

Bringing the Marks on a Whiteboard to Electronic Life

Eric Saund

Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304
saund@parc.xerox.com

Abstract. This paper discusses our implementation and experience with a camera-based whiteboard scanner. The *ZombieBoard* system (so called because it brings to electronic life the marks on a whiteboard) is built into both the physical environment and the information space, while augmenting and linking the two. Computer vision underlies two key technology components. First, *image mosaicing* is used to obtain high-resolution images of large surfaces using relatively low-resolution cameras. Second, real time activity analysis and line drawing analysis enable a *Diagrammatic User Interface* whereby commands are issued to the system by drawing on the whiteboard itself. The system has been in routine use at our research center for two years and has demonstrated the value of this approach to linking whiteboards with the electronic document world.

Keywords. whiteboard scanning, *ZombieBoard*, image mosaicing, Diagrammatic User Interface, digital office.

1 Motivation

Few creative workplaces lack a whiteboard or chalkboard. Whiteboard-scale surfaces afford pacing, gesticulating, sharing with large and small groups, and stepping back to get a look at the big picture. Small office whiteboards support conversations, lists, and notes. Medium size conference room whiteboards participate in presentations and group collaborations. Large whiteboard walls maintain organizational reference material including schedules, timetables, and assignment postings.

For example, Figure 1 shows the latter stages of a meeting in which a group of eight people has used a conference room whiteboard to work out a series of steps required to introduce a new device to the local network. The discussion has raised many issues that filled two whiteboards. To document the deliberations and decisions reached, this information needs to be posted to a web site of meeting minutes.

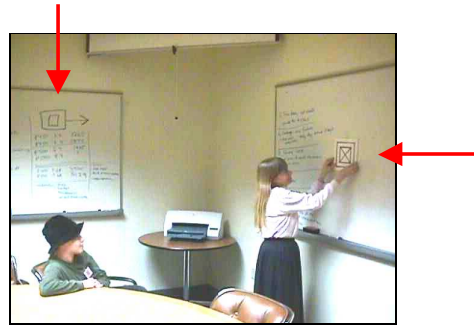


Figure 1. User issuing a “Scan Whiteboard” command.

To facilitate this kind of work process in our research center, the Figure shows the group leader using the system described in this paper to capture the whiteboard contents as an electronic document. A pan/tilt camera in the ceiling has (with the help of some computation) constructed a high-resolution picture of the left whiteboard. The group leader has drawn a symbol (the box with an arrow pointing to the right) that commands the camera to now point at the right whiteboard, and she is in the process of placing a pre-printed command button on this board, which the camera will interpret as the command to next scan this surface. A print of this whiteboard image is delivered a few minutes later by the printer on the table behind her. An electronic image is also placed in a temporary file directory associated with this conference room. Later in the afternoon, when the group leader summarizes the meeting’s accomplishments on a web page, she can type while referencing the whiteboard image printout, or she can directly include the electronic image. Before this technology was introduced, she would write a big “Do Not Erase” message on the whiteboard until she got around to transcribing the contents (often hastily and incompletely). Of course this practice disrupted use of the whiteboard for later meetings.

Several commercial devices exist today for whiteboard image capture (e.g. SMARTBoard, Softboard, Tegrity, Mimeo, VideoBrush). These are dominated by online, or stylus-tracking approaches which provide a record of the time course of strokes added to the whiteboard. Stylus-tracking approaches offer advantages in speed, but often involve specially-fitted whiteboards of limited size, bothersome apparatus and procedures for dealing with the pen, unreliability in detecting when the pen is touching the board, and a requirement to fidget with technology in order to turn the system on, connect it to a computer, etc.

By contrast, optical, or camera-based whiteboard scanning offers to transparently retrofit any existing whiteboard of any size. It provides “what you see is what you get” data acquisition: not only are pen strokes captured, but also posters, documents, and post-it notes---anything on the board. Even people present at the meeting are “photographed” if present in front of the whiteboard surface. For these reasons, we have explored the camera-based approach to instrumenting buildings for whiteboard capture.

This project is part of a larger effort to build office appliances that bring computationally-enabled functionality out into the physical world, but in a “calm” setting (Weiser and Brown, 1996, Black et al, 1998). By mounting a pan/tilt camera in the ceiling, the physical machinery for whiteboard capture is moved out of the way, to a place where it can double as a user interface input device. To obtain sufficient resolution, multiple zoomed-in snapshots are assembled into a larger composite or mosaic image. This is done using a serial port connection to a commercial camera (e.g. a Sony EVI-D30) allowing computer control of pan, tilt, zoom, and focus. Capture and processing of the needed zoomed-in frames is fully automated; a “photograph” of the whiteboard appears on a nearby printer a few minutes after a scan command is issued.

The primary mode for initiating the whiteboard scan function is through a *Diagrammatic User Interface* (DUI). A computer vision system continually monitors activity in front of the whiteboard and watches for users to draw and annotate “buttons” indicating commands and their associated parameters. Whereas a physical button is the most straightforward mechanism for getting a machine to do something, a diagram is often the most effective way of indicating spatial symbolic data. Stafford-Fraser described an initial exploration of this idea (Stafford-Fraser, 1996).

Our system, called “ZombieBoard” (because it brings to electronic life the lifeless ink marks on a whiteboard), has successfully been in routine use at our research center since the spring of 1997, and during that time has delivered over one thousand images to everyday users. Section 2 of the paper provides a brief technical overview of the system and motivations for the design choices made. Section 3 offers observations on the use of the system and its effectiveness in integrating the whiteboard component of the physical and electronic document worlds.

2 System Overview

As a device intended to augment physical space by connecting it to the computational world, a whiteboard scanner system faces the following issues: (1) getting the computational system to perform the desired function; (2) providing a user interface appropriate to the physical setting; and (3) integrating with the computational world.

2.1 Image Capture

The automatic construction of image mosaics from multiple overlapping images has recently become a popular outgrowth of computer vision research (e.g. Szeliski, 1994; Irani, et al, 1995; Capel and Zisserman, 1998). The basic mosaicing problem is to determine image transformation parameters (e.g. pure translation, affine, true perspective) for all component snapshots that will align snapshots' overlapping regions without showing seams.

Existing approaches to this problem can be classified by a handful of properties. ZombieBoard was designed with its specific instrumented whiteboard application in mind. Accordingly, its place in the design space is as follows:

Still frame vs. Video: Several image mosaicing systems take as input video sequences characterized by a large number of image frames possessing large frame-to-frame overlap. To mitigate burden on computing resources, ZombieBoard uses instead a still-frame approach where snapshots overlap one another by about 30%. A typical 72" by 45" whiteboard will use about 18 snapshots.

Signal or Feature: Most image mosaicing systems search for image-to-image transformations that will minimize some pixelwise, or signal-level, cost function. For efficiency of computation in the sparsely-textured domain of whiteboard images, ZombieBoard uses a feature-matching approach whereby salient image features occurring in the whiteboard's contents are automatically detected and used to align overlapping snapshots.

Painting versus Global Alignment: The most straightforward image mosaicing methods treat the destination image as a "canvas". Each source image is "painted" in one at a time by registering with the existing destination image as it has been painted in thus far. This technique is prone to severely distorted resulting images because small errors in registrations tend to build up and compound one another as more frames are added. ZombieBoard is among the approaches that perform "batch" registration whereby all snapshots are aligned with each of its neighbors in an iterative global optimization process.

One dimensional versus two-dimensional mosaics. "Painting" methods work best in constructing one-dimensional mosaics where the set of images is a simple pan of the target scene. ZombieBoard is among the mosaicing methods that assemble composites representing up-and-down motion of the camera as well as right-to-left.

General versus controlled camera positioning: At the time a ZombieBoard installation is created, the pan/tilt camera's location and orientation with respect to the whiteboard is calibrated. This information is used to advantage when the whiteboard is scanned. The calibration information is used to bootstrap an initial "dead-reckoning" estimate of the image transform parameters required to construct a seamless mosaic.

Blending versus color normalization. Most image mosaicing systems hide seams by blending overlapping images. Because ZombieBoard is tailored to whiteboard scenes, we are able to simplify the blending step by applying a color normalization algorithm to each snapshot before combining them into the final mosaic. This color normalization involves segmenting "white" (whiteboard) regions and using these to estimate illumination which is then used to normalize intensities of ink regions.

A result image is shown in Figure 2. We typically provide whiteboard scans at a resolution of 30 dots/inch, but obviously this is controlled by the zoom factors of the snapshot layout.

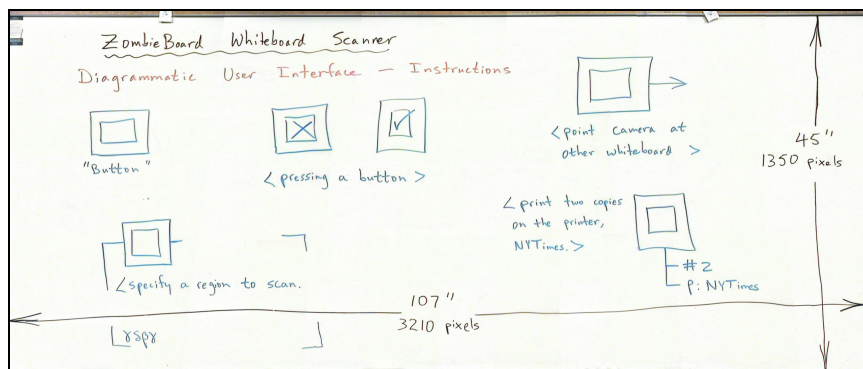


Figure 2. Whiteboard image obtained by automatic mosaicing of 19 snapshots from a computer-controlled pan/tilt video camera. Contents of this whiteboard represent the diagrammatic protocol designed for the Diagrammatic User Interface.

2.2 Diagrammatic User Interface

The simplest user interface to a whiteboard scanner would be a big “Print” button on the wall above a printer where the output would appear. One and only one function would be readily apparent to the user: scan and print the output on that printer. In any sophisticated cooperative environment however, one would want a more complex array of functionality. What if several copies were desired, one for each participant in the meeting? Or how would one specify an arbitrary electronic destination such as an email address or file directory? Could one command that only a subregion of the whiteboard be scanned?

While the simplicity of a big Print button remains attractive, the push-button as a command modality is severely limited. We seek to expand the interaction space and enable an open-ended set of commands.

Alternatives include keyboards, speech, and gesture interfaces. We perceive each as potentially viable but possessing serious impediments as well. Keyboards allow entry of arbitrary text by those who are inclined to type, but they require an accompanying console display that is expensive and ungainly in many whiteboard settings. More seriously, whiteboard work inherently occurs not through intricate finger movements, but at a physical scale of human arm and body movements, often in a social setting. To turn away from the other participants and attend to a keyboard/console interface in order to operate a piece of technology breaks the rhythm and dynamic of a meeting session. Speech recognition has reached commercial viability for some applications, but strongly favors high quality acoustic input which is difficult to achieve in an average whiteboard setting. Gestural input

with a stylus is appropriate for whiteboard-scale interactive display surfaces, but not for the lower-tech ordinary whiteboard.

The ZombieBoard project has chosen to explore the notion of a Diagrammatic User Interface for several reasons. First, a diagrammatic interface befits the medium. Diagrams or drawings are one of the principal representation types people put on whiteboards. Second, a diagrammatic interface presents a calm mode of interaction. The technology is hidden and unobtrusive to the user. A statically available drawing can be created at the user’s own pace, and edited to their satisfaction. Third, diagrammatic interfaces can be expressive and flexible. Whiteboard marks can represent both symbolic information and spatial references. Finally, when a camera is already present as with an optical whiteboard scanner, a diagrammatic UI leverages the existing imaging and computing infrastructure.

The space of Diagrammatic User Interface command and interaction protocols is vast and ripe for exploration. The underlying technology of line drawing analysis is immature, especially where hand-drawn diagrams are concerned, so for any given application a protocol must be designed under constraints of both the needs and accessibility to users and the algorithmic capacities of machine vision systems.

Figure 2 shows the diagrammatic command conventions we have chosen; of course others are possible. The ZombieBoard Diagrammatic User Interface (or DUI, pronounced “dew-we”), is designed around the conceptual notion of a “button” which the user draws to gain the system’s attention that a command is being issued. A button consists of a pair of nested squares. This pattern is easy to draw, rare to occur among common whiteboard material in most domains, and relatively easy to recognize by our line drawing analysis module. The button can be “pushed” by drawing an X or check mark inside. This amounts to issuing a “GO” or “SCAN” command.

Also, a button can be annotated to elaborate and parameterize the command. In our testbench prototype, button annotations include: (1) drawing an arrow to cause the camera to point at another whiteboard in the room; (2) encircling a region of the board to scan; (3) indicating symbolic information such as the number of copies to be printed, file directory, fax, and email destinations for the electronic image, and ink colors to be omitted from the output image. Of these, only the first has been included in the deployed system to date.

When not engaged in collecting zoomed-in snapshots of the whiteboard, the camera is zoomed back to view the entire whiteboard area. Detection and interpretation of diagrammatic commands is performed using images captured at this relatively low resolution.

Technically, the ZombieBoard DUI consists of two main functional modules. First a real-time Activity Analysis module filters an image stream to extract subimages that could possibly represent a diagrammatic command. In general, to analyze in detail the markings on a whiteboard is a compute-expensive job, even when the system is looking only for a stereotypical pattern such as the key Nested Box Button. Relatively little of the raw input stream is new material drawn on the whiteboard though; most of the time, most of the input images contain whiteboard material previously seen and

analyzed, or else people engaged in whiteboard work. The system design therefore employs an activity analysis filter whose function is to pass to the next stage images only of newly modified persistent image content exemplified by material newly written on the whiteboard.

The second functional module of the DUI performs line drawing analysis to interpret any visible commands by extracting and analyzing the spatial pattern of whiteboard markings. Due to unconstrained imaging geometry and tremendous variability by people in drawing even simple figures such as the Nested Box Button, line drawing analysis must be extremely tolerant to deviations from the prototypical geometrical shapes. Our approach is based on perceptual grouping (Saund, 1990; Saund and Moran, 1994). The incoming line drawing image is subjected to center-surround filtering, thresholding, and thinning. Curvilinear lines are collected by tracing, and perceptually salient corners are found by a multiscale corner detection algorithm (see Saund, 1993). The result is a set of primitive curve element tokens representing relatively straight curvilinear contour segments. These tokens in turn undergo a series of grouping operations designed to make explicit spatial structure such as extended curvilinear arcs, corners, parallels, nested corners, and finally, the nested box.

In operation, when running on a Sparc 20 the DUI normally responds to hand-drawn commands within 5 to 10 seconds. This amount of delay is very significant and understandably annoying to users and we anticipate improving response time through the use of faster computing and frame-grabbing hardware.

2.3 Connection to the Electronic World

Increasingly, knowledge work is done online in the context of electronic document representations which may or may not exist on paper. Hardcopy prints of a whiteboard's contents are extremely useful because they may be carried away, copied, filed, and so forth. But it is equally important to provide access to whiteboard documents from the online world. We therefore provide a web-based user interface to ZombieBoard in addition to the Diagrammatic User Interface. A few clicks at a web browser takes the user to a page from which they can select among the nineteen or so ZombieBoard installations in the building.

An installation's web page provides two basic functions. Users can perform a whiteboard scan remotely, and they can access a gallery of images previously scanned in that conference room or office.

Images in the gallery can be cropped, copied to files, and sent to printers. At present, conversion from bitmap form to digital ink that be edited by sketch editing programs such as Tivoli (Pedersen, et al, 1993) is something we have technology for but not as yet attached to the deployed ZombieBoard system.

The ZombieBoard system is built with a client/server architecture using the ILU distributed object system (ILU, 1991). Whiteboard scanning command and image processing operations are published services which could be accessed from other networked devices such as laptop computers and PDAs, although this capability has not yet been exploited.

3 Use and Effectiveness

Deployment: ZombieBoard has been deployed in our research center among approximately 150 scientists and support personnel for over two years. In that time the number of installations has grown to twelve in conference rooms, four in group working spaces, one in an open area, and two in private offices, for a current total of nineteen. The system is used on average 10-20 times per week, sometimes at the conclusion of a meeting, sometimes several times during a meeting or work session. Many of these uses involve multiple scans.

Use: Several groups conduct weekly meetings or study groups in which the whiteboard is the focal point of the work and ZombieBoard scanning is performed religiously. Some of the most ardent users are not researchers but technical support personnel whose whiteboard work includes planning meetings, to-do assignments, and schedules that must be consulted and distributed for days and weeks after the meeting. Two of the conference rooms have very large whiteboards that on occasion get entirely filled with material. Our observation is that the system is used more rarely in these rooms, but its value when needed is proportionately greater.

Electronic images: A subset of users whose document work practices occur primarily online do not use the printed output at all, but rely on the electronic image of their scan that is automatically stored in a public file directory. Images are stored in jpeg format and typically consume 300 KBytes of memory to store the full-resolution color image plus two thumbnail images for browsing. On account of file space, older scans in the public file space are expunged after four months. Some groups maintain online shared repositories to organize and provide access to their project's documents. In these cases, ZombieBoard scans are typically copied from the shared public directory to the group's own file directories where they can be kept indefinitely.

Privacy: To assure privacy for meetings held behind closed doors, every ZombieBoard installed in a conference room is equipped with a simple pull-down shade that blocks the whiteboard from camera view. These are indeed used on occasion, indicating that the presence of a camera can raise people's awareness and privacy concerns.

User interface: Feedback to the user is an important component of any user interface, and the deployed ZombieBoard does not as yet adequately address this issue. After issuing a diagrammatic "GO" command by drawing a button, users can tell that a whiteboard scan has begun by observing the camera panning and tilting. But to divert

one's attention to notice this camera activity is distracting and inappropriate in a meeting situation. Furthermore, the 5-10 second delay in response while the DUI processes the image is so slow as to be disruptive. The delay problem can be eliminated through improvements to the algorithm and faster computers, but a better feedback mechanism is required to indicate the system's status.

In some installations, we have therefore experimented with audio feedback in the form of audio icons and background sounds played at an unobtrusive volume level. Sounds indicate that the command has been recognized, that image snapshots are being collected (and therefore people may want to stand clear of the board), when snapshot collection is complete, and finally when processing is done and the image has been sent to the printer.

Another issue arises with the mechanics of drawing a "GO" button. Due to the imaging conditions, reliable line drawing recognition depends on dark marks. Half-dried out markers under poor lighting conditions are virtually invisible to the DUI. Users, who are unfamiliar with the fact that their eyes are a lot better than the cameras', rapidly but justifiably become impatient when the system fails to recognize a weakly-drawn button. For this reason, we provide a card pre-printed with the "GO" symbol that sticks magnetically to a metal whiteboard. Slapping this button on the board is truly as easy and in most cases nearly as reliable as pushing a physical button.

Reliability: In our experience, the principle form of system-level failure occurs when, after a meeting, a user goes to the printer room and does not find his or her printout *not* because the whiteboard was not scanned, but because the printer is jammed or out of paper. Another failure mode occurs when users from across the building know how to use ZombieBoard but don't know where the printer is to find the output. This information is printed on an instruction sheet posted on the wall, but few users are prone to read instructions. No matter what the cause, *any* form of failure reduces users' confidence that their whiteboard work will be saved and they can safely erase the board. For this reason, we have recently begun deploying low-cost inkjet printers in each conference room equipped with a ZombieBoard so users will have their output on the spot.

4 Conclusion

The notion of Ubiquitous Computing opens a vista of alternative visions for augmented environments that support individual and group work. In this spirit, the ZombieBoard whiteboard scanner places cameras unobtrusively in front of whiteboards in order to link these physical document media with the computational world. Two component technologies borrowed from the field of computer vision---high-resolution scanning through image mosaicing, and Diagrammatic User Interfaces---have demonstrated their effectiveness through two years of routine use in a real-user setting. Many system-level and design options present themselves for exploration, and many opportunities remain for improvement. But we believe that this example demonstrates a powerful and realistically viable approach by which

computationally-enhanced cooperative environments are beginning to come into fruition.

Acknowledgements

Many people deserve thanks for their contributions to the ZombieBoard project. Dietmar Aust, Bikram Bakshi, Ron Frederick, David Goldberg, Josh Kesselman, and Dan Lerner contributed to the working programs. For valuable suggestions, feedback and technical support I thank Andy Berlin, Michael Black, Becky Burwell, Todd Cass, David Fleet, Bill Janssen, John Lamping, Jim Mahoney, Tom Moran, David Marimont, Karin Petersen, Ken Pier, Dan Swinehart, and many early adopters who offered helpful encouragement.

References

1. Black, M., Berard, F., Jepson, A., Newman, W. Saund, E., Socher, G, Taylor, M. (1998) The Digital Office: Overview *AAAI Spring Symposium on Intelligent Environments* Stanford University (March, 1998), 1-6.
2. Capel, D., and Zisserman, A. (1998). Automated Mosaicing with Super-resolution and Zoom. (*Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '98)*). Santa Barbara, CA (June 23-25). 885-891.
3. ILU: Inter-Language Unification (1991-1999) <http://www.parc.xerox.com/pub/ilu/ilu.html>
4. Irani, M., Anandan, P., Hsu, S. (1995). Mosaic-Based Representations of Video Sequences and Their Applications. *Proc. 5th Int. Conference on Computer Vision*, Cambridge, MA (June 20-23, 1995). IEEE Press. 605-611.
5. Mimeo, *Virtual Ink, Inc.* <http://www.virtual-ink.com>.
6. Pedersen, E., McCall, K., Moran, T., Halasz, F. (1993) Tivoli: An Electronic Whiteboard for Informal Workgroup Meetings. *Proceedings of the InterCHI93 Conference on Human Factors in Computer Systems*. ACM, New York.
7. Saund, E. (1990). Symbolic Construction of a 2D Scale-Space Image. *IEEE TPAMI*, 12:8, 817-830.
8. Saund, E. (1993). Identifying Salient Circular Arcs on Curves. *CVGIP: Image Understanding*, 58:3, 327-337.
9. Saund, E., and Moran, T. (1994). A Perceptually-Supported Sketch Editor. *Proc. ACM Symposium on User Interface and Software Technology (UIST '94)*. 175-184.
10. SMARTBoard, *SMART Technologies, Inc.* <http://www.smarttech.com>
11. Softboard, *Microfield Graphics*, <http://www.micg.com>
12. Stafford-Fraser, Q. (1996). BrightBoard: A Video-Augmented Environment, *Proc. ACM Conf. on Human-Computer Interaction (CHI '96)*.
13. Szeliski, R. (1994). Image Mosaicing for Tele-Reality Applications. *Second IEEE Workshop on Applications of Computer Vision (WACV '94)*. Sarasota, FL.
14. Tegrity, Inc. <http://www.tegrity.com>.

15. VideoBrush, *PictureWorks Technology, Inc.* <http://www.pictureworks.com>
16. Weiser, M., and Brown, J.S. (1996). Designing Calm Technology, *PowerGrid Journal* v. 1.01. <http://www.powergrid.electricity.com/1.01>.