

Vision-Based Projection-Handwriting Integration in Classroom

Weihong Li
Graduate Center
City University of New York
New York, NY 10016
wli@gc.cuny.edu

Hao Tang
City College
City University of New York
New York, NY 10031
tang@cs.ccny.cuny.edu

Zhigang Zhu
Graduate Center & City College
City University of New York
New York, NY 10031
zhu@cs.ccny.cuny.edu

Abstract

A projector based presentation and a whiteboard handwriting capture system, when monitored by a camera, could provide a better means to present a lecture in a classroom and later to deliver more effective e-lectures for off-campus students. When the PowerPoint® (PPT) slides are projected on the whiteboard where the instructor annotates, modifies, or expands the PPT presentation, geometrically registering high-resolution images from the two sources yields high quality digital presentations. Since the two sources do not share the same materials, one with printing notes and the other with handwriting notes, a digital camera is used as a bridge to align the two sources, as well as to record the whole lecture. We present a hybrid registration approach to align the PPT and handwriting images and then propose a coarse-to-fine content matching method to register handwriting contents captured by the video camera and the whiteboard capture system. We also show an integrated user-customized presentation system to provide students better lecturing context. Experimental results are presented to validate our approach, and future research directions are also discussed.

1. Introduction

A domain of primary importance in the future of web-based technology and digital libraries is distance and electronic education (e-learning). Multimedia materials from classrooms and seminars are rich sources of information. Today's first generation e-learning systems primarily adopt a "record-and-playback" approach, which does not leverage the processing capabilities (during live capture, after-capture post-processing, and during later user interaction with archived materials) that we believe will underlie the next generation of more automated, flexible, and interactive e-learning systems.

Using a portable presentation system with cost-effective multimedia sensors, we have developed the

following basic components for our Virtualized Classroom project - automated data collection, intelligent media processing and integration algorithms, and user-customized presentation interface designs. In this paper, we will mainly focus on alignment and integration of images with PPT printing notes and images with whiteboard handwriting contents. So the result is an ability to dynamically allow the instructor add handwritten material generated in a different medium (i.e., the whiteboard) onto the projected slides. This is done in high resolution using low-resolution video as the alignment mechanism. Our approach is fully automated, without any special requirements for the instructor using the system. This approach has further potential towards content-based media integration with handwriting recognition and content-based video representation.

The organization of the paper is as follows. Some related work is discussed in Section 2. In Section 3 we give a brief description of our classroom sensors, and the automated data collection and synchronization mechanism. Section 4 gives an overview of our hybrid media integration approach. Section 5 presents algorithms for registering the low-resolution video with high-resolution PPT slide images. In Section 6 we describe our content matching method to align video and whiteboard handwriting images. Section 7 presents several experimental results of registering high-resolution PPT images and whiteboard handwriting images via the low-resolution video. Section 8 presents our user-customized presentation interface. Then, we conclude our paper and introduce some of our future work in content-based media integration in Section 9.

2. Related Work

Projects focused on record and playback technologies include: the Georgia Tech Classroom 2000 (eClass) [1], the CMU Just-In-Time system [7], and the UMass RIPPLES / MANIC [15], among others. Automated production is a major issue for these e-lecture systems, because to improve the quality of presentations many systems require significant manual effort in analog data collection, digitization, and

synchronization. We note that some systems have introduced forms of automatic production. Cornell's lecture browser [11] includes lecturer tracking, slide change detection and segmentation, and matching slide projections with digital slides. A precursor to Cornell's lecture browser, UC Berkeley BMRC Lecture Browser [3] has several additional features including audio search capabilities, bookmarks and an interactive whiteboard for online students. Auto-Auditorium [2] uses multiple cameras that automatically switch based on context. However we are not aware any work trying to automatically registering PPT presentations with handwriting contents.

Another class of related work is automated keystone correction [13, 16] and projector array alignment [6] using video cameras and computer vision techniques. We share the similar ideas of using a video camera and the homography of the planar projection. However, our task is more challenging. The keystone correction could be easily implemented in real-time since only the perspective distortion needs to be removed for keystone correction. In the work of automated projector array alignment, the registration is only among video images and is an off-line calibration step. In our application, we need to dynamically register three different types of images (PPT slide, handwriting, and video images) on-line.

3. Media Capture and Synchronization

The commonly used classroom and seminar presentation tools are a multimedia projector with PPT slides, and blackboards /whiteboards. In order to digitize the classroom handwriting contents to create e-lectures, we use a low-cost Mimio[®] digital whiteboards system [10] to substitute for the use of the blackboard.

Therefore, the "sensors" we are using in our portable presentation system for a classroom setting are a PPT slide capturer, a Mimio Virtual Ink handwriting system, and a digital video camera (to capture classroom video/audio). All the sensors are controlled by the laptop used for the presentation and can be easily managed by the instructor/lecturer using the system [4].

The PowerPoint *slide capturer* was modified from the Berkeley PPT Recording Add-In [3, 12]. We added the start date and time information of the presentation under recording in order to synchronize the recorded PPT slides with the accompanying whiteboard pages. The pop-up dialog-box of the PPT slide capturer add-in (top-right in Fig. 1) is activated automatically when the instructor starts his/her PPT presentation. In addition to the pages of slides in one of the image formats (e.g. JPEG), a PPT log file is automatically generated with timing information and the titles of all the slides in the presentation.

The Mimio Virtual Ink ultrasonic position capture system consists of a capture bar, color-coded marker sleeves and an electronic eraser (top-left in Fig. 1). The system is capable of recording handwriting and drawing contents on a normal whiteboard of 2.4 m x 1.2 m (8 ft x 4 ft), with 100 dpi resolution. The capture bar is a two-foot ultrasonic tracking array positioned along the upper-left edge of the whiteboard. The electronic marker sleeves transmit an ultrasonic signal to the capture bar, which triangulates the pen's position on the board as the instructor writes. The whiteboard presentation is saved in a series of html files, one html file (with timing information) for each whiteboard page saved in a JPEG image. The system can also be used as a remote mouse in front of the whiteboard.

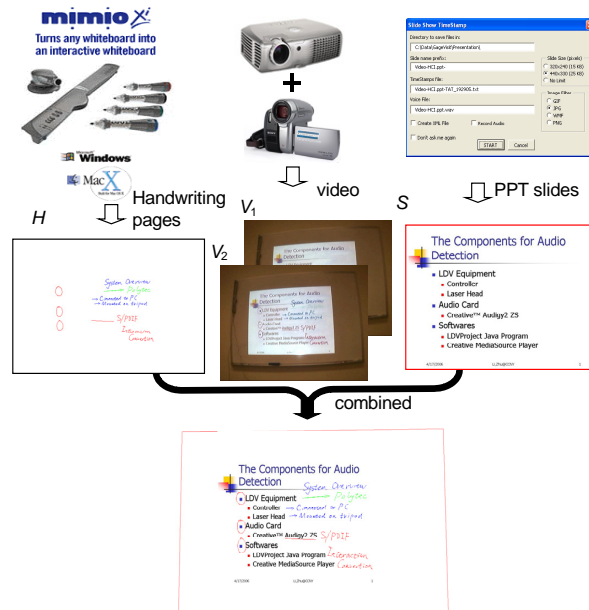


Fig. 1. Multi-sensor data collection, synchronization and integration in a classroom setting.

For the best use of the above sensors, we assume that the instructor will use a projector to project PPT slides on a whiteboard. The handwriting contents written on the whiteboard could be on top of the slide projections. A video camera is used to automatically collect audio/video of the classroom activities, mainly the projector, whiteboard and the instructor.

After the instructor sets up the sensors and starts the presentation, everything is automatically saved for him/her. The synchronization of the PPT slides, whiteboard pages and audio/video streams is enabled by a simple stream synchronization algorithm which uses the timing information in the PPT log file, the whiteboard log files, and the video/audio stream [4].

4. Media Integration Overview

The PPT slide presentation and the whiteboard system when used together could provide a better means to present a lecture in a classroom. When the PPT slides are projected on the whiteboard, where the instructor annotates, modifies or expands the PPT presentation, we need to geometrically register the images from the two sources to create high quality electronic presentations. If all the devices (the slide projector and the whiteboard system) could remain stationary during the lecture, we could require the instructor to mark at least the four corners of the PPT projection area in the whiteboard as a “calibration” step (Fig. 2) [4]. The whiteboard image with the four markers is captured by the Mimio Virtual Ink and a corresponding PPT slide image is captured by PPT slide capturer. Since both of them are images of the same plane (the whiteboard), we can use a projective transformation to register the two images with (at least) four point pairs without using the video camera.

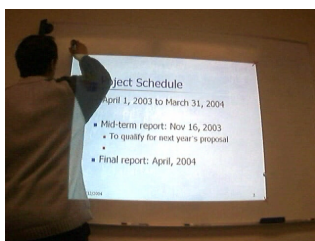


Fig. 2. Slide and whiteboard image registration with projection border markers

However, this simple approach has several drawbacks: (1) The occlusion of the instructor to the projections usually causes difficulty to accurately mark the projector’s four corners. (2) Sometimes the corners are too high to be reached; in other occasions, the instructor may forget to mark the four corners. (3) It is hard to keep all the devices stationary with a portable representation; if either the projector or the whiteboard capture bar moves, the instructor has to mark the four points again.

This paper presents a completely automatic and “mind-free” approach. Since the two sources (PPT slides and handwriting pages) do not share the same contents, we use the low-cost digital camera (which is also used to capture the classroom video) as a bridge to align the two source images. This is possible since video images include both the PPT slide projections and the handwriting contents on the whiteboard. Without a camera, the method of marking four points at any locations does not work since the marks can only be recorded in Mimio pages, but not the corresponding PPT slides.

As an on-going research, we are also working on analyzing and presenting lecturer and students' actions in the classroom scenario. The recognition of the activities in the classroom could provide a better interaction between lecturer and students and human-computer interaction. As a by-product, the activity recognition is also helpful for data collection and media integration described in this paper, e.g., to save the Mimio page, re-do calibration when needed, etc.

Here is the overview of our approach. First, the corresponding PPT slides, handwriting pages, and video frames are matched up via the synchronization mechanism in the following manner (Fig. 1). When the instructor changes slides, a new slide image S is saved by the PPT slide capturer and a slide-only video frame V_1 is extracted. When a new handwriting page H is saved by the Mimio Virtual Ink, the timing information will locate this page with the right slide S , and the video frame V_2 with both slide projection and handwriting contents will be saved automatically. It is usually the case that before and after handwriting, there is a short time period when no occlusion occurred. Thus the handling of occlusion problem is avoided by using this observation to obtain V_1 and V_2 . Second, the slide S and the handwriting page H are registered by way of the video frames V_1 and V_2 . The slide-video registration is implemented by finding the boundary of the projection area in the video frame V_1 . The video-whiteboard registration is fulfilled by matching the handwriting contents saved in the Mimio page H and the video frames (V_1 and V_2).

Even though we register images on a planar whiteboard, we face the following challenging issues: (1) The images are from three completely different sources (sensors) – direct image copies of PPT slides, video images of the slide projections and handwriting inks, and the handwriting strokes recorded by the localizing sensor of the Mimio Virtual Ink. (2) There are unknown perspective distortions among the three sources. The PPT projections are subject to keystone distortion. The camera is placed in a convenient location; therefore perspective distortion may be unavoidable. In the following two sections, we discuss the two main parts of our registration approach, namely slide-video registration and whiteboard-video registration. Then we will present several slide-whiteboard integration results in Section 7, when the devices move and keystone distortion exists.

5. Slide and Video Registration

First we assume that the projection area is always in the camera’s field of view (FOV). In the current implementation, the size of a digital PPT slide image is 720x540, the size of a Mimio page is 720x480, and the size of a video frame is 640x480. The image

resolutions are comparable; however, the projection area is only a small portion of the entire whiteboard that should be visible in a video frame. Hence, the resolution of the fonts in the video frames is much lower than in the digital slides. In addition, the illumination of the slide projections is different from the ideal digital slides, and it is not uniform. Radial distortion may exist as well as perspective distortion in the video frames. Thus, directly matching the fonts in the corresponding PPT digital slides and their projections captured by the camera is a non-trivial task. Therefore, after we remove the radial distortion of the digital camera, we register the video frames with PPT slides by using *a priori* knowledge of the projection area, e.g., the projection boundary.

5.1. Camera calibration for distortion removal

In order to remove the radial distortion (e.g. in image V_1 of Fig. 3a), we apply the calibration method proposed in [8]. The radial distortion is modeled as

$$\begin{aligned} x' &= x(1 + k_1 r^2 + k_2 r^4) \\ y' &= y(1 + k_1 r^2 + k_2 r^4) \end{aligned} \quad (1)$$

where k_1, k_2 are the coefficients of the radial distortion, (x, y) and (x', y') are the normalized image coordinates before and after distortion removal, and $r = (x^2 + y^2)^{1/2}$. Fig. 3b shows the image after removing the radial distortion. Clearly, the boundaries of the projection areas become straight in the rectified image V_1' .

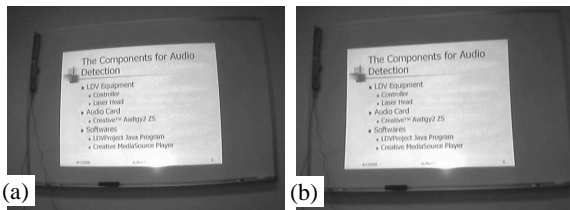


Fig. 3. Video images (a) before distortion removal (V_1) and (b) after distortion removal (V_1').

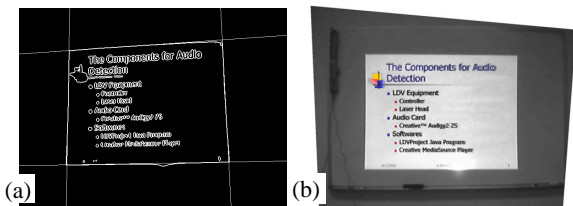


Fig. 4. (a) Boundary detection and fitting (b) Video-slide matching

5.2. Slide-video registration algorithm

The boundary of the projection area in the rectified video frame V_1' (Fig. 3b) needs to be detected in order to register it to the corresponding digital slide image S . Since the projection area is usually significantly

brighter than other areas, we first generate a binary image from the video frame. Then we use the Laplacian-of-Gaussian (LoG) operator to obtain the edges (zero-crossings). Finally we use the Hough Transform to extract four boundary lines and calculate the coordinates of the four corners of the projection area (Fig. 4a).

A 3×3 projective transformation matrix A_1 is calculated between the quadrilateral formed by four boundary lines and the rectangle frame boundary of the PPT slide image. The projective mapping relation from the video frame point x_v to the slide image point x_s is represented by

$$x_s \cong A_1 x_v \quad (2)$$

Fig. 4b shows the video and PPT alignment result that uses the projective mapping matrix A_1 . The video frame is transformed to the PPT slide coordinates so that the orthogonal view of PPT slide remains. The perspective distortion of the original video frame is obvious from the shape of transformed video frame. Note that only the fonts (but not the background) on the PPT slide are superimposed on the transformed video frame for showing the accuracy of the alignment.

Sometimes, very cluttered notes in a slide will affect the boundary extraction results using the Hough Transform. Therefore, we have tried simple ways to remove the PPT print notes. For example, one approach is to scan the edge map of the video frame from its four borders in order to only keep the boundary edge pixels. Since we use the Hough Transform for boundary extraction, we do not need to have the edges of the complete projection boundary.

In some other cases, the projection boundary might not be obvious due to the use of difficult background colors in PPT presentations. Our ongoing work includes developing methods to match between the layouts of the PPT printing notes in the video frame and the PPT digital slide, as in [11].

6. Whiteboard and Video Registration

The matching of the whiteboard handwriting page and the video frame is more challenging since the camera cannot “see” the invisible frame boundary of the whiteboard page. Therefore we propose to match the handwriting contents from the Mimio Virtual Ink pages and the video frames. The Mimio Virtual Ink records the locations (the “virtual” ink) of the pens in working, while the video records the real inks. Due to the different recording methods as well as several steps in obtaining handwriting contents from the video frames, the thickness of the strokes is different in images from the two sources. Second, the strokes in the Mimio Virtual Ink are usually complete while those from the video frames are not. Third, the scales of the fonts in the two images are different. This section will

discuss how to extract and match handwriting contents from these two completely different recording sources.

6.1. Feature extraction

The feature extraction is to obtain the handwriting contents by subtracting the video frame with PPT projection only (V_1') from video frame with both PPT projection and Mimio handwriting (V_2'). Since the difference image, V_d , shows obvious differences in places with handwriting contents, a simple thresholding process can reveal most of the handwriting contents (Fig. 5a). Note the contents are not complete in the difference image. For example, the handwriting notes “Polytec” in green (see Fig. 1) in the center of the image V_d are almost missing.

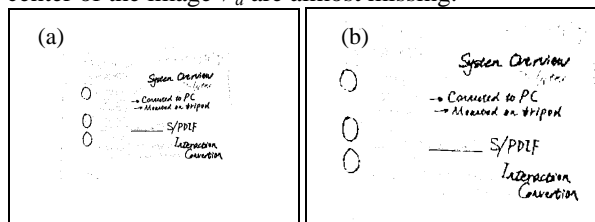


Fig. 5. Handwriting contents from video difference. (a) Difference image V_d (b) Rectified difference image V_h

Using the projective mapping matrix in Eq. (2), we can rectify the video handwriting contents to align with the digital PPT slide so that the camera perspective distortion is removed (Fig. 5b), if the keystone effect (another perspective distortion) of the slide projection is not significant, as in this example. Note that this is valid since the handwriting contents are written within the slide projection area. The rectified video difference handwriting image is denoted as V_h for convenience.

6.2. Board-video registration algorithm

We have developed a two-step content matching approach for registering handwriting contents from the two different sources. In the initial matching step, we try to find some robust features to translate and scale the images in order to roughly align the two. In the following refinement step, we use the connected components in the Mimio page as content-based matching primitives to find the best matches in the video handwriting image. As a result, a projective transformation matrix obtained to align the Mimio page H with the rectified video frame V_h , which has been aligned with the PPT digital slide image.

(1). Initial matching

In order to better match the two handwriting contents from the two sources, with different thickness of the strokes, we first run a thinning algorithm as proposed in [14, 9] on both video and Mimio images to get skeletons of the handwriting (Fig. 6a and 6b). Then, the centroids of both thinned images are

calculated. The thinning process is useful to obtain the centroid with the least effect from the unbalanced stroke thickness. The difference in the locations of the two centroids, C and c , in the video and the Mimio images respectively, gives the translation between two images: $t = C - c$.

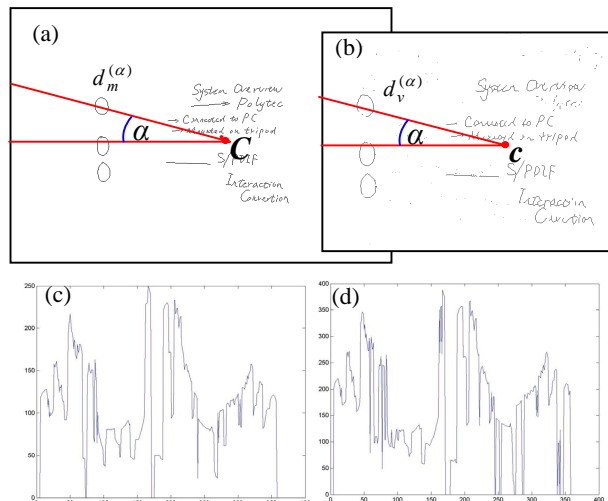


Fig. 6. (a) Mimio note skeleton S_m (b) Video note skeleton S_v (c) Polar profile for the Mimio skeleton (d) Polar profile of the video skeleton

To determine scale factor between these two images, we generate a centroid-centered polar profile for each image by using the distances of all the outmost stroke points in all directions from the centroid. Fig. 6c and 6d show the polar profiles, $\{d_m^{(\alpha)}\}$ and $\{d_v^{(\alpha)}\}$, in the Mimio image and the video image, respectively, $\alpha = 1, 2, \dots, 360$. Then the scale factor s is given by

$$s = \frac{\sum_{\alpha=1}^{360} d_v^{(\alpha)}}{\sum_{\alpha=1}^{360} d_m^{(\alpha)}} \quad (3)$$

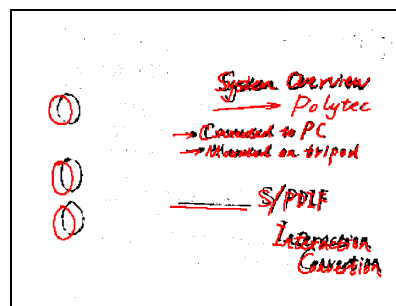


Fig. 7. Initial matching result image (Video + Mimio). Translation $t = (71, 11)$; scale $s = 1.523$

After image translation and scaling, the two images are roughly aligned (Fig. 7), which gives a good start for the refinement step. Note that the 1.523 scaling factor (Fig. 7) indicates that the fonts in the rectified video handwriting image are 1.523 times the size of the

fonts in the Mimio handwriting page. Since the video image has been scaled in order to be aligned with the PPT digital slide, this scaling factor is consistent with the ratio of resolutions of the PPT slides and the Mimio pages, with the similar image sizes but different fields of view. Therefore, we translate and scale the original Mimio handwriting page (before thinning) to match up with the video image (and hence the PPT slide).

(2). Fine registration

We once again run the thinning algorithm on the transformed Mimio image to generate the handwriting skeleton image. However, in order to make the match robust to small rotation and perspective distortion, we run a morphological dilation operation [5] to turn the handwriting fonts to a thickness of about 3 pixels on both the video and Mimio skeleton images. Then, we extract from both images connected components as the matching primitives. Each connected component could be a single letter, part of a letter, a continuously figure or includes several letters. The connected components in the Mimio page and in the video frame do not have one-to-one correspondence (Fig. 8). A rectangular bounding box indicates the corresponding connected component inside the box. Since the connected components in the Mimio and the video images are different, and in the latter are incomplete, we do not use connected components as match primitives in the video image. Instead, we use the connected components in the Mimio page as *content-based matching templates* and search for the right matches in the video image by an adaptive template matching method based on normalized cross-correlation. This method is efficient since only one match per connected component (i.e., handwriting notes) needs to be searched for locally.

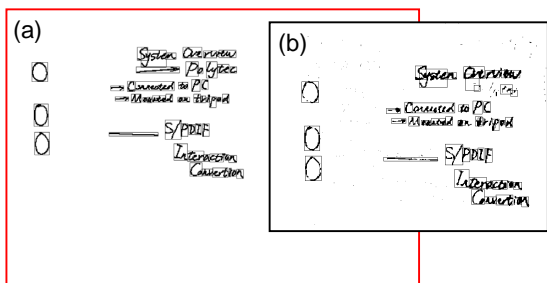


Fig. 8. Fine matching primitives (a) Matching primitives with rectangles in Mimio; (b) correspondence regions in video

First, we select those connected components of the Mimio page whose size is sufficiently large for a robust match. Then, we use the pixel pattern under the bounding box of each selected primitive as the matching template to search for the best matched rectangle region in the video handwriting image using the normalized cross-correlation measures (Fig. 8b).

The size of the searching range depends on the accuracy of the initial match.

We choose those matches whose maximum normalized correlation values are above a threshold (e.g., 0.3) to calculate the final projective transformation. In order to keep the image resolution as high as possible, we only perform the geometric transformation once from the original Mimio handwriting page to the corresponding PPT image. Since we used a translated and scaled Mimio image for the video-whiteboard matching, we calculate the coordinates of the centers of the matching rectangles in the original Mimio handwriting page using the scale s and translation t . Thus, the projective transformation matrix A_2 between a point x_h in the original Mimio page and the corresponding point x_s in the rectified video image (i.e. the original PPT slide) can be obtained by using a least square method with the matches. Therefore the transformation is

$$x_s \cong A_2 x_h$$

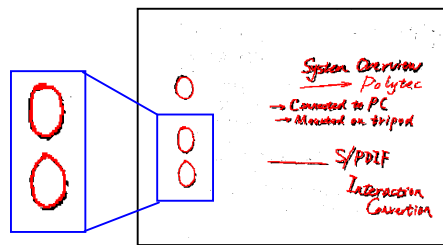


Fig. 9. Fine registration result (Mimio contents in red and video contents in black)

Fig. 9 shows the fine alignment of the handwriting contents from video and Mimio, respectively, using the projective transformation matrix. The transformation matrix A_2 is used to register the Mimio handwriting page to the corresponding PPT slide image (Fig. 10). Note that about one-pixel misalignment can be found, e.g., in the close-up shown in Fig. 9. However, for our application, the accuracy is sufficient for a good integration, as illustrated in Fig. 10.

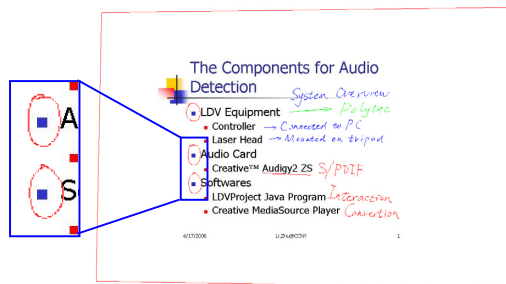


Fig. 10. Mimio and PPT registered image

7. Slide-Board Integration Results

This section presents a few examples of the slide-whiteboard integration results. The transformation matrix A_2 is used to register the Mimio handwriting page to the corresponding PPT slide image. In the integrated page, the orthogonal view of the PPT slide remains; the handwriting and drawing contents in the Mimio page are aligned and superimposed on the PPT slide inside and outside the projection area. Fig. 10 shows the integrated result of the example we have shown in the previous sections. Note that circling is accurately aligned with the printing text. The red boundary for the Mimio page shows that the PPT projection area is almost orthogonal. Fig. 11 shows the integration result for the following slide that uses the same projective transformation matrix of Fig. 10, since none of the devices (the camera, the projector and the whiteboard system) moves. Note that handwriting contents could be out of the projection area in Fig. 11. The time needed for each calibration is about 2 seconds on a Pentium 1.2GHz PC.

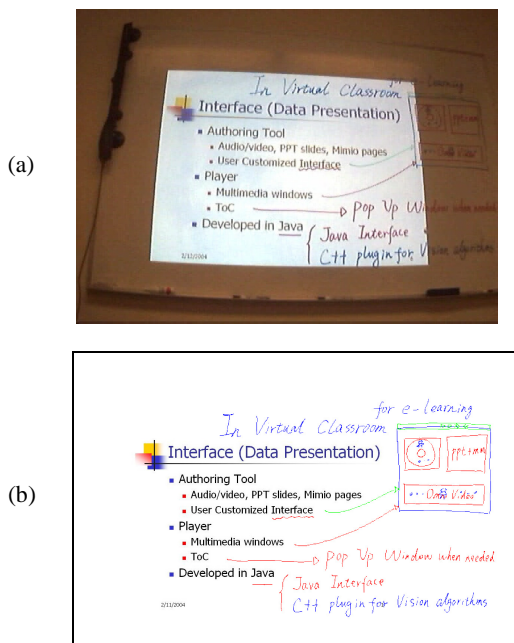


Fig. 11. Another slide (a) the video frame with both PPT and handwriting contents (b) PPT-whiteboard registration result.

8. User-Customized Presentation

Our current implementation of the Virtualized Classroom project includes a Virtualized Classroom Presentation System (VCPS) [4], which is designed as both an authoring tool and a presentation interface for different kinds of lectures, and has a user-selectable interface. The VCPS (developed in Java) includes two parts: the VCPS Creator and Player. The user (an instructor or a student) who uses the system can

customize the presentation by using the VCPS Creator to include different media (video, audio, slide, handwriting pages and/or integrated pages) in windows with user-selected sizes and positions. After the user opens a new Creator page, there will be a floating frame that has checkboxes on it. He/She will use these checkboxes to add or remove presentation components. Each time the user clicks on one of the checkboxes, a popup window will appear directing him/her to click on a certain file so that the program can load the proper information. In the VCPS Player, all the media contents are synchronized by using the timing information (from the PPT log file, the whiteboard html files, and the video/audio timing information). Using a pop-up table of contents (ToC), we can make full use of the space in the player. Fig. 12 shows a snapshot of the VCPS Player interface with windows of a PPT slide, handwriting page, video and integrated slide, and a pop-up ToC.

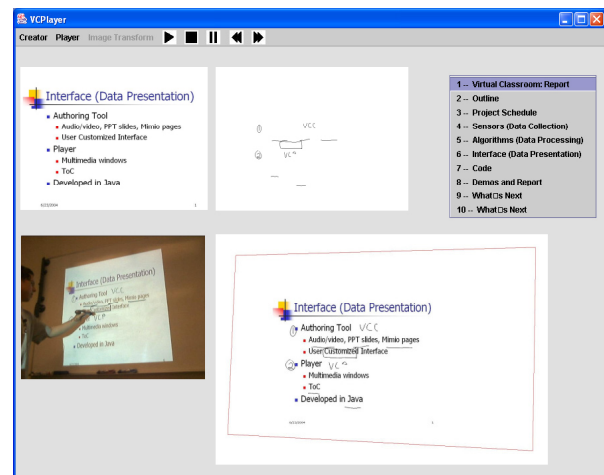


Fig. 12. VCPS Player interface: integrated presentation of video, PPT and whiteboard pages.

9. Conclusions and Discussions

This paper presents a hybrid approach to register images from completely different sources for the purpose of e-lecture production. The two different sources, PPT slide images and the handwriting contents, which do not share the same contents but are presented on the same planar object (the whiteboard) are connected by a video camera that can see both of them. We have implemented several techniques to achieve a fast and robust integration of different media. We use the domain knowledge in the classroom presentation (i.e. the illumination of slide projection, the timing information for both slide changes and handwriting page creation) to align the camera views with the digital slides, and to extract handwriting contents from clustered video scenes. We present a coarse-to-fine content matching method to align

handwriting contents captured by the video camera and the digital whiteboard system. Experimental results are presented to validate our approach.

We realize it is important to have a usability study of our Virtualized Classroom system in real classroom use, including robustness and recoverability from the system failure. We are actively pushing this at both CCNY and UMass-Amherst. Here, we discuss several on-going research issues that enable fast presentation creation, active navigation and natural interaction in the Virtualized Classroom.

(1). Registering slides and instructor images

The position and gestures of the instructor are a very effective way to attract the attention of students and help them to recall what they have learned in the class. For example, the lecturer may point to the items in the digital slides. In the Virtualized Classroom, actual video images of the lecturer will be merged in real time with the digital slides, so that a student will perceive the natural spatial relation between the lecturer and the visual aids. We have done some preliminary research on automatic instructor extraction from the video streams using computer vision techniques to perform content-based video compression and slide-video integration.

(2). Recognizing classroom events

Recognizing important events in classroom scenario could be very useful for both lecturers and students, online or offline. Motion events caused by the movement of either the projector, or the camera or the Mimio capture bar provide clues for our system to re-do the calibration. Detecting the event of an instructor's always staying in front of the slide projection area indicates an occlusion for students, hence a reminder might be friendly delivered to the instructor to step away for a while. The detection of this event could also be useful for the system to capture key frames for slide-board registration. Also, a reminder could be very helpful for an instructor to save the page when an event of erasing white board handwriting is detected. We are trying to model all these events into a Bayesian network framework to provide more efficient e-lectures for both online and offline applications.

10. Acknowledgements

This work is partially supported by NSF under Grant No. CNS-0551598 and Grant No. CNS-0424539, the CUNY Institute for Software Development and Design (CISDD) and the CUNY Graduate Research Technology Initiative (GRTI). We would like to thank Profs. Rick Adrion, Jim Kurose and Edward Riseman at UMass-Amherst for the fruitful discussions and collaboration in this project. Our thanks are also given to Dr. Parviz Kermani for

sharing with us his experience and tools used in his e-learning lecture navigator (ELN) at the Polytechnic University of New York, Mr. Chad McKittrick for developing the VC Presentation System, and Mr. Robert Hill for proofreading an early manuscript, both at the City College of New York.

11. References

- [1] Gregory D. Abowd, Classroom 2000: An Experiment with the Instrumentation of a Living Educational Environment, *IBM Systems Journal*, Special issue on Pervasive Computing, 38(4), 508-530, Oct. 1999
- [2] M. Bianchi, "AutoAuditorium: a Fully Automatic Multi-Camera System to Televisе Auditorium Presentations, *Proceedings of the Joint DARPA/NIST Smart Spaces Workshop*, Gaithersburg, MD, July 1998.
- [3] BMRC Lecture Browser download at <http://bmerc.berkeley.edu/frame/projects/lb/>
- [4] Z. Zhu, C. McKittrick and W. Li, Virtualized Classroom – automated production, media integration and user-customized presentation, *CVPR'04 Workshop on Multimedia Data and Document Engineering (MDDE04)*, July, 2004, Washington DC.
- [5] Kenneth R. Castleman, *Digital Image Processing*, 1996. pp 471–473
- [6] H. Chen, R. Sukthankar, G. Wallace, K. Li. Scalable Alignment of Large-Format Multi-Projector Displays Using Camera Homography Trees. *Proceedings of Visualization*, 2002
- [7] R. B. Dannenberg, "Are Just-In-Time Lectures Effective At Teaching?" <http://www.jitl.cs.cmu.edu/effectiv.htm>
- [8] Z. Zhang, A flexible new technique for camera calibration, *IEEE-PAMI*, 22(11), pp. 1330-1334, 2000.
- [9] W. Xu, C. Wang, CGT: a Fast Thinning Algorithm Implemented On a Sequential Computer. *IEEE trans. Sys, Man, Cyb.* Oct 1987
- [10] Mimio Virtual Ink, <http://www.mimio.com/>
- [11] S. Mukhopadhyay and B. Smith, Passive Capture and Structuring of Lectures, *Proc. ACM Multimedia*, Orlando, FL, October 1999
- [12] PINY ELN: e-Learning Lecture Navigator, <http://westlabp4-7.poly.edu/webpage/index.htm>
- [13] R. Raskar and P. Beardsley, A Self Correcting Projector, In *Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR) 2001*, Hawaii, Dec 2001
- [14] Azrel Rosenfel and Advinash C. Kak, *Digital Picture Processing*. 1981, Vol 2
- [15] M. Stern, J. Steinberg, H.I. Lee, J. Padhye, J. Kurose, MANIC: Multimedia Asynchronous Networked Individualized Courseware, *Proc. Educational Multimedia and Hypermedia*, 1997.
- [16] R. Sukthankar, R. Stockton, M. Mullin. Smarter Presentations: Exploiting Homography in Camera-Projector Systems. *Proceedings of International Conference on Computer Vision*, 2001