Projected Light Displays Using Visual Feedback

James M. Rehg[1], **Matthew Flagg**[1], **Tat-Jen Cham**[2], **Rahul Sukthankar**[3], **Gita Sukthankar**[3]

[1]College of Computing
Georgia Institute of Technology
Atlanta, GA 30332
{rehg,mflagg}@cc.gatech.edu

[2]School of Computer Engineering
Nanyang Technological University
Singapore 639798
astjcham@ntu.edu.sg

[3]Cambridge Research Lab
HP Labs
Cambridge, MA 02142
{Rahul,Gita}.Sukthankar@hp.com

Abstract

A system of coordinated projectors and cameras enables the creation of projected light displays that are robust to environmental disturbances. This paper describes approaches for tackling both geometric and photometric aspects of the problem: (1) the projected image remains stable even when the system components (projector, camera or screen) are moved; (2) the display automatically removes shadows caused by users moving between a projector and the screen, while simultaneously suppressing projected light on the user. The former can be accomplished without knowing the positions of the system components. The latter can be achieved without direct observation of the occluder. We demonstrate that the system responds quickly to environmental disturbances and achieves low steady-state errors.

1 Introduction

The increasing affordability and portability of high quality projectors has generated a surge of interest in projector-camera systems. Recent examples include the construction of seamless multi-projector video walls [1, 2, 3, 4], real-time range scanning [5] and immersive 3-D virtual environment generation [6]. In most of these previous systems, cameras are used to coordinate the aggregation of multiple projectors into a single, large projected display. In constructing a video wall, for example, the geometric alignment and photometric blending of overlapping projector outputs can be accomplished by using a camera to measure the keystone distortions in projected test patterns and then appropriately pre-warping the projected images. The result is a highly scalable display system, in contrast to fixed format displays such as plasma screens.

In addition to their utility in creating wall-sized displays, projector-camera systems can also be used to create ubiquitous, interactive displays using the ordinary visible surfaces in a person's environment. Displays could be conveniently located on tabletops, nearby walls, etc. Users could reposition or resize them using simple hand gestures. Displays could even be "attached" to objects in the environment or be made to follow a user around as desired. This could be particularly compelling as means to augment the output capabilities of handheld devices such as PDAs. In order to realize this vision, two challenging sensing problems must be solved: (1) Determining where and how to create displays based on input from the user. (2) Creating stable displays in the presence of environmental disturbances such as occlusions by the user and changes in ambient light.

In this paper we examine the visual sensing challenges that arise in creating interactive occlusion-free displays using projected light in real-world environments. The first challenge is to allow the user to interactively define the region containing the display, and then automatically calibrate the cameras and projectors to that region. The second challenge is to maintain the stability of the display in the face of environmental disturbances such as occlusions. Specifically, there are two problems that arise in front-projection systems when a user passes between the projectors and the display surface: (1) Shadows cast on the display surface due to the occlusion of one or more projectors by the user, (2) Bright light projected on the user, which is often a source of distraction and discomfort. We present a solution to these two challenges that does not require accurate 3-D localization of projectors, cameras, or occluders, and avoids the need for accurate photometric calibration of the display surface. The key is a display-centric camera feedback loop that rejects disturbances and unmodeled effects.

Our system uses multiple, conventional projectors which are positioned so that their projections overlap on the selected display surface. It produces shadow-free displays even in the presence of multiple, moving occluders. Furthermore, projector light cast on the occluders is suppressed without affecting the quality of the display. The result is shown in Figure 4.

The two classes of problems to be addressed in our system are: (i) geometric, and (ii) photometric. The geometric problems relate to computation of the spatial correspondences between pixels in the projectors and the projected display on the screen. The projectors should be accurately and automatically calibrated to the screen, to the camera and to each other. The calibration should enable the images in each projector to be pre-warped so as to create a desired projected display that is aligned with the screen. It should be possible to control the display area on the screen in real-time. The photometric issues are the accurate and fast computation of the desired pixel intensities in each projector so as to eliminate shadows and suppress occluder illumination. This involves occlusion detection based on camera input and correctly adapting the projector output to achieve the necessary goals. These two classes of problems are addressed in sections 2 and 3 respectively.

2 Autocalibration of Cameras and Projectors

In a multi-projector system, several projectors are positioned so that their outputs converge onto a display surface S (see Figure 2). 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:

$$\left(\begin{array}{c} xw \\ yw \\ w \end{array} \right) = \left(\begin{array}{ccc} p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 \\ p_7 & p_8 & p_9 \end{array} \right) \left(\begin{array}{c} X \\ Y \\ 1 \end{array} \right), \quad (1)$$

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

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]

2.1 Real-Time Calibration

A key issue for the robustness of the projector-camera system is the ability to recalibrate the homographies quickly if either the camera or the projector are moved. In addition, a basic question is how to specify the location of the display. We now describe a real-time calibration system which addresses both of these concerns. The system uses a set of four fiducial marks on a display surface such as a wall or table to define the four corners of the desired projected area. Since walls tend to be light colored, we have found that any small dark target, such as a poker chip, can serve as a fiducial. By positioning the targets appropriately on the display surface, the user can identify the desired display area. Through visual tracking of both the positions of the four markers and the corners of the quadrilateral formed by the projector output, the appropriate transformation can be computed.

[1] Hough-transform line-fitting [8] locates the edges of the quadrilateral, and its corner coordinates are given by intersecting these lines.

The geometric relationship between the detected corners of the projector quadrilateral and the location of the markers determines a homography that aligns the projector output with the markers. The image coordinates of the four markers fully specify the homography between camera and screen $T_{c,s}$. The homography between each projector and the screen $T_{P_i,s}$ can be recovered using the equation:

$$T_{P_i,s} = T_{c,P_i}^{-1} T_{c,s}, \quad (2)$$

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

In some applications the positions of the projector or camera, as well as the positions of the markers, may change over time. We can view each of these changes as disturbances that perturb the calibrated relationship between the cameras, projectors, and display. In this instance, disturbance rejection can be easily accomplished by tracking the quadrilateral corners and marker positions in real-time, and updating the warping parameters appropriately. Note that the dynamics in this case are extremely fast, since the only limit on the speed at which the projector output can be changed is the overall system bandwidth.

2.2 Experimental Results

We performed three experiments to evaluate the ability of the visual feedback loop to compensate for disturbances to the projector and camera positions and the positions of the fiducial markers. Our system can perform tracking and disturbance rejection at 10 Hz.

The first experiment tested the ability of the system to compensate for changes in the location of the markers, resulting in a rescaling of the projected display. Figure 1(a) and (b) shows the result for two different marker configurations. Note the automatic rescaling of the display region, so that its corners remain aligned with the markers. In each image, the boundary of the projected area can be seen as a large pale quadrilateral which contains the display region. In this test the projectors and camera remained stationary, and so the location of the projector quadilateral in the image does not change (i.e. T_{c,P_i} did not change). These images were captured with a second camera located in the audience, which was not used in autocalibration.

In the second experiment, we kept the camera and markers fixed and changed the location of the projector. Small disturbances in the projector orientation can induce large changes in the display. The result is illustrated in Figure 1(c) and (d) for two positions of the projector. As desired, the configuration of the display region on the wall as defined by the fixed markers is unaffected by movement of the projector. Note also that there is no change in the marker positions between frames, as expected (i.e. $T_{c,s}$ did not change).

The third experiment tested the response of the system when the position of the camera was changed and the projector and markers remained fixed. In this situation, there is no change in the homography between the projector and the display (i.e. $T_{P_i,s}$ does not change). However, the image locations of both the marker positions (in $T_{c,s}$) and the
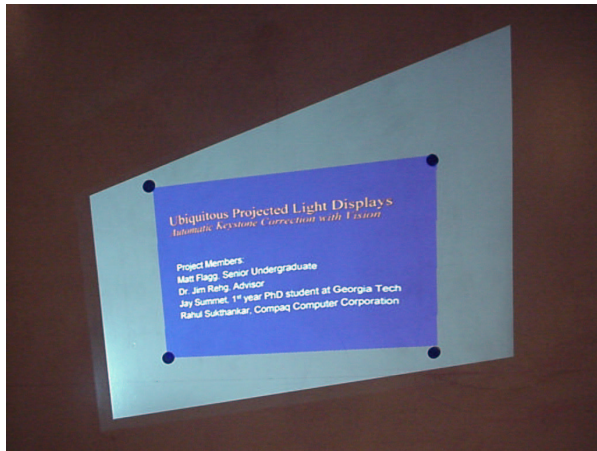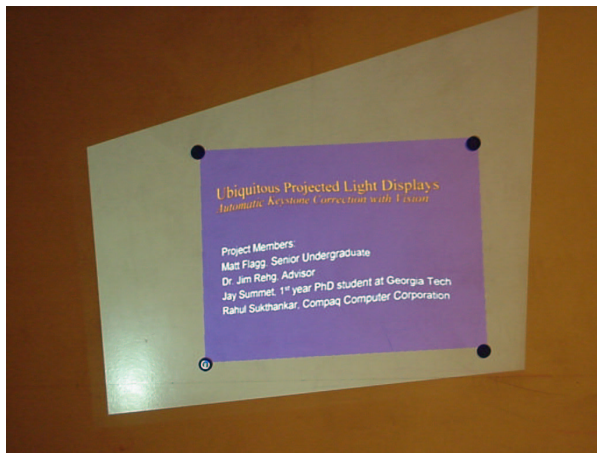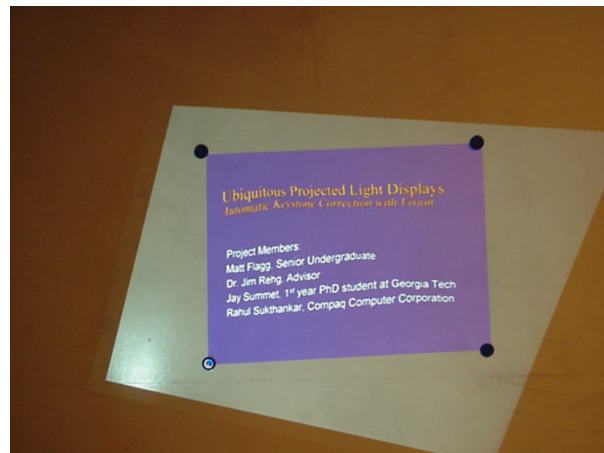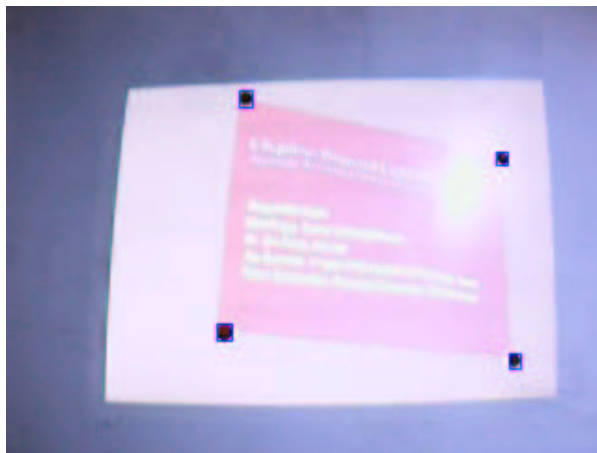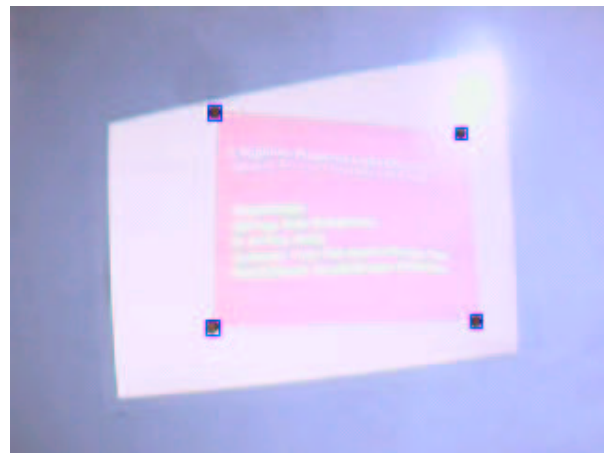
Figure 1. (a) and (b): The effect of a change in the marker configuration on the system is shown at two different time instants. The four markers define an interface by which the user can control the size and location of the display. **(c) and (d):** The effect of a change in the projector position is shown at two different time instants. The projector quadrilateral changes while the display defined by the markers does not. **(e) and (f):** The effect of a change in the camera position is shown at two different time instants. The entire image is distorted, but the display continues to fill the region defined by the markers.

quadrilateral corners (in T_{c,P_i}) will change as the camera moves. Figure 1(d) and (e) illustrates the result for two camera configurations. Once again the display is unaffected, as desired. These images where captured by the camera which is used in autocalibration.

3 Shadow Elimination and Occluder Light Suppression

In this section we describe a system which handles real-time photometric compensation using visual feedback. The system comprises a number of projectors which are aimed at a screen such that their projection regions overlap and a camera which is positioned such that can view the entire screen. During normal functioning, the system displays a high quality, dekeystoned image on the screen. When users walk between the projectors and the screen, shadows are cast on the screen. These shadows can be classified as *umbral* when all projectors are simultaneously occluded, or *penumbral* when at least one projector remains unoccluded. The system eliminates all penumbral shadows cast on the screen,[2] as well as suppressing projector light falling on the occluders. This enables the system to continue presenting a high quality image without projecting distracting light on users. See Figure 2 for the setup.
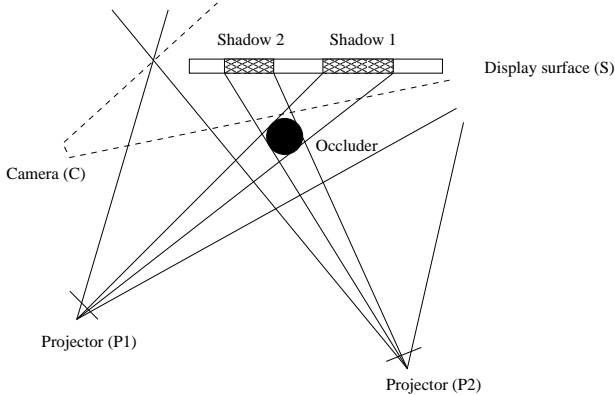

Figure 2. An overhead view of the multi-projector display system. Several projectors (P_1, P_2) are placed such that their projection areas converge onto the display surface (S). A camera (C) is positioned so that S is clearly visible in its field of view. The projectors combine their pre-warped outputs to create a single high-quality image on S based on computed homographies. The system is able to dynamically compensate for penumbral shadows and suppress projected light on occluders.

[2] By definition, pixels in an umbral shadow are blocked from every projector and cannot be removed. Umbral shadows can be minimized by increasing the number of projectors and by mounting the projectors at highly-oblique angles.

3.1 Photometric Framework

After the projectors have been geometrically aligned, we can easily determine which source pixels from the projectors contribute to the intensity of an arbitrary screen pixel. In the following analysis, we assume that the contributions are at some level additive. Given N projectors, the observed intensity Z_t of a particular screen pixel at time t may be expressed by

$$Z_t = C\left(A + \sum_{i=1}^{N} k_{i,t} S_i(I_{i,t})\right), \qquad (3)$$

where I_{jt} is the corresponding source pixel intensity set in projector j at time t, $S_j(\cdot)$ is the projector to screen intensity transfer function, A is the ambient light contribution which is assumed to be time invariant, $C(\cdot)$ is the screen to camera intensity transfer function and k_{jt} is the *visibility ratio* of the source pixel in projector j at time t. See Figure 3.
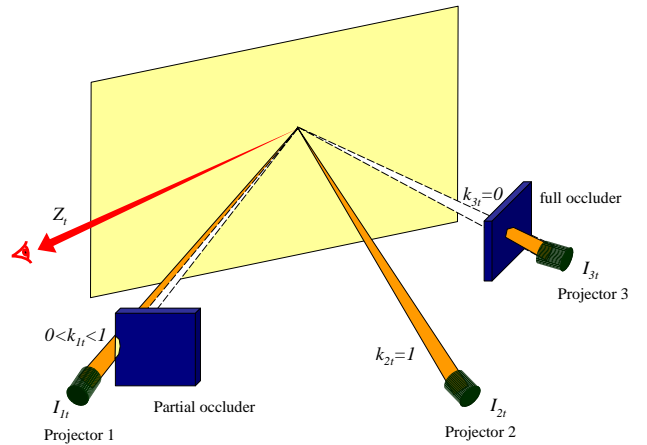

Figure 3. Photometric framework. This diagram illustrates equation (3), in which the observed display intensity Z_t is related to the combination of projector source pixels I_{jt} and the corresponding visibility ratios k_{jt}. The visibility ratios vary accordingly with non-occlusion, partial and full occlusion.

When occluders obstruct the paths of the light rays from some of the projectors to the screen, Z_t diminishes and shadows occur. This situation is quantitatively modeled via the visibility ratios, which represent the proportion of light rays from corresponding source pixels in the projectors that remain unobstructed.

Mathematically, the desired intensity of a particular screen pixel may be represented by Z_0 (obtained in an initialization phase). As an occluder is introduced in front of projector k to create penumbral shadows, the visibility ratio k_{jt} decreases, such that $k_{jt} < 1$. Hence $Z_t < Z_0$. These deviations in the screen can be detected via a pixel-wise image difference between current and reference camera images to locate shadow artifacts.

3.2 Iterative Photometric Compensation

Our system handles occluders by

1. compensating for shadows on the screen by boosting the intensities of unoccluded source pixels; and

2. removing projector light falling on the occluder by blanking the intensities of occluded source pixels.

As in [9], the change in the intensity of each source pixel in each projector is controlled by the alpha value associated with the pixel:

$$I_{jt} = \alpha_{jt} I_0, \qquad (4)$$

where I_0 is the original value of the source pixel (i.e. pixel value in the presentation slide) and is the same across all projectors, while $\alpha_{jt}, 0 < \alpha_{jt} < 1$ is the time-varying, projector-dependent alpha value. The alpha values for the source pixels in one projector is collectively termed the alpha mask for the projector.

Shadows should be eliminated by adjusting the alpha masks for all projectors such that $|Z_t - Z_0|$ is minimized. Additionally, alpha values for occluded source pixels should be set to zero in order to supress projector light falling on the occluder. This can be done iteratively with the aid of the visual feedback from the camera.

3.3 Components of the Visual Feedback Rule

Eliminating shadows involves increasing values for corresponding source pixels. The shadow elimination (SE) component of the system is based on

$$(\Delta\alpha_{jt})_{\text{SE}} = -\gamma(Z_t - Z_0), \qquad (5)$$

where $\Delta\alpha_{jt} = \alpha_{j(t+1)} - \alpha_{jt}$ is change of α_{jt} in the next time-frame, and γ is a proportional constant. This component is a simple proportional control law.

Suppressing the projector light falling on the occluders involves diminishing the source pixels corresponding to the occluded light rays. We determine whether a source pixel is occluded by determining if changes in the source pixel have resulted in changes in the screen pixel. However, since there are N possible changes of source pixel intensities from N projectors but only one observable screen intensity, we need to probe by varying the source pixels in different projectors separately. This cyclical probing results in a serial variation of the projector intensities.

The light suppression (LS) component of the feedback rule is based on

$$(\Delta\alpha_{jt})_{\text{LS}} = -\beta \frac{\Delta\alpha_{j(t-N)}^2}{\Delta Z_t^2 + \epsilon}, \qquad (6)$$

where $\Delta Z_t = Z_t - Z_{t-N}$ is the change in the screen pixel intensity caused by the change of alpha value $\Delta\alpha_{j(t-N)}$ in the previous time frame when projector j is active, and β is a small proportional constant and ϵ is a small positive constant to prevent a null denominator.

The rationale for (6) is that if the change in α_{jt} results in a corresponding-sized change in Z_t, the subsequent change in α_{jt} will be relatively minor (based on a small β). However if a change in α_{jt} does not result in a change in Z_t, this implies that the source pixel is occluded. The denominator of (6) approaches zero and α_{jt} is strongly reduced in the next time frame. Hence occluded source pixels are forced to black.

Note that the system must be able to discover when a pixel which was turned off due to the presence of an occluder is available again, due to the occluder's disappearance. This requirement is smoothly incorporated into our algorithm.

The complete iterative feedback rule is obtained by combining (5) and (6) to get

$$\Delta\alpha_{jt} = (\Delta\alpha_{jt})_{\text{SE}} + (\Delta\alpha_{jt})_{\text{LS}}. \qquad (7)$$

The alpha values are updated within limits such that

$$\alpha_{jt} = \begin{cases} 1, & \text{if } \alpha_{jt} + \Delta\alpha_{jt} > 1, \\ 0, & \text{if } \alpha_{jt} + \Delta\alpha_{jt} < 0, \\ \alpha_{jt} + \Delta\alpha_{jt}, & \text{otherwise.} \end{cases} \qquad (8)$$

3.4 System Details

During the initialization phase of its operation (when the scene is occluder-free) the system projects each presentation slide and captures a reference image per slide with the camera. During normal operation, the system camera continuously acquires images of the projected display which may contain uncorrected shadows. The comparison between the observed images and the reference image facilitates the computation of the alpha masks for individual projectors through (7). These are merged with the presentation slide in the screen frame of reference, followed by further warping into the projector frame of reference. These projected images from all projectors optically blend to form the actual screen display.

Note that the cost of shadow elimination is the use of redundant projectors. This means that at any point in time there are pixels on one or more projectors that are not being utilized because they fall outside the display surface or are occluded. We feel this is a small price to pay, particularly in comparison to the large costs, in either expense and required space, for other display technologies such as rear projection or plasma. Fortunately, portable projectors are becoming increasingly affordable as their image quality improves and their weight decreases.

Some images of the system in action are shown in Figure 4. The images demonstrate the difference between shadow elimination alone and in combination with occluder light suppression.

A benefit of using a visual feedback system (as opposed to an open-loop approach based on an accurate photometric model) is that the system is surprisingly robust. For instance, if one of the projectors in the multi-projector array were to fail, the remaining projectors would automatically brighten their images to compensate. Furthermore, the overall brightness of the entire multi-projector array can be changed simply by adjusting the camera aperture.

Figure 4. Comparison between different projection systems. These images were taken from an audience memeber's viewpoint: (a) Single projector; (b) Two aligned projectors, passive; (c) Two aligned projectors with shadow elimination only; (d) Two aligned projectors with shadow elimination and occluder light suppression. Note that the harsh shadow in (a) is replaced by softer double shadows in (b). Shadows are completely removed in (c) and (d). However, the user's face is brightly illuminated with projected light in (c). This blinding light is completely suppressed in (d).

4 Related Work

Research in the area of camera-assisted multi-projector displays is becoming more popular, particularly in the context of seamless video walls [1, 10, 2, 3, 4, 6]. Two previous papers [9, 11] presented solutions to the shadow elimination problem for forward-projection systems. In more recent unpublished work [12], we present a preliminary version of the occluder light suppression system described in Section 3. This problem is technically much more challenging because it requires the ability to determine *which* rays of projected light are being occluded. We believe our results in occluder light suppression are unique.

A simple camera feedback system, related to the one presented here, was used by [4] to adjust projector illumination for uniform blending in the overlap region of a video wall. In [13] a projector-mirror system is used to steer the output of a single projector to arbitrary locations in an environment. The Shader Lamps system [14] uses multiple projectors and a known 3-D model to synthesize interesting visual effects on 3-D objects. The geometric 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. In the Tele-graffiti system [15], a camera is used to track the motion of a flat display surface and automatically maintain the alignment of a projected image with the moving display. This system shares our goal of real-time geometric compensation, but lacks interactive control of the display window and the ability to compensate for occlusions.

In their work on the "Magic Board", Coutaz et al. [16] describe a system in which a user can control the size of a projected display window by manipulating poker chips which are tracked by a camera. In this work, however, the projector and screen are carefully aligned in advance, and the detected poker chips specify window coordinates in a pre-calibrated and fixed reference frame. Our work extends this paradigm to include autocalibration and the ability to adapt to changes in the positions of the camera and projector.

5   Conclusions and Future Work

Visual feedback is a key component in developing robust, interactive projected light displays. Camera-based sensing of the environment makes it possible to compensate in real-time for both geometric and photometric effects. Visual tracking of both fiducial markers and the corners of the projector output supports real-time autocalibration and makes the system robust to changes in the position of the projector, camera, and screen. It also permits the user to specify the desired screen location by positioning fiducial marks on the display surface. In addition, a photometric feedback rule makes it possible to eliminate shadows and the illumination of occluding objects in a multi-projector configuration.

In the future, we plan to extend the system in several ways. In addition to increasing the frame rate at which the system operates, we will incorporate multiple cameras into the visual feedback loop. This will enable the system to work reliably even when a camera is occluded. We are also developing user-interface techniques for controlling and adjusting virtual displays using hand gestures. In particular, we are exploring shadow detection as a means to support touch-based interaction with the projected light display.

References

[1] R. Yang, D. Gotz, J. Hensley, H. Towles, and M. Brown, "PixelFlex: A reconfigurable multi-projector display system," in *Proceedings of IEEE Visualization*, 2001.

[2] M. Hereld, I. Judson, and R. Stevens, "Introduction to building projection-based tiled displays," *Computer Graphics and Applications*, vol. 20, no. 4, 2000.

[3] 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*, vol. 20, no. 4, 2000.

[4] R. Surati, *A Scalable Self-Calibrating Technology for Seamless Large-Scale Displays*. PhD thesis, Department of Electrical Engineering and Computer Science, Massachussetts Institute of Technology, 1999.

[5] O. Hall-Holt and S. Rusinkiewicz, "Stripe boundary codes for real-time structured light range scanning of moving objects," in *Proceedings of International Conference on Computer Vision*, 2001.

[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 International Conference on Computer Vision*, 2001.

[8] D. Ballard and C. Brown, *Computer Vision*. Prentice-Hall, 1982.

[9] R. Sukthankar, T.-J. Cham, and G. Sukthankar, "Dynamic shadow elimination for multi-projector displays," in *Proceedings of Computer Vision and Pattern Recognition*, 2001.

[10] A. Majumder and G. Welch, "Computer graphics optique: Optical superposition of projected computer graphics," in *Proceedings of Eurographics Workshop on Rendering*, 2001.

[11] C. Jaynes, S. Webb, R. M. Steele, M. Brown, and W. B. Seales, "Dynamic shadow removal from front projection displays," in *Proc. IEEE Visualization*, 2001.

[12] T.-J. Cham, R. Sukthankar, J. M. Rehg, and G. Sukthankar, "Shadow elimination and occluder light suppression for multi-projector displays," in *Submitted to Proc. IEEE Visualization*, 2002.

[13] C. Pinhanez, "The Everywhere display," in *Proceedings of Ubiquitous Computing*, 2001.

[14] R. Raskar, G. Welch, and K.-L. Low, "Shader Lamps: Animating real objects with image-based illumination," in *Proceedings of Eurographics Workshop on Rendering*, 2001.

[15] N. Takao, J. Shi, and S. Baker, "Tele-graffiti," Tech. Rep. CMU-RI-TR-02-10, Carnegie Mellon University, March 2002.

[16] J. Coutaz, J. L. Crowley, and F. Berard, "Things that see: Machine perception for human computer interaction," *Communications of the ACM*, vol. 43, March 2000.