



Dynamic Shadow Elimination for Multi-Projector Displays

Rahul Sukthankar Tat-Jen Cham Gita Sukthankar

Cambridge Research Laboratory

Technical Report Series

CRL 2001/4

June 2001

Dynamic Shadow Elimination for Multi-Projector Displays

Rahul Sukthankar

Tat-Jen Cham

Gita Sukthankar

June 2001

Abstract

A major problem with interactive displays based on front-projection is that users cast undesirable shadows on the display surface. This situation is only partially-addressed by mounting a single projector at an extreme angle and pre-warping the projected image to undo keystone distortions. This paper demonstrates that shadows can be muted by redundantly-illuminating the display surface using multiple projectors, all mounted at different locations. However, this technique does not eliminate shadows: multiple projectors create multiple dark regions on the surface (penumbral occlusions). We solve the problem by using cameras to automatically identify occlusions as they occur and dynamically adjust each projector's output so that additional light is projected onto each partially-occluded patch. The system is self-calibrating: relevant homographies relating projectors, cameras and the display surface are recovered by observing the distortions induced in projected calibration patterns. The resulting redundantly-projected display retains the high image quality of a single-projector system while dynamically correcting for all penumbral occlusions. Our initial two-projector implementation operates at 3 Hz.

©Compaq Computer Corporation, 2001

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of the Cambridge Research Laboratory of Compaq Computer Corporation in Cambridge, Massachusetts; an acknowledgment of the authors and individual contributors to the work; and all applicable portions of the copyright notice. Copying, reproducing, or republishing for any other purpose shall require a license with payment of fee to the Cambridge Research Laboratory. All rights reserved.

CRL Technical reports are available on the CRL's web page at
<http://crl.research.compaq.com>.

Compaq Computer Corporation
Cambridge Research Laboratory
One Cambridge Center
Cambridge, Massachusetts 02142 USA

1. Introduction

Camera-projector systems have recently been employed for a variety of applications including the construction of seamless multi-projector video walls [3, 5, 8], real-time range scanning [2] and immersive 3-D virtual environment generation [6]. These systems typically exploit camera-projector synergy in two ways: (1) by casting light into the environment to observe the resulting change in scene appearance; (2) by modifying the projected images in response to camera input. For instance, a video wall can calibrate off-center projectors by projecting a known test-pattern from each projector, observing the keystone distortions in the pattern (using the camera), and correcting the distortion by appropriately pre-warping the image sent to each projector — so that each projected image appears undistorted on the display surface. In a standard video wall, the projectors are mounted so that the projection areas of adjacent projectors overlap slightly. Unsightly seams in the overlap zone are eliminated by precise projector registration and appropriate blending of the contributions from adjacent projectors.

This paper introduces a new application for camera-projector systems where multiple projectors are used not to increase the display size, but to generate redundant illumination over the display surface. A multi-projector display with shadow elimination could provide a good alternative to expensive rear-projection systems. As shown in Figure 1, the projectors are placed at extreme angles but oriented so that their projection areas overlap significantly. By appropriately pre-warping the images sent to each projector, the system generates a sharp, keystone-corrected image in the overlap zone. Redundant illumination makes the display resistant to occlusions: the content in a partially-occluded region is readable as long as one projector maintains an unblocked light path. Unfortunately, the occlusion still causes a shadow in that region (visible as a darker patch). This paper describes a system that automatically detects and dynamically eliminates these shadows so that the display surface appears shadow-free even in the presence of multiple, moving occluders (see Figure 3).

2. System Overview

The technical challenges involved in building a shadow elimination system for multi-projector displays fall into the following three categories. First, the system must accurately align the projected images on the display surface; poor image alignment creates double-images and keystoneing. Second, each occluder can create multiple shadows on the display surface. Shadows where all of the light sources are occluded are termed *umbral*, and those where at least one light source is unoccluded are termed *penumbral*. By definition, the system cannot control lighting within an umbra so we strive to avoid umbral occlusions by positioning projectors so that the display is illuminated from several different directions. All penumbral occlusions can be corrected by projecting additional light into the region. Obviously, this light must come from one of the non-occluded projectors, and the system must determine how to correct each penumbral occlusion. Third, the system must precisely adjust projector output to compensate for each occlusion. If too little light is added, the shadows will remain visible; if too much light is used, over-illumination artifacts will be created. The shadow boundaries must be treated carefully since humans are very sensitive to edge artifacts. As the occluders move, the system should respond swiftly to changes in the shadowed region. The following subsections describe the system and present our solutions to these issues.

2.1. Automatic Multi-Projector Alignment

As shown in Figure 1, several projectors are placed so that their projection areas all converge onto a single display surface S . The goal is to combine the light from the projectors to create a single, sharp image on S . Clearly, one cannot simply project the same raw image simultaneously through the different projectors; not only does a given point on S correspond to very different pixel locations in each projector, but the image produced on S from any single projector will be distorted (since the projectors are off-center to S).

We assume that: the positions, orientations and optical parameters of the camera and projectors are unknown; camera and projector optics can be modeled by perspective transforms; the projection screen is flat. Therefore, the various transforms between camera, screen and projectors can all be modeled as 2-D planar homographies:

$$\begin{pmatrix} xw \\ yw \\ w \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 \\ p_7 & p_8 & p_9 \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix},$$

where (x, y) and (X, Y) are corresponding points in two frames of reference, and $\vec{p} = (p_1 \dots p_9)^T$ (constrained by $|\vec{p}| = 1$) are the parameters specifying the homography. These parameters can be determined from as few as four point correspondences, using the technique described in [7].

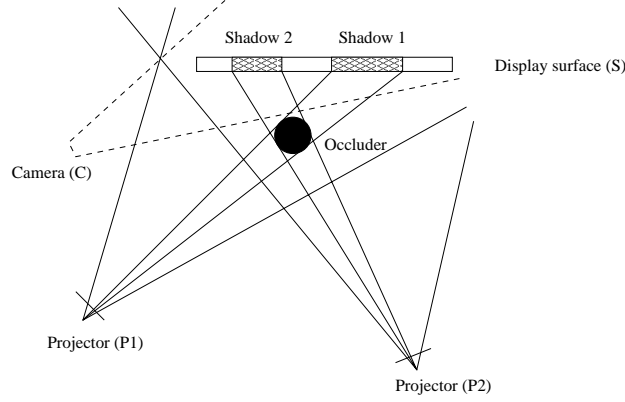


Figure 1: An overhead view of the multi-projector display system. Several projectors (P_1 , P_2) are placed such that their projection areas converge onto the display surface (S). A camera (C) is positioned so that S is clearly visible in its field of view. Once the homographies relating P_i , C , S are automatically determined, the projectors combine their pre-warped outputs to create a single high-quality image on S . Each occluder in the environment may cast up to one shadow per projector on the display surface. The shadow elimination system locates these penumbral occlusions in the camera image and dynamically adjusts the projected images to make the shadows disappear.

The homography for each camera-projector pair T_{c,P_i} can be determined by projecting a rectangle from the given projector into the environment. The coordinates of the rectangle's corners in projector coordinates (x_i, y_i) are known *a priori*, and the coordinates of the corners in the camera frame (X_i, Y_i) are located using standard image processing techniques.¹

The display area is either automatically determined using the camera, or interactively specified by the user. The former case requires the display surface to be a white, rectangular projection screen against a contrasting background. Such a screen shows up clearly in the camera image as a bright quadrilateral, and can be unambiguously identified by the automatic calibration process. The corners of this quadrilateral fully specify the homography between camera and screen $T_{c,s}$, and the homography between each projector and the screen $T_{P_i,s}$ can be recovered using the equation:

$$T_{P_i,s} = T_{c,P_i}^{-1} T_{c,s},$$

where the homographies on the right hand side of the equation are all known.

Alternatively, the user can interactively specify the display area by manipulating the outline of a projected quadrilateral until it appears as a rectangle of the desired size on the display surface. This directly specifies the homography between the selected projector and the screen $T_{P_i,s}$; the outline of the selected rectangle can then be detected in the camera image as discussed above to determine the camera to screen homography $T_{c,s}$.

The projector-screen homographies $T_{P_i,s}$ model the geometric distortion (keystone warping) that is induced when an image is projected from an off-center projector P_i . This distortion can be corrected by projecting a *pre-warped* image, generated by applying the inverse transform $T_{P_i,s}^{-1}$ to the original image.² Since $T_{P_i,s}^{-1} T_{P_i,s} = I$, one can see that the pre-warping also aligns the images from different projectors so that all are precisely projected onto S . Applying the homographies derived from camera images, a multi-projector array can thus be efficiently configured to eliminate keystone distortions and double images on the display surface.

2.2. Occlusion detection

Rather than locating occluders by tracking objects in the environment, the system focuses exclusively on detecting artifacts on the display surface. These can occur for either of two reasons. First, uncorrected penumbral occlusions appear as darker regions in a camera image that can be corrected by projecting additional light into the region. Second, artifacts may be caused by over-illumination of the display area, and occur most often when an occluding object (whose shadows had been eliminated) moves away suddenly. These bright spots are corrected by reducing the light intensity in the region. Our algorithm makes no assumptions are made about the locations, sizes or shapes of occluders.

¹Hough-transform line-fitting [1] locates the edges of the quadrilateral, and its corner coordinates are given by intersecting these lines.

²In our current system, this pre-warp is efficiently implemented using the texture-mapping operations available in standard 3-D graphics hardware.

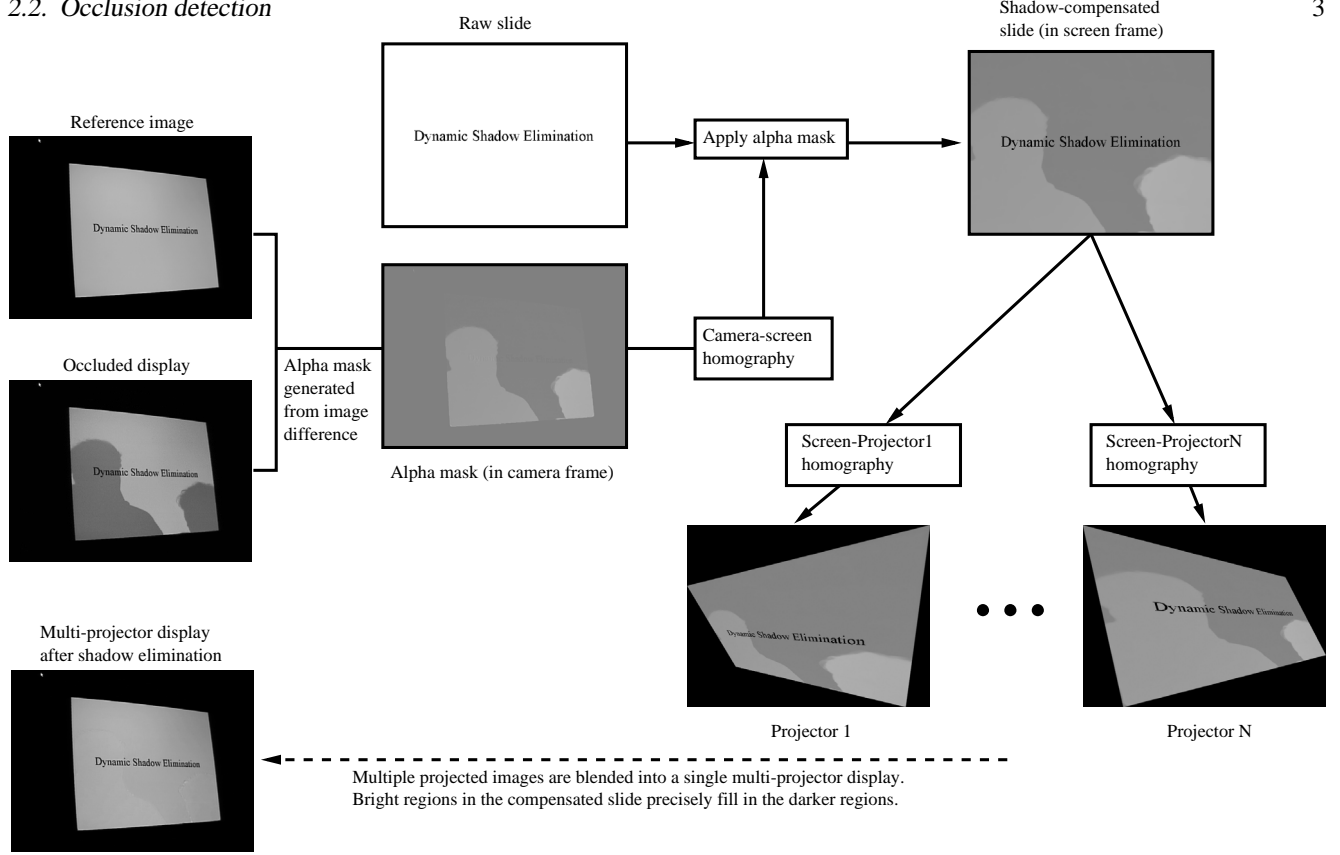


Figure 2: This diagram summarizes the occlusion detection and shadow elimination algorithms. The images in the left column were taken by the system camera during operation. The two penumbral occlusions caused by the person blocking both projectors are identified and corrected to create a shadow-free display (bottom left). See text for details.

Figure 2 illustrates the algorithm. During its initialization phase (when the scene is occluder-free) the system projects each presentation slide and captures several camera images of the projected display. These images are pixel-wise averaged to create a reference image for that slide, and this image represents the desired state of the display (Figure 2, top left). The goal of occlusion detection is to identify regions in the current image that deviate from this ideal state. During operation, the system camera acquires a current image of the projected display which may contain uncorrected shadows. For example, the image shown in Figure 2 (center left) has two dark regions, corresponding to the two penumbrae cast by one person standing in front of the display (each projector creates one shadow).

Since the display surface remains static, a pixel-wise image difference between current and reference camera images can be used to locate shadow and over-compensation artifacts. To reduce the effects of camera noise, we apply a 5×5 spatial median filter to the difference image. A negative value in a difference image pixel means that the corresponding patch on the screen was under-illuminated in the current image. This information is represented in terms of an alpha mask (α_t), which when applied to the current camera image, should bring it closer to the reference image. Alpha values range from 0 (dark) to 255 (bright), and the mask is initialized to 128 at $t=0$. The use of the alpha mask is described in the next section. The alpha mask is updated at every time-step using the following simple feedback system:

$$\alpha_t(x, y) = \alpha_{t-1}(x, y) - \gamma (I_t(x, y) - I_0(x, y)),$$

where I_t is the camera image at time t , I_0 is the reference image, and γ is a system parameter (set to 0.25 in our implementation). For a static scene, the alpha mask converges to a stable fixed point in a very short period of time.

A noteworthy point about our system is that all of the projectors in the multi-projector system use the *same* alpha mask for shadow removal, as described in the following section.

2.3. Shadow elimination

The alpha mask (described above) integrates the previous state of the shadow correction, and information from the current difference image. However, since it was computed in the camera frame of reference, it must be transformed into the screen frame of reference before it can be applied; this is done using the camera-screen homography $T_{c,s}$ discussed in Section 2.1.

Applying the alpha mask to the current slide is straightforward: the warped alpha mask simply replaces the alpha channel of the slide image. The slide is then pre-warped for each projector (using its particular screen-to-projector homography) and displayed.

It is surprising that projecting the *same* slide from all projectors correctly eliminates *all* of the penumbral shadows! This can be explained by the following argument. Consider the two-penumbra shadow configuration generated by the two-projector, one-occluder system shown in Figures 1 and 2 (the result holds for all penumbrae generated by any number and configuration of projectors and occluders). From P1's perspective, the left high-alpha region falls precisely on the left penumbra (Shadow2) while the right high-alpha region harmlessly over-illuminates the occluder. From P2's perspective, the left high-alpha region falls on the occluder (without effect) and the right one corrects for the right penumbra (Shadow1). Thus, both projectors are able to use the same slide to eliminate shadows.

Since we do not have a good photometric model of the environment, we cannot predict precisely how much light is needed to remove a shadow. However, the iterative feedback loop used to update the alpha mask allows us to avoid this problem: the system will continue adding light to shadowed regions until the region appears as it did in the reference image. The drawback to such an iterative technique is that the alpha mask can require several iterations to converge; in practice, shadows are eliminated in approximately 3 iterations.

3. Results

The experiments described in this section are based on the following implementation. Images are acquired using an NTSC camera (640×480) attached to a PCI digitizer; the output subsystem consists of two Compaq MP-2800 XGA-resolution (1024 × 768) DLP microportable projectors driven by a dual-headed graphics card; the software runs on a standard workstation, and utilizes the MMX-enhanced routines from the Intel Open Source Computer Vision Library [4] for efficient image processing. The projectors are positioned on either side of a whiteboard and the camera is mounted at an extreme angle to ensure occlusion-free coverage of the display surface. The location, orientation and optical parameters of the camera and projectors are unknown.

We evaluated the shadow elimination system on three criteria: (1) image quality; (2) responsiveness; (3) ease of use. These experiments are discussed below.

3.1 Image quality

The primary goal of the multi-projector display system is to project high-quality images that remain shadow-free in the presence of occluders. In multi-projector display blending, any inaccuracies in the homography estimation lead to misalignment in projected images creating glaring artifacts such as double-images. Figure 3 shows views of the same slide projected on single- and multi-projector displays; the multi-projector display is as sharp as the single-projector display and the effect of shadows is dramatically reduced. With dynamic shadow elimination, (Figure 3, right) the remaining penumbral occlusions completely disappear.

These qualitative results were confirmed by experiments comparing our shadow elimination system's performance with baseline one- and two-projector systems. A series of slides were displayed using the three systems, first without occluders and then with static occluders in the scene. Images of the display region were captured using a tripod-mounted camera. We computed a sum-squared-difference (SSD) of the grey-level intensities (over the pixels corresponding to the display region in the camera image) for each slide pair (with and without occluders). Table 1 summarizes these results. As expected, the (umbra) shadows in the single-projector display are the major source of error. We see a significant improvement with two projectors since the majority of occlusions become penumbral. Finally, the best results are achieved when these penumbral occlusions are actively eliminated.

3.2 Responsiveness

Since our shadow elimination algorithm uses an iterative technique, we wanted to confirm that the method converges quickly to a stable value without excessive oscillation or overshoot. Figure 4 shows a trace of the SSD between the current and

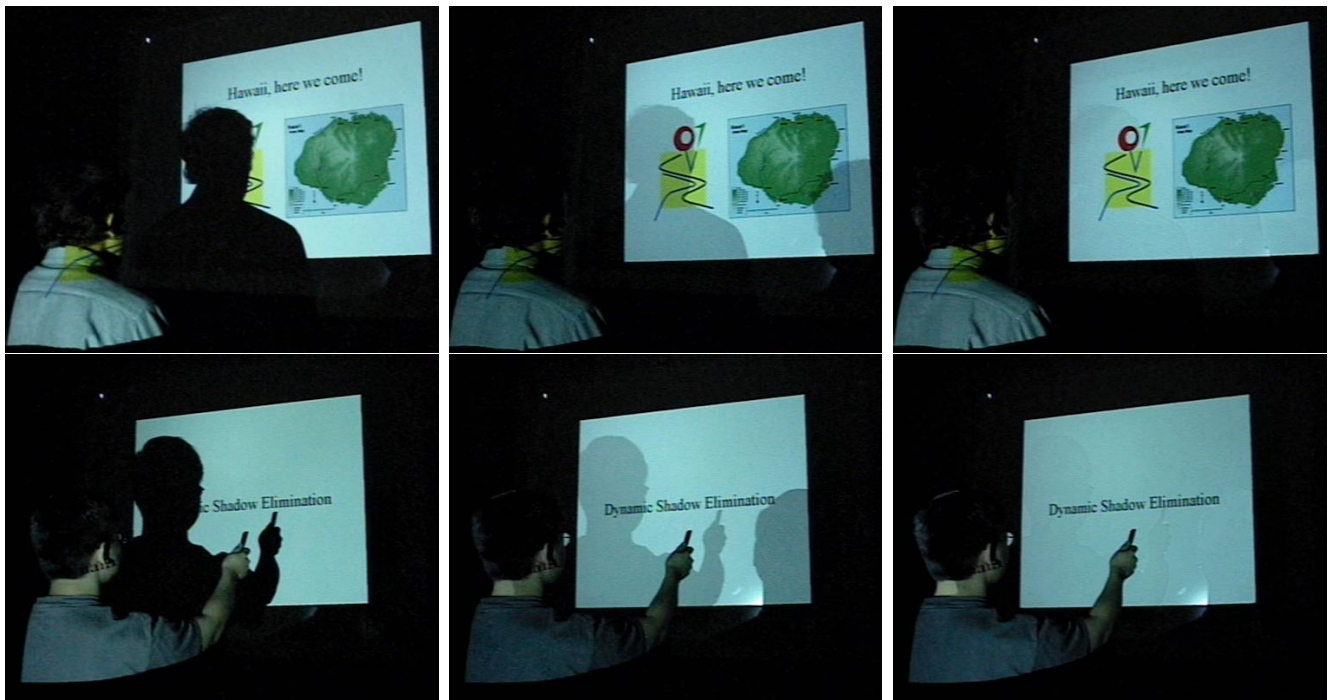


Figure 3: Comparison between different projection systems. Left: in a single projector display, a single occluder creates a sharp shadow on the display. Center: in this multi-projector display, the outputs from two off-center projectors are blended over the entire surface to create an equally sharp image; the occluder now creates two darker regions rather than a single shadow. This is an improvement over the single-projector system since the information in the occluded regions (e.g., the CVPR logo) is now visible, albeit faintly. Right: the shadow elimination system locates these blemishes and removes them by boosting the light intensity in those regions. The resulting shadow-free display looks as good as the original image.

Table 1: SSD error between an unoccluded reference image and the same scene with a person occluding both projectors. A two-projector display markedly improves the display by muting the shadows. The shadow-elimination system removes the remaining artifacts and results in an almost-perfect image (see Figure 3).

	SSD error over entire display
One projector	381.19×10^6
Two projectors	56.15×10^6
Shadow elimination	21.19×10^6

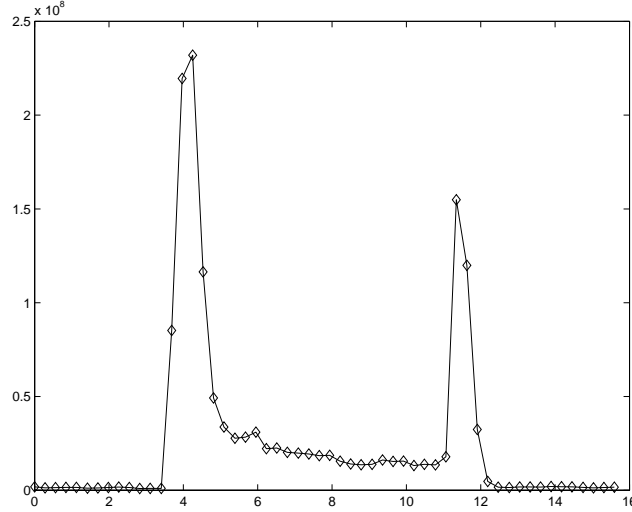


Figure 4: A trace of the SSD error (difference between current and reference camera images over the display region) during the operation of the shadow elimination system. The occluder is introduced at $t=4$ and abruptly removed at $t=11.5$. The shadow elimination system restores the observed display intensities in three iterations. The SSD error is non-zero while the occluder is moving due to system response lag.

reference images during one such experiment. As expected, there is no difference between reference and current images until $t=3.5$, when an occluder is suddenly introduced into the scene, at which point, the SSD spikes up. Once the occluder slows down³ ($t=4$), the shadow-elimination algorithm is able to reduce the SSD error near its original levels. The remaining error is mainly due to “halo” effects (the occluder’s leading edge creates a shadow while trailing edge creates an inverse-shadow). At $t=11$, the occluder abruptly leaves the scene, momentarily leaving inverse-shadows on the areas of the display that were previously occluded. However, within three iterations, the system removes these artifacts and returns to its unoccluded state.

A benefit of using a feedback system for shadow elimination (as opposed to a photometric model approach) is that the system is surprisingly robust. For instance, if one of the projectors in the multi-projector array fails, the alpha mask will uniformly increase, and the remaining projectors will brighten their images to compensate. Furthermore, the overall brightness of the entire multi-projector array can be changed simply by adjusting the camera aperture.

3.3 Ease of use

A major concern with multi-projector displays is that they typically require tedious manual calibration and alignment. Ease of use was of particular interest to us because we wanted to exploit the portability of lightweight (3 lb) projectors with the goal of deploying the system at a presentation site within minutes.

The system offers two options for vision-based self-calibration (in addition to the backup manual alignment). The first is fully-automatic: when the display surface can be unambiguously identified in the camera image, the system automatically derives the projector-screen homography. Projector-camera homographies are always derived automatically (by projecting a known test-pattern into the environment). Thus, in good conditions, the multi-projector display system can calibrate itself without human intervention.

The second self-calibration mode is semi-automatic, and is employed when the user wishes to interactively specify a particular display region. The system observes the user’s selection in the camera image and automatically adjusts projector homographies so that the image appears in the desired region.

Either calibration mode is significantly faster than manually specifying the transforms needed to align multiple projectors.

³The current implementation of our system updates at 3 Hz, and cannot correct for fast-moving occluders.

4. Related Work

7

The authors are not aware of any prior work in active shadow elimination for projected displays. Research in the area of camera-assisted multi-projector displays is becoming more popular, particularly in the context of seamless video walls [3, 5, 6, 8]. However, none of these have explored the idea of using multiple projectors to create redundantly-illuminated displays. A simple feedback system, similar to the one presented here, was used by [8] to adjust projector illumination for uniform blending in the overlap region of a video wall. The self-calibration techniques used in this paper were adopted from [7], where they were applied to the task of automatic keystone correction for single projector systems.

5. Conclusion

We have developed a practical solution to the problem of shadows on front-projection systems. By redundantly-illuminating the display area with multiple self-calibrating projectors, the majority of occlusions become penumbral. These can be actively corrected by adjusting the output of the projectors to create a high-quality, shadow-free display.

We are extending the system in the following areas. First, we are exploring shadow detection as an interface medium that enables users to interact with a projected display as if it were a touch screen. Second, we are developing a new shadow-elimination algorithm that removes light from occluding objects in addition to removing shadows. This would enable a presenter to stand in front of a multi-projector display without being blinded by any of the projectors. Finally, we are optimizing our system so that it will run at frame rate. Such an implementation would provide an affordable, high-quality alternative to expensive rear-projection display systems.

Acknowledgments

The authors would like to thank Jim Rehg, David Hsu, T. M. Murali and Mike Jones for their valuable contributions.

References

- [1] D. Ballard and C. Brown. *Computer Vision*. Prentice-Hall, 1982.
- [2] O. Hall-Holt and S. Rusinkiewicz. Stripe boundary codes for real-time structured light range scanning of moving objects. In *Proceedings of ICCV*, 2001.
- [3] M. Hereld, I. Judson, and R. Stevens. Introduction to building projection-based tiled displays. *Computer Graphics and Applications*, 20(4), 2000.
- [4] Intel Corporation. Open source computer vision library. <<http://www.intel.com/research/mrl/research/opencv/>>.
- [5] K. Li, H. Chen, Y. Chen, D. Clark, P. Cook, S. Daminakis, G. Essl, A. Finkelstein, T. Funkhouser, A. Klein, Z. Liu, E. Praun, R. Samanta, B. Shedd, J. Singh, G. Tzanetakis, and J. Zheng. Building and using a scalable display wall system. *Computer Graphics and Applications*, 20(4), 2000.
- [6] R. Raskar, M. Brown, R. Yang, W. Chen, G. Welch, H. Towles, B. Seales, and H. Fuchs. Multi-projector displays using camera-based registration. In *Proceedings of IEEE Visualization*, 1999.
- [7] R. Sukthankar, R. Stockton, and M. Mullin. Smarter presentations: Exploiting homography in camera-projector systems. In *Proceedings of ICCV*, 2001.
- [8] R. Surati. *A Scalable Self-Calibrating Technology for Seamless Large-Scale Displays*. PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 1999.

