MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

A Self-Correcting Projector

Ramesh Raskar, Paul Beardsley

TR2001-46 December 2001

Abstract

We describe a calibration and rendering technique for a projector that can render rectangular images under keystoned position. The projector utilizes a rigidly attached camera to form a stereo pair. We describe a very easy to use technique for calibration of the projector-camera pair using only black planar surfaces. We present an efficient rendering method to pre-warp images so that they appear correctly on the screen, and show experimental results.

IEEE Computer Vision and Pattern Recognition (CVPR) 2001, Hawaii, Dec 2001

Submitted May 2001, revised and released Septemeber 2001.

IEEE Computer Vision and Pattern Recognition (CVPR) 2001, Hawaii, Dec 2001

A Self-Correcting Projector

Ramesh Raskar Paul Beardsley

{raskar, pab}@merl.com

Mitsubishi Electric Research Laboratories (MERL),

Cambridge, MA 02139, USA

Abstract

We describe a calibration and rendering technique for a projector that can render rectangular images under keystoned position. The projector utilizes a rigidly attached camera to form a stereo pair. We describe a very easy to use technique for calibration of the projector-camera pair using only black planar surfaces. We present an efficient rendering method to pre-warp images so that they appear correctly on the screen, and show experimental results.

Keywords: *Projector, camera, calibration, homography.*

1 Introduction

A data projector is a dual of a camera, and the image projection process can be expressed using the standard pinhole camera model. Thus far, however, projectors have received little attention in the computer vision or projective geometry community. In this paper, we present a technique to calibrate a projector, and associated camera, by exploiting planar homographies. We then demonstrate a method for correctly displaying rectangular images on a planar surface even under oblique projection.

The internal and external parameters of a projector can be expressed in the same way as those of a camera. But there are two main differences between a projector and a camera. First and most obviously, projectors cannot view the illuminated surface. Thus, while it is easy to calibrate a projector given the correspondence between six or more 2D projector pixels and corresponding 3D points on a known target object [Faugeras93], this requires a tedious manual process to select those projector pixels which illuminate the 3D points. To automate this or any other semi-automatic method, it is clear that a sensing device such as a camera must be used. The second main difference between a projector and camera is that a traditional assumption about simplified camera models, that the principal point is close to the image center, is not valid for projectors. Most projectors use an off-axis projection. When they are set on a table or mounted upside-down on a ceiling, the image is projected through the upper or the lower half of the lens, respectively. Hence, the principal point is vertically shifted.

In this paper, we apply techniques for the epipolar geometry of a pair of cameras to a projector-camera system. Then we describe a rendering technique which pre-warps the projected image so that it appears correctly on the screen. The calibration technique is designed to be very simple to use, and is influenced by recent trends to employ planar patterns to calibrate cameras [Zhang99][Str-May99]. A novelty of the system here is that we calibrate using blank planes onto which calibration patterns are projected.

There would be potential benefits to attaching two (or more) cameras to a projector, and using stereo, but we are investigating a

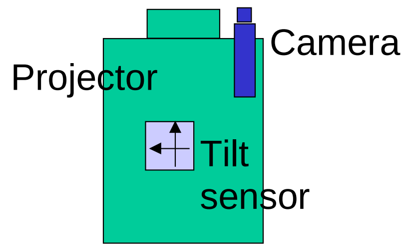

Figure 1. The projector is augmented with a camera and a tilt sensor. (Top) a schematic diagram (bottom) our prototype, with attached camera on the left.

system in which only a single camera is rigidly attached to the projector. The primary reasons for this is cost, which would be incurred not only because of the extra camera, but also through the extra requirements on processing power and bandwidth. Using two cameras would also incur more extensive modifications to the design and volume of an existing projector.

1.1 Self-correcting projection

The purpose of our self-correcting projector is to generate a rectangular image of known aspect ratio, even when aimed at an arbitarily inclined planar surface. Furthermore, for a vertical or near-vertical display surface, the sides of the rectangle should be aligned appropriately with the world vertical. Our approach is unique due to two features. First, we can project on blank walls i.e. without fixed markers to create two-dimensional Euclidean frame of reference. Second, we use a single self-contained device. Hence, we wish to avoid the use of devices that are external to the projector, such as calibrated cameras mounted at known locations (e.g. [Raskar99][Surati99]). To allow self-correcting projection on blank screens, we wish to avoid placing physical markings on the screen or explicitly detecting the extents of the pre-defined display rectangle (e.g. [Raskar00][Sukthan01]). We created the simple prototype in Figure 1 by taking a standard commercial projector and rigidly mounting a low-cost camera and tilt sensors to it. The system is calibrated with the technique described in

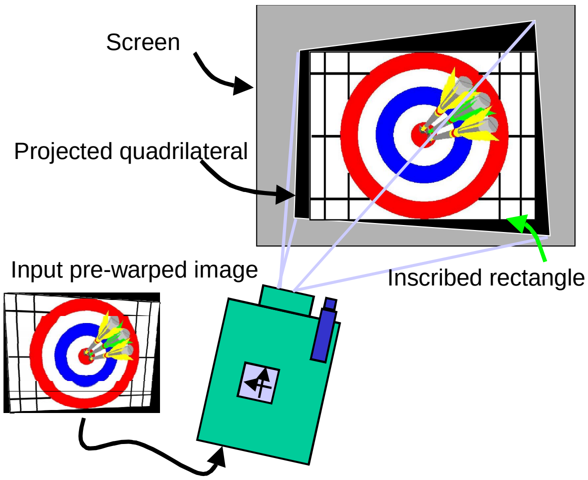

Figure 2. Keystoned projector generates a quadrilateral on the display plane. A pre-warped image creates a correct view inside the inscribed rectangle.

Section 2. The rendering technique is described in Section 3, and involves pre-warping of images using the homography between projector and display surface, so that the projected image appears correct.

2 Calibration

Our current approach is to carry out a full calibration for the projector-camera system, with computation of the intrinsic parameters for both devices plus their relative pose. The rendering process will utilize this information in conjunction with gravity based tilt sensing (which provides the world vertical direction), as described in section 3.

The intrinsics of the camera are obtained from multiple views of a planar calibration pattern, using the techniques described in [Zhang99][Str-May99]. The remaining calibration requires two or more homographies between the projector and camera. To obtain a homography, the projector displays a calibration pattern on a blank planar surface, and four or more point (or line) correspondences between the projector image and the camera image are used to compute the homography for that plane [Hartley97]. This is repeated for different positions of the system relative to the plane, to obtain two or more distinct homographies.

2.1 Projective reconstruction

First, we review how a pair of homographies can be used to construct a scene up to a projective transformation. Given two cameras, viewing points on a 3D plane Π, the point positions in the two images are related by a 3x3 homography matrix H. If m_1 and m_2 are projections of a 3D point M which belongs to Π, then

$$m_2 \sim= H\, m_1$$

where $m1$ and $m2$ are homogeneous coordinates and $\sim=$ means equality up to scale. Given two distinct planes, and hence two distinct homographies, the epipoles e_1 and e_2 in the two images can be computed using the generalized eigenvalue equation

$$(e_1 \sim=) \qquad k\, H_1^{-1}\, e_2 \;=\; H_2^{-1}\, e_2$$

where k is an unknown scalar [Johan99].

The perspective projective matrices of the two cameras can then be defined as follows, for a projective reconstruction of the scene,

$$P_{1p} = [I \mid 0] \qquad P_{2p} = [H \mid e_2]$$

where, H is one of the homographies [Faugeras93]. In our system, P_{1p} defines the projection matrix for the camera and P_{2p} the projection matrix for the projector.

2.2 Euclidean reconstruction

In this section, the projection matrices are upgraded to give a Euclidean reconstruction. The goal is to find a 4x4 matrix G_p such that

$$P_{1e} = P_{1p}\, G_p \sim= A_1\, [\, I \mid 0\,]$$

$$P_{2e} = P_{2p}\, G_p \sim= A_2\, [R \mid -R\, t]$$

where, A_1 is a 3x3 matrix describing the camera intrinsics, A_2 is a 3x3 matrix describing the projector intrinsics, and R and t define the relative pose between the camera and the projector up to unknown scale.

In fact A_1 is known and can be factored out, so the goal is to find matrix G such that

$$P_{1e} = P_{1p}\, G \sim= [\, I \mid 0\,] \qquad (1)$$

$$P_{2e} = P_{2p}\, G \sim= A_2\, [R \mid -Rt]$$

The most general form of A_2 would involve five intrinsic parameters: focal length, aspect ratio, principal point and skew angle (we currently ignore radial distortion). It is reasonable to assume that a projector has unity aspect ratio and zero skew. However, an assumption that the principal point is close to the image center, common for cameras, is not true for projectors where the principal point usually has a large vertical offset from the image center (Figure 3). Thus A_2 has the form,

$$A_2 = [\; f \;\; 0 \;\; 0; \;\; 0 \;\; f \;\; d; \;\; 0 \;\; 0 \;\; 1\;]$$

From (1), G is of the form

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ n1 & n2 & n3 & 1 \end{bmatrix}$$

(Here the vector $n = [n1\ n2\ n3]^T$ defines the plane at infinity in the projective coordinate frame.)

If G' denotes the first three columns of G, then it follows

$$P_{2p}\, G' \sim= A_2\, R$$

Hence

$$P_{2p}\, G'\, G'^T\, P_{2p}^T \sim= A_2\, R\, R^T\, A_2^T = A_2\, A_2^T$$

This leads to

$$P_{2p} \begin{bmatrix} I_{3x3} & n \\ n^T & n^T n \end{bmatrix} P_{2p}^T \cong A_2 A_2^T \qquad (2)$$

where

$$K_2 = A_2 A_2^T = [\ f^2 \ \ 0 \ \ 0; \ \ 0 \ \ (f^2+d^2) \ \ d; \ \ 0 \ \ d \ \ 1\ ]$$

Equation (2) can be used to generate three constraints on the three unknowns of n. Two of the constraints ($K_2(1,2) = 0$ and $K_2(1,3) = 0$) are linear in $n1$, $n2$, $n3$, and $n^T n$. The third constraint ($K_2(2,2) - K_2(1,1) - K_2(2,3)^2 = 0$) is quadratic. Hence, it is possible to express $n1$, $n2$, and $n3$ in terms of $n^T n$. Using the quadratic constraint ($n1^2 + n2^2 + n3^2 = n^T n$) generates four solutions for the three unknowns of n. Each solution is then used with equation (2) to compute f and d. Finally, equation (1) is used to compute R and t. Physically impossible solutions (e.g. solutions in which observed scene points are behind the camera) are eliminated to give a single solution for the true A_2, R and t. A similar approach to the above, but using different assumptions for the intrinsic parameters, for the generation of constraints on calibration parameters, is found in [Xu00].

3 Rendering

Our goal is to display a rectangular image with the aspect ratio 4:3 that is aligned with the world horizontal and vertical directions. Using a calibrated pair of a camera and a projector, we find the relationship between an illuminate planar surface and the projected image. Then we compute the largest axis aligned rectangle inscribed in the projected quadrilateral on the plane and display strictly inside this rectangle (Figure 2 and 4). The remaining pixels remain black.

The reconstruction of the illuminated quadrilateral, however, is valid up to a similarity transform. Pure translation and scale do not affect the relationship between the illuminated quadrilateral and the computed inscribed rectangle. But, the three degrees of freedom for rotation guide the choice of the inscribed rectangle. We do not have a notion of the world horizontal or the vertical direction purely from the camera-projector pose. Hence, we add two gravity based tilt sensors to find the roll and elevation of the projector, i.e. the orientation of the projector image array out of the vertical plane. We exploit the fact that the illuminated surface is usually a vertical wall or screen. If the flat screen is not in a vertical plane, we can still get a rectangular image by ignoring the measured tilt angles.

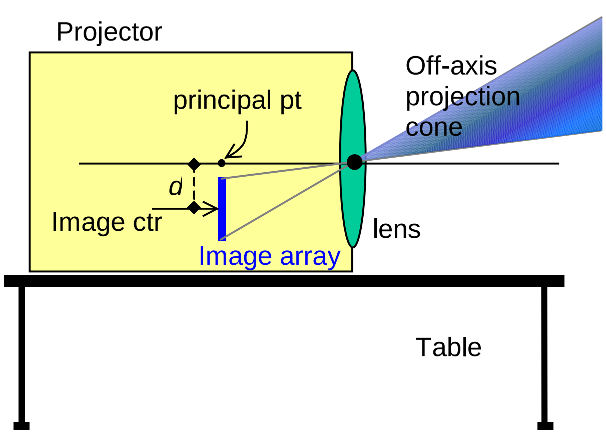

Figure 3 : The projector optics is designed for image projection in a off-axis cone so that a projector can be kept on a table. The principal point and image center are separated by a large vertical shift.

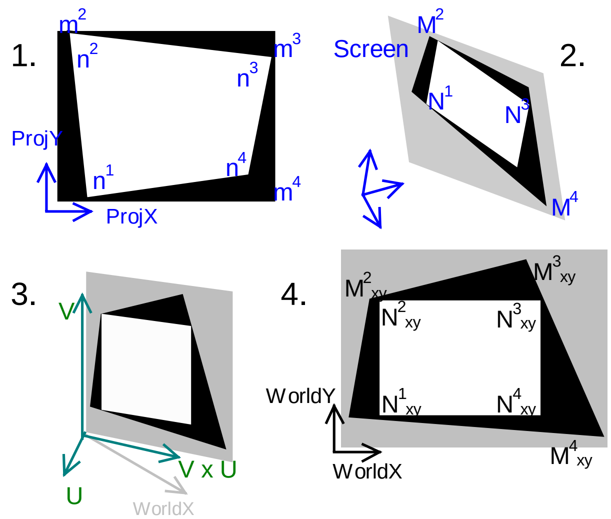

Figure 4 : The relationship between projected image and illuminated quadrilateral. The homography is used for pre-warping.

Consider a world coordinate system where the ground plane is parallel to the x-z plane, the illuminate vertical wall is parallel to the x-y plane, and center of projection of the projector is the origin. In the following discussion, we first discuss how to find the axis aligned inscribed rectangle in the world coordinates and then describe the rendering process.

3.1 Display region

Finding inscribed rectangle for a quadrilateral is a 2D problem. The illuminated quadrilateral is specified in projector coordinate system. Hence, we need to rotate the quadrilateral using the correct notion of the world horizontal and vertical direction. Then we need to transform the quadrilateral into a convenient plane parallel to x-y plane, so that the rectangular display region can be computed using simply the x and y coordinates. The steps are as follows.

The 3D coordinates of the corners of the quadrilateral illuminated by the projector are computed by triangulation. Given corresponding points m_1 and m_2 in the camera and the projector image, respectively, the 3D coordinates of the point M are $z_1\, m_1$ and $z_2\, m_2$, where z_1 and z_2 are the depths.

$$z_2\, m_2 = R(z_1\, m_1) + t$$

The equation

$$[R\, m_1 - m_2]\, [z_1\, z_2]^T = -t$$

leads to the solution

$$[z_1\ z_2]^T = -(D^TD)^{-1}D^T\ t$$

where, $D = [R\ m_1 - m_2]$.

Let us denote the corner points of projector framebuffer by m^i, $i=1,..,4$ (Figure 4.1) and the corresponding corners of the illuminated quadrilateral in projector coordinate system by M^i, $i=1,..,4$ (Figure 4.2). The rotation of the projector image array out of the vertical plane, R_{Tilt}, is measured by tilt sensors. R_{Tilt} can be used to transform the quadrilateral corners into points that are aligned with the horizon and lie in a vertical plane (Figure 4.3). When 3D points are aligned with the horizon, the y-coordinate represents distance from an arbitrary horizontal plane.

$$M^i_{pi} = R_{Tilt}\ M^i$$

Next, we transform the vertical plane into a plane parallel to the vertical x-y plane (Figure 4.4). This is essentially a single rotation about the world y-axis. The rotation matrix can be computed as follows. The plane normal is

$$U = (M^2_{pi} - M^1_{pi}) \times (M^3_{pi} - M^1_{pi})$$

The vertical vector in the plane is $V = [0\quad 1\quad 0]^T$. Hence the horizontal vector in the plane is $V \times U$. The rotation matrix then is

$$R_{pi\text{-}xy} = [\ (V \times U)^T;\quad V^T;\quad U^T]$$

The computed points on the plane parallel to x-y plane are:

$$M^i_{x\text{-}y} = R_{pi\text{-}xy}\ M^i_{pi} = R_{pi\text{-}xy}\ R_{Tilt}\ M^i$$

If the points M^i_{pi} are on the plane $[U^T, -s]$, then the z coordinate for all four $M^i_{x\text{-}y}$ is equal to s. Ignoring the z coordinate, we get a quadrilateral in 2D, specified by the x and y coordinates (Figure 4.4). We are now ready to compute the largest axis aligned rectangle of aspect ratio 4:3 inscribed in the quadrilateral $M^i_{x\text{-}y}$. We assume that at least one vertex of the rectangle lies on the edges of the quadrilateral. We discretize the edges of the quadrilateral and solve for the largest rectangle by testing for each possible position of the rectangle vertex. Let us say the corners of the inscribed rectangle are $N^i_{x\text{-}y}$, $i=1,..,4$.

3.2 Pre-warping using homography

The pre-warping required for the projected image is defined by the homography between pixel coordinates of the corners of projector image and pixel coordinates of the projection of the chosen inscribed rectangle in projector image space.

The coordinates of the rectangle corner in projector coordinate system, N^i, can be easily computed using

$$N^i = (R_{pi\text{-}xy}\ R_{Tilt})^{-1}\ N^i_{x\text{-}y}$$

The pre-warping can be achieved using two different methods.

In the first method, we compute projection of N^i in projector image space

$$n^i \sim= A_2\ N^i$$

Then compute the 3x3 homography matrix $H_{m\text{-}n}$ mapping projector image corners to projection n^i of corners of the chosen display area.

$$n^i \sim= H_{m\text{-}n}\ m^i$$

To pre-warp an image, load the image as a texture and texture map a unit rectangle in x-y plane (i.e. with extents [-1:1,-1:1]) . Since $n^i \sim= H_{m\text{-}n}\ A_2\ M^i$, use the projection matrix

$$H_{m\text{-}n}\ A_2\ [\ I\ |\ 0]$$

to render the appropriately pre-warped transformation of the input image. The technique to transform the 3x3 projection matrix into a 4x4 matrix used in traditional graphics pipeline is described in [Raskar00]. One needs to take care of transformation of depth buffer values when homography matrix is involved.

We introduce a much simpler second technique for pre-warping. Note that $n^i \sim= A_2\ N^i$. Hence, we load the image as a texture and texture map the inscribed rectangle N^i, specified in projector coordinate system. Use the projection matrix

$$A_2\ [\ I\ |\ 0]$$

for rendering.

4 Implementation

We implemented the system using a Mitsubishi X-80 projector, a very low cost Logitech Quickcam Pro camera and a 2-axis tilt sensor by Crossbow (Figure 1). The projector resolution is 1024x768 and camera resolution is 640x480. The values from tilt sensor are averaged to improve the accuracy. The calibration results show that, the focal length of the projector can be recovered with high accuracy (verified by repeatability). The computed vertical shift of the principal point shows greater variation, but this does not seem to affect the alignment of the projected pre-warped image.

Using a current generation graphics card (ATI Radeon) on a PC, we have implemented a rendering system that loads images, updates the texture memory and displays pre-warped images in real time. The rendering technique of perspectively correct texture mapping exploits the high quality texture filtering available on the graphics cards. As seen in Figure 5, we can skew the projector with respect to the horizontal flat surface and still correct the image so that it appears aligned and rectangular with correct aspect ratio. For comparison, we show the outline of the keystoned image in the bottom window. The vertical edges in the pre-warped image (in this case a PowerPoint slide) appear to be parallel to the edges of the whiteboard.

5 Conclusion

This paper has described a projector, augmented with a camera and tilt-sensing, which can perform physically correct display onto planar surfaces of arbitrary inclination. The system applies a corrective warp to projected data so that the result is a true rectangle of known aspect ratio, correctly aligned with the world vertical and horizontal directions. The resulting rendering is real-time. This functionality could only be realized automatically by incorporating computer vision into the device. As far as we know, there are no easy to use methods available to calibrate a projector. Our two main contributions are (i) a technique to calibrate a projector using blank planes and (ii) a complete corrective warping technique using tilt sensors and homographies.

The traditional use of projectors has been for group presentations. But new ideas such as the 'office of the future' have indicated how projected information can be incorporated into the everyday environment. Furthermore, portable computing devices are become more and more compact. Having the ability to easily setup such devices with a projector allows a larger display area when required, to aid interaction, or for sharing of data with other people. Quick setup and ease-of-use of a projector becomes essential if it is intended for casual, transient use on table-tops, walls and other surfaces.

Finally, our current research includes further work on the calibration process, and extension of the current ideas to handle automatic mosaicing of data [Raskar02] from multiple projectors.

Figure 5. The keystone correction. (Top) A projector is casually placed so that the image array is not vertical (Middle) The projector displays a quadrilateral on a vertical screeen (Bottom) The correction based on feedback from the tilt sensors and the computed homography allows us to correct the image in a few seconds.

References

[Faugeras93] O. Faugeras. "Three-Dimensional Computer Vision: A Geometric Viewpoint" MIT Press, Cambridge, Massachusetts, 1993.

[Hartley97] Hartley, R., "In defense of the eight-point algorithm", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.19, No.6, pp.580-593, 1997.

[Johan99] Johansson, B., "View synthesis and 3D reconstruction of piecewise planar scenes using intersection lines between the planes", Proc. 7th Int. Conference on Computer Vision, pp.54-59, Corfu, Greece, Sept., 1999.

[Raskar99] 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 IEEE Visualization, pages 161-68, San Francisco, October 1999.

[Raskar00] Raskar, R., "Immersive planar display using roughly aligned projector", In IEEE Virtual Reality, March 2000.

[Raskar02] R.Raskar, J. Chai, J. vanBaar. "Low cost projector mosaics", To appear in Proceedings of Fifth Asian Conference on Computer Vision, January 2002.

[Str-May99] Sturm, P. and Maybank, S., "On plane-based camera calibration: A general algorithm, singularities, applications", Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp.432—437, Fort Collins, Colorado, June 1999.

[Sukthan01] R. Sukthankar, R. Stockton, and M. Mullin., "Smarter presentations: Exploiting homography in camera-projector systems", In Proceedings of International Conference on Computer Vision, 2001.

[Surati99] R. Surati. "Scalable Self-Calibrating Display Technology for Seamless Large-Scale Displays", PhD thesis, Massachusetts Institute of Technology, 1999.

[Xu00] Xu, Gang, Terai Jun-Ichi and Shum, Heung-Yeung., "A linear algorithm for camera self-calibration, motion and structure recovery for multi-planar scenes from two perspective images", Proc. of IEEE Conference on Computer Vision and Pattern Recognition, 2000.

[Zhang99] Zhang, Z., "Flexible camera calibration by viewing a plane from unknown orientations", Proc. 7th Int. Conference on Computer Vision, pp. 666-673, 1999.