

Rectified Catadioptric Stereo Sensors *

Joshua Gluckman and Shree K. Nayar
Department of Computer Science
Columbia University
New York, NY 10027

Abstract

It has been previously shown how mirrors can be used to capture stereo images with a single camera, an approach termed catadioptric stereo. In this paper, we present novel catadioptric sensors which use mirrors to produce rectified stereo images. The scan-line correspondence of these images benefits real-time stereo by avoiding the computational cost and image degradation due to resampling when rectification is performed after image capture. First, we develop a theory which determines the number of mirrors that must be used and the constraints on those mirrors that must be satisfied to obtain rectified stereo images with a single camera. Then we discuss in detail the use of both one and three mirrors. In addition, we show how the mirrors should be placed in order to minimize sensor size for a given baseline, an important design consideration.

1 Introduction

Catadioptric systems are optical systems that consist of a combination of mirrors and lenses [4]. As demonstrated by several researchers, catadioptrics can be used to design stereo sensors that use only a single camera [11] [3] [5] [9] [15] [12] [2]. Although the systems described by these researchers use a variety of different mirror shapes and configurations the underlying motivation is the same. By using multiple mirrors, scene points can be imaged from two or more viewpoints while using only a single camera.

Single camera stereo has several advantages over traditional two-camera stereo. Because only a single camera and digitizer are used, system parameters such as spectral response, gain, and offset are identical for the stereo pair. In addition, only a single set of internal calibration parameters needs to be determined. Perhaps most important is that single camera stereo simplifies data acquisition by only requiring a single camera and digitizer and no hardware or software for synchronization.

Real-time stereo systems, whether catadioptric or two-camera, require images to be rectified prior to stereo matching. A pair of stereo images is rectified if the epipolar lines are aligned with the scanlines of the images. When properly aligned, the search for correspondence is simplified and thus real-time performance can be obtained [1]. Once the epipolar geometry of a stereo system is determined, rectifying transformations can be applied to the im-

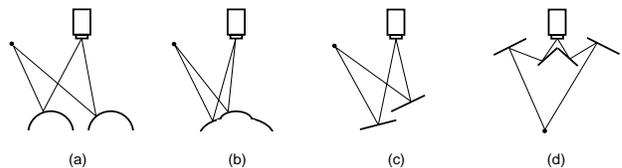


Figure 1: Single camera stereo systems using a variety of mirrors. By imaging two reflections of a scene point, the 3-d location can be determined from a single camera. (a) Two spherical mirrors. (b) Two stacked convex mirrors. (c) Two planar mirrors. (d) Four planar mirrors.

ages [14] [6] [8] [13]. However, rectifying in this manner has two disadvantages for real-time stereo. Applying transformations to the images at run-time is both computationally costly and degrades the stereo data due to the resampling of the images.

An alternative to rectifying the images at run-time is to ensure that the *geometry* of the stereo system produces rectified images. With two-camera stereo this is accomplished by removing any rotation between the two cameras, aligning the direction of translation with the scan lines of the cameras and using identical internal parameters for the two cameras (a difficult task). The geometric requirements of a rectified catadioptric stereo system are not trivial and have not been studied. The purpose of this paper is twofold. First, to develop the constraints that must be satisfied to ensure rectified images, and second to describe an automated tool for placing the mirrors such that sensor size is minimized for a given baseline. These results can be used to design and build novel compact stereo sensors.

1.1 Previous work

Several researchers have demonstrated the use of both curved and planar mirrors to acquire stereo data with a single camera. Curved mirrors have been primarily used to capture a wide field of view. One of the first uses of curved mirrors for stereo was in [11], where Nayar suggested a wide field of view stereo system consisting of a conventional camera pointed at two specular spheres (see Figure 1(a)). A similar system using two convex mirrors, one placed on top of the other, was proposed by Southewell *et al.* [15] (see Figure 1(b)). Later, Nene and Nayar presented several different catadioptric stereo configurations using a single camera with parabolic, elliptic and hyperbolic mirrors [12].

Several others have also investigated the use of planar

*This work was supported in parts by a DARPA/ONR MURI Grant (N00014-95-1-0601), an NSF National Young Investigator Award, and a David and Lucile Packard Fellowship.

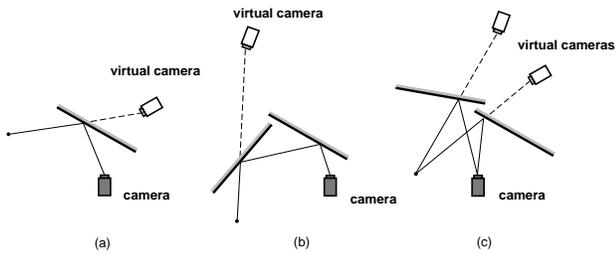


Figure 2: Image formation with mirrors. (a) When a planar mirror reflects a scene point, the image formed is from a virtual camera found by reflecting the real camera about the plane containing the mirror. (b) When two mirrors reflect the scene point, the virtual camera is found by applying two reflections. (c) If the two mirrors produce two reflections of the scene point then a stereo image is obtained.

mirrors to design single camera stereo sensors. A sensor designed by Goshtasby and Gruver used two planar mirrors connected by a hinge centered in the field of view of the camera [3]. Gluckman and Nayar demonstrated how two mirrors in an arbitrary configuration can be self-calibrated and used for single camera stereo [2] (see Figure 1(c)). Stereo systems using four planar mirrors were proposed by both Inaba *et al.* [5] and Mathieu and Devernay [9] (see Figure 1(d)). By imaging an object and its mirror reflection, a stereo image can also be obtained using only a single mirror [10] [16].

In all of these systems, the stereo images are not rectified, therefore the images must be transformed at run-time prior to stereo matching. One exception is a system described by Lee *et al.* that uses *prisms* rather than *mirrors* to acquire rectified stereo images from a single camera [7]. Although prisms are an interesting alternative to mirrors it is not clear that compact sensors with sufficient baseline can be designed.

1.2 Background

Before describing the requirements for rectified catadioptric stereo we explain image formation with planar mirrors. As Figure 2(a) shows, the image formed when a mirror reflects a scene point is the same perspective image taken by a *virtual camera* located on the opposite side of the mirror. The location of the coordinate system of the virtual camera relative to the coordinate system of the real camera is found by applying a reflection transformation. If we represent the mirror with the normal \mathbf{n} and the distance d measured from the real camera center, the reflection transformation \mathbf{D} is found to be

$$\mathbf{D} = \begin{bmatrix} \mathbf{I} - 2\mathbf{nn}^T & 2d\mathbf{n} \\ \mathbf{0} & 1 \end{bmatrix}.$$

The transformation \mathbf{D} between the real and virtual camera coordinate systems is a combination of a rigid transformation and a switch from a left to a right handed (or vice-versa) coordinate system. Also note that a reflection

transform is its own inverse:

$$\mathbf{D}\mathbf{D} = \mathbf{I}.$$

When two mirrors (see Figure 2(b)) reflect a scene point, the virtual camera is found by applying two consecutive reflection transformations. As shown in [2] the resulting transformation represents a planar rigid motion, meaning the direction of translation is orthogonal to the axis of rotation. It was also shown that the axis of rotation is $(\mathbf{n}_1 \times \mathbf{n}_2)$, where \mathbf{n}_1 and \mathbf{n}_2 are the normals of the two mirrors. In the two mirror case there is no switch from a left to right handed system because the two switches cancel out.

For each additional mirror the virtual camera is found by applying another reflection transformation. In general, if the number of mirrors is odd then the resulting transformation switches the handedness of the coordinate system, thus producing a mirror image of the scene.

As shown in Figure 2(c), if the field of view is split such that different mirrors reflect the scene onto different portions of the imaging plane then the scene is imaged from multiple virtual cameras, and thus a stereo image is obtained. Now, we will examine how mirrors can be used to obtain a *rectified* stereo image.

2 How many mirrors are needed?

To produce rectified images, a stereo system must meet several requirements. There must be no relative rotation between the two cameras, the translation must be parallel to the scanlines of the image plane, and the internal parameters of the two cameras must be identical. For catadioptric stereo the last requirement is met because only a single camera is used. To ensure the first two requirements, the mirrors must satisfy

$$\mathbf{D}_1 \dots \mathbf{D}_i \dots \mathbf{D}_m = \begin{bmatrix} 1 & 0 & 0 & b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where b is the baseline, m is the number of mirrors used and \mathbf{D}_i is the reflection transformation produced by the i^{th} mirror. The mirrors are ordered as shown in Figure 3 and each mirror is defined in a coordinate system attached to the left virtual camera, where the x-axis is along the scanlines and the z-axis is in the direction of the optical axis.

Although satisfying (1) is sufficient to ensure rectification there is one caveat. Because we split the field of view of the real camera between the two virtual cameras we must guarantee that the fields of view properly overlap. Figure 4 shows a four mirror system where the two virtual cameras are rectified but do not share a common field of view. When the field of view is split between two different systems of mirrors, each virtual camera sees half of the field of view of the real camera. Clearly, in practice the two half fields of view must overlap.

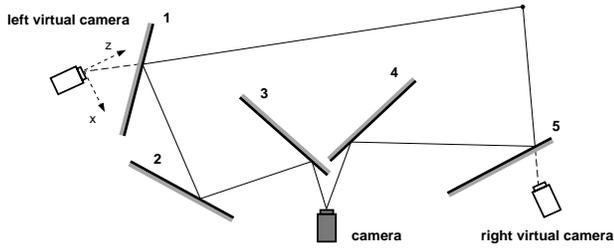


Figure 3: The relative orientation between the left and right virtual cameras is found by applying consecutive reflection transformations in the order defined by the numbers in the above figure.

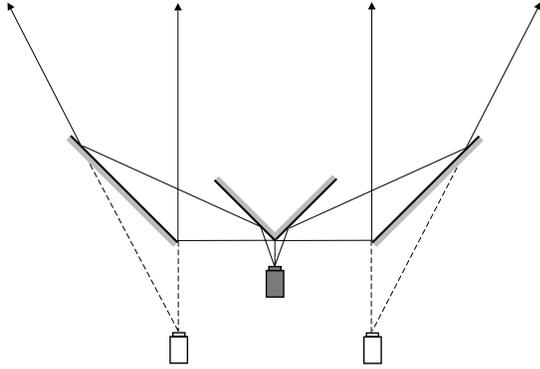


Figure 4: When two sets of parallel mirrors are used a rectified stereo system can be constructed. However, this solution is not practical because the two virtual cameras do not share a common field of view.

As shown in Figure 5, if rectified there are three possible configurations for the two virtual cameras, depending upon the number of mirrors used. However, only the one on the left leads to a practical stereo system. To obtain the configuration on the left, one of the half fields of view must be reflected relative to the other and therefore an odd number of mirrors must be used. We enforce this by flipping the direction of the x-axis

$$\mathbf{D}_1 \dots \mathbf{D}_i \dots \mathbf{D}_m = \begin{bmatrix} -1 & 0 & 0 & b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

It is straightforward to show that for any number of odd mirrors a solution exists. For $m = 1$ the mirror normal $\mathbf{n}_1 = [1, 0, 0]^T$ leads to a solution. For all odd m , ($m > 1$) a trivial solution can be obtained by adding $\frac{m-1}{2}$ pairs of identical reflection transformations; because a reflection transformation is its own inverse each pair will cancel out.

Although there are many solutions to (2) most are not physically realizable due to occlusions and intersecting mirrors. Next, we will discuss possible solutions using one and three mirrors. Five or more mirrors can be used however these systems are complex and their advantages are unclear.

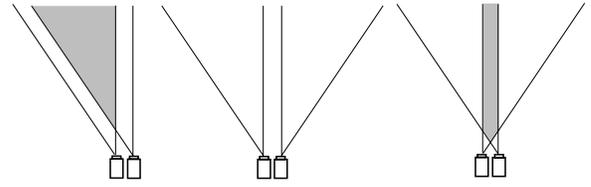


Figure 5: Once rectified, there are three possible configurations of the virtual cameras. However, only the one on the left leads to a practical solution. The middle configuration has no overlapping field of view and the configuration on the right only sees a narrow beam.

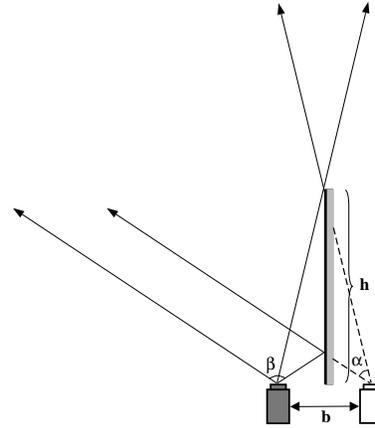


Figure 6: To obtain a rectified image with a single mirror, the normal of the mirror must be parallel to the scanlines of the camera. Note that the field of view of the right virtual camera will be limited by the finite size of the mirror.

3 Single mirror rectified stereo

To obtain a rectified image with a single mirror, the plane containing the mirror must satisfy

$$\mathbf{D}_1 = \begin{bmatrix} \mathbf{I} - 2\mathbf{n}_1\mathbf{n}_1^T & 2d_1\mathbf{n}_1 \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

where \mathbf{n}_1 and d_1 are the normal and distance of the mirror from the camera coordinate system. For this to be satisfied, $\mathbf{n}_1 = [1, 0, 0]^T$. Thus, the only one mirror solution occurs when the normal of the mirror is parallel to the scanlines of the camera (the x-axis) as shown in Figure 6. The stereo system will remain rectified for any distance d_1 , however the baseline b will change as $b = 2d_1$.

The advantage of this solution is its simplicity. However, because a finite mirror must be used the field of view of the virtual camera is limited by the angle the mirror subtends with respect to the virtual camera (see Figure 6). The field of view α is related to the baseline b and the length of the mirror h as

$$\alpha = \arctan\left(\frac{2h}{b}\right) - \frac{\pi}{2} + \frac{\beta}{2}. \quad (4)$$

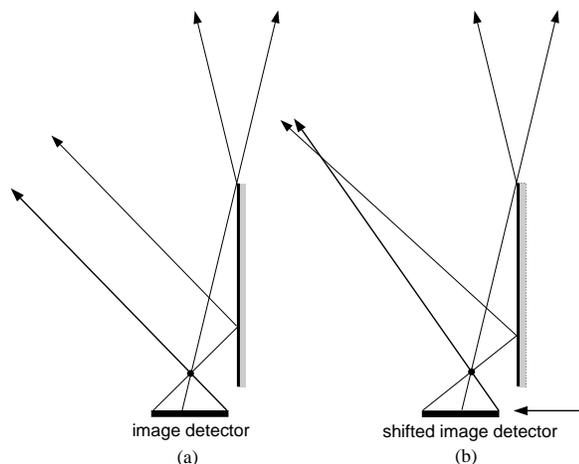


Figure 7: Shifting the image detector to reduce asymmetry in the stereo field of view. (a) When a single mirror is used the left camera view uses a larger portion of the image detector than the right camera view. (b) The asymmetry can be removed by shifting the image detector with respect to the imaging lens (center of projection).

Therefore, if a large baseline is desired then a large mirror must be employed or else the field of view will be severely limited.

Because the mirror is finite a field of view asymmetry exists between the real and virtual camera. A larger portion of the image detector is used by the real camera. As shown in Figure 7, this asymmetry can be removed by shifting the image detector with respect to the center of projection of the camera. For applications where the scene of interest lies close to the camera this has the benefit of increasing the viewing volume close to the stereo system.

4 Three mirror rectified stereo

We can overcome the limitations of single mirror rectified stereo by incorporating additional mirrors. With three mirrors we can ensure that the field of view is equally shared between the two virtual cameras (see Figure 8). Furthermore, a large baseline can be obtained using relatively small mirrors. However, we can not arbitrarily place the three mirrors. As we will show, to obtain rectified stereo five constraints between the mirrors and the camera must be satisfied. From (2) we know that the mirrors must be placed such that

$$\mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_3 = \begin{bmatrix} -1 & 0 & 0 & b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (5)$$

Using (5) we will first derive three constraints by determining how the mirrors must be placed so that there is no rotation between the virtual cameras. Then we will show the remaining two constraints that arise from requiring the

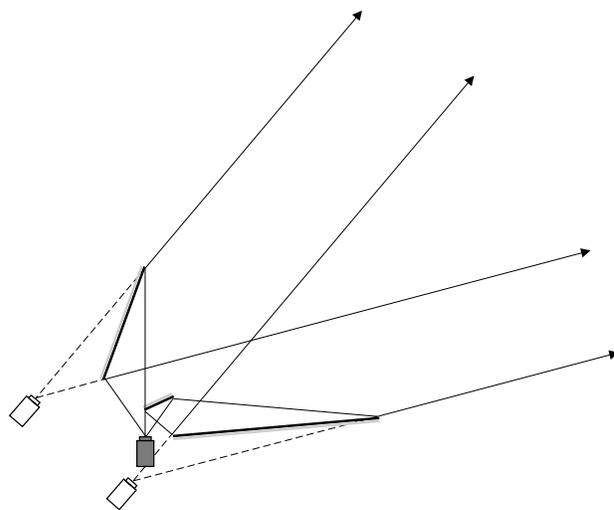


Figure 8: Three mirror rectified stereo. By using three mirrors a rectified stereo system can be designed where the field of view of each virtual camera is half the field of view of the real camera.

direction of translation to be parallel to the scanlines (the x-axis).

4.1 Rotation constraints

The upper left 3×3 block of the matrix in (5) is the requirement that there be no rotation between the virtual cameras:

$$\mathbf{R}_1 \mathbf{R}_2 \mathbf{R}_3 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (6)$$

where

$$\mathbf{R}_i = [\mathbf{I} - 2\mathbf{n}_i \mathbf{n}_i^T]. \quad (7)$$

Recall that $\mathbf{R}_1 \mathbf{R}_2$ is a rotation matrix with a rotational axis defined by $\mathbf{n}_1 \times \mathbf{n}_2$ [2]. This implies that

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{R}_3 (\mathbf{n}_1 \times \mathbf{n}_2) = (\mathbf{n}_1 \times \mathbf{n}_2). \quad (8)$$

It is straightforward to show that

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{R}_3 (\mathbf{n}_3 \times [1, 0, 0]^T) = (\mathbf{n}_3 \times [1, 0, 0]^T). \quad (9)$$

Therefore, to satisfy (6) it is necessary that

$$(\mathbf{n}_1 \times \mathbf{n}_2) = \lambda (\mathbf{n}_3 \times [1, 0, 0]^T). \quad (10)$$

The scale factor λ can be removed by taking the dot product with \mathbf{n}_1 or \mathbf{n}_3 giving us two independent constraints,

$$\mathbf{n}_1 \cdot (\mathbf{n}_3 \times [1, 0, 0]^T) = 0 \quad (11)$$

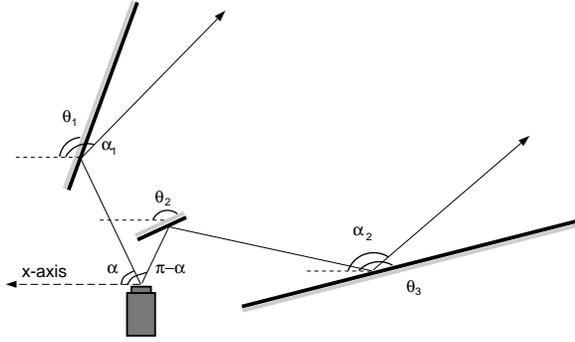


Figure 9: The mirrors must be angled such that two rays at angles α and $\pi - \alpha$ are parallel after being reflected by the mirrors, that is $\alpha_1 = \alpha_2$. Doing so will ensure that there is no rotation between the two virtual cameras.

and

$$\mathbf{n}_3 \cdot (\mathbf{n}_1 \times \mathbf{n}_2) = 0. \quad (12)$$

This implies that the normals of the three mirrors \mathbf{n}_1 , \mathbf{n}_2 and \mathbf{n}_3 and the x-axis are all co-planar. When the normals are co-planar, the mirrors rotate the virtual cameras about a common axis. Of course, we still need to orient the mirrors so that the rotational angle is cancelled. But now we can simplify our analysis by working in the two dimensions of the plane containing the normals and the x-axis. In 2-D the mirrors are represented by lines where θ_i is the angle the i^{th} mirror normal makes with the x-axis of the real camera. For simplicity we have moved the x-axis to the real camera.

If we consider a ray leaving the camera center at angle α and a corresponding ray leaving the camera at an angle $\pi - \alpha$ then the mirrors must be angled such the two reflected rays are parallel, thus ensuring there is no rotation between the virtual cameras (see Figure 9). After being reflected by the mirror oriented at angle θ_1 , the angle of the left ray is

$$\alpha_1 = 2\theta_1 - \alpha + \pi. \quad (13)$$

The angle of the right ray α_2 after reflection by the two mirrors is

$$\alpha_2 = 2\theta_3 - 2\theta_2 - \alpha + \pi. \quad (14)$$

The two rays are parallel if $\alpha_1 = \alpha_2$. Therefore the rotation is cancelled if

$$\theta_3 - \theta_2 = \theta_1. \quad (15)$$

We can express (15) in terms of the normals of the mirrors as,

$$\mathbf{n}_3 \cdot \mathbf{n}_2 = \mathbf{n}_1 \cdot [1, 0, 0]^T. \quad (16)$$

To summarize, if the normals of the mirrors satisfy the three constraints (11), (12) and (16) then there will be no rotation between the two virtual cameras.

4.2 Translation constraints

So we have three constraints that must be satisfied to rectify the virtual cameras. However, we also need to ensure that the direction of translation between the virtual cameras is along the x-axis.

For this we will need to examine the translational part of (5), that is

$$2d_3\mathbf{R}_1\mathbf{R}_2\mathbf{n}_3 + 2d_2\mathbf{R}_1\mathbf{n}_2 + 2d_1\mathbf{n}_1 = \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix}. \quad (17)$$

After multiplying through with $\mathbf{R}_3\mathbf{R}_2\mathbf{R}_1$ and substituting (6) we get

$$2d_3\mathbf{R}_3\mathbf{n}_3 + 2d_2\mathbf{R}_3\mathbf{R}_2\mathbf{n}_2 + 2d_1 \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{n}_1 = \begin{bmatrix} -b \\ 0 \\ 0 \end{bmatrix}. \quad (18)$$

Next, by substituting (7) for the \mathbf{R}_i we get

$$-2d_3\mathbf{n}_3 - 2d_2\mathbf{n}_2 + 4(\mathbf{n}_3 \cdot \mathbf{n}_2)d_2\mathbf{n}_3 + 2d_1 \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{n}_1 = \begin{bmatrix} -b \\ 0 \\ 0 \end{bmatrix}. \quad (19)$$

Since the baseline b is arbitrary, we have only two linear constraints on d_1 , d_2 and d_3 , namely

$$-n_{3y}d_3 + (2(\mathbf{n}_3 \cdot \mathbf{n}_2)n_{3y} - n_{2y})d_2 + n_{1y}d_1 = 0 \quad (20)$$

and

$$-n_{3z}d_3 + (2(\mathbf{n}_3 \cdot \mathbf{n}_2)n_{3z} - n_{2z})d_2 + n_{1z}d_1 = 0, \quad (21)$$

where $\mathbf{n}_i = [n_{ix}, n_{iy}, n_{iz}]$.

We have now derived the 5 constraints for three mirror rectified stereo. To summarize, if the normals of the three mirrors are co-planar with the x-axis, the angles between the mirrors satisfy (16) and the distances to the mirrors are chosen such that (20) and (21) are satisfied then the two virtual cameras will be rectified.

4.3 Sensor design via optimization

For many stereo applications the compactness of the sensor is important. In this section we describe an automated tool for catadioptric stereo sensor design. Given design parameters such as baseline, field of view, and size of the real camera we compute the optimal placement of the mirrors such that the virtual cameras are rectified and sensor size is minimized.

To simplify the optimization, we assume the plane containing the normals of the mirrors is the x-z plane, meaning there is no tilt between the camera and the mirrors. Now, each of the three mirrors is represented by a line, so there are only six parameters: θ_1 , θ_2 , θ_3 , d_1 , d_2 and d_3 . Because we are restricted to the x-z plane, (11), (12) and (20)

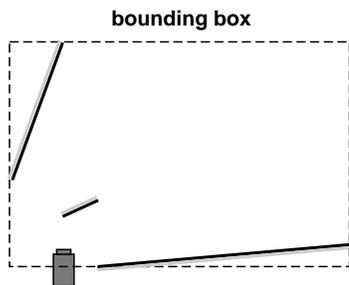


Figure 10: We search for the mirror locations that satisfy the rectification constraints and minimize the perimeter of the bounding box of the mirrors, for a given baseline.

are satisfied, and thus there are only two rectification constraints, one on the angles of the mirrors (15) and one on the distances (21). Given a desired baseline b we have one more constraint from (19)

$$2n_{3x}d_3 + (2n_{2x} - 4(\mathbf{n}_3 \cdot \mathbf{n}_2)n_{3x})d_2 + 2n_{1x}d_1 = b. \quad (22)$$

Three constraints on six parameters leaves three free parameters. To optimize these parameters some criteria for sensor size must be chosen. One simple measure is the perimeter of the bounding box of the mirrors (see Figure 10).

To find the best configuration we search through all possible locations (θ_1, d_1) of the first mirror and possible angles θ_2 for the second mirror. The remaining three parameters, the distance d_2 to the second mirror, the distance d_3 to the third mirror and the angle θ_3 of the third mirror are found by solving (15), (21) and (22), where $n_{ix} = \cos \theta_i$ and $n_{iz} = \sin \theta_i$. For each set of computed parameters we determine the end points of the mirrors by tracing the optical axis and the limiting rays of the field of view and intersecting them with the mirrors. Once the end points are found the perimeter of the bounding box is computed. The design that minimizes the perimeter is chosen.

We only admit solutions where the mirrors do not occlude each other. Thus we ensure that the ray \mathbf{r}_2 , in Figure 11, does not intersect the mirrors \mathbf{M}_2 and \mathbf{M}_3 and the ray \mathbf{r}_3 does not intersect \mathbf{M}_2 . In addition, we only consider solutions where the real camera does not see itself and thus we ensure the ray \mathbf{r}_1 is a minimum distance c from the camera center of projection (note that c is proportional to the baseline b and determined from the size of the camera).

Figure 12 shows two optimal configurations for a camera with a 60° field of view. As shown in 12(a), when $c = 0$ the second mirror is infinitesimally small and located at the center of projection of the camera. Figure 12(b) shows the optimal solution when $c = 0.1$. As both c and the field of view increase the optimal sensor size also increases. Note that we may use the same approach to find the optimal sensor for some other size criterion, such as the area of the bounding box.

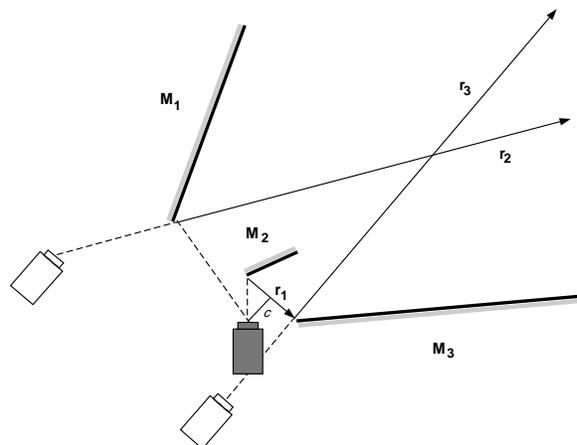


Figure 11: It is important that the reflected scene rays are not occluded by the mirrors. Therefore, we ensure that the ray \mathbf{r}_2 does not intersect the mirrors \mathbf{M}_2 and \mathbf{M}_3 and the ray \mathbf{r}_3 does not intersect \mathbf{M}_2 . If we have some notion of the size of our camera then we can also ensure that the ray \mathbf{r}_1 is at a minimum distance c from the camera center of projection.

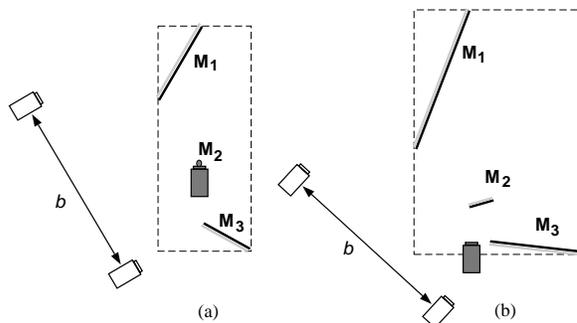


Figure 12: Optimized stereo sensors for baseline b . (a) If we do not consider the size of the camera, $c = 0$, then the optimal solution has one mirror at the center of projection. (b) This configuration is optimal when $c = 0.1$. For scale, the baseline between the virtual cameras is shown.

5 Experimental results

We have used the constraints introduced in the preceding sections to build both one and three mirror stereo sensors. As in Figure 13(a), we can capture rectified stereo images using a single mirror by placing the mirror parallel to the optics of the camera. We found that aligning the mirror by hand was adequate to obtain a depth map of the scene (see Figure 14). To demonstrate that the stereo image is rectified the correspondence search is only performed along the scanlines. The depth map was computed on a 640×480 image using both SSD and normalized cross-correlation with a 15×15 size window. In the single mirror case we found that normalized cross-correlation is beneficial because of intensity differences introduced by the reflection of scene rays at acute angles with the mirror (recall that the reflectance of a mirror falls slightly as a function

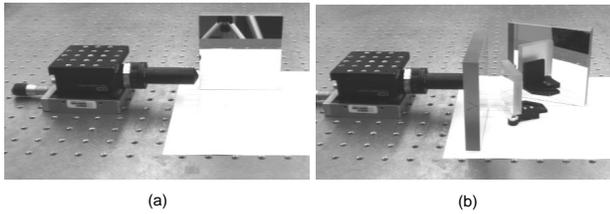


Figure 13: One and three mirror rectified stereo systems using a Sony XC-75 camera with a Computar 4mm pinhole lens. (a) A single mirror is placed such that the mirror normal is perpendicular to the optical axis of the imaging system. (b) Three mirrors are placed in a rectified configuration.

of the angle of incidence).

To construct a three mirror system we used the optimal solution described in the previous section and scaled it so that the baseline was 10 cm. Using a drawing tool we printed the location of the three mirrors and camera center of projection on a piece of paper (see Figure 13(b)). We used a Computar 4mm pinhole lens so that we could accurately place the pinhole of the lens over the desired position marked on the paper.

Figure 15(a) shows several images taken by the three mirror system. 15(b) and 15(c) show depth maps computed by searching along the scanlines using both SSD and normalized cross-correlation. For the three mirror case, we found that normalizing the data did not improve the results. The three mirror system does not suffer from the acute angles of incidence which are encountered when using a single mirror. Therefore, the extra computational cost of normalized cross-correlation can be avoided.

6 Conclusion

In this paper, we have shown how to design a class of novel stereo sensors. By avoiding the need for synchronization, rectification and normalization of the data these sensors are well-suited for real-time applications. Currently, we are using the ideas presented here to build a variety of compact stereo devices.

References

- [1] O. Faugeras, B. Hotz, H. Mathieu, T. Vieville, Z. Zhang, P. Fau, E. Theron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy. Real-time correlation-based stereo: algorithm, implementation and application. Technical Report 2013, INRIA Sophia Antipolis, 1993.
- [2] J. Gluckman and S. Nayar. Planar catadioptric stereo: geometry and calibration. In *Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition*, 1999.
- [3] A. Goshtasby and W. Gruver. Design of a single-lens stereo camera system. *Pattern Recognition*, 26(6):923–937, 1993.
- [4] E. Hecht and A. Zajac. *Optics*. Addison-Wesley, 1974.
- [5] M. Inaba, T. Hara, and H. Inoue. A stereo viewer based on a single camera with view-control mechanism. In *Proceedings of the International Conference on Robots and Systems*, July 1993.
- [6] F. Isgro and E. Trucco. Projective rectification without epipolar geometry. In *Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition*, 1999.
- [7] D. Lee, I. Kweon, and R. Cipolla. A biprism stereo camera system. In *Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition*, 1999.
- [8] C. Loop and Z. Zhang. Computing rectifying homographies for stereo vision. In *Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition*, 1999.
- [9] H. Mathieu and F. Devernay. Systeme de miroirs pour la stereoscopie. Technical Report 0172, INRIA Sophia-Antipolis, 1995. in French.
- [10] H. Mitsumoto, S. Tamura, K. Okazaki, N. Kajimi, and Y. Fukui. 3d reconstruction using mirror images based on a plane symmetry recovery method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(9):941–945, 1992.
- [11] S. Nayar. Robotic vision system. United States Patent 4,893,183, Aug. 1988.
- [12] S. Nene and S. Nayar. Stereo with mirrors. In *Proceedings of the 6th International Conference on Computer Vision*, Bombay, India, January 1998. IEEE Computer Society.
- [13] M. Pollefeys, R. Koch, and L. VanGool. A simple and efficient rectification method for general motion.
- [14] L. Robert, M. Buffa, and M. Herbert. Weakly-calibrated stereo perception for rover navigation. In *Proceedings of the 5th International Conference on Computer Vision*, 1995.
- [15] D. Southwell, A. Basu, M. Fiala, and J. Reyda. Panoramic stereo. In *Proceedings of the Int'l Conference on Pattern Recognition*, 1996.
- [16] Z. Zhang and H. Tsui. 3d reconstruction from a single view of an object and its image in a plane mirror. In *International Conference on Pattern Recognition*, 1998.

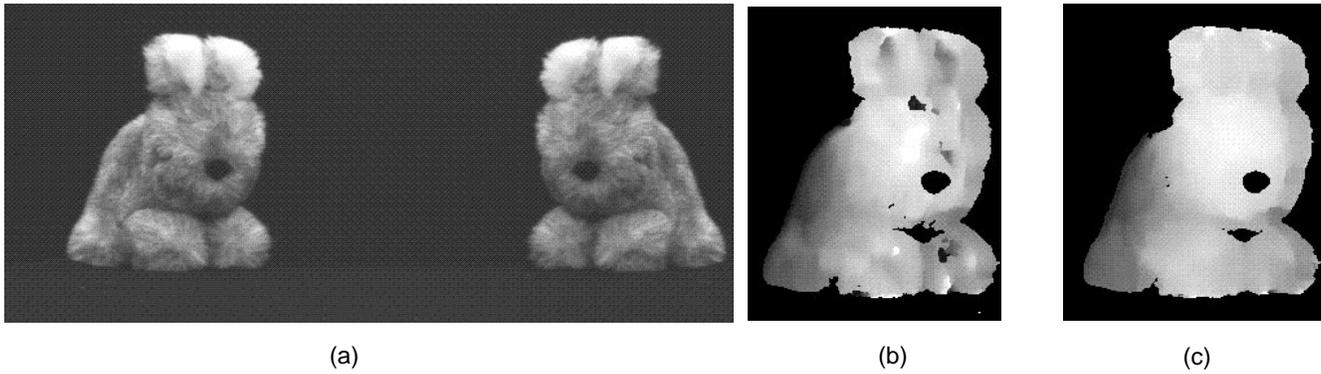


Figure 14: Single mirror rectified stereo. (a) An image (cropped for display) captured by the single mirror stereo system. Note that the right side of the image is reflected. To demonstrate that the images are rectified we perform stereo matching along the scanlines of the image after removing the reflection. (b) and (c) are the depth maps computed using SSD and normalized cross-correlation with a 15×15 window. Normalized cross-correlation performs slightly better due to the intensity differences introduced by the reflection of scene rays at acute angles with the mirror. Depth is not computed for background pixels.

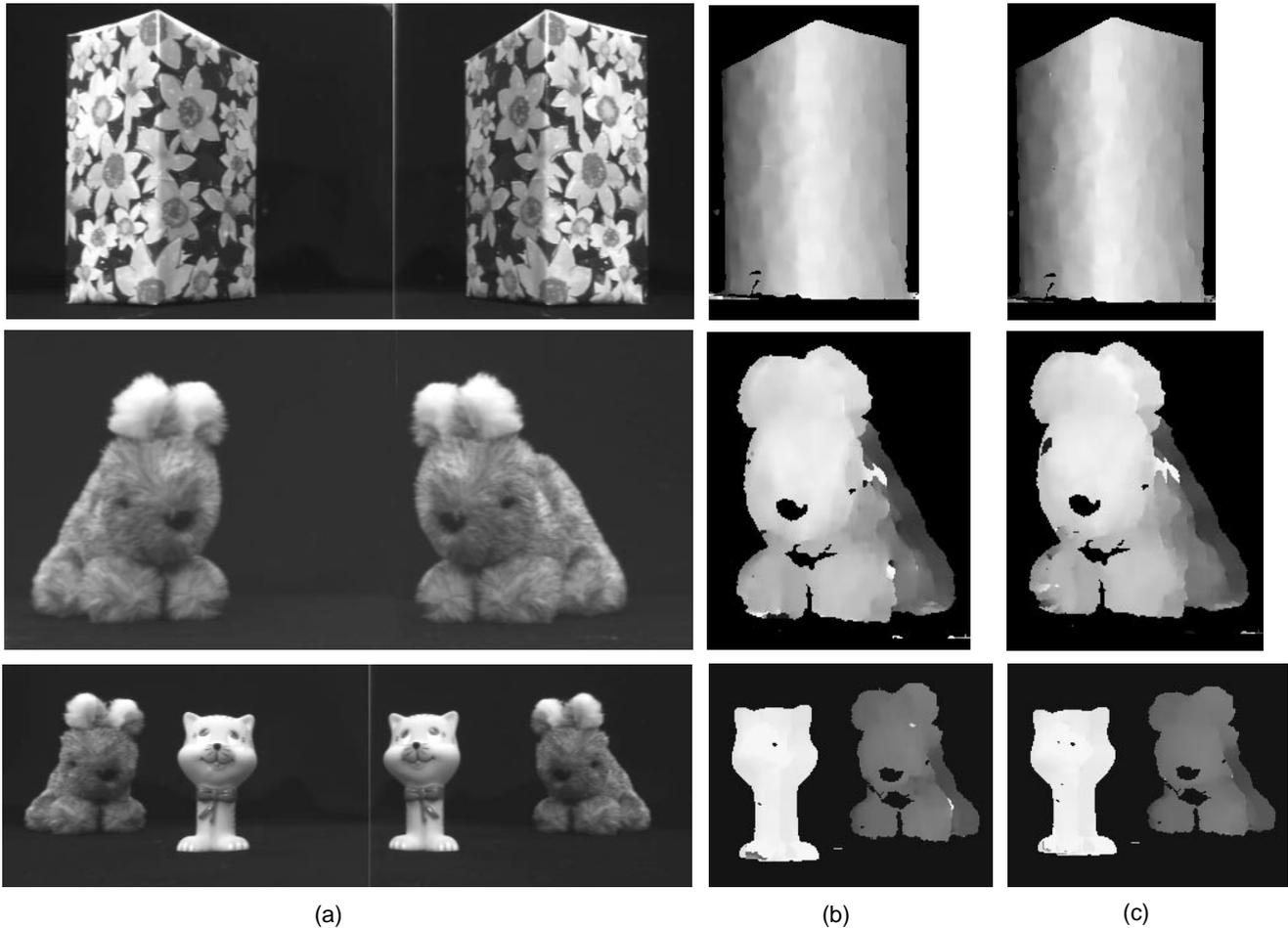


Figure 15: Images and depth maps using three mirror rectified stereo. (a) Three 640×480 images (cropped for display) captured using three mirrors. Before stereo matching along the scanlines is performed the right half of each image is flipped to remove the reflection. (b) and (c) are the depth maps obtained using SSD and normalized cross-correlation with a 15×15 window. Note that the difference between the two is negligible thus emphasizing the fact that when only a single camera is used SSD is sufficient for stereo matching. Depth is not computed for the background pixels.