Registration Techniques for Using Imperfect and Partially Calibrated Devices in Planar Multi-Projector Displays

Ezekiel Bhasker, Ray Juang, *Student, IEEE*, and Aditi Majumder, *Member, IEEE*

Abstract— Multi-projector displays today are automatically registered, both geometrically and photometrically, using cameras. Existing registration techniques assume pre-calibrated projectors and cameras that are devoid of imperfections such as lens distortion. In practice, however, these devices are usually imperfect and uncalibrated. Registration of each of these devices is often more challenging than the multi-projector display registration itself. To make tiled projection-based displays accessible to a layman user we should allow the use of uncalibrated inexpensive devices that are prone to imperfections. In this paper, we make two important advances in this direction. First, we present a new geometric registration technique that can achieve geometric alignment *in the presence of severe projector lens distortion* using a relatively inexpensive low-resolution camera. This is achieved via a closed-form model that relates the projectors to cameras, in planar multi-projector displays, using rational Bezier patches. This enables us to geometrically calibrate a 3000×2500 resolution planar multi-projector display made of 3×3 array of nine severely distorted projectors using a low resolution (640×480) VGA camera. Second, we present a photometric self-calibration technique for a projector-camera pair. This allows us to photometrically calibrate the same display made of nine projectors using a *photometrically uncalibrated* camera. To the best of our knowledge, this is the first work that allows geometrically imperfect projectors and photometrically uncalibrated cameras in calibrating multi-projector displays.

Index Terms—Geometric calibration, photometric calibration, tiled displays.

---◆---

1 INTRODUCTION

Multi-projector displays are popular today for many applications such as visualization, entertainment, training, and simulation. Contemporary researchers envision future workspaces with *ubiquitous pixels* rendered by multi-projector displays of various scales and forms aiding users in collaboration, interface, and visualization [9, 12, 16, 17, 24, 25, 30, 33, 36].

To enable such readily accessible multi-projector displays, automated camera-based registration techniques that register the imagery coming from multiple projectors, both geometrically and photometrically, have been proposed [5, 6, 13, 19, 20, 21, 23, 26, 27, 28, 29, 34, 37, 38]. More recently, Bhasker, et al. [2], proposed a distributed network of compact projector-camera units that can self-register without any user input, relieving the user of installation and maintenance responsibilities. However, all these methods do not handle imperfections, such as lens distortions, that are common in inexpensive commodity devices. They assume pre-calibrated devices (*i.e.,* devices whose geometric and photometric parameters are known) leaving the responsibility of device calibration to the user. Device calibration is often more challenging than multi-projector display registration and demands a very educated user. This difficulty in deployment has quarantined tiled projection-based displays to national laboratories and universities despite huge advances in automated camera-based registration techniques.

The work in this paper makes initial advances by developing geometric and photometric registration techniques that allow imperfect uncalibrated devices to be used for the most common cases of planar displays. The ultimate goal would be to allow the use of projectors with *severe* geometric and photometric distortion and still achieve acceptable geometric and photometric registration. We achieve this goal

- *Ezekiel Bhasker is with University of California, Irvine, E-mail: ebhasker@ics.uci.edu.*
- *Ray Juang is with University of California, Irvine, E-mail: rjuang@ics.uci.edu.*
- *Aditi Majumder is with University of California, Irvine, E-mail: majumder@ics.uci.edu.*

using cameras that are neither geometrically nor photometrically registered. The paper by Majumder and Stevens [21] is the only work to date in this direction and presents initial methods to handle large spatial variation in intensity within and across projectors. In this paper, we make further advances to these unexplored domains. These are: (a) achieving geometric registration while using projectors that have considerable geometric imperfections (*e.g.,* lens distortions); (b) achieving photometric registration with a photometrically uncalibrated camera.

1.1 Using Geometrically Imperfect Projectors

We present a new geometric registration technique that can correct for severe lens distortion using a relatively low-resolution camera for planar multi-projector displays. This method relies on a new parametric geometric model relating the projector to the camera used for registering the display. The relationship is modeled in a compact and efficient manner using a rational Bezier patch.

Projection technology is advancing at a tremendous pace. Pocket projectors that can fit on one's palm are no longer a myth but a reality [1]. Industry initiatives are finding ways to build short-throw, unbreakable projectors that use LEDs as light sources so that they can be embedded in cell phones. It is evident that lens distortion will be significant in such low-cost projectors. Our method will enable the use of such low-cost projectors to build commodity tiled displays.

Further, in designing a compact multi-projector display or changing an existing setup to increase space utility, a short-throw lens is required to decrease the throw distance of the projectors. These lenses currently require several optical elements to reverse the high lens distortion created by short focal lengths. Hence, they are cost-prohibitive (around $2000 - $6000) and at least 5-6 times the cost of regular projectors (around $800 - $1500). Our work handles severe lens distortions, thus paving the way for using inexpensive lenses on projectors, just as they are used on cameras today. This allows for even more compact multi-projector display designs.

1.2 Using Photometrically Uncalibrated Cameras

Existing methods for photometric calibration of a multi-projector display often use a camera that is first calibrated using high dynamic range images [8, 26]. This involves taking images in an outdoor environment where one does not have control of motion, illumination, etc. This calibrated camera is then used to recover the photometric proper-

ties of the projectors which are then modified appropriately to achieve a photometrically seamless multi-projector display [21, 26].

In this paper, we present a photometric self-calibration technique for a projector-camera pair that allows us to register a multi-projector display photometrically using an uncalibrated camera. To the best of our knowledge, this is the first work that estimates both the intensity transfer function and the spatial variation in intensity of projectors simultaneously using a photometrically uncalibrated camera.

2 RELATED WORK

In this section, we review existing geometric and photometric registration techniques and relate them to the techniques we present in this paper.

2.1 Geometric Registration

Geometric registration of planar multi-projector displays entails reconstructing two functions: (a) the function that relates the individual projector coordinates to the camera coordinates; and (b) the function that relates the camera coordinates to the global screen coordinates. Most geometric registration techniques devised for planar displays so far fall into the following categories.

Linear methods assume linear models for cameras and projectors. They relate the projector, camera, and the screen coordinates by linear matrices called homographies [5, 34]. However, linear models do not account for geometric non-linearities, such as lens distortion, and cannot achieve accurate alignment in their presence. To achieve acceptable alignment in such cases, projectors are manually adjusted to the 'sweet-spot' in zoom level where the non-linear distortions become negligible.

Piecewise linear methods address these geometric non-linearities by using a piecewise linear function, essentially a triangulation, to relate the projector, camera and screen coordinates to each other. Though this achieves reasonable geometric accuracy, a dense triangulation is required to sample the parameter space adequately [37]. Thus, a high-resolution (5-6 Megapixel) camera is required to register even relatively small displays made of 4-6 projectors.

Hereld, et. al. [13], use a *non-linear model* where a closed-form cubic function handles some of the geometric non-linearities. This method has two limitations. First, a simple cubic polynomial cannot capture the perspective projection between the projector and the camera adequately. As a result, this method assumes a close-to-rectangular array of projectors resulting from on-axis projection. Such an arrangement is relatively easy to achieve manually in a rear-projection system, but not in front-projection system where the projector is on the ceiling projecting on a screen placed on the ground. Second, this work is not motivated by accurate analysis of the kind of non-linearities seen in projectors and hence cannot handle higher order non-linearities.

Our method incorporates the non-linear distortions of the projector into the function that relates the projector to other entities, such as the camera and screen, using a single accurate closed-form function – the rational Bezier patch. This allows combination of distortions arising from completely different reasons (*e.g.,* lens distortion and key-stoning) into a single unified representation. Further, it results in an efficient camera-based geometric registration for planar, multi-projector displays, that can handle significant lens distortion.

2.2 Photometric Registration

Photometric properties of a camera/projector involve its intensity transfer function, or ITF (more commonly called the gamma function), and the spatial intensity fall-off in its field-of-view (more commonly called the vignetting effect). *Photometric registration* of projection-based displays requires estimation of the photometric properties of the projectors using a camera. To do this, the camera is first photometrically registered. The high dynamic range imaging technique [8, 22] can be used to estimate the camera's ITF by taking images outdoors. Currently, there is no simple automated way of estimating the camera's vignetting effect. The registered camera is usually set to a narrow aperture when using it to estimate the photometric properties of the projectors in a tiled display. The techniques in the work by Raij,

et al. [26], are then applied to estimate the projectors' ITF using the registered camera. Using the projectors with known ITF, the techniques described in the papers by Majumder and Stevens [20, 21] are then used to estimate the spatial intensity fall-off of the projector, and to register the multiple projectors photometrically using the estimated photometric properties. No method currently exists that can use a photometrically uncalibrated camera and recover both the ITF and the spatial intensity fall-off of the projectors simultaneously. Our method in this paper estimates all the different photometric parameters of both devices in a photometrically uncalibrated projector-camera pair. It is accomplished in a controlled indoor setting without using any other physical props. Photometric registration techniques that use these parameters [20, 21, 38] can then be utilized to achieve photometric seamlessness in the multi-projector display.

3 USING GEOMETRICALLY IMPERFECT PROJECTORS

In this section, we present our geometric registration technique that corrects for projector geometric imperfections, such as lens distortion. We use a camera that is corrected for lens distortion [3]. First, we develop a new model that relates a single projector to the camera in a planar multi-projector display using a rational Bezier patch. The projector lens distortion is modeled by an affine-invariant *non-rational* Bezier patch. For a projector-camera setup with a planar display, this function is followed by a homography that relates the projector to the camera and to the screen linearly. We use a perspective-invariant *rational* Bezier patch to model the combination of these non-linear and linear functions. Thus, all the different distortions in a planar projection-based display (*e.g.,* key-stoning, radial, and tangential distortions) are combined into a single closed-form function. To the best of our knowledge, this is the first work that models both linear and non-linear distortions of a projector-camera system in a unified manner.

To geometrically register multiple distorted projectors, we sparsely sample the rational Bezier function relating the projector coordinates with the camera and screen coordinates, and estimate the function parameters. These parameters are then used to warp the input image to the projector appropriately resulting in a geometrically aligned image. The sparse sampling proves adequate for acceptable geometric registration since the rational Bezier offers a piecewise curve representation (which is more suited for lens distortions that transform lines to curves), rather than a piecewise linear representation. This sparse sampling allows us to use a low-resolution camera to achieve geometric registration for even medium sized displays made of 6-16 projectors.

3.1 Model for Single Projector

We first consider the simple case of a single projector projecting onto a planar screen and being observed by a single camera corrected for lens distortion. Our goal is to design an accurate parametric function $\mathscr{F}$ that defines the relationship between the projector pixels (x, y) and the camera pixels (u, v), *i.e.,* $(u, v) = \mathscr{F}(x, y)$. Ideally, the function $\mathscr{F}$ has two components. The first is the *lens distortion function $\mathscr{R}$* that relates the undistorted projector coordinates (x, y) to the distorted coordinates on the display screen (x_d, y_d). The second component is the well-known *homography $\mathscr{H}$* that relates the distorted coordinates to camera coordinates (u, v).

$$(u, v) = \mathscr{H}(x_d, y_d) = \mathscr{H}(\mathscr{R}(x, y)) = \mathscr{F}(x, y) \qquad (1)$$

We would like to model $\mathscr{F}$ by a single compact parametric function that accurately captures the effect of the concatenation of these two functions, $\mathscr{H}$ and $\mathscr{R}$, and is amenable to efficient data fitting computation.

Modeling $\mathscr{R}$ with Non-rational Bezier Patches: The classic lens distortion model [4, 10] consists of two independent distortions: *radial distortion* (*e.g.,* barrel or pin-cushion) and *tangential distortion*.

Radial distortion is usually modeled by

$$x_d = x + (x - x_p)(k_1 r^2 + k_2 r^4 + k_3 r^6) = x + \rho_x, \qquad (2)$$

$$y_d = y + (y - y_p)(k_1 r^2 + k_2 r^4 + k_3 r^6) = y + \rho_y, \qquad (3)$$

Case No.	Lens Distortion Parameters			% Error For Bezier of degree					
	Center (x_c, y_c)	Radial (k_1, k_2, k_3)	Tangential (p_1, p_2)	2	3	4	5	6	7
(a)	$(0.5, 0.5)$	$(-0.35, 0, 0)$	$(0, 0)$	5.78	ε				
(b)	$(0.5, 0.5)$	$(0, 0, 0)$	$(0.1, 0.1)$	ε					
(c)	$(0.5, 0.5)$	$(-0.35, -0.35, 0)$	$(0, 0)$	9.00	0.91	0.56	ε		
(d)	$(0.6, 0.55)$	$(-0.35, -0.13, 0)$	$(0.05, 0.05)$	1.9	0.14	0.03	ε		
(e)	$(0.6, 0.55)$	$(-0.35, -0.13, -0.016)$	$(0.05, 0.05)$	2.0	0.17	0.038	0.0012	0.00022	ε

Table 1. This table shows the accuracy of our fit for lens distortion functions using non-rational Bezier patches of different degree. Due to the large number of lower degree terms, note that in case (e) distortion of higher degrees (*e.g.,* degree 7) can be fitted to reasonable accuracy by a lower degree Bezier (*e.g.,* degree 4). ε represents the convergence threshold used which is strictly less than 10^{-11}.
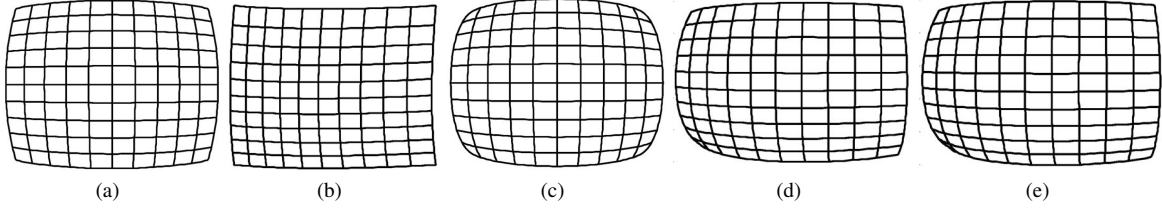

Fig. 1. Images showing the fives cases of distortion mentioned in Table 1.

where (x_d, y_d) and (x, y) are the distorted and undistorted image coordinates respectively, $r = \sqrt{(x - x_p)^2 + (y - y_p)^2}$ is the radial distance from the principal center (x_p, y_p) in the undistorted image, and k_i, $1 \leq i \leq 3$ are the radial distortion coefficients. Negative values of k_i result in barrel distortion while positive values result in pin-cushioning. The principal center is a point in the image that is unchanged by radial distortion. In general, the principal center (x_p, y_p) need not be at the center of the image but is usually close to the center.

The tangential distortion is modeled by

$$x_d = x + 2p_1xy + p_2r^2 + 2p_2x^2 = x + \tau_x, \quad (4)$$

$$y_d = y + 2p_2xy + p_1r^2 + 2p_1y^2 = y + \tau_y \quad (5)$$

where $r = \sqrt{(x - x_p)^2 + (y - y_p)^2}$ and p_1 and p_2 are the tangential distortion parameters.

Radial and tangential distortion are independent of each other and hence their effects can be combined into a comprehensive lens distortion equation, as

$$(x_d, y_d) = (x + \rho_x + \tau_x, y + \rho_y + \tau_y). \quad (6)$$

Estimating the unknown radial and tangential distortion coefficients involves solving complex non-linear optimization problems. Hence, several simplifying assumptions are usually made which are often not true in a real system. The most common simplifications are to assume that the principal point coincides with the center of the image and to ignore the higher degree terms [3, 7, 15, 32, 35, 39].

We model this lens distortion by *non-rational Bezier patches* instead. A non-rational Bezier patch of degree k and dimension d is defined by $(k + 1)^d$ control points in a d-dimensional space. In our case, we consider a two-dimensional patch where $d = 2$. The Bezier patch is a tensor product patch where the control points, P_{ij}, are weighted by blending functions that are products of independent one-dimensional Bernstein polynomials. There are $k + 1$ Bernstein polynomials of degree k given by

$$b_{nk}(t) = \frac{k!}{n!(k-n)!} t^n (1-t)^{k-n} \quad (7)$$

where $0 \leq n \leq k$ and $0 \leq t \leq 1$. The non-rational Bezier patch can then be defined from these polynomials as

$$\mathcal{N}\mathcal{B}(x, y) = \sum_{i=0}^{k} \sum_{j=0}^{k} P_{ij} b_{ik}(x) b_{jk}(y) \quad (8)$$

$$= \sum_{i=0}^{k} \sum_{j=0}^{k} P_{ij} B_{ijk}(x, y), \quad (9)$$

where B_{ijk} is the tensor product of the Bernstein polynomials b_{ik} and b_{jk} and $0 \leq (x, y) \leq 1$. Hence, in the context of the single projector-camera pair for planar displays, the projector lens distortion function $\mathcal{R}(x, y)$ is represented by

$$(x_d, y_d) = \mathcal{R}(x, y) = \mathcal{N}\mathcal{B}(x, y) \quad (10)$$

Our choice of using a non-rational Bezier patch is guided by the following rationale.

1. The non-rational Bezier patch provides a unified framework to represent a large range of lens distortions. More severe lens distortions can be included by increasing the *degree* of the Bezier patch. For example, a bicubic Bezier patch may be sufficient for standard narrow FOV lenses, but a Bezier patch of degree 4 may be required for the more severe distortions of wide FOV lenses. All of these are handled by the non-rational Bezier patch without affecting the underlying computational framework.

2. When Equation 10 is expanded and compared with Equation 6, one finds that the non-rational Bezier patch has many more higher-order cross terms, making it a more flexible model. For example, it is possible for short-throw lens projectors to have a mix of both barrel and pin-cushion distortions. Such cases are better handled by a more flexible model such as the Bezier patch.

3. The principal center creates a particular nuisance in traditional lens distortion estimation techniques, and forces most practical systems to assume it to be at the center of the image. When using a non-rational Bezier patch, the principle center is encoded within the control points themselves, and need not be estimated seperately.

4. Like lens distortion, non-rational Bezier patches are also affine-invariant and hence provide an accurate model.

Table 1 and Figure 1 present results on the accuracy of our Bezier patch model. We consider a dense sampling of the (x, y) space, distort it using various traditional lens distortion parameters, and fit a non-rational Bezier patch to the distorted points. To evaluate the accuracy of the fitted Bezier, we find the average Euclidian distance error between fitted and the distorted data as a percentage of the distance from
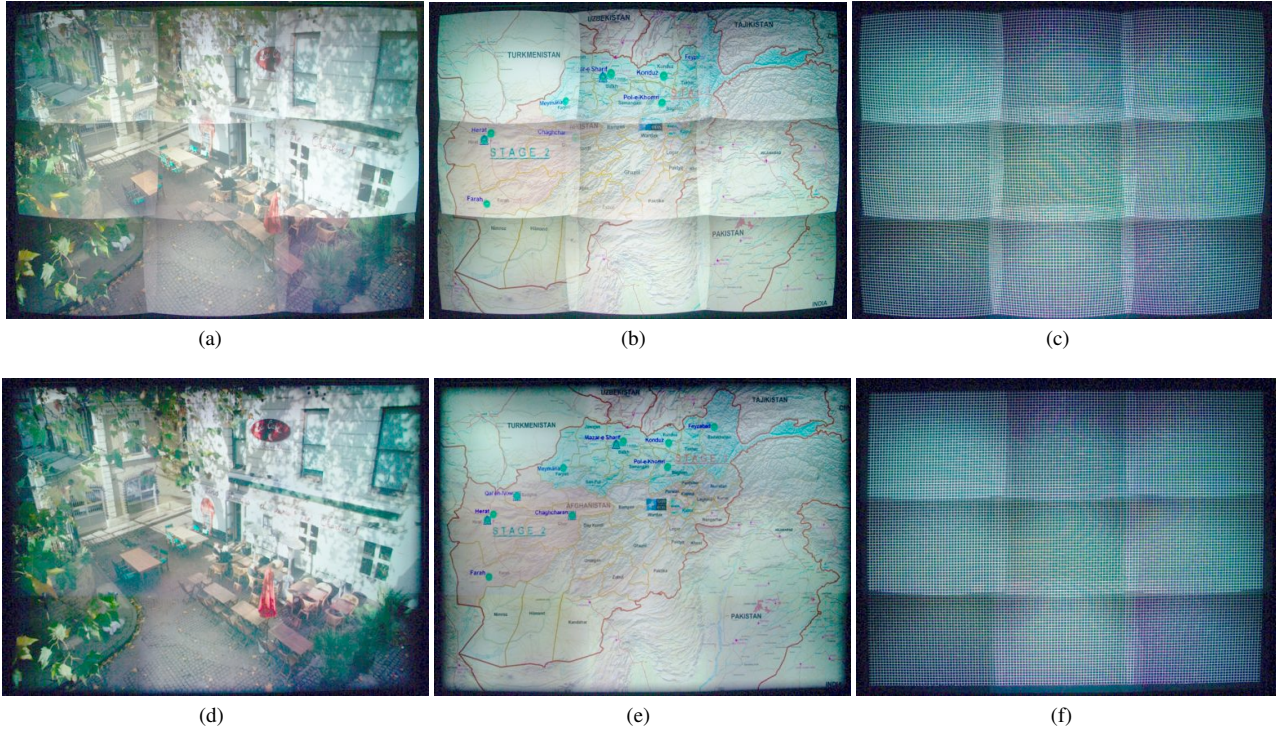
Fig. 2. Results of geometric registration shown on our display which is a 3×3 array of nine projectors. The geometric registration is achieved using bicubic rational Bezier patches with sparse correspondences (8×6 for a 1024×768 projector). (a, b, c) Before registration; (d, e, f) After registration. The grid image in the last column is made of lines that are one pixel wide. To retain clarity of the grids in the overlap region, photometric registration is not applied to the grid image in (c) when generating the registered image (f).

the principal center. Note that even when higher order distortions are present (*e.g.*, x^7), lower degree Bezier patches (*e.g.*, bicubic) can provide a good fit. This is due to the presence of a large number of lower-order cross-terms, which are absent in the traditional lens distortion model.

Modeling $\mathscr{HR}$ with Rational Bezier Patches: Non-rational Bezier patches are affine-invariant but not perspective-invariant. Relating the projector to the camera not only involves a lens distortion but also a perspective projection. In the case of planar displays, this is a homography $\mathscr{H}$, a simpler 3×3 transformation induced by the planar display screen. So, to model the combined effects of $\mathscr{H}$ and $\mathscr{R}$ in Equation 1, we desire a function that retains all the properties of the Bezier and is also perspective-invariant. Retaining all the properties of Bezier patch assures retaining the accurate model of non-linearities due to distortion. Invariance under perspective projection assures the additional modeling of the perspective projection between the projector and the camera. To our advantage, a generalized form of Bezier patches called rational Bezier patches offer us a function that has all the properties of a non-rational Bezier and is also perspective-invariant. In addition to the 2D control points of Bezier patches, rational Bezier patches have a scalar weight associated with each control point, and the patch is defined as a weighted sum of the blended control points:

$$\mathscr{B}(x,y) = \frac{\sum_{i=0}^{k} \sum_{j=0}^{k} P_{ij} w_{ij} B_{ijk}(x,y)}{\sum_{i=0}^{k} \sum_{j=0}^{k} w_{ij} B_{ijk}(x,y)}. \tag{11}$$

The weights provide this function with the capability to sample the parameter space in a non-uniform (but controlled) manner. This makes it perspective-invariant. Hence, we propose the use of a rational Bezier patch to estimate the function $\mathscr{F} = \mathscr{HR}$.

3.2 Geometric Registration of Multiple Projectors

Geometric registration of a multi-projector display requires estimating two functions. The first function maps the projector coordinates (x,y) to the camera coordinates (u,v). We use a rational Bezier patch $\mathscr{B}$ for

this purpose. The second function maps the camera coordinates (u,v) to the normalized screen coordinates (p,q). The presence of a planar screen allows us to use a homography $\mathscr{H}_S$ for this mapping. These functions are estimated as follows.

Parameter Estimation: To estimate $\mathscr{H}_S$, we require four correspondences between the camera and screen coordinates, (u,v) and (p,q). The four correspondences result in a system of linear equations that can be solved to recover the homography matrix. To find the correspondences, we identify four points in the camera image that correspond to the four corners of the screen. This can be done automatically by attaching fiducials to the display. Since our display is nearly rectangular, we set all the projectors to display white light and use simple image processing to detect the extreme corners of the lighted regions to find the camera coordinates that correspond to the four corners of the normalized screen space.

To estimate the parameters of the rational Bezier patch, $\mathscr{B}$, we find correspondences between projector pixels (x,y) and camera pixels (u,v). Let us assume N correspondences, *i.e.*, (x_c,y_c) corresponding to (u_c,v_c), $0 \le c \le N-1$. The unknowns are the control points in the camera's space, $P_{ij} = (u_{ij}, v_{ij})$, and their scalar weights w_{ij}. From the known correspondences, Equation 11 can be written for each correspondence as follows:

$$(u_c,v_c) \sum_{i=0}^{k} \sum_{j=0}^{k} w_{ij} B_{ijk}(x_c,y_c) - \sum_{i=0}^{k} \sum_{j=0}^{k} (u_{ij},v_{ij}) w_{ij} B_{ijk}(x_c,y_c) = 0. \tag{12}$$

Assuming $U_{ij} = u_{ij} w_{ij}$ and $V_{ij} = v_{ij} w_{ij}$, Equation 12 yields two equations for each correspondence, as follows:

$$u_c \sum_{i=0}^{k} \sum_{j=0}^{k} w_{ij} B_{ijk}(x_c,y_c) - \sum_{i=0}^{k} \sum_{j=0}^{k} U_{ij} B_{ijk}(x_c,y_c) = 0. \tag{13}$$

$$v_c \sum_{i=0}^{k} \sum_{j=0}^{k} w_{ij} B_{ijk}(x_c,y_c) - \sum_{i=0}^{k} \sum_{j=0}^{k} V_{ij} B_{ijk}(x_c,y_c) = 0. \tag{14}$$
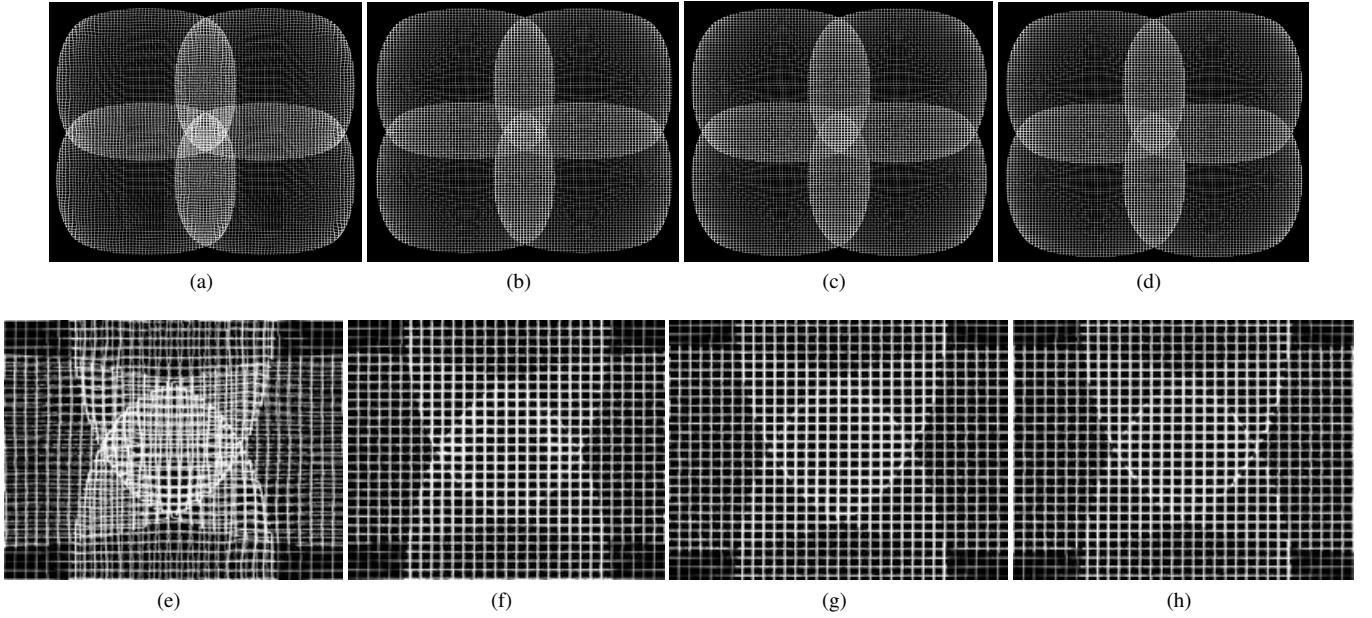
Fig. 3. Geometric registration of projectors with severe *barrel* distortion: using a linear piecewise method with sparse - 48 samples (a) and dense - 432 samples (b) correspondences; using rational *bicubic* Bezier patches with sparse - 48 samples (c) and dense - 432 samples (d) correspondences. (e,f,g,h) provides a zoomed in view of the projector overlap regions in (a,b,c,d) respectively. It is evident that piecewise linear methods with sparse correspondences lead to severe mismatches. The results using rational Bezier are similar, even when the sampling is reduced by an order of magnitude. In fact, at places, the results from using rational Bezier patches with sparse sampling is better than the results from using a dense sampling for the piecewise linear method. This is evident as the grid lines are thicker in the overlap region in (b) due to slight errors.

The unknowns for estimating a two-dimensional rational Bezier patch of degree k are $(k+1)^2$ control points and weights, which result in a total of $3(k+1)^2$ unknowns. The weights are variables that depend on the control points in a non-linear fashion. So, the above equations result in a non-linear least-squares fitting problem that can be solved efficiently using the Levenberg-Marquardt gradient descent optimization technique [31]. As an initial guess to this optimization, we fit the control polygon tightly around the set of correspondences (u_c, v_c) in the camera space and constrain the weights to be strictly greater than 0.

To find the correspondences between projector and camera coordinates, (x, y) and (u, v), we display a rectangular grid of Gaussian blobs with known coordinates on each projector. These are captured by a calibrated camera in a series of non-overlapping images. We use a 2D-stepping procedure that depends on a user to identify the top-left blob and its immediate right and bottom neighbors in camera space. With this minimal three point user input, we can design a stepping procedure that (a) estimates the rough position of the next blob in scan-line order, and (b) searches for the correct blob position using the nearest windowed center-of-mass technique [11]. If this is not possible for more extreme projector distortions, one can binary-encode the blobs and project them in a time sequential manner to recover the exact ids of the detected blobs and find the correspondences [28, 37].

Image Correction: To create a geometrically seamless and undistorted display, we warp the image for each projector as follows: For each projector coordinate (x, y), we first compute the corresponding camera coordinate (u, v) using the rational Bezier patch $\mathcal{B}$. Next, we compute the corresponding normalized screen coordinates (p, q) using the homography $\mathcal{H}_S$. The generated coordinates (p, q) are used as indices into the image being rendered on the tiled display. Bilinear interpolation is used to assign the final value of each projector pixel. Thus, the correction is achieved by the mapping

$$(p, q) = \mathcal{H}_S(u, v) = \mathcal{H}_S(\mathcal{B}(x, y)). \tag{15}$$

3.3 Implementation and Results

We have applied our method on our 8×6-foot display made of 3×3 array of nine projectors. Our projectors have relatively large throw ra-

tios and hence do not reveal major lens distortions. Instead of choosing the cost prohibitive option of buying nine short-throw lenses, we chose to simulate the distortion digitally by distorting the input images to the projectors.

The registration is done offline and takes a couple of minutes. This generates the parameters of the rational Bezier for each projector, which are then stored to be used during image correction. We have implemented the image correction in real-time using modern GPUs through Chromium – an open-source distributed rendering engine for PC clusters [14]. A module for Chromium is written that first pre-computes the coordinate-mappings of all pixel coordinates using the stored parameters of the rational Bezier. This results in a per-pixel projector to screen lookup table. A fragment program for the pixel shader quickly maps pixels from the projector coordinate space to the screen coordinate space during rendering.

Figure 2 shows the results. The advantage of using an appropriate parametric function lies in an accurate alignment even with a sparse sampling of the function. We can achieve accurate alignment of our relatively high resolution display (3000×2100) by using a low-resolution VGA (640×480) camera. Notice that even with a low-resolution camera we get sub-pixel accuracy for fine grids (one pixel wide lines) and text. Figure 3 shows that we can achieve accuracy comparable to existing methods (that use much higher resolution cameras) even when our sampling is sparser by an order of magnitude. To achieve perceptual photometric seamlessness, we used the work by Majumder, et al. [21]. To make the quality of the geometric registration visible in the grid images, we did not register the display photometrically for Figure 2(c) and Figure 3. So, these images, though geometrically aligned, still show photometric seams. The real-time rendering is demonstrated in the submitted video. However the video, being in VGA resolution, cannot capture the high resolution details.

4 USING PHOTOMETRICALLY UNCALIBRATED CAMERAS

In this section, we present a method for photometric registration of multi-projector displays using a photometrically uncalibrated camera. Projector photometric parameters that include its ITF and spatial intensity variation, need to be estimated to instrument existing photometric registration techniques [21]. We use an uncalibrated camera (whose
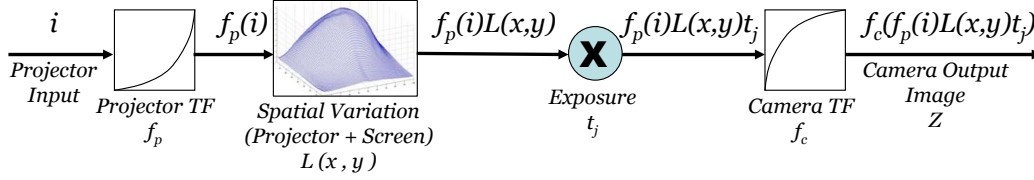
Fig. 4. The transformation process of an image as it passes through a projector-camera system.

ITF is unknown) to extract these projector photometric parameters. In the process, we also recover the camera ITF and hence calibrate the camera photometrically. To suppress the spatial intensity variation, such as vignetting effect, of the camera, like most previous work in scene reconstruction [8, 23], we operate the camera at a narrow aperture setting where it approaches the ideal pinhole model.

4.1 Model for a Single Projector

We assume a geometrically calibrated projector-camera system where a pixel (u,v) in the camera coordinate system is related to a pixel (x,y) in the projector coordinate system by a linear or non-linear warp $G(x,y) = (u,v)$. G can be determined by any geometric calibration method, including the one presented in the previous section [38].

Consider a spatially uniform grayscale input to the projector. Let the grayscale level be denoted by i. As per the model presented in the paper by Majumder and Gopi [18], the uniform image is first transformed by a spatially invariant input transfer function of the projector, f_p, to create a spatially uniform output, $f_p(i)$. Next, the projector optics introduce a spatially-dependent but input-independent intensity variation. This is further modulated by the screen reflectance/transmittance function. We denote this combined spatially dependent attenuation factor due to projector lens and screen as $L(x,y)$. Thus, the image produced after passing through the projector lens and screen is $f_p(i)L(x,y)$. The light from the screen then reaches the camera. The amount of light accepted by the camera is scaled by its exposure time t_j, where j indexes different exposure times. The different exposures are instrumented by changing the shutter speed of the camera. This produces an image $f_p(i)L(x,y)t_j$ that passes through the camera optics. Assuming negligible camera vignetting at narrow apertures, the image is transformed only by the input transfer function of the camera, f_c, to generate the grayscale value recorded by the camera Z. Thus, Z is a function of the input i, the exposure time index j, and the spatial coordinates (x,y). This is illustrated in Figure 4. The final equation is

$$Z(i,j,x,y) = f_c(f_p(i)L(x,y)t_j). \qquad (16)$$

For cameras, the intensity transfer function is monotonic [8], and hence invertible. Note that the same is not true for projectors [20]. Assuming invertible f_c, the above equation becomes

$$f_c^{-1}(Z(i,j,x,y)) = f_p(i)L(x,y)t_j. \qquad (17)$$

Taking the natural logarithm of both sides we get,

$$ln f_c^{-1}(Z(i,j,x,y)) = ln f_p(i) + ln(L(x,y)) + ln(t_j). \qquad (18)$$

To simplify the notation, we define $h_c = ln f_c^{-1}$ and $h_p = ln f_p$. The above equation then becomes

$$h_c(Z(i,j,x,y)) = h_p(i) + ln(L(x,y)) + ln(t_j) \qquad (19)$$

where i ranges over the grayscale inputs, j ranges over the exposure times, and (x,y) ranges over the spatial coordinates of the projector. In this equation, Z and t_j are known while h_p, h_c and L are unknown. Varying i and t_j results in different values of Z for each pixel (x,y). We can use this to setup a system of linear equations. We recover h_p, h_c and L that best satisfy Equation 19 in a least-squares sense. Note that recovering h_p and h_c involves solving the functions for a finite number of samples in the complete range of input values. The recovered parameters are illustrated in Figure 5.

Improving Efficiency: The system of linear equations achieved by Equation 19 is very large. Let R_i and R_Z each be the cardinality of the set of input samples for h_p and h_c respectively, P be the total number of projector pixels, and T be the number of exposures. This results in PTR_i equations and $P + R_i + R_Z$ unknown variables for the system of equations defined by Equation 19. Typically $R_i = R_Z = 256$ and $P = 1,000,000$ (assuming common XVGA projector resolution). Since there are multiple exposures for each input in the range R_i, the size of the linear system is on the order of at least a few million equations and cannot be solved efficiently.

To address this inefficiency we use a limited number of pixels in the projector space to solve for h_p and h_c. This is possible since h_p and h_c are both spatially constant. To ensure a sufficiently over-determined system, the criteria $PTR_i > P + R_i + R_z$ should be satisfied. For $R_i = R_Z = 256$ and $T = 6$, a choice of 100 for P is more than adequate. So, we first subsample L to a resolution of 10×10 pixel and use this to setup a smaller linear system of equations to solve for h_p and h_c.

With the estimated h_p and h_c, we substitute these into Equation 19 and quickly solve for $L(x,y)$ at the various projector coordinates, by rewriting Equation 19 as

$$ln(L(x,y)) = h_c(Z(i,j,x,y)) - h_p(i) - ln(t_j) \qquad (20)$$

Handling Noise: Our system does exhibit considerable random noise arising not only from the devices (camera and projector) but also from the screen. In particular, we use a rear projection screen with relatively high gain which has been shown to generate considerable random noise [20]. This has an adverse effect on the signal to noise ratio, especially for low projector input i or low camera output Z. Thus, estimating h_p and h_c directly from Equation 19 and L from Equation 20 results in noisy parameter estimation. We take several measures to reduce the effect of this random noise.

To assure smooth h_p and h_c while solving the system of equations, we minimize the error function

$$E = \sum_{j \in T} \sum_{(x,y) \in P} \sum_{i \in R_i} [(h_c(Z) - h_p(i) - ln(L(x,y)) - ln(t_j)]^2 \qquad (21)$$

$$+ \lambda \left(\sum_{Z \in R_Z} h_c''(Z)^2 + \sum_{i \in R_i} h_p''(i)^2 \right). \qquad (22)$$

The first term assures that the solution arises from the set of equations given by Equation 19 in a least-squares sense. The second term is a smoothing constraint on the curvature of h_p and h_c, given by their second derivatives. In the discrete domain, we use the Laplacian operator to find the curvature of h_p and h_c, i.e., $h_p''(i) = h_p(i-1) - 2h_p(i) + h_p(i+1)$. The scale factor λ weights the smoothness term relative to the data fitting term and should be chosen based on the amount of noise in Z. We choose $\lambda = 10$.

However, note that the first term in Equation 22 gives equal weights to all recorded camera values Z, and the second term gives equal weights to all i and Z. For the first term, we want to emphasize images with higher signal to noise ratio i.e. images for higher i with higher recorded values Z. For the second term, we want to emphasize smoothing of h_p and h_c in the lower ranges of i and Z respectively where low signal to noise ratio would result in a noisier estimate. To
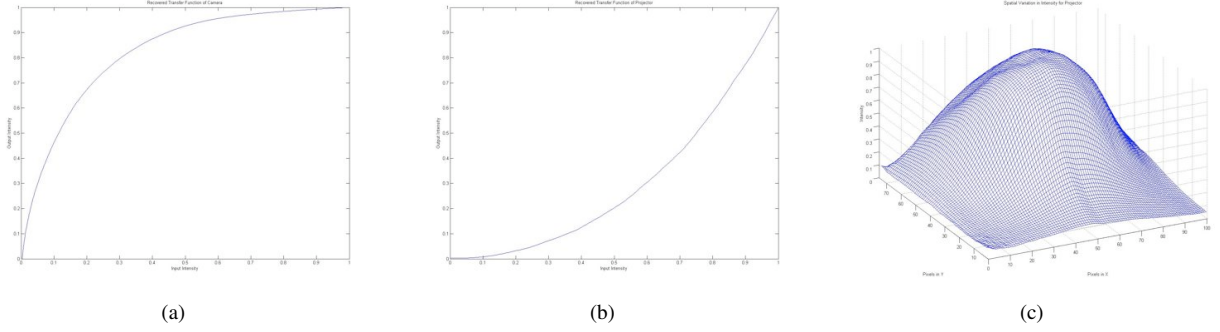
Fig. 5. (a) The estimated camera input transfer function f_c. (b) The estimated projector input transfer function f_p. (c) The estimated spatial intensity variation due to projector, screen, and camera L.

achieve this, we weight the first and second term of the error function in Equation 22 as follows,

$$E = \sum_{j \in T} \sum_{(x,y) \in P} \sum_{i \in R_i} [w_c(Z)(h_c(Z) - h_p(i) - ln(L(x,y)) - ln(t_j)]^2 \quad (23)$$

$$+ \lambda \left(\sum_{Z \in R_Z} [(Z_{max} - w_c(Z))h_c''(Z)]^2 + \sum_{i \in R_i} [(i_{max} - w_p(i))h_p''(i)]^2 \right). \quad (24)$$

where w_p and w_c are the weighting functions corresponding to the projector input and the recorded camera values, respectively, and Z_{max} and i_{max} are the maximum camera output and projector input respectively. To give the higher intensities greater confidence we use linear weighting functions, $w_c(Z) = Z$ and $w_p(i) = i$.

Similarly, noise can have adverse effects when estimating L using Equation 20 after h_p and h_c are recovered. Usually the image captured for input i at a particular exposure does not yield unsaturated outputs at all spatial pixel locations (x,y). So, we need to use different images for reconstruction of L at different spatial locations. This will yield an L with different signal to noise ratio at different spatial regions depending on the image that was used to reconstruct L in any spatial region. Even if a rare situation occurs when one finds a single image (at a particular input and exposure) where all the pixels are unsaturated, L will still be noisy due to the inherent random noise of the projector-screen-camera system.

Averaging multiple observations is the standard way to reduce random noise significantly [11]. Using this fact, we reduce noise in L at any pixel (x,y) by solving it from multiple images and finding a weighted mean of the multiple solutions. The weighting function should emphasize higher i and Z to reduce the effect of random noise. We achieve this by a weighting function $w_L(Z,i) = w_c(Z)w_p(i)$. The solution can now be derived by modifying Equation 20 as

$$ln(L(x,y)) = \frac{\sum_{j \in T} \sum_{i \in R_i} w_L(Z,i)[h_c(Z) - h_p(i) - ln(t_j)]}{\sum_{j \in T} \sum_{i \in R_i} w_L(Z,i)}, \quad (25)$$

resulting in clean estimated parameters as illustrated in Figure 5.

4.2 Photometric Registration of Multiple Projectors

For photometric registration of a multi-projector display, we estimate the f_p and $L(x,y)$ of each projector independently using the same camera. Note that in order to recover the relative intensity levels of the projectors, we use the unnormalized $L(x,y)$ estimated independently from each projector. Using the extracted parameters from each projector, we use the photometric registration method proposed by Majumder and Stevens [21] to generate a perceptually seamless display.

4.3 Implementation and Results

We used a Kodak DCS ProSLR/n camera and projected 32 flat grayscale fields with intensity levels uniformly sampled from 0 to 255. For each intensity level, 15 exposures were taken. The camera was set to a narrow aperture, f/32. The data collection took 25 minutes. To reduce the collection time, we tried reducing the number of exposures. Empirically, the minimum number of exposures needed before quantization effects became visible was eight. The exposures used, however, need to be well distributed amongst the range of available camera exposures. The program was written in C++ and utilizes Matlab and the OpenCV library. It takes about 11-15 minutes to register each projector on a Pentium 4, 2.8 GHz PC.

The photometric registration in Figure 2 (d) and (e) are achieved by using a photometrically uncalibrated camera. Additional results with projectors that do not exhibit lens distortion are shown in Figure 6. Here a linear geometric registration method is used [5]. These results show that a photometrically uncalibrated camera can be used irrespective of the kind of geometric registration method being used. When using the photometric registration technique of Majumder and Stevens [21] we use a low photometric uniformity parameter to retain the high contrast of the display as much as possible. This results in some perceivable seams near the boundaries of the bottom left and right projectors, which are significantly dimmer than the other projectors in our display. However, these seams exist even when the photometric registration is carried out with a calibrated camera, showing that they are artifacts of the photometric registration method itself and are not due to imperfect projector parameter estimation.

5 CONCLUSION

This paper presents new registration techniques to handle imperfect and uncalibrated devices in planar multi-projector displays. The four relevant cases include allowing (a) photometrically imperfect projectors, (b) geometrically imperfect projectors, (c) a photometrically uncalibrated camera, and (d) a geometrically uncalibrated camera. Our work addresses (b) and (c) in particular, while (a) is addressed by existing literature [21]. Our work allows severely distorted projectors to be used which opens up the possibility of mounting inexpensive lenses on projectors, just as with cameras, leading to very short-throw projectors. Using a photometrically uncalibrated camera allows the photometric registration to be conducted within a controlled indoor environment. Further, we have proposed the first method that estimates all projector/camera photometric parameters simultaneously. However, case (d) above of achieving geometric registration in multi-projector displays using a geometrically uncalibrated camera is still a challenging problem.

Further, our methods are within the realm of centralized architecture. Current distributed architectures [2] are very amenable to easy deployment, simple maintenance and hence more affordable multi-projector displays. Our goal is to extend our developed methodologies in this paper to the distributed framework of a network of projector-camera systems [2]. This poses a significant challenge which, if resolved, can be instrumental in realizing commodity tiled projection-based displays. All of these methodologies are stepping stones towards the vision of *ubiquitous pixels* in future workspaces.

Fig. 6. Images of our 3×3 array of nine projectors. (a,c) The geometrically registered display without any photometric registration. (b,d) Photometric registration achieved by using an uncalibrated camera.

REFERENCES

[1] Mitsubishi pocketprojector led dlp projector. Technical report, http://www.mitsubishi-presentations.com/proj_pocket.asp.

[2] E. S. Bhasker, P. Sinha, and A. Majumder. Asynchronous distributed calibration for scalable reconfigurable multi-projector displays. *IEEE Transactions of Visualization and Computer Graphics (TVCG)*, 12(5):1101–1108, 2006.

[3] J. Y. Bouguet. Camera calibration toolbox for matlab. Technical report, http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

[4] D. Brown. Decentering distortion of lenses. *Photometric Engineering*, 32(3):444 – 462, 1966.

[5] H. Chen, R. Sukthankar, G. Wallace, and K. Li. Scalable alignment of large-format multi-projector displays using camera homography trees. *Proceedings of IEEE Visualization*, 2002.

[6] Y. Chen, D. W. Clark, A. Finkelstein, T. Housel, and K. Li. Automatic alignment of high-resolution multi-projector displays using an uncalibrated camera. *Proceedings of IEEE Visualization*, 2000.

[7] K. Cornelis, M. Pollefeys, and L. V. Gool. Lens distortion recovery for accurate sequential structure and motion recovery. *European Conference in Computer Vision*, pages 186 – 200, 2002.

[8] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. *Proceedings of ACM Siggraph*, pages 369–378, 1997.

[9] T. Disz, M. Papka, and R. Stevens. Ubiworld: An environment integrating virtual reality. *Heterogeneous Computing Workshop, Geneva, Switzerland*, 1997.

[10] J. Fryer and D. Brown. Lens distortion for close-range photogrammetry. *Photogrammetric Engineering and Remote Sensing*, 52(1):51 – 58, 1986.

[11] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison Wesley, 1992.

[12] M. Hereld, I. Judson, and R. Stevens. Introduction to building projection-based tiled display systems. *IEEE Computer Graphics and Applications*, 2000.

[13] M. Hereld, I. R. Judson, and R. Stevens. Dottytoto: A measurement engine for aligning multi-projector display systems. *Argonne National Laboratory preprint ANL/MCS-P958-0502*, 2002.

[14] G. Humphreys, M. Houston, R. Ng, R. Frank, S. Ahem, P. Kirchner, and J. Klosowski. Chromium : A stream processing framework for interactive rendering on clusters. *ACM Transactions of Graphics*, 2002.

[15] H. G. Jung, Y. H. Lee, P. J. Yoon, and J. Kim. Radial distortion refinement by inverse mapping-based extrapolation. *Proceedings to 18th International Conference on Pattern Recognition(ICPR)*, 2006.

[16] J. Lai, A. Levas, P. Chou, C. Pinhanez, and M. Viveros. Bluespace: Personalizing workspace through awareness and adaptability. *International Journal of Human Computer Studies*, 57(5):415 – 428, 2002.

[17] K. Li, H. Chen, Y. Chen, D. W. Clark, P. Cook, S. Damianakis, G. Essl, A. Finkelstein, T. Funkhouser, A. Klein, Z. Liu, E. Praun, R. Samanta, B. Shedd, J. P. Singh, G. Tzanetakis, and J. Zheng. Early experiences and challenges in building and using a scalable display wall system. *IEEE Computer Graphics and Applications*, 20(4):671–680, 2000.

[18] A. Majumder and M. Gopi. Modeling color properties of tiled displays. *Computer Graphics Forum*, June 2005.

[19] A. Majumder, Z. He, H. Towles, and G. Welch. Achieving color uniformity across multi-projector displays. *Proceedings of IEEE Visualization*, 2000.

[20] A. Majumder and R. Stevens. Color nonuniformity in projection-based displays: Analysis and solutions. *IEEE Transactions on Visualization and Computer Graphics*, 10(2), March–April 2003.

[21] A. Majumder and R. Stevens. Perceptual photometric seamlessness in tiled projection-based displays. *ACM Transactions on Graphics*, 24(1), January 2005.

[22] T. Mitsunaga and S. Nayar. Radiometric self calibration. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1999.

[23] S. K. Nayar, H. Peri, M. D. Grossberg, and P. N. Belhumeur. A projection system with radiometric compensation for screen imperfections. *Proceedings of IEEE International Workshop on Projector-Camera Systems*, 2003.

[24] C. Pinhanez. The everywhere displays projector: A device to create ubiquitous graphical interfaces. *Proceedings of Ubiquitous Computing, Atlanta, Georgia*, 2001.

[25] C. S. Pinhanez, F. C. Kjeldsen, A. Levas, G. S. Pingali, M. E. Podlaseck, and P. B. Chou. Ubiquitous interactive graphics. Technical Report RC22495 (W0205-143), IBM Research, 2002.

[26] A. Raij, G. Gill, A. Majumder, H. Towles, and H. Fuchs. PixelFlex2: A Comprehensive, Automatic, Casually-Aligned Multi-Projector Display. *IEEE International Workshop on Projector-Camera Systems*, 2003.

[27] R. Raskar. Immersive planar displays using roughly aligned projectors. In *Proceedings of IEEE Virtual Reality 2000*, 1999.

[28] R. Raskar, M. Brown, R. Yang, W. Chen, H. Towles, B. Seales, and H. Fuchs. Multi projector displays using camera based registration. *Proceedings of IEEE Visualization*, 1999.

[29] R. Raskar, J. van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines. ilamps: Geometrically aware and self-configuring projectors. *ACM Transactions on Graphics*, 22(3), 2003.

[30] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future: A unified approach to image based modeling and spatially immersive display. In *Proceedings of ACM Siggraph*, pages 168–176, 1998.

[31] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, 1995.

[32] J. Salvi, X. Armangu, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35:1617 – 1635, 2002.

[33] P. Steurer and M. B. Srivastava. System design of smart table. *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, 2003.

[34] R. Sukthankar, R. Stockton, and M. Mullin. Smarter presentations: Exploiting homography in cameraprojector systems. *Proceedings of International Conference on Computer Vision (ICCV)*, 2001.

[35] R. Swaminathan and S. K. Nayar. Nonmetric calibration of wide-angle lenses and polycameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1172 – 1178, 2000.

[36] J. Underkoffler, B. Ullmer, and H. Ishii. Emancipated pixels: Real-world graphics in the luminous room. *Proceedings of ACM Siggraph*, 1999.

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

[38] R. Yang, A. Majumder, and M. Brown. Camera based calibration techniques for seamless multi-projector displays. *IEEE Transactions on Visualization and Computer Graphics*, 11(2), March-April 2005.

[39] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientation. *IEEE International Conference on Computer Vision*, pages 666 – 673, 1999.