

Multi-Projector Display with Continuous Self-Calibration

Jin Zhou

Liang Wang

Amir Akbarzadeh*

Ruigang Yang

Graphics and Vision Technology Lab (GRAVITY Lab)
Center for Visualization and Virtual Environments, University of Kentucky
1 Quality Street, Suite 800, Lexington, KY, 40507-1464, USA

Abstract

Most existing calibration techniques for multi-projector display system require that the display configuration remain fixed during the display process. We in this paper present a new approach to continuously re-calibrate the projection system to automatically adapt to the display configuration changes, *while the multi-projector system is being used without interruption*. By rigidly attaching a camera to each projector, we argue the projector with sensing capability and use the camera to provide *online* close-loop control. In contrast to previous auto or continuous projector calibration solutions, our approach can be used on surfaces of arbitrary geometry and can handle both projector and display surface movement, yielding more flexible system configuration and better scalability. Experimental results show that our approach is both accurate and robust.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Digitizing and scanning; Display algorithms; I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Camera calibration; Imaging geometry; I.4.9 [Image Processing and Computer Vision]: Applications—Image Display

Keywords: multi-projector Display, projector-camera system, continuous calibration

1 Introduction

Multi-projector display is an effective tool for achieving large-scale seamless imagery. The advancement of this technique is very useful to a variety of applications such as scientific visualization, virtual environments and entertainment. Research on multi-projector display has recently experienced somewhat of a new era due to the increasing computer performance and the diminishing cost and size of digital light projectors and cameras.

While multi-projector display systems offer several advantages over other display options, the system calibration is often a tedious undertaking. To achieve seamless image alignment both intrinsic and extrinsic parameters of the projectors and a geometric description of the display surface need to be accurately known. Moreover, re-calibration is required to maintain geometric alignment once the previous calibration is destroyed by external factors, e.g. the movement of a projector. Such unexpected cases can happen from time to time in some special display environment such as vehicles or ships. Frequent system re-calibration will not only interrupt the display process but also consume significant effort.

* Amir Akbarzadeh is currently with Microsoft Live Labs.



Figure 1: Our augmented projector, which has an rigidly attached firewire digital camera on it.

To alleviate the aforementioned limitations we in this paper present a technique that is able to automatically recalibrate a projector over an arbitrary-shaped display surface. The approach is able to run in near real time (less than one second), which avoids interrupting the display of imagery. By rigidly attaching a video camera to a projector we develop a hardware platform with sensing capability. **The key idea of automatic calibration is to use cameras to provide online close-loop control. Note that this is different from most projector-camera calibration procedures, which are typically treated as an off-line process.** With a camera, computer vision techniques are leveraged to determine the necessary correction function for each projector so that continuous self-calibration can be achieved. The proposed technique is robust and accurate. In addition, the desirable scalability property and automatic reconfigurability make multi-projector systems more practicable for different users or applications.

1.1 Relation to previous work

This work is related to a sizable body of literature on projector-based display systems, which are too vast to summarize here. Instead, we will report only approaches that consider the problem of automatic calibration which requires little or no user interaction, as it is the focus of this paper.

[Yang et al. 2001] presents a spatially reconfigurable multi-projector display system which is optimized for planar or near planar display surfaces. Similar to our approach, the authors use cameras for closed-loop calibration. The system provides the flexibility and versatility to change display layout for different applications, allowing users to easily create, save and restore a multitude of display layouts efficiently. [Raij and Pollefeys 2004] proposes an automatic method for defining the display area on a plane, removing the need for physical fiducials and measurement of the area defined by them. Planar auto-calibration can be used to determine the intrinsics of an array of projectors projecting on a single plane. The camera, projectors and display plane are then reconstructed using a relative pose estimation technique for planar scenes. **Raskar et al. [Raskar et al. 2003]** have investigated how to use projectors in a flexible way in everyday settings. Their basic display unit is a projector with sensors, computation, and networking capability. It can create a seamless display that adapts to the surfaces or objects

being projected on. Display surfaces with complex geometries, e.g. curved surface, can be handled. In [Yang and Welch 2001], features in the imagery being projected are matched between a pre-calibrated projector and camera to automatically determine the geometry of the display surface. However the estimation algorithm works in an iterative manner and is not suitable for continuous correction in real time. Instead of matching features across images there are active techniques where calibration aids are embedded into the user imagery. For instance, [Cotting et al. a; Cotting et al. b] embed imperceptible calibration patterns into the projected images. The approach takes advantage of the micro-mirror flip sequence in Digital Light Processing (DLP) projectors and slightly modifies the per-pixel intensity to let the synchronized camera capture the desired pattern. However, one major drawback of the approach is that it requires a portion of projector's dynamic range to be sacrificed which will in turn cause a degradation of the imagery being projected.

Our work is most related to [Johnson and Fuchs 2007], which also demonstrates the ability to continuously calibrate a projector on arbitrary display surface without modifying the projected imagery. By matching features between the user image stored in the frame buffer and the projected image captured by a stationary camera, the approach re-estimates the pose of the projector lively. However due to the difficulties existing in the matching between camera and projector images (e.g. differences in resolution, large camera-projector baseline and radiometric effects etc), the paper obtains the feature correspondences indirectly by matching between the camera and an image generated to predict what the camera will see. A radiometric model for generating this predicted image is presented in the paper. Our approach is fundamentally different than theirs in that we attach each projector with its own camera while the method in [Johnson and Fuchs 2007] uses a single camera to observe the entire display. Our distributed architecture is more robust to system configuration changes (the camera has to be stationary in [Johnson and Fuchs 2007]) and more flexible and scalable. The cost of a camera is small compared to the projector cost. In addition since we have multiple cameras in our setup, we can directly match features between camera pairs, eliminating the need for radiometric calibration.

2 Approach

As shown in Figure 1, a projector-camera pair forms the basic building block of the proposed self-correction display system. The camera is rigidly attached to the projector and its field of view is wider than that of the projector, making it possible to capture the entire area illuminated by the projector. In addition, in order to acquire depth information with a single projector-camera pair there should be a reasonable baseline between the projector and camera. The display system consists of multiple such projector-camera unit connected to a computing platform (e.g. a PC). Inter-system communication is accomplished through networking.

For the scope of this paper, we focus on only the geometric aspect of calibration. Our system is designed to deal with display surfaces with arbitrary complex geometry. In this most general case the 3D model of the display surface as well as projector poses within a common coordinate system need to be known to achieve large seamless imagery from multiple projectors. To correct the interruption caused by display configuration changes, e.g. certain projector-camera pair is moved, the projector-camera unit is further required to be able to perform online self-calibration to obtain a new pose estimate.

Figure 2 illustrates the main modules of our display system. The projector-camera calibration module estimates the intrinsic param-

eters of the projector and camera, as well as their relative pose. The next module is to acquire the shape of the display surface so that each projector-camera pair obtains its own surface description (e.g. a mesh). In this module, the relative pose between each pair is also estimated. This completes the one-time off-line calibration procedure and the display is ready for use. In the run time, the online calibration module monitors the display configuration and estimates the relative pose between each projector-camera pair. Once a projector movement is detected, corresponding adjustments are made automatically to correct the rendering process.

2.1 Projector-camera calibration

In our system each projector-camera pair, denoted as $\{P_i, C_i\}$, forms an active stereo pair S_i . The objective of the projector-camera calibration is to calibrate S_i both intrinsically and extrinsically. That means after calibration their intrinsic parameters (lens distortion, focal length, image center etc.) are known. In addition, the relative pose of the camera with respect to the projector is also obtained. We assume this relative pose does not change. Note that this is a reasonable assumption since the camera can be integrated inside the projector by the design, which can be easily fixed. Without loss of generality (WLOG), we choose P_i to be the reference view which defines the local coordinate frame.

While camera can be calibrated easily using the techniques proposed in [Zhang 1999], the projector cannot directly observe any image so it has to be calibrated indirectly. In this paper we adopt a two-phase approach to accurately calibrate the projector-camera pair. WLOG we assume there are two projector-camera units in our display system and use S_0 and S_1 to refer them respectively. It is a valid assumption since being a multi-projector display system, there are always more than one projector-camera pairs available.

First of all the two cameras $\{C_0, C_1\}$ are calibrated within a unified coordinate frame using the traditional camera calibration method by viewing a plane from multiple unknown orientations [Zhang 1999].

We thereby obtain their intrinsic parameters as well as the camera matrices representing camera poses. Next we use the calibrated cameras to calibrate the projectors $\{P_0, P_1\}$. During the calibration of P_0 we let P_0 project a pattern (e.g. a chessboard pattern) onto a planar surface. By letting both cameras C_0 and C_1 observe the projected pattern we can locate the 3D coordinates of the projected features (e.g. the corners on the chessboard pattern) using linear triangulation method [Hartley and Zisserman 2000]. Since the corresponding 2D image coordinates of the 3D features on the projected pattern are known we can solve P_0 's projection matrix from these 2D-3D correspondences. To overcome the data noise and achieve better accuracy we in practice change the surface's position and orientation in 3D to span a large depth range. 2D-3D correspondences obtained from different surface configurations are used together to achieve a least square solution using the linear method (as will be further described in section 2.4.1). Given the fact that the above calibration procedure is accomplished within a unified coordinate frame the relative pose (i.e. the translation and rotation) between the projector and camera can be easily derived from their projection matrices.

This calibration approach can be naturally extended to system containing more than two projector-camera units. Projector-camera pairs that are next to each other can be calibrated together via the above method.

2.2 Geometric calibration

We accomplish the geometric calibration by observing projected structured light patterns with each active stereo pair S_i . This pro-

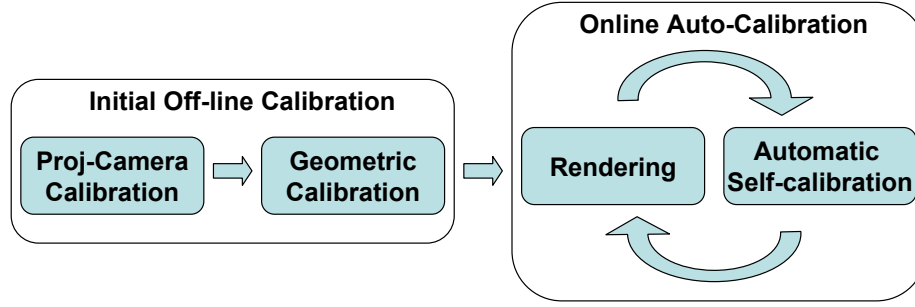


Figure 2: Main modules of our current system. The projector-camera calibration module estimates the intrinsic parameters of the projector and camera, as well as their relative pose. The Geometric calibration module is to acquire the surface shape and estimate the relative pose between each projector-camera pair. In the run time, our online auto-calibration block monitors the projector pose and make adjustment to the rendering automatically.

cess yields a set of 2D correspondences between P_i and C_i that allows us to reconstruct the geometry of the display surface. Repeating this procedure for each S_i effectively scans the entire display surface in 3D. However, by now each D_i is registered in a different coordinate frame because D_i is defined in the local coordinate frame of P_i .

Since the field of view of C_i is larger than the projector's, it can see not only the region illuminated by its associated projector P_i , but also any overlapping portion from an adjacent projector P_j . Using the point correspondences acquired in the overlapping region between two display surfaces D_i and D_j , a rigid transformation consisting of a rotation and a translation can be computed to bring D_i and D_j into alignment. The same transformation also unifies P_i and P_j into the same coordinate frame. WLOG, we can simply define P_0 as the global reference coordinate frame, to which all other projectors are transformed via a series of rigid transformations.

The above approach yields flexible geometric description of the display surface and multi-projector alignment. In addition, this technique is scalable, allowing immersive displays to be deployed in a wide-range of environments.

2.3 Rendering

For a moving viewer in an arbitrary display environment, a necessary warping function between each projector and the desired image must be dynamically computed as the viewpoint changes. In our system we adopt the algorithm presented by Raskar et al. [Raskar et al. 1999]. Figure 3 shows a diagram of this two-pass rendering approach. The desired image is rendered in the first pass. This image is then served as a projective texture and projected from the viewer's point of view onto a 3D model of the display surface. The textured 3D model is rendered from the view point of the projector in the second pass. When displayed, this second rendered image will appear as the desired image to the viewer.

2.4 Automatic self-calibration

After projector-camera and geometric calibration the system is in a calibrated state. During run-time various factors can ruin the calibration and lead to incorrect image alignment. The two most common cases are projector movement and display surface movement. To detect calibration failures, each camera in the projector-camera pair continuously observes the display surface and compares it against the predicted scene (previously captured frame). If the captured image agrees with the predicted image, obviously no system configuration has changed. That is the projector, the camera,

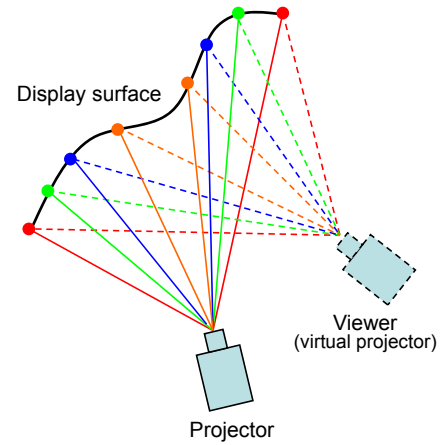


Figure 3: Two-pass rendering algorithm on an arbitrary display surface. The first pass (dashed line) renders the desired image to be observed from the user's view point. This rendered image is then used as a projective texture onto the display surface. The textured display surface is then rendered from the projector's point of view constituting the second pass render (solid line).

and the display surface remains static. Otherwise if inconsistency occurs, self-calibration is required to bring the system back to a new calibrated state.

However, without special hardware support it is nontrivial for each projector-camera pair to understand the exact cause of the inconsistency. For example either the projector movement or the display surface movement can result in the same observation. Fully distinguishing such ambiguities is outside the scope of this paper, we therefore make the following simple assumptions in our current implementation:

- The system knows whether the inconsistency is caused by projector movement or display surface movement. In the future, this can be solved by attaching a vibration sensor to the projector-camera pair.
- If the inconsistency is caused by projector movement then the system knows which projector-camera pair has been moved.
- If the inconsistency is caused by display surface movement we assume the surface is planar.

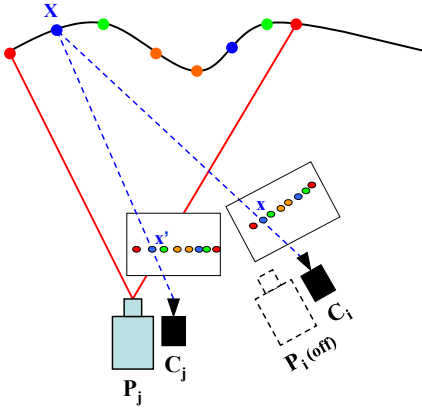


Figure 4: After S_i has been moved P_i is turned off and we let C_i observe the scene illuminated by P_i 's neighboring projector P_j . $\mathbf{x} \leftrightarrow \mathbf{x}'$ is a pair of 2D feature correspondence. Since P_j and C_j remain calibrated the valid 3D information of point \mathbf{X} can be acquired from S_j 's surface mesh. 2D-3D correspondence $\mathbf{x} \leftrightarrow \mathbf{X}$ is therefore established.

2.4.1 Projector movement

As illustrated in Figure 4, after the movement of the projector-camera pair S_i has been detected, P_i will be turned off (project an image with zero intensity everywhere). C_i is used to observe the feature points projected by other projectors. Since the geometry of the display surface and other projectors remain calibrated, these features have valid 3D information, e.g. depth values. By matching features across images we indirectly obtain a set of 2D-3D point correspondences from which we can estimate the current pose of C_i .

In our implementation 2D features are extracted from each captured image via the Harris corner detector [Harris and Stephens 1988]. Features detected on the image of C_i , say I_i , are compared against features on other images based on patch similarity. A 11×11 local image patch centered at each feature pixel is used to describe that feature. After matching using Normalized Cross Correlation (NCC), features without reliable match are discarded and each remaining feature on I_i has one matched correspondence on other images. As described previously, since other projector-camera pairs are still calibrated, for feature points in I_i we can obtain their corresponding 3D points on the display surface mesh.

To compute the 3 by 4 projection matrix \mathbf{P} of the camera which maps world points to image coordinate we begin with a simple linear algorithm given a set of N 2D-3D point correspondences $\mathbf{x} \leftrightarrow \mathbf{X}$. The transformation is given by the equation $\mathbf{x} = \mathbf{P}\mathbf{X}$. This is an equation involving homogeneous vectors therefore the 3-vector \mathbf{x} and $\mathbf{P}\mathbf{X}$ are not necessarily equal. They have the same direction but may differ in magnitude by a scale factor. By expressing this equation in terms of the vector cross product as $\mathbf{x} \times \mathbf{P}\mathbf{X} = 0$. Writing $\mathbf{x} = (x, y, 1)^T$, the cross product may then be given explicitly as

$$\begin{bmatrix} 0^T & -\mathbf{X}^T & y\mathbf{X}^T \\ \mathbf{X}^T & 0^T & -x\mathbf{X}^T \\ -y\mathbf{X}^T & x\mathbf{X}^T & 0^T \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = 0, \quad (1)$$

where \mathbf{P}^{iT} is a 4-vector representing the i -th row of \mathbf{P} .

Since the three equations of (1) are linear dependent we choose to

use only the first two equations

$$\begin{bmatrix} 0^T & -\mathbf{X}^T & y\mathbf{X}^T \\ \mathbf{X}^T & 0^T & -x\mathbf{X}^T \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = 0. \quad (2)$$

In theory at a minimum 6 such correspondences can give us a solution, but to overcome the matching outliers we in practice use **RANSAC** based robust estimation to select a number of K inliers from all N correspondences. We further obtain a $2K \times 12$ matrix \mathbf{A} by stacking up the equation (2) for each correspondence classified as inlier. The null-space of \mathbf{A} , which is a least square solution of $\mathbf{A}\mathbf{P} = 0$, is thereafter served as an initial guess to a Levenberg-Marquardt algorithm that further iteratively improves the calibration accuracy [Hartley and Zisserman 2000]. Since C_i and P_i are rigidly attached P_i 's calibration can be solved from C_i 's projection matrix \mathbf{P} .

Since we can recover the new projector-camera pose in a single frame, we can in theory achieve calibration at video rate. In practice, however, we found out that our simple hardware setup without projector-camera synchronization needs a delay of several frames for the moved projector to be reset. This again shows the importance of engineering the projector and the camera as an integrated system.

2.4.2 Planar display surface movement

In the case that the display surface has been moved, the projector-camera pairs remain static but the geometric description of the surface needs to be redetermined. If we own no prior knowledge about the surface geometry (i.e. the surface has arbitrary shape) the structured light technique has to be applied again to recover the 3D surface mesh. Unfortunately re-scanning is not practical here since it will discontinue the display process. By assuming the display surface is planar this problem is simplified because we can leverage the plane fitting technique to reconstruct the display surface given a few 2D-2D feature correspondences. Again we use Harris feature detector to locate the 2D image features. Matched 2D correspondences are triangulated to obtain a set of 3D points. An optimal plane that describes the surface geometry is fitted from the point cloud using the standard **RANSAC** based plane fitting [Quirk et al. 2006].

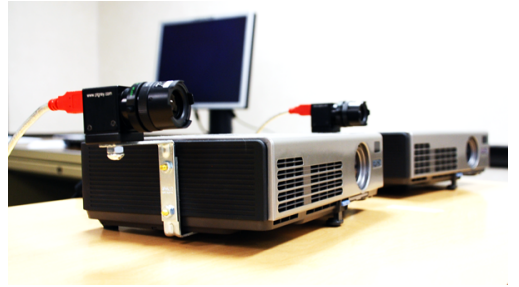


Figure 5: System prototype. Two projector-camera pairs are driven by an Intel 2.4GHz PC with dual graphic cards.

3 Results

We have implemented all the projector calibration methods described in section 2. Figure 5 shows our experiment setup. Currently our system prototype consists of two projector-camera pairs. Both of them are driven by a PC equipped with 2.4GHz CPU and

dual graphics cards. The firewire PointGrey Flea camera's resolution is 1024×768 and its field of view is about 60 degrees.

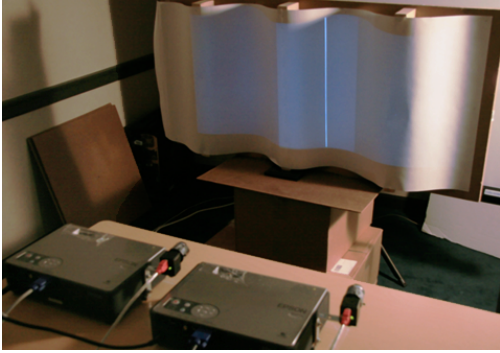


Figure 6: Scanning of the display surface using structured light approach.

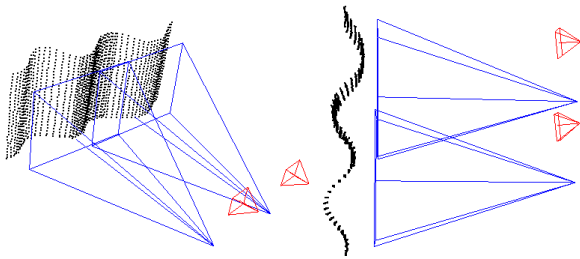


Figure 7: The reconstructed surface geometry rendered from two different perspective. The black dots represent the 3D points sampled on the display surface. Each blue frustra represents a projector and each red one represents a camera.

Figure 6 shows the structured light scanning procedure on a curved screen. In a typical horizontal configuration of the projector-camera pair the projector projects a moving vertical line across its field of view. Based on the epipolar geometry, the camera can determine the lighted 3D points on the display surface via triangulation. This scanning process takes a few seconds for each projector-camera pair. After the projectors' coordinate frames are unified we further render the reconstructed surface in 3D for verification and visualization purpose. In Figure 7 we show the estimated geometry of the display surface rendered from two different perspectives. As we can see the two local meshes are aligned fairly well and accurately approximate the surface shape.

Figure 8 and 9 show some projected images from our online system displaying on a planar and a curved display surfaces respectively. Note that the alpha-blending is turned off in these experiments to better highlight the image alignment accuracy. In both cases we can move the projector-camera pair almost arbitrarily during the display, as long as the camera can still see part of the other projector's image. The entire process is fully automatic and the correction is instantaneous. The selected areas contain boundaries of the overlapping regions. The zoom-in views demonstrate the satisfactory accuracy of our online self-calibration algorithm.

We in Figure 10 show some results for the planar display surface movement experiment. In this experiment we assume the projector-camera pairs remain static and only the display surface's configuration is changed. The first (left) image is the original display configuration and the two projected images are well aligned. As can be seen in the second (middle) image the display surface has been rotated by about 30 degrees, and without adaptive correction

the projected images are no longer aligned. The third (right) image shows the corrected projection after continuous calibration. The display surface geometry is re-estimated by fitting a plane using matched 3D features.

Currently the speed performance of our system prototype is severely limited by the projector-camera synchronization problem mentioned earlier. After the projector or display surface has been moved we basically let the camera sleep 0.2 seconds before capturing an image for feature matching in order to avoid the motion blur caused by projector movement. Besides this factor, the performance also depends on the number of features the system is able to detect and match. We test our system using ordinary imagery as those shown in figures 8, 9 and 10. Usually 100 to 300 matched feature pairs are detected at each frame and the average performance of the display correction is about 3.4 Hz. Note that if the 0.2 seconds delay can be avoided by introducing better synchronization mechanism the system is able to run in near video frame-rate.

4 Conclusion and future work

We have presented a new technique that is able to continuously recalibrate a projector given arbitrary-shaped display surface, without user intervention and obvious display interruption while the multi-projector system is being used for real work. By rigidly attaching a camera to a projector we argue the projector with sensing capability. After the relative pose between the projector and camera is pre-calibrated, feature correspondences matched between cameras are utilized to estimate the new projector pose to adapt to display configuration changes such as the movement of a projector. In contrast to previous proposed solutions, our approach can be used on surfaces of complex geometry and requires neither degrading the projected imagery quality nor additional radiometric calibration procedure. When planar display surface is used, our system can also compensate for the movement of the surface by assuming the projectors remain stationary.

Looking into the future, we would like to better engineer the projector-camera pair. In particular, we would like to add more sensors to the unit to allow it to differentiate between projector movement and display surface movement. We also want to synchronize the pair, making it possible to provide real-time pose correction with dynamic contents. By treating a projector and a camera as one single coherent unit, these intelligent projectors will be able to collectively form a high-resolution seamless display that can 1) be automatically calibrated with little or no user intervention and 2) continuously monitor the changes in the display configuration and make appropriate correction to maintain the seamless projection imagery.

Acknowledgements

This work is supported in part by US National Science Foundation grant IIS-0448185, and a grant from US Department of Homeland Security.

References

- COTTING, D., NAEF, M., GROSS, M., AND FUCHS, H. Embedding imperceptible patterns into projected images for simultaneous acquisition and display. In *Proceedings of IEEE/ACM International Symposium on Mixed and Augmented Reality*, 100–109.
- COTTING, D., ZIEGLER, R., GROSS, M., AND FUCHS, H. Adaptive instant displays: Continuously calibrated projections using



Figure 8: First row: the image alignment results on a planar display surface. Second row: zoom-in views of the scene.



Figure 9: First row: the image alignment results on a curved display surface (the one shown in Figure 6). Second row: zoom-in views of the scene.

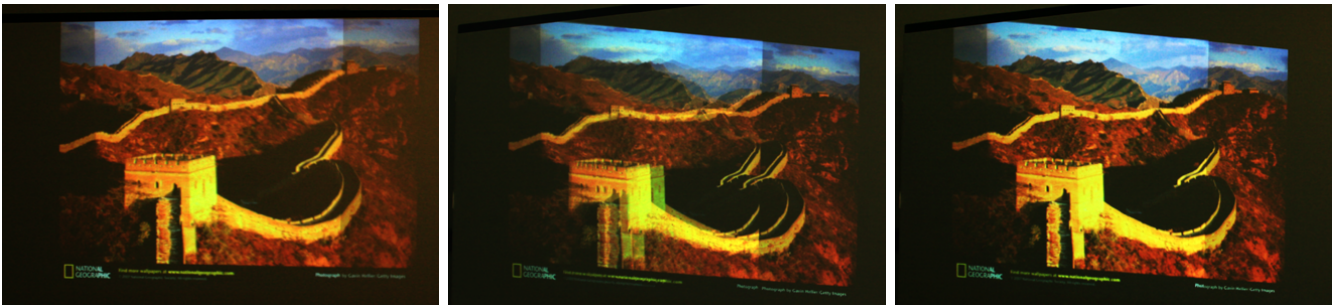


Figure 10: Results for the planar display surface movement experiment. The first (left) image is the original display configuration and the two projected images are well aligned. In the second (middle) image the display surface has been rotated by about 30 degrees, and without adaptive correction the projected images are no longer aligned. The third (right) image shows the corrected projection after continuous calibration.

per-pixel light control. In *Proceedings of Eurographics*, 705–714.

HARRIS, C., AND STEPHENS, M. 1988. A combined corner and edge detector. In *Proceedings of the 4th Alvey Vision Conference*, 147–151.

HARTLEY, R., AND ZISSERMAN, A. 2000. *Multiple View Geometry in Computer Vision*. Cambridge University Press.

JOHNSON, T., AND FUCHS, H. 2007. Real-time projector tracking on complex geometry using ordinary imagery. In *Proceedings of the IEEE Virtual International Workshop on Projector-Camera Systems*.

QUIRK, P., JOHNSON, T., SKARBEZ, R., TOWLES, H., GYARFAS, F., AND FUCHS, H. 2006. Ransac-assisted display model reconstruction for projective display. In *Proceedings of the IEEE Virtual Reality Workshop on Emerging Display Technologies*.

RAIJ, A., AND POLLEFEYS, M. 2004. Auto-calibration of multi-projector display walls. In *Proceedings of the International Conference on Pattern Recognition*.

RASKAR, R., BROWN, M. S., YANG, R., CHEN, W.-C., WELCH, G., TOWLES, H., SEALES, B., AND FUCHS, H. 1999. Multi-projector displays using camera-based registration. In *Proceedings of the IEEE Visualization*, 161–168.

RASKAR, R., VAN BAAR, J., BEARDSLEY, P., WILLWACHER, T., RAO, S., AND FORLINES, C. 2003. ilamps: Geometrically aware and self-configuring projectors. In *Proceedings of the ACM Siggraph*.

YANG, R., AND WELCH, G. 2001. Automatic and continuous projector display surface calibration using every-day imagery. In *Proceedings of the 9th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*.

YANG, R., GOTZ, D., HENSLEY, J., TOWLES, H., AND BROWN, M. S. 2001. Pixelflex: A reconfigurable multi-projector display system. In *Proceedings of the IEEE Visualization*, 177–188.

ZHANG, Z. 1999. Flexible camera calibration by viewing a plane from unknown orientations. In *Proceedings of the International Conference on Computer Vision*, 666–673.