, a 4 megapixel digital still camera. The camera provides images with 2272 x 1704 pixel resolution, equivalent to 31.6 dpi for a 6' x 4' whiteboard. The camera is controlled via USB by a computer, which utilizes the available SDK for the camera to communicate. It takes a picture roughly every 5 seconds, which is as fast as possible for the PowerShot G2. Software specifies camera parameters on a per-shot basis. Pictures are saved as JPEGs and sent to the computer via USB. The computer is a Pentium III 800 MHz dual CPU PC. It takes approximately 20 minutes of processing for every hour of session time. The input image sequence requires 34 MB/hour using motion JPEG compression, while the audio requires 15 MB/hour to store 16-bit 11 KHz mono mp3 audio.

To analyze the captured images, the WCS first divides the frame into a grid of rectangular cells of approximately 1.5" x 1.5" inches. The frame is represented by a 3D matrix of cell images. Analysis using cells requires much less computational power compared to a per-pixel analysis. Once the frame is divided into cells, the system uses an algorithm to prepare key frames in a readable fashion. The algorithm is as follows:

1. **Rectify** the whiteboard region of every image in the sequence
2. **Extract** the whiteboard background color
3. **Cluster** the cell images throughout the sequence for the same cell. If two images are considered to be the same, they are clustered in the same group.
4. **Classify** each cell image as a stroke, a foreground object, or the whiteboard
5. **Filter** the cell images both spatially and temporally to refine the classification results.
6. **Extract** the key frame images using the classification results
7. **Color-balance** the key frame images

1. To rectify the whiteboard region, they use a one-time calibration step where the corners of the board are selected to manually define the frame.

2. The whiteboard color is computed by finding cells with the brightest luminance over time and small variance. During this step, cells which are impeded during the capture process (for example, by a person standing in front of the board) act as outliers. These are detected and accounted for using a least-median-squares technique.

3. Cell images are clustered together if they are determined to be identical. This is determined by a modified Normalized Cross-Correlation algorithm and the use of the Mahalanobis distance calculation.

4. Cells are classified as whiteboard, stroke, or foreground cells. Heuristically, whiteboard cells are uniform in color, either white or grey. Stroke cells are mostly white or grey with one or two primary colors mixed in. Foreground cells do not fit either of these descriptions. Essentially, it determines if a cell image is the same as a whiteboard cell image, very similar to a whiteboard cell

image (stroke), or completely different (foreground object).

5. Filtering is used to better define cells. The system's spatial filter changes totally isolated foreground cells to stroke cells, since typically foreground cells are clustered (i.e. a person). Stroke cells touching foreground cells are also changed to foreground cells, unless said cells are also filtered temporally. The temporal filter looks at all cell images for a certain cell over time. If two cell images match, all cell images in between are deemed to be the same image as well. This filters out foreground objects.

6. Key frames are considered a summary of the entire presentation. The WCS defines key frames as frames that capture all information on the board just before an erase. They should only contain whiteboard and stroke cells, the whiteboard color should be uniform and the pen colors should be distinct. Image reconstruction is used to ensure the quality of the key frames. If the cell image is a whiteboard or stroke, its own image is used. If it's a foreground object within the span of a stroke, it is replaced with the stroke image from the neighboring frame. Otherwise, it is replaced by a whiteboard cell image. Key frames are only present before large erases, so they are only identified if at least 20% of the total stroke count is erased. The WCS divides the presentation into chapters, with each chapter being concluded by a key frame.

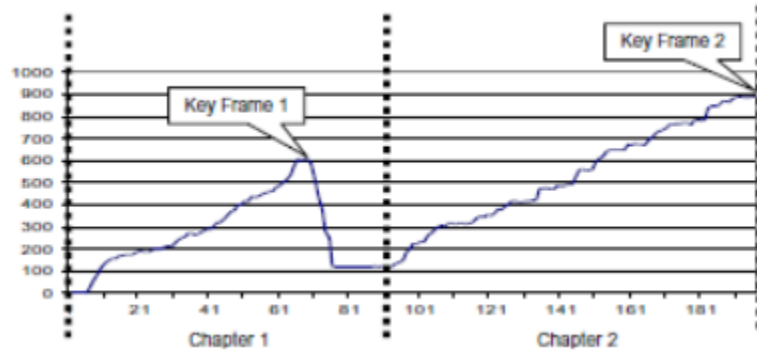


Figure 4: Plot of number of stroke cells vs. time. Key frames are shown at peaks just before stroke count decreases dramatically.

7. Color balance is done to make the background uniformly white and increase the color saturation of the pen strokes. It also reduces image noise. A final image is presented in Figure 2, below.

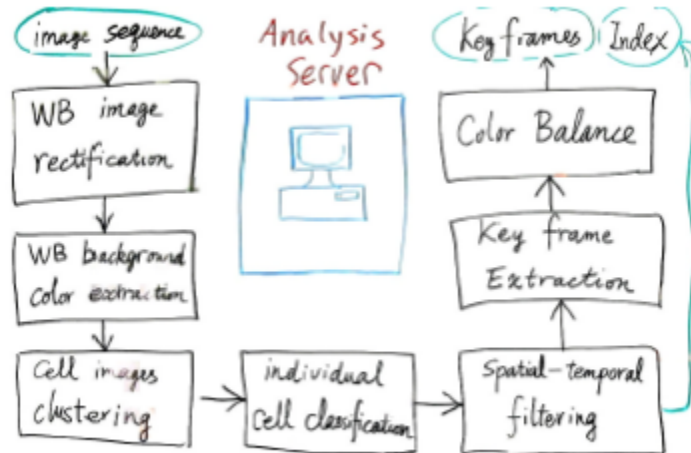


Figure 5: Example of a key frame as processed by the WCS

Browsing software allows viewing of key frames and audio playback. Users can zoom in/out and cut and paste the image to other documents. Double clicking on pen stroke cells plays audio at the time of the time stamp of the cell. Stroke cells will then be displayed as time progresses in the audio to their respective time stamps.

Access to the browsing software is via a token-based access model, where an access URL is emailed to all meeting participants.

Limitations

For proper functionality, the WCS requires a constant use environment. The whiteboard must have a constant color, constant lighting, and the camera must have constant exposure settings for the system to work properly. Most of these factors are standard for a normal whiteboard environment. However, the research team suggests that if a constant environment could not be guaranteed, the system might still work if a patch of known color were placed in the camera's field of vision, so the system could adjust settings on the fly based on the changes to the constant color patch.

At the time of publication, the team noted that the frame rate of cameras on the market was a limiting factor for the system. Since this research is 10 years old, this is no longer as limiting - a quick search of competitively priced cameras returned cameras with over twice the frame rate as the PowerShot G2 used here. Other problems with the system are less likely. If a person in a white shirt stands perfectly still in front of the board for a long period of time, the system may not deduce that the person is in the foreground. This is interesting to note, but not a likely occurrence in a lecture environment.

Relevance to ProPANE

This system, as a whole, is very similar to our final goals for the BU ProPANE system. Especially useful to us is the key frame idea, since professors using the system will not want to sort through hundreds of images in order to find just a few important ones. The idea of breaking the board up into cells for analysis is also very interesting to us, as it greatly reduces the processing requirements. However, it is not perfect for our project. The WCS is purely a *whiteboard* capture system, whereas our system should work for black boards as well. Additionally, we are not focusing on any audio capture or playback for our system.