Conformal Rectification of Omnidirectional Stereo Pairs

Christopher Geyer
Department of EECS, UC Berkeley
Berkeley, CA 94720
cgeyer@eecs.berkeley.edu

Kostas Daniilidis
GRASP Laboratory, U. of Pennsylvania
Philadelphia, PA 19104
kostas@cis.upenn.edu

Abstract

A pair of stereo images are said to be rectified *if corresponding image points have the same y-coordinate in their respective images. In this paper we consider the rectification of two omnidirectional cameras, specifically two parabolic catadioptric cameras. Such systems consist of a parabolic mirror and an orthographically projecting lens. We show that if the image coordinates are represented as a point z in the complex plane, then the rectification is specified by $coth^{-1} z$. This rectification is shown to be conformal, in that it is locally distortionless, and furthermore, it is unique up to scale and transformation. We show an experiment in which two real images have been rectified and a stereo matching performed.*

§ 1 Introduction

Consider the two projections x and x' in two arbitrary cameras, each of which is the projection of a single point X in space. As we know, as X varies over all of three-dimensional space, the pair (x, x') does not fill up the space of all possible pairs of image points. Such a pair must satisfy the *epipolar constraint*, which is determined by the difference in the positions and orientations of the two cameras, as well as their intrinsic parameters, e.g. focal length, etc. Given a likely feature at x_0 it is the task of *stereo matching* to find the point x' which is most likely to be the projection of the same point as that of x_0. Typically some metric based on the similarity of the neighborhoods of these points, such as cross-correlation, is used to determine a likely match. Fortunately, one only needs to search among those pairs (x_0, x') satisfying the epipolar constraint. In the case of an ideal perspective camera, such a locus is a line, an *epipolar line*, and so the search is one-dimensional. The set of all epipolar lines in one image intersects in a single point.

To perform dense stereo matching, this procedure is usually performed at every point in the two images. Therefore many methods have been investigated to increase the efficiency of this procedure. The most common approach is to perform a transformation of the images called a *stereo rectification*. The goal of such a procedure is to ensure that all of the points on one epipole line lie on the same row in the transformed image. In addition, this transformation can be performed so that points on corresponding epipolar lines have the same y-coordinate in each transformed image. Such a procedure offers the following advantages. First, the stereo matching can be performed without the overhead of determining the epipolar lines which determine the range to search. Second, there is no need to perform a complicated warping of neighborhoods on which to perform the cross-correlation or other similarity metric. In the case of perspective images, the rectification can be performed so that the transformed images could have been perspective images, indeed the rectifying transformations are homographies.

In this paper we are interested in performing a rectification of two *uncalibrated* parabolic catadioptric cameras. Such a sensor is a combination of a radially symmetric mirror with a parabolic profile and an orthographically projecting camera [11]. Their primary advantage is their large field-of-view, often in excess of $180°$, and as a result have applications in robotics [15] and telepresence [13]. Under ideal conditions, such as true orthographic projection and a perfect mirror, this system has a single effective viewpoint at the focus of the paraboloid. This is a realistic assumption equivalent to the common assumption that a perspective camera obeys the pinhole projection model. In particular, it has been shown that small deviations from the ideal do not grossly affect estimation of the epipolar geometry.

Stereo rectification has long been used by photogrammetrists [14]. In the vision community, [1, 9] were among the first works to consider rectification in software. Recent work has been done to choose the best rectification using a given metric, e.g. the most area-preserving [8], the least skew inducing [2]. In [7] a system consisting of a pair of *planar* catadioptric sensors is designed in which the images are automatically rectified, no warping of the images is necessary. We are not aware, though, of literature on the stereo rectification of parabolic catadioptric images.

The rectification proposed in this paper is motivated by an observation of the geometry induced by such catadioptric cameras. It can be shown that the parabolic catadioptric projection is conformal. By this we mean that small neighborhoods of any image point always look rectified. In particular, any angle defined on a sphere centered at the viewpoint will be preserved when it is projected to the image plane. This follows because the projection from such a sphere to the image plane is equivalent to stereographic projection [4], i.e. the projection from the sphere's north pole, which is conformal [12] and has been shown to be an advantage in computer vision [3]. This suggests considering rectifying transformations which preserve the conformality of the projection.

The geometry induced by the camera forces us to consider new techniques for rectification. In particular, with the method we propose, at the end of the procedure the resulting image is *not* equivalent to either a parabolic catadioptric or perspective projection. Regarding the first possibility, no such transformation exists because the locus of points satisfying the epipolar constraint are always *epipolar circles*. Any equivalent parabolic projection will have this property; there are no parabolic projections where the epipolar curves become a pencil of parallel lines. With respect to a perspective projection, this is not possible to achieve without having calibrated the catadioptric camera. Furthermore, it would not be wise to do so because of the wide field of view — an equivalent perspective image would have to be infinite in extent and would still lose information.

The transformation we consider is based on the *bipolar coordinate system* [10] and is easily represented as a transformation of the complex plane. In particular, if the image point (x, y) is represented by the complex point $z = x + i$, then the transformation will be roughly of the form $w = \coth^{-1} z$. This function is analytic in a region of $\mathbb{C}$ which excludes the epipoles and so it is conformal. We will show that this is the *only* conformal rectification of a parabolic catadioptric image.

§ 2 Parabolic catadioptric projection

The projection model is not key to our discussion but we include it nevertheless. It is assumed that the projection of any point in $\mathbb{R}^3$ is given by [4]:

$$p(x,y,z) \quad = \quad \begin{pmatrix} c_x + \frac{fx}{-z+\sqrt{x^2+y^2+z^2}} \\ c_y + \frac{fy}{-z+\sqrt{x^2+y^2+z^2}} \end{pmatrix} .$$

Note that unlike perspective projections, projections of points on opposite sides of the viewpoint are not equal. In particular, $p(x, y, z) \neq p(-x, -y, -z)$.

Briefly, consider the set S^2 containing those points in $\mathbb{R}^3$ satisfying $x^2 + y^2 + z^2 = 1$, i.e., the unit sphere. Assume that a point $X = (x, y, z)$ lies on the sphere and that its projection is (u, v). We can write the space coordinates in terms of the image coordinates as follows:

$$\left(\frac{2u}{1+u^2+v^2}, \frac{2v}{1+u^2+v^2}, \frac{-1+u^2+v^2}{1+u^2+v^2} \right), \qquad (1)$$

where $p|_{S^2}$ denotes the restriction of the domain of p to S^2. In effect, we have provided a rational parameterization of the sphere minus the north pole. Compute the Jacobian $J = \partial p|_{S^2}{}^{-1}/\partial(u, v)$ and notice that it satisfies $JJ^T = \lambda I$. Therefore the Jacobian from the sphere to the image plane is a similarity transformation. Consequently, parabolic catadioptric projections are *conformal* from the sphere to the image plane, that is, locally they preserve angles.

§ 3 Perspective epipolar geometry

The purpose of this section is to review the *epipolar geometry* of perspective sensors. In the perspective case the epipolar geometry can be derived from the properties of the bilinear form which corresponding image points satisfy, namely, that if $x_1, x_2 \in \mathbb{P}^2$ are projections of the same point X in space, then

$$x_2{}^T F\, x_1 \quad = \quad 0 \qquad (2)$$

where F is the *fundamental matrix* (for perspective cameras). The fundamental matrix is a 3×3 matrix which has rank equal to two. The epipolar geometry of the stereo pair is encoded in this matrix.

To each image is associated an *epipole*, which is the projection of the viewpoint of the opposite camera. If the viewpoints are equal then the epipoles are undefined and furthermore the epipolar geometry is degenerate. The epipoles e_1 and e_2 can be found from the nullspaces of F as follows:

$$F\, e_1 \;=\; 0 \;\text{ and }\; F^T e_2 \;=\; 0\,,$$

If x_1 is held constant in (2), then the locus of points x_2 satisfying the constraint is called an *epipolar line*. This line specifies the set of all possible correspondences, and for example the epipolar line in the second image is the set of all x satisfying

$$\ell_2^T x \;=\; 0 \;\text{ where }\; \ell_2 \;=\; F x_1\,.$$

Similarly $\ell_1 = F^T x_2$ is an epipolar line in the first image. Evidently every epipolar line in a given image contains the epipole of that image.

Rectification in perspective images is performed by finding two homographies H_1 and H_2 which:

1. take e_1 to $H_1 e_1 = (1, 0, 0)$ and e_2 to $H_2 e_2 = (1, 0, 0)$ both of which are on the line at infinity (ℓ_∞), this has the effect of mapping each pencil of epipolar lines to a set of parallel lines; and

2. are such that if $x_2^T F x_1 = 0$ then

$$(H_1 x_1)_2/(H_1 x_1)_3 = (H_2 x_2)_2/(H_2 x_2)_3 ,$$

in other words the transformed points have the same y-coordinate in the image plane.

Combinations of H_1 and H_2 compatible with F and these two conditions are not unique. Several conditions have been proposed which choose optimal homographies based on their effects on the images, e.g., choosing homographies which preserve area as much as possible [8] or which minimize the introduction of skew [2].

§ 4 Parabolic catadioptric epipolar geometry

Systems whose projections are given by (1) exhibit the property that the image of any line in space is an arc of a circle. Consequently, in parabolic catadioptric images, epipolar lines become *epipolar circles*. Furthermore, since the projections of antipodal points are separate, the epipole in one perspective image becomes two epipoles in parabolic images, one for each side of the viewpoint. The a pencil of epipolar lines in the perspective image becomes a system of coaxal circles in the parabolic image. All circles in such a system have centers which are collinear and all intersect in two points, which in this case are the epipoles.

To determine these properties one examines the epipolar constraint for parabolic catadioptric cameras. It can be shown that if x_1 and x_2 are the parabolic catadioptric projections of the same point X in space then they must satisfy

$$\tilde{x}_2^T F \, \tilde{x}_1 \;\; = \;\; 0 \, . \qquad (3)$$

for some *catadioptric fundamental matrix* F where if $x = (u, v)$ then

$$\tilde{x} \;\; = \;\; (2u, 2v, u^2 + v^2 - 1, u^2 + v^2 + 1) \, .$$

Note that this is the same as (1), but expressed in homogeneous coordinates; $\tilde{x}$ lies on the unit sphere in projective space. This paper is not about the properties of F nor its estimation, so we will limit the discussion to its most basic properties; for information on its properties see [5, 6]. It has been shown that F is a 4×4 matrix whose rank is equal to two, and it can be written as $F = A_2^T E A_1$ where E is a perspective essential matrix and A_1 and A_2 are two 3×4 matrices.

Since F is rank two, its left and right two-dimensional nullspaces represent lines in projective space. For a valid F, these lines must intersect the sphere in two points which are the representations of the epipoles. In particular two *epipole pairs* $(e_{1,1}, e_{1,2})$ and $(e_{2,1}, e_{2,2})$ satisfy

$$F \, \tilde{e}_{1,j} \;\; = \;\; 0 \;\; \text{and} \;\; F^T \tilde{e}_{2,j} \;\; = \;\; 0 \, .$$

The reason for the two points is that the viewpoint of the other camera may be known to lie on a line through the viewpoint of the first camera, however, the side of the line on which the other viewpoint may be found is ambiguous.

Suppose that x_1 is constant and that x is chosen so as to satisfy the equation

$$\gamma_1^T \tilde{x} \;\; = \;\; 0 \;\; \text{where} \;\; \gamma_1 \;\; = \;\; F \, \tilde{x}_1 \, ,$$

The locus of all such x can be found to be a circle. We call such a circle an epipolar circle, and in analogy with the perspective case we let γ_1 be its representation. Apparently γ_1 is the representation of a plane in $\mathbb{P}^3$. Its intersection with the sphere, i.e. the range of $\tilde{x}$ over all x, is a circle on the sphere which determines some circle in the plane.

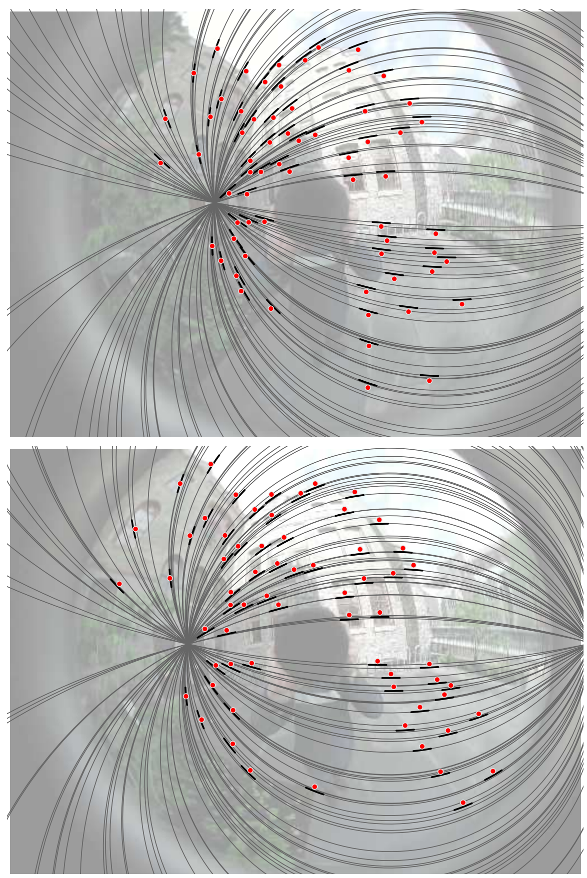

Figure 1: A parabolic catadioptric stereo pair. Points highlighted are use to determine the matrix F which determine the epipolar circles, also highlighted so as to make clear on which epipolar circle a given point lies. The epipoles are the intersection of all the epipolar circles.

We now state a formula determining the correspondence between two epipolar circles. Since F is rank two, it can be written as $F = u_1 v_1^T + u_2 v_2^T$ for some $u_i, v_i \in \mathbb{R}^4$. Let $W = u_1 u_2^T - u_2 u_1^T$. If x is a point in the first image equal

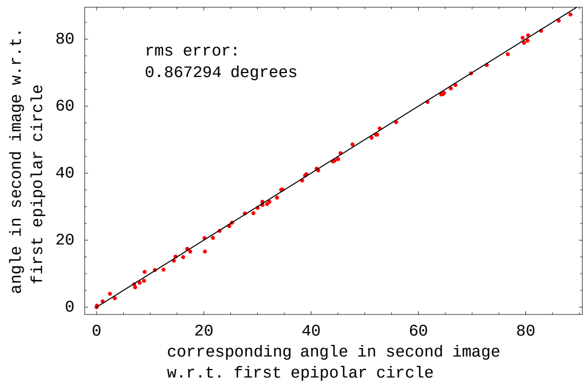

Figure 2: The angles between epipolar circles determined from point correspondences.

to neither $e_{1,1}$ nor $e_{1,2}$, then it can be shown that $\tilde{\gamma} = W\tilde{x}$ and $\tilde{\eta} = F\tilde{x}$ are corresponding epipolar circles.

Now suppose that γ_1 and γ_2 represent any two epipolar circles in one image. Let $\theta = \angle(\gamma_1, \gamma_2)$ be the the angle of intersection between the two circles represented by γ_1 and γ_2. This angle obeys the following equation:

$$\cos^2\theta \;=\; \left(\gamma_1^T Q\,\gamma_2\right)^2 / \left(\gamma_1^T Q\,\gamma_2\,\gamma_2^T Q\,\gamma_2\right) .$$

where $Q = \mathrm{diag}(1, 1, 1, -1)$ is the projective quadratic form of the unit sphere. Let η_1 and η_2 be the corresponding epipolar circles, respectively. It can be shown that F encodes the properties of epipolar circles such that

$$\angle(\gamma_1, \gamma_2) \;=\; \angle(\eta_1, \eta_2)\,. \tag{4}$$

Consider the case where we know γ_1, γ_2 and η_1. We can directly determine η_2 without F because it must satisfy (4).

In *uncalibrated* parabolic catadioptric images, therefore, the epipolar geometry is completely encoded by four points and one angle. In particular, the two epipolar point pairs uniquely determine the systems of coaxal circles in each image, and an angle can be used to determine the correspondence between two epipolar circles in each image. This contrasts with perspective epipolar geometry, in which the correspondence between epipolar lines is given by a homography of $\mathbb{P}^1$, which leaves only the cross-ratio invariant.

To illustrate the parabolic epipolar geometry, we show in Figure 1 a parabolic catadioptric stereo pair. The parabolic catadioptric fundamental matrix has been estimated for this pair using the highlighted points. The epipolar circles which go through the points are drawn and we indicated which epipolar circle a given point lies on by drawing the circle in bold in the neighborhood of the point. As a verification of the claim that the angles between epipolar circles are equal, we have plotted in Figure 2 the pairs $(\angle(\gamma_0, W\tilde{x}_i), \angle(\eta_0, F\tilde{y}_i))$ for all corresponding point pairs (x_i, y_i). Ideally this should be a straight line to indicate that (4) is satisfied; as we can see this is very nearly the case.

§ 5 Catadioptric rectification

Because of the compactness of notation it is natural to use the complex plane as a representation of the image plane. In this case we represent the point (u, v) by $z = u + iv$. Let z be any image point in one image and w in the other such that they satisfy the epipolar constraint. We suppose that z and w are transformed as follows:

$$z' \;=\; f(z) \;\text{ and }\; w' \;=\; g(w)\,.$$

The transformations rectify the stereo pair if the y-coordinates are constant over the images of corresponding epipolar circles. This condition is expressed by the equation

$$\operatorname{im} f(z) \;=\; \operatorname{im} g(w)$$

for all corresponding z and w. This is the bare minimum that we could hope to achieve in a rectification.

Unlike in the perspective case, we cannot perform a rectification using homographies. The closest known equivalent to homographies for parabolic projections are the group of Möbius transformations in the complex plane or equivalently Lorentz transformations of $\tilde{x}$. The former transforms an image point (u, v) represented by the complex number $z = u + iv$ via the function $f(z) = (az + b)/(cz + d)$. All such transformations, however, preserve circles. Therefore a coaxal system of circles remains a coaxal system of circles and so they do not have the capability of producing a set of parallel lines from a system of coaxal circles. This should not be seen as a limit which could be overcome by some extension to planar homographies, they too must preserve images of lines and therefore epipolar circles.

We propose an alternative transformation based on the *bipolar coordinate system* [10]. A bipolar coordinate system consists of a pair of two mutually orthogonal systems of coaxal circles. For the moment we consider only one image and suppose that e_1 and e_2 are the two epipoles. It can be verified that the set of all z such that the angle

$$\theta_0 \;=\; \angle e_1\, z\, e_2 \tag{5}$$

is constant is a circle through e_1 and e_2, and furthermore any circle is uniquely determined by some angle. The angle θ_0 differs by a constant from the angle of intersection between the line through e_1 and e_2 and the circle it defines. Similarly, the set of all z such that the ratio

$$r_0 \;=\; d(z, e_1) \,/\, d(z, e_2)$$

is constant is a circle orthogonal to any of the circles defined by (5), where here $d(\cdot, \cdot)$ denotes the distance between two complex points.

We can turn this into a transformation of the complex plane if we set $w' = \rho + i\theta$ where $\rho = \log r$, for then,

$$w' \;=\; \log\left|\frac{z-e_1}{z-e_2}\right| + i\arg\frac{z-e_1}{z-e_2} \;=\; \log\frac{z-e_1}{z-e_2}\,.$$

Another representation of this transformation can be chosen since $\coth^{-1} = \frac{1}{2}\log z + 1/z - 1$, so that alternatively we have

$$w' \;=\; 2\coth^{-1}\frac{e_1+e_2-2z}{e_1-e_2}\,.$$

Since $\coth^{-1}$ is analytic — it satisfies the Cauchy-Riemann equations [12] — this rectification is conformal. The range of this transformation is from $-\infty$ to ∞ in the real component, and $-\pi$ to π in the imaginary component.

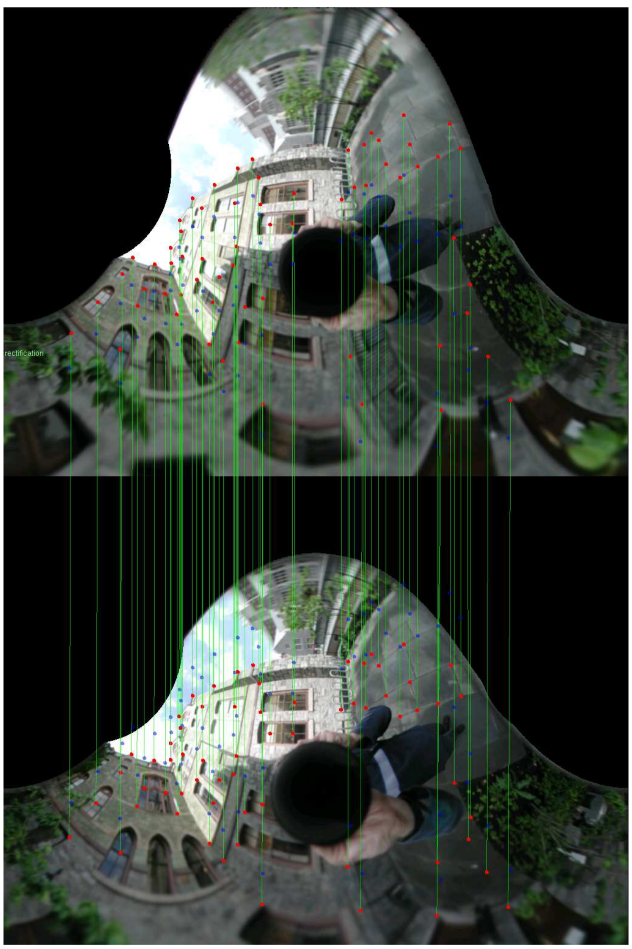

Figure 3: Here we show the rectified images rotated by $90°$ and plot the lines joining points in the two images to demonstrate that they do indeed lie in the same row (here column).

Returning to two images let $e_{i,j}$ be all four epipoles. Assume that z_0 and w_0 satisfy the epipolar constraint. It is not sufficient to apply (6) since z_0' and w_0' will not necessarily have the same imaginary components. The difference of their imaginary components, however, will be some constant θ_0 independent of the choice of corresponding z_0 and w_0. In particular for some θ_0,

$$\theta_0 + \operatorname{im} z_0' \quad \cong_{2\pi} \quad \operatorname{im} w_0'\,.$$

where as usual $a \cong_{2\pi} b$ if and only if 2π divides $a - b$. To find θ_0 from the correspondence simply solve for θ_0 in the congruence equation. Observe that

$$i\theta + 2\coth^{-1} z \;=\; 2\coth^{-1}\frac{-z(e^{i\theta}+1)+e^{i\theta}-1}{z(e^{i\theta}-1)+e^{i\theta}-1}\,,$$

up to a congruence of 2π in the imaginary part. This leads to two equations:

$$z' \;=\; 2\coth^{-1}\frac{-z(e^{i\theta}+1)+e_{1,2}\,e^{i\theta}+e_{1,1}}{z(e^{i\theta}-1)-e_{1,2}\,e^{i\theta}+e_{1,1}}\,,$$
$$w' \;=\; 2\coth^{-1}\frac{e_{2,1}+e_{2,2}-2w}{e_{2,1}-e_{2,2}}\,,$$

which summarize the rectification equations of the stereo pair. This transformation can be applied directly to an image represented in the complex plane. One must exclude circles around the epipoles since the epipoles are a singularity in the mapping. The range and resolution are determined by the positions of the epipoles.

We now show that there are no other conformal rectifications. Consider again only one image and consider a transformation g which rectifies an already rectified image, which would be the same as inverting the current rectification and applying some other rectification. The function g must satisfy

$$\operatorname{im} g(x+iy) \;=\; g_u(x,y) + i g_v(y)\,.$$

In other words, its imaginary component ought not depend on the real component of its argument. Suppose that g is twice differentiable and that it satisfies the Cauchy-Riemann equations:

$$\frac{\partial g_u}{\partial x} \;=\; \frac{\partial g_v}{\partial y} \;\text{ and }\; \frac{\partial g_v}{\partial x} \;=\; -\frac{\partial g_u}{\partial y}\,.$$

Clearly $\partial g_v/\partial x = 0$. Using the Riemann-Cauchy equations we find that $\partial^2 g_u/\partial x\partial y = 0$ and that $\partial^2 g_u/\partial y\partial x = g_v''(y)$. Therefore $g_v(y) = ax + b$ for some real a and b, and similarly for g_u. Therefore g is no more than a translation and scale.

§ 6 Experiment and conclusion

We demonstrate the proposed rectification by applying it to the images shown in Figure 1. The results of the rectification are shown in Figure 3. We draw lines between the corresponding features in the rectified images to demonstrate

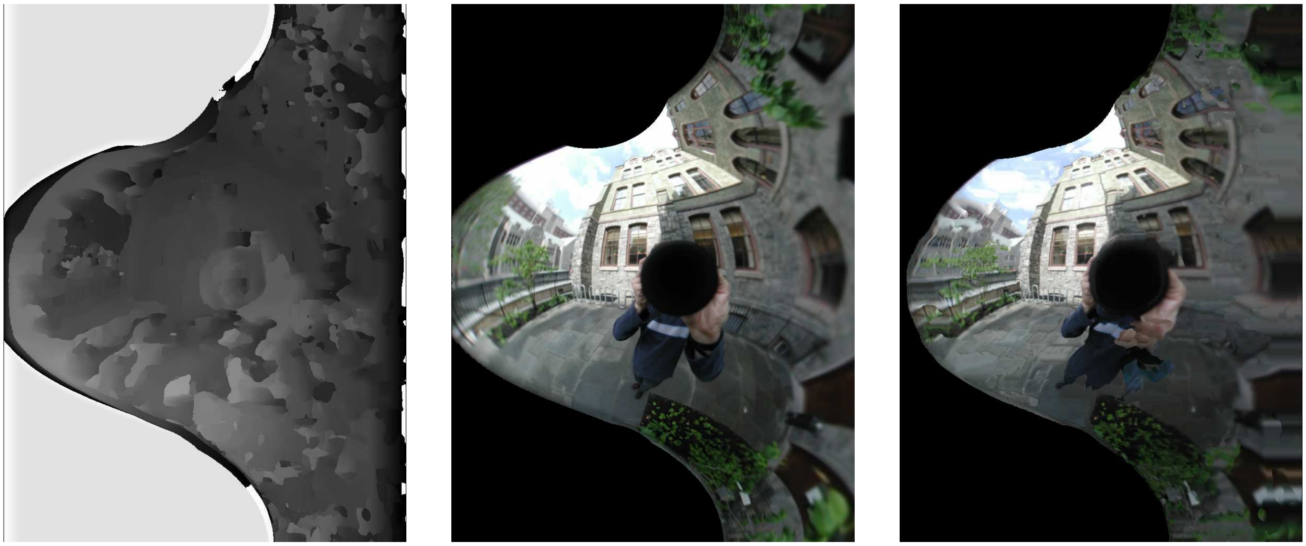

Figure 4: Left: A disparity image is estimated with a correlation based algorithm on the rectified images. Middle: One of the rectified images. Right: A warping of the other image (the bottom image in Figure 3) into the middle image based on the disparity map.

the the features are on the same row of the rectified images. We have used these rectified images to perform a correlation based disparity estimation, the results are shown in the left of Figure (4). Using this disparity map we warp the bottom image of Figure 3 into the top image. The original top image is duplicated in the middle of Figure 3 and the warped image is shown on the right. Ideally the two images would look exactly the same, in areas of low texture this is not quite the case, though it does perform well in those regions where it was easier to obtain a correspondence.

In this paper we have demonstrated a technique for the rectification of parabolic catadioptric stereo pairs. The rectifying transformation is the unique conformal transformation able to rectify the images. We have demonstrated the procedure on two real images and given a demonstration of disparity estimation from the rectified images based on cross-correlation.

References

[1] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Santa Margerita, Italy, May 23-26, G. Sandini (Ed.), Lecture Notes in Computer Science 588, Springer-Verlag, Berlin et al.*, pages 563–578, 1992.

[2] O. Faugeras, B. Hortz, H. Mathieu, T. Viéville, Z. Zhang, P. Fua, E. Théron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proj. Real-time correlation based stereo: algorithms, implementation and application. Technical report 2013, INRIA Sophia Antipolis, 1993.

[3] M. Fleck. Perspective projection: the wrong imaging model. Technical report 95-01, University of Iowa, 1995.

[4] C. Geyer and K. Daniilidis. Catadioptric projective geometry. *International Journal of Computer Vision*, 43:223–243, 2001.

[5] C. Geyer and K. Daniilidis. Structure and motion from uncalibrated catadioptric views. In *IEEE Conf. Computer Vision and Pattern Recognition*, Hawaii, Dec. 11-13, 2001.

[6] C. Geyer and K. Daniilidis. Properties of the catadioptric fundamental matrix. In *Proc. Seventh European Conference on Computer Vision*, pages 140–154, Copenhagen, Denmark, 2002.

[7] J. Gluckman and S. Nayar. Rectifi ed catadioptric stereo sensors. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 224–236, Hilton Head Island, SC, June 13-15, 2000.

[8] J. Gluckman and S. Nayar. Rectifying transformations that minimize resampling effects. In *IEEE Conf. Computer Vision and Pattern Recognition*, Hawaii, Dec. 11-13, 2001.

[9] R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *IEEE Conf. Computer Vision and Pattern Recognition*, Champaign, IL, June 15-18, 1992.

[10] G. Korn and T. Korn. *Mathematical Handbook for Scientists and Engineers*. McGraw-Hill, New York, 1968.

[11] S. Nayar. Catadioptric omnidirectional camera. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 482–488, Puerto Rico, June 17-19, 1997.

[12] T. Needham. *Visual Complex Analysis*. Clarendon Press, Oxford, 1997.

[13] Y. Onoe, K. Yamazawa, H. Takemura, and N. Yokoya. Telepresence by real-time view-dependent image generation from omnidirectional video streams. *Computer Vision and Image Understanding*, 71:588–592, 1998.

[14] C. Slama, editor. *Manual of Photogrammetry*. American Society of Photogrammetry, Falls Chuch, VA, 1980.

[15] Y. Yagi, S. Kawato, and S. Tsuji. Real-time omnidirectional image sensor (copis) for vision-guided navigation. *IEEE Trans. on Robotics and Automation*, 10:11–22, 1994.