Multiview Radial Catadioptric Imaging for Scene Capture

Sujit Kuthirummal   Shree K. Nayar
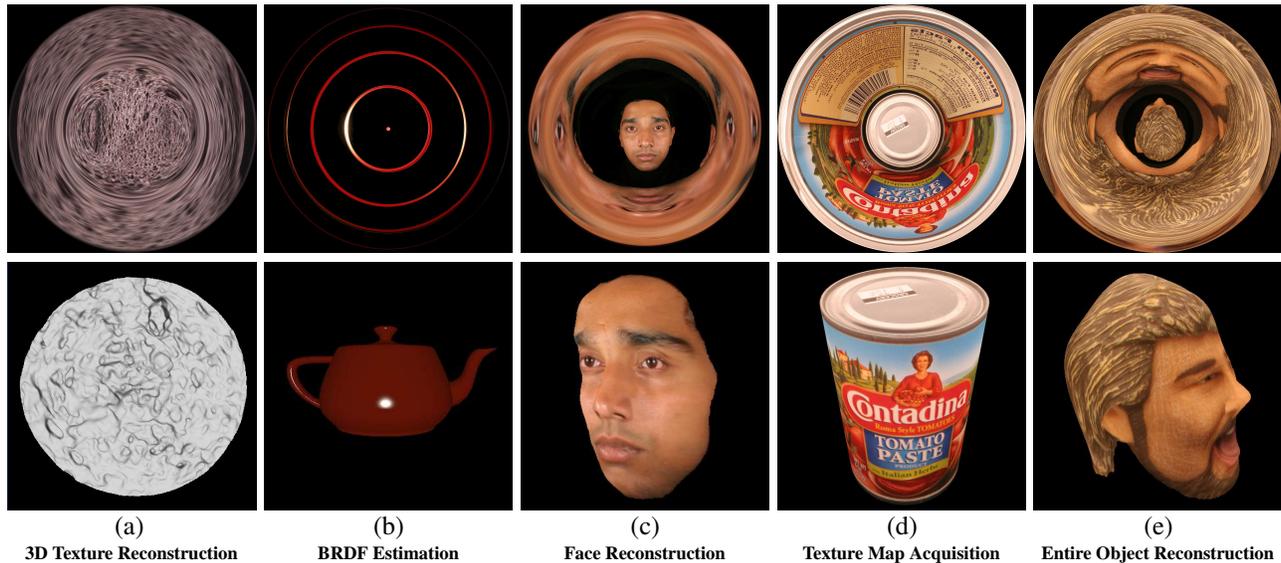
Columbia University *

Figure 1: Top row: Images captured by multiview radial imaging systems. Bottom row: Scene information recovered from the images in the top row. (a) The 3D structure of a piece of bread is recovered. (b) The analytic BRDF model parameters for red satin paint are estimated and used to render a teapot. (c) The 3D structure of a face is recovered. (d) The texture map (top and all sides) of a cylindrical object is captured. (e) The complete geometry of a toy head is recovered. For the results in (a-d) only a single image was used and for (e) two images were used.

Abstract

In this paper, we present a class of imaging systems, called *radial imaging systems*, that capture a scene from a large number of viewpoints within a single image, using a camera and a curved mirror. These systems can recover scene properties such as geometry, reflectance, and texture. We derive analytic expressions that describe the properties of a complete family of radial imaging systems, including their loci of viewpoints, fields of view, and resolution characteristics. We have built radial imaging systems that, from a single image, recover the frontal 3D structure of an object, generate the complete texture map of a convex object, and estimate the parameters of an analytic BRDF model for an isotropic material. In addition, one of our systems can recover the complete geometry of a convex object by capturing only two images. These results show that radial imaging systems are simple, effective, and convenient devices for a wide range of applications in computer graphics and computer vision.

CR Categories: I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Imaging geometry,Reflectance; I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Stereo

Keywords: radial imaging, multiview imaging, catadioptric imaging, 3D reconstruction, stereo, BRDF estimation, texture mapping.

1 Multi-Viewpoint Imaging

Many applications in computer graphics and computer vision require the same scene to be imaged from multiple viewpoints. The tradi-

tional approach is to either move a single camera with respect to the scene and sequentially capture multiple images [Levoy and Hanrahan 1996; Gortler et al. 1996; Peleg and Herman 1997; Shum and He 1999; Seitz and Kim 2002], or to simultaneously capture the same images using multiple cameras located at different viewpoints [Kanade et al. 1996; Kanade et al. 1997]. Using a single camera has the advantage that the radiometric properties are the same across all the captured images. However, this approach is only applicable to static scenes and requires precise estimation of the camera's motion. Using multiple cameras alleviates these problems, but requires the cameras to be synchronized. More importantly, the cameras must be radiometrically and geometrically calibrated with respect to each other. Furthermore, to achieve a dense sampling of viewpoints such systems need a large number of cameras – an expensive proposition.

In this paper, we develop a class of imaging systems called *radial imaging systems* that capture the scene from multiple viewpoints instantly within a single image[1]. As only one camera is used, all projections of each scene point are subjected to the same radiometric camera response. Moreover, since only a single image is captured, there are no synchronization requirements. Radial imaging systems consist of a conventional camera looking through a hollow rotationally symmetric mirror (e.g., a truncated cone) polished on the inside. The field of view of the camera is folded inwards and consequently the scene is captured from multiple viewpoints within a single image. As the results in Figure 1 illustrate, this simple principle enables radial imaging systems to solve a variety of problems in computer graphics and computer vision. In this paper, we demonstrate the use of radial imaging systems for the following applications:

Reconstructing Scenes with Fewer Ambiguities: One type of radial imaging system captures scene points multiple times within an image. Thus, it enables recovery of scene geometry from a single

*e-mail: {sujit,nayar}@cs.columbia.edu

---

[1]Although an image captured by a radial imaging system includes multiple viewpoints, each viewpoint does not capture a 'complete' image of the scene, unlike the imaging systems proposed in [Unger et al. 2003; Levoy et al. 2004].

image. We show that the epipolar lines for such a system are radial. Hence, unlike traditional stereo systems, ambiguities occur in stereo matching only for edges oriented along radial lines in the image – an uncommon scenario. This inherent property enables the system to produce high quality geometric models of both fine 3D textures and macroscopic objects, as shown in Figures 1(a) and 1(c), respectively.

Sampling and Estimating BRDFs: Another type of radial imaging system captures a sample point from a large number of viewpoints in a single image. These measurements can be used to fit an analytical Bidirectional Reflectance Distribution Function (BRDF) that represents the material properties of an isotropic sample point, as shown in Figure 1(b).

Capturing Complete Objects: A radial imaging system can be configured to look all around a convex object and capture its complete texture map (except possibly the bottom surface) in a single image, as shown in Figure 1(d). Capturing two such images with parallax, by moving the object or the system, yields the complete geometry of the object, as shown in Figure 1(e). To our knowledge, this is the first system with such a capability.

In summary, radial imaging systems can recover useful geometric and radiometric properties of scene objects by capturing one or at most two images, making them simple and effective devices for a variety of applications in graphics and vision. It must be noted that these benefits come at the cost of spatial resolution – the multiple views are projected onto a single image detector. Fortunately, with the ever increasing spatial resolution of today's cameras, this shortcoming becomes less significant. In our systems we have used 6 and 8 megapixel cameras and have found that the computed results have adequate resolution for our applications.

2 Related Work

Several mirror-based imaging systems have been developed that capture a scene from multiple viewpoints within a single image [Southwell et al. 1996; Nene and Nayar 1998; Gluckman et al. 1998; Gluckman and Nayar 1999; Han and Perlin 2003]. These are specialized systems designed to acquire a specific characteristic of the scene; either geometry or appearance. In this paper, we present a complete family of radial imaging systems. Specific members of this family have different characteristics and hence are suited to recover different properties of a scene, including, geometry, reflectance, and texture.

One application of multiview imaging is to recover scene geometry. Mirror-based, single-camera stereo systems [Nene and Nayar 1998; Gluckman and Nayar 1999] instantly capture the scene from multiple viewpoints within an image. Similar to conventional stereo systems, they measure disparities along a single direction, for example along image scan-lines. As a result, ambiguities arise for scene features that project as edges parallel to this direction. The panoramic stereo systems in [Southwell et al. 1996; Gluckman et al. 1998; Lin and Bajcsy 2003] have radial epipolar geometry for two outward looking views; i.e., they measure disparities along radial lines in the image. However, they suffer from ambiguities when reconstructing vertical scene edges as these features are mapped onto radial image lines. In comparison, our systems do not have such large panoramic fields of view. Their epipolar lines are radial but the only ambiguities that arise in matching and reconstruction are for scene features that project as edges oriented along radial lines in the image, a highly unusual occurrence[2]. Thus, radial imaging systems are able to compute the structures of scenes with less ambiguity than previous methods.

Sampling the appearance of a material requires a large number of images to be taken under different viewing and lighting conditions. Mirrors have been used to expedite this sampling process. For example, Ward [1992] and Dana [2001] have used curved mirrors to capture in a single image multiple reflections of a sample point that

correspond to different viewing directions for a single lighting condition. We show that one of our radial imaging systems achieves the same goal. It should be noted that a dense sampling of viewing directions is needed to characterize the appearance of specular materials. Our system uses multiple reflections within the curved mirror to obtain dense sampling along multiple closed curves in the 2D space of viewing directions. Compared to [Ward 1992; Dana 2001], this system captures fewer viewing directions. However, the manner in which it samples the space of viewing directions is sufficient to fit analytic BRDF models for a large variety of isotropic materials, as we will show. Han and Perlin [2003] also use multiple reflections in a mirror to capture a number of discrete views of a surface with the aim of estimating its Bidirectional Texture Function (BTF). Since the sampling of viewing directions is coarse and discrete, the data from a single image is insufficient to estimate the BRDFs of points or the continuous BTF of the surface. Consequently, multiple images are taken under different lighting conditions to obtain a large number of view-light pairs. In comparison, we restrict ourselves to estimating the parameters of an analytic BRDF model for an isotropic sample point, but can achieve this goal by capturing just a single image. Our system is similar in spirit to the conical mirror system used by Hawkins et al. [2005] to estimate the phase function of a participating medium. In fact, the system of Hawkins et al. [2005] is a specific instance of the class of imaging systems we present.

Some applications require imaging all sides of an object. Peripheral photography [Davidhazy 1987] does so in a single photograph by imaging a rotating object through a narrow slit placed in front of a moving film. The captured images, called periphotographs or cyclographs [Seitz and Kim 2002], provide an inward looking panoramic view of the object. We show how radial imaging systems can capture the top view as well as the peripheral view of a convex object in a single image, without using any moving parts. We also show how the complete 3D structure of a convex object can be recovered by capturing two such images, by translating the object or the imaging system in between the two images.

3 Radial Imaging Systems

To understand the basic principle underlying radial imaging systems, consider the example configuration shown in Figure 2(a). It consists of a camera looking through a hollow cone that is mirrored on the inside. The axis of the cone and the camera's optical axis are coincident. The camera images scene points both directly and after reflection by the mirror. As a result, scene points are imaged from different viewpoints within a single image.

The imaging system in Figure 2(a) captures the scene from the real viewpoint of the camera as well as a circular locus of virtual viewpoints produced by the mirror. To see this consider a radial slice of the imaging system that passes through the optical axis of the camera, as shown in Figure 2(b). The real viewpoint of the camera is located at O. The mirrors m_1 and m_2 (that are straight lines in a radial slice) produce the two virtual viewpoints V_1 and V_2, respectively, which are reflections of the real viewpoint O. Therefore, each radial slice of the system has two virtual viewpoints that are symmetric with respect to the optical axis. Since the complete imaging system includes a continuum of radial slices, it has a circular locus of virtual viewpoints whose center lies on the camera's optical axis.

Figure 2(c) shows the structure of an image captured by a radial imaging system. The three viewpoints O, V_1, and V_2 in a radial slice project the scene onto a radial line in the image, which is the intersection of the image plane with that particular slice. This radial image line has three segments – JK, KL, and LM, as shown in Figure 2(c). The real viewpoint O of the camera projects the scene onto the central part KL of the radial line, while the virtual viewpoints V_1 and V_2 project the scene onto JK and LM, respectively. The three viewpoints (real and virtual) capture only scene points that lie on that particular radial slice. If P is such a scene point, it is imaged thrice (if visible

---

[2]In our systems, ambiguities arise for vertical scene edges only if they project onto the *vertical radial line* in the image.
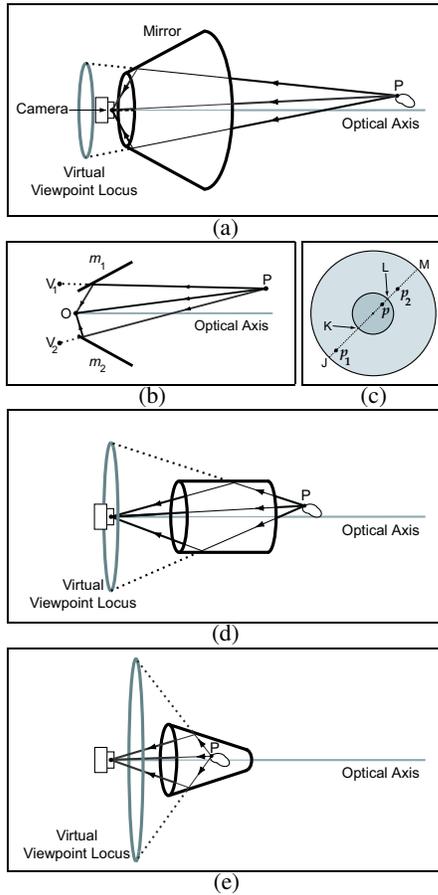
Figure 2: (a) Radial imaging system with a cone mirrored on the inside that images the scene from a circular locus of virtual viewpoints in addition to the real viewpoint of the camera. The axis of the cone and the camera's optical axis are coincident. (b) A radial slice of the system shown in (a). (c) Structure of the image captured by the system shown in (a). (d) Radial imaging system with a cylinder mirrored on the inside. (e) Radial imaging system with a cone mirrored on the inside. In this case, the apex of the cone lies on the other side of the camera compared to the system in (a).

to all three viewpoints) along the corresponding radial image line at locations p, p_1, and p_2, as shown in Figure 2(c). Since this is true for every radial slice, the epipolar lines of such a system are radial. Since all radial image lines have three segments (JK, KL, and LM) and the lengths of these segments are independent of the chosen radial image line, the captured image has the form of a donut. The camera's real viewpoint captures the scene directly in the inner circle, while the annulus corresponds to reflection of the scene – the scene as seen from the circular locus of virtual viewpoints.

Varying the parameters of the conical mirror in Figure 2(a) and its distance from the camera, we obtain a continuous family of radial imaging systems, two instances of which are shown in Figures 2(d) and 2(e). The system in Figure 2(d) has a cylindrical mirror. The system in Figure 2(e) has a conical mirror whose apex lies on the other side of the camera compared to the one in Figure 2(a). These systems differ in the geometric properties of their viewpoint loci and their fields of view, making them suitable for different applications. However, the images that they all capture have the same structure as in Figure 2(c).

Multiple circular loci of virtual viewpoints can be generated by choosing a mirror that reflects light rays multiple times before being captured by the camera. For instance, two circular loci of virtual viewpoints are obtained by allowing light rays from the scene

to reflect atmost twice before entering the camera. In this case, the captured image will have an inner circle, where the scene is directly imaged by the camera's viewpoint, surrounded by two annuli, one for each circular locus of virtual viewpoints. Later we show how such a system with multiple circular loci of virtual viewpoints can be used.

In this paper, for the sake of simplicity, we restrict ourselves to radial imaging systems with conical and cylindrical (which is just a special case) mirrors, which appear as lines in the radial slices. It should be noted that in general the mirrors only have to be rotationally symmetric; they can have more complex cross-sections.

4 Properties of a Radial Imaging System

We now analyze the properties of a radial imaging system. For simplicity, we restrict ourselves to the case where light rays from the scene reflect at most once in the mirror before being captured by the camera. In Section 5.3, we will analyze a system with multiple reflections. For illustration, we will use Figure 3 which shows a radial slice of the system shown in Figure 2(a). However, the expressions we derive hold for all radial imaging systems including the ones shown in Figures 2(d) and 2(e). A cone can be described using three parameters – the radius r of one end (in our case, the end near the camera), its length l, and the half-angle β at its apex, as shown in Figure 3(a). The complete system can be described using one more parameter – the field of view (fov) 2θ of the camera[3]. To differentiate between the configurations in Figures 2(a) and 2(e), we use the following convention: if the cone's apex and the camera lie on the same side of the cone, $\beta \geq 0$; else $\beta < 0$. Therefore, for the systems shown in Figures 2(a), (d), and (e), $\beta > 0$, $\beta = 0$, and $\beta < 0$, respectively.

The near end of the cone should be placed at a distance $d = r\cot(\theta)$ from the camera's real viewpoint so that the extreme rays of the camera's fov graze the near end, as shown in Figure 3(a). Such a d would ensure that the entire fov of the camera is utilized.

4.1 Viewpoint Locus

In Section 3 we saw that radial imaging systems have a circular locus of virtual viewpoints. We now examine how the size and location of this circular locus varies with the parameters of the system. Since the system is rotationally symmetric, we can do this analysis in 2D by determining the location of the virtual viewpoints in the radial slice shown in Figure 3(a). The virtual viewpoints V_1 and V_2 in a radial slice are the reflections of the camera's real viewpoint O produced by the mirrors m_1 and m_2, respectively. The distance of the virtual viewpoints from the optical axis gives the radius v_r of the circular locus of virtual viewpoints, which can be shown to be

$$v_r = 2r\cos(\beta)\sin(\theta - \beta)\csc(\theta). \qquad (1)$$

The distance (along the optical axis) of the virtual viewpoints from the real viewpoint of the camera is the distance v_d between the circular locus of virtual viewpoints and the camera's real viewpoint:

$$v_d = -2r\sin(\beta)\sin(\theta - \beta)\csc(\theta). \qquad (2)$$

It is interesting to note that when $\beta > 0$, as in the system shown in Figure 2(a), $v_d < 0$, implying that the virtual viewpoint locus is located behind the real viewpoint of the camera. In configurations with $\beta = 0$, as in Figure 2(d), the center of the circular virtual viewpoint locus is at the real viewpoint of the camera. Finally, the circular locus moves in front of the camera's real viewpoint for configurations with $\beta < 0$, as in the one shown in Figure 2(e).

The length of the cone determines how many times light rays from the scene reflect in the mirror before being captured by the camera. Since in this section we consider systems that allow light rays from the scene to be reflected at most once, from Figure 3(a) it can be shown that the length l of the cone should be less than l', where

$$l' = 2r\cos(\beta)\cos(\theta - 2\beta)\csc(\theta - 3\beta). \qquad (3)$$

For ease of analysis, from this point onwards we assume that $l = l'$.

---

[3]The field of view of a camera in a radial imaging system is the minimum of the camera's horizontal and vertical fields of view.
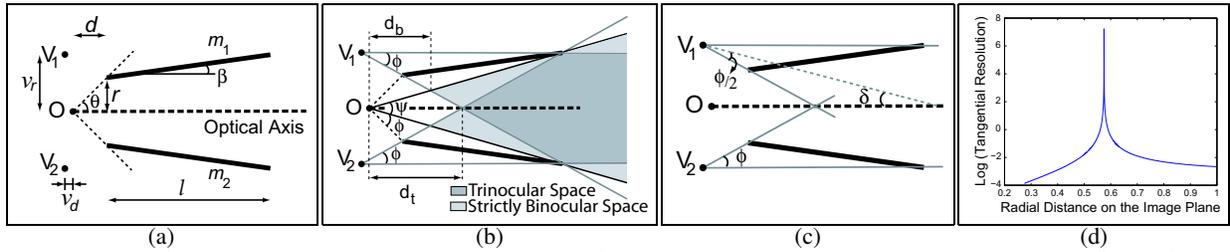
Figure 3: Properties of a Radial Imaging System. (a) Radial slice of the imaging system shown in Figure 2(a). (b) The fields of view of the viewpoints in a radial slice. (c) The orientation of a virtual viewpoint in a radial slice. (d) The tangential resolution of an image captured by an imaging system with $\beta = 12°$, $r = 3.5$ cm, and $\theta = 45°$ for a scene plane parallel to the image plane located at a distance of 50 cm from the camera's real viewpoint. The radial distance is measured on the image plane at unit distance from the camera's real viewpoint.

4.2 Field of View

We now analyze how the fov of the viewpoints in a radial slice depend on the parameters of the imaging system. Consider the radial slice shown in Figure 3(b). As we can see, the fov ϕ of a virtual viewpoint is the portion of the fov of the camera that is incident on the corresponding mirror and is given by

$$\phi = \arctan(\frac{2\cos(\theta - 2\beta)\sin(\theta)\sin(\theta - \beta)}{\sin(\theta - 3\beta) + 2\sin(\theta)\cos(\theta - 2\beta)\cos(\theta - \beta)}). \quad (4)$$

Therefore, the effective fov ψ of the real viewpoint of the camera is the remaining portion of the camera's fov, which is

$$\psi = 2(\theta - \phi). \quad (5)$$

The number of image projections of any given scene point equals the number of viewpoints in the corresponding radial slice that can 'see' it. This in turn depends on where the scene point lies. If a scene point lies in the trinocular space – area common to the fovs of all viewpoints in a radial slice – it is imaged thrice. On the other hand, if a point lies in the binocular space – area common to the fovs of at least two viewpoints – it is imaged at least twice. Figure 3(b) shows the trinocular and binocular spaces. The scene point in the trinocular space closest to O is obtained by intersecting the fovs of the virtual viewpoints. This point lies at a distance

$$d_t = r\sin(2\theta - 2\beta)\csc(\theta)\csc(\theta - 2\beta) \quad (6)$$

from O. Similarly, by intersecting the effective fov of the camera's real viewpoint and the fov of a virtual viewpoint, we obtain the distance of the two scene points in the binocular space closest to O as

$$d_b = r\sin(2\theta - 2\beta)\cos(\theta - \phi)\csc(\theta)\csc(2\theta - 2\beta - \phi). \quad (7)$$

Examining the expression for d_t tells us that for systems with $\beta > 0$ (Figure 2(a)), the trinocular space exists only if $\theta > 2\beta$. On the other hand, in configurations with $\beta \leq 0$ (Figures 2(d) and 2(e)), the fovs of all viewpoints in a radial slice always overlap. Note that the binocular space exists in all cases.

We define the orientation of a virtual viewpoint as the angle δ made by the central ray in its fov with the optical axis, as shown in Figure 3(c). It can be shown, using simple geometry, that δ is given by

$$\delta = (\theta - \frac{\phi}{2} - 2\beta)t. \quad (8)$$

Here, $t = 1$, if the central rays of the virtual viewpoint fovs meet in front of the camera's real viewpoint, i.e., the fovs converge, and $t = -1$ otherwise. It can be shown that when $\beta \leq 0$, the virtual viewpoint fovs always converge. When $\beta > 0$, the fovs converge only if $\theta > 3\beta$.

4.3 Resolution

We now examine the resolution characteristics of radial imaging systems. For simplicity, we analyze resolutions along the radial and tangential directions of a captured image separately. As described in Section 3, a radial line in the image has three segments – one for each viewpoint in the corresponding radial slice. Therefore, in a radial line the spatial resolution of the camera is split among the three viewpoints. Using simple geometry, it can be shown that on a radial

image line, the ratio of the lengths of the line segments belonging to the camera's real viewpoint and a virtual viewpoint is $\frac{\cos(\theta)}{\cos(\theta - 2\beta)}$.

We now study resolution in the tangential direction. Consider a scene plane Π_s parallel to the image plane located at a distance w from the camera's real viewpoint. Let a circle of pixels of radius ρ_i on the image plane image a circle of radius ρ_s on the scene plane Π_s; the centers of both circles lie on the optical axis of the camera. We then define tangential resolution, for the circle on the image plane, as the ratio of the perimeters of the two circles $= \rho_i/\rho_s$. If a circle of pixels on the image plane does not see the mirror, its tangential resolution is $1/w$ (assuming focal length is 1). To determine the tangential resolution for a circle of pixels that sees the mirror, we need to compute the mapping between a pixel on the image plane and the point it images on the scene plane. This can be derived using the geometry shown in Figure 3(a). From this mapping we can determine the radius ρ_s of the circle on the scene plane Π_s that is imaged by a circle of pixels of radius ρ_i on the image plane. Then, tangential resolution is $\rho_i/\rho_s =$

$$\frac{\rho_i \sin(\theta)(\cos(2\beta) + \rho_i \sin(2\beta))}{2r\sin(\theta - \beta)(\cos(\beta) + \rho_i \sin(\beta)) - w\sin(\theta)(\rho_i \cos(2\beta) - \sin(2\beta))}.$$

Note that tangential resolution is depth dependent – it depends on the distance w of the scene plane Π_s. For a given w, there exists a circle of radius ρ_i on the image plane, which makes the denominator of the above expression zero. Consequently, that circle on the image plane has infinite tangential resolution[4], as it is imaging a single scene point – the scene point on Π_s that lies on the optical axis. This property can be seen in all the images captured by radial imaging systems shown in Figure 1. In Section 5.3 we exploit this property to estimate the BRDF of a material using a single image. The tangential resolution for a particular radial imaging system and a chosen scene plane is shown in Figure 3(d).

We have built several radial imaging systems which we describe next. The mirrors in these systems were custom-made by Quintesco, Inc. The camera and the mirror were aligned manually by checking that in a captured image the circles corresponding to the two ends of the mirror are approximately concentric. In our experiments, we found that very small errors in alignment did not affect our results in any significant way.

5 Cylindrical Mirror

We now present a radial imaging system that consists of a cylinder mirrored on the inside. Such a system is shown in Figure 2(d). In this case, the half-angle $\beta = 0$.

5.1 Properties

Let us examine the properties of this specific imaging system. Putting $\beta = 0$ in Equations 1 and 2, we get $v_r = 2r$ and $v_d = 0$. Therefore, the virtual viewpoints of the system form a circle of radius $2r$ around the optical axis centered at the real viewpoint of the camera. It can be shown from Equations 4 and 5 that, in this system, the fov ϕ of

---

[4]In practice, tangential resolution is always finite as it is limited by the resolution of the image detector.
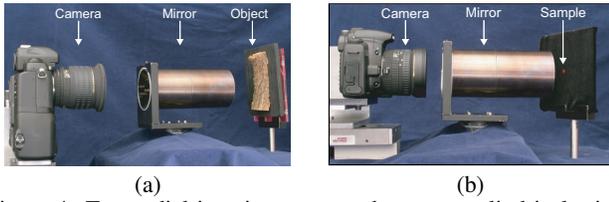
(a)   (b)

Figure 4: Two radial imaging systems that use a cylindrical mirror of radius 3.5 cm and length 16.89 cm. (a) System used for reconstructing 3D textures that has a Kodak DCS760 camera with a Sigma 20mm lens. (b) System used to estimate the BRDF of a sample point that has a Canon 20D camera with a Sigma 8mm Fish-eye lens.

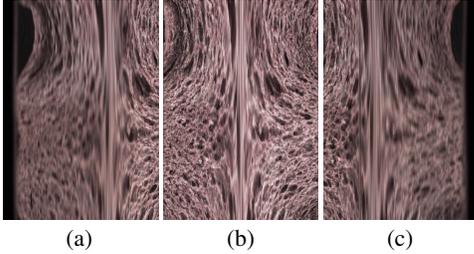(a)   (b)   (c)

Figure 5: The left (a), central (b), and right (c) view images constructed from the captured image shown in Figure 1(a).

the virtual viewpoints is always smaller than the effective fov ψ of the real viewpoint of the camera. Another interesting characteristic of the system is that the fovs of its viewpoints always converge. As a result, it is useful for recovering properties of small nearby objects. Specifically, we use the system to reconstruct 3D textures and estimate the BRDFs of materials.

5.2   3D Texture Reconstruction and Synthesis

A radial imaging system can be used to recover, from a single image, the depth of scene points that lie in its binocular or trinocular space, as these points are imaged from multiple viewpoints. We use a radial imaging system with a cylindrical mirror to recover the geometry of 3D texture samples.   Figure 4(a) shows the prototype we built. The camera captures 3032×2008 pixel images. The radial image lies within a 1791×1791 pixel square in the captured image. In this configuration, the fovs of the three viewpoints in a radial slice intercept line segments of equal length i.e., 597 pixels on the corresponding radial image line. An image of a slice of bread captured by this system is shown in Figure 1(a). Observe that the structure of this image is identical to that shown in Figure 2(c).

Let us now see how we can recover the structure of the scene from a single image. To determine the depth of a particular scene point, its projections in the image, i.e., corresponding points, have to be identified via stereo matching. As the epipolar lines are radial, the search for corresponding points needs to be restricted to a radial line in the image. However, most stereo matching techniques reported in literature deal with image pairs with horizontal epipolar lines [Scharstein and Szeliski 2002]. Therefore, it would be desirable to convert the information captured in the image into a form where the epipolar lines are horizontal. Recall that a radial line in the image has three parts – JK, KL, and LM, one for each viewpoint in the corresponding radial slice (See Figure 2(c)). We create a new image called the central view image by stacking the KL parts of successive radial lines. This view image corresponds to the central viewpoint in the radial slices. We create similar view images for the virtual viewpoints in the radial slices – the left view image by stacking the LM parts of successive radial lines and the right view image by stacking the JK parts. To account for the reflection of the scene by the mirror the contents of each JK and LM lines are flipped. Figure 5 shows the three 597×900 view images constructed from the captured image shown in Figure 1(a). Observe that the epipolar lines are now horizontal. Thus, traditional stereo matching algorithms can now be directly applied.
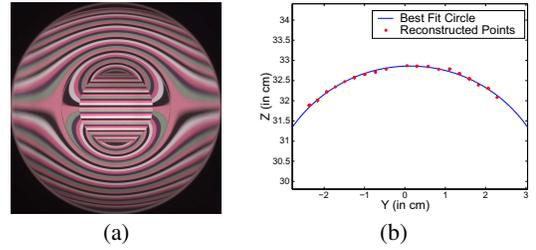

(a)   (b)

Figure 6: Determining the reconstruction accuracy of the cylindrical mirror system shown in Fig 4(a). (a) Captured image of the inside of a section of a hollow cylinder. (b) Some reconstructed points and the best fit circle corresponding to the vertical radial line in the image. (See text for details.)

For the 3D reconstruction results in this paper, we used a window-based method for stereo matching with normalized cross-correlation as the similarity metric [Scharstein and Szeliski 2002]. The central view image (Figure 5(b)) was the reference with which we matched the left and right view images (Figures 5(a) and 5(c)).   The left and right view images look blurry in regions that correspond to the peripheral areas of the captured image, due to optical aberrations introduced by the curvature of the mirror. To compensate for this, we took an image of a planar scene with a large number of dots. We then computed the blur kernels for different columns in the central view image that transform the 'dot' features to the corresponding features in the left and right view images. The central view image was blurred with these blur kernels prior to matching. This transformation, though an approximation, makes the images similar thereby making the matching process more robust. Once correspondences are obtained, the depths of scene points can be computed. The reconstructed 3D texture of the bread sample – a disk of diameter 390 pixels – is shown in Figure 1(a).

To determine the accuracy of the reconstructions obtained, we imaged an object of known geometry – the inside of a section of a hollow cylinder of radius 3.739 cm. The captured image is shown in Figure 6(a), in which the curvature of the object is along the vertical direction. We reconstructed 145 points along the vertical radial image line and fit a circle to them, as shown in Figure 6(b). The radius of the best-fit circle is 3.557 cm and the rms error of the fit is 0.263 mm, indicating very good reconstruction accuracy.

Figures 7 (a,b) show another example of 3D texture reconstruction – of the bark of a tree. Since we now have both the texture and the geometry, we can synthesize novel 3D texture samples. This part of our work is inspired by 2D texture synthesis methods [Efros and Leung 1999; Efros and Freeman 2001; Kwatra et al. 2003] that, starting from an RGB texture patch, create novel 2D texture patches. To create novel 3D texture samples, we extended the simple image quilting algorithm of Efros and Freeman [2001] to operate on texture patches that in addition to having the three (RGB) color channels have another channel – the z value at every pixel[5].

The 3D texture shown in Figure 7(b) was quilted to obtain a large 3D texture patch, which we then wrapped around a cylinder to create a tree trunk. This trunk was then rendered under a moving point light source and inserted into an existing picture to create the images in Figures 7(c) and 7(d). The light source moves from left to right as one goes from (c) to (d). Notice how the cast shadows within the

---

[5]To incorporate the z channel, we made the following changes to [Efros and Freeman 2001]: (a) When computing the similarity of two regions, for the RGB intensity channels, we use Sum-of-Squared Differences (SSD), while for the z channel, the z values in each region are made zero-mean and then SSD is computed. The final error is a linear combination of intensity and z-channel errors. (b) To ensure that no depth discontinuities are created when pasting a new block into the texture, we do the following. We compute the difference of the means of the z values in the overlapping regions of the texture and the new block. This difference is used to offset z values in the new block.
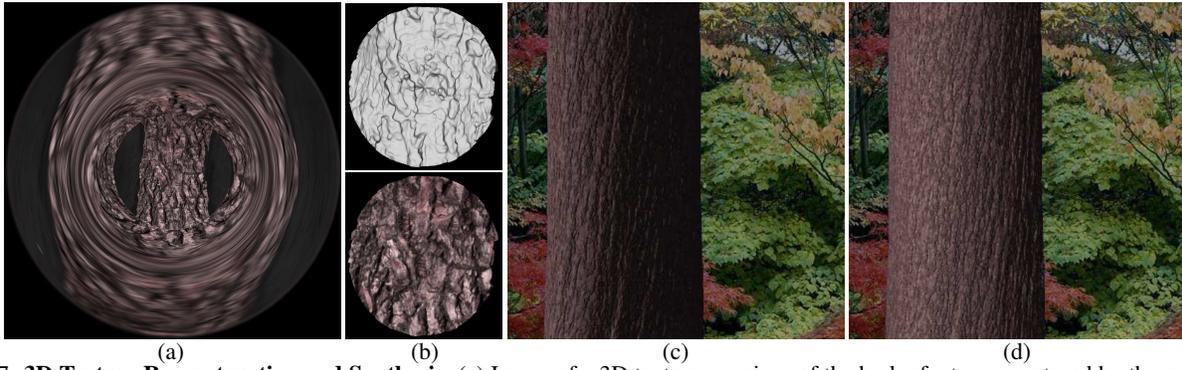
| (a) | (b) | (c) | (d) |

Figure 7: **3D Texture Reconstruction and Synthesis.** (a) Image of a 3D texture – a piece of the bark of a tree – captured by the cylindrical mirror imaging system shown in Figure 4(a). (b) Shaded and texture mapped views of the reconstructed piece of bark. (c-d) The reconstructed 3D texture was used to synthesize a large 3D texture sample which was then wrapped around a cylinder to create a tree trunk. This trunk was rendered under a moving point light source (left to right as one goes from c to d) and then inserted into another image.

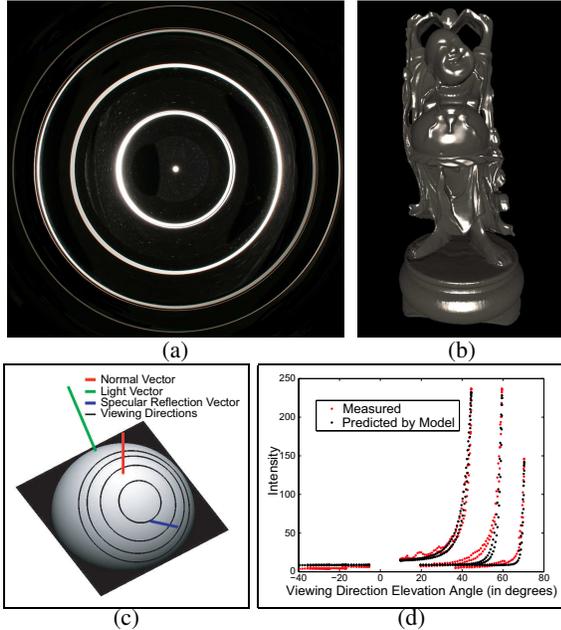| (a) | (b) |

| (c) | (d) |

Figure 8: **BRDF Sampling and Estimation.** (a) Image of a metallic paint sample captured by the cylindrical imaging system shown in Figure 4(b). (b) A model rendered with the metallic paint BRDF estimated from (a). (c) Plot showing the sample normal, light source direction, and the viewing directions for the image in (a). (d) Plot comparing the measured radiances in the red channel for different viewing directions, with those predicted by the fitted analytical model.

bark of the tree differ in the two images.

To recover the geometry of 3D textures Liu et al. [2001] apply shape-from-shading techniques to a number of images taken under different illumination conditions. These images also come in handy at the time of texture synthesis as they can be used to impart view dependent effects to the appearance of the new texture. In contrast, we capture both the texture and the geometry of a 3D texture in a single image. However, since we have only one image of the sample and do not know its material properties, we implicitly make the assumption that the sample is Lambertian when we perform 3D texture synthesis.

5.3 BRDF Sampling and Estimation

We now show how a radial imaging system can be used to estimate the parameters of an analytic BRDF model for an isotropic material. We make the observation that points on the optical axis of a radial imaging system lie on all radial slices. Hence, if we place a sample point on the optical axis of the system, it is imaged by all viewpoints. In fact, such a point is imaged along a circle on the image plane – the

tangential resolution for that circle is infinite. We can get more viewpoints by letting light rays from the sample point reflect in the mirror multiple times before being captured by the camera. As discussed earlier, this would result in the sample point being imaged from several circular loci of virtual viewpoints. It can be shown that the minimum length of the cylinder that is needed for realizing n circular loci of virtual viewpoints is given by $l_n = 2(n-1)r\cot(\theta)$, $n > 1$. The virtual viewpoints of this system form concentric circles of radii $2r$, $4r$, $\cdots$, $2nr$.

Our prototype system, whose camera captures 3504×2336 pixel images, is shown in Figure 4(b). The radial image lies within a 2261×2261 pixel square in the captured image. Figure 8(a) shows an image of a metallic paint sample taken by this system. As one can see, the sample is imaged along four concentric circles, implying that it is viewed from four circular loci of virtual viewpoints. We placed the sample and a distant point light source such that the radiance along the specular angle was measured by at least one viewpoint[6].

To understand the viewing directions that image the sample point, consider Figure 8(c), which shows the hemisphere of directions centered around the normal of the sample point. The four virtual viewpoint circles map to concentric circles on this hemisphere. Note that one of the viewing circles intersects the specular angle. The radiance measurements for these viewing directions and the fixed lighting direction are then used to fit an analytical BRDF model. We use the Oren-Nayar model for the diffuse component and the Torrance-Sparrow model for the specular component. Due to space constraints we are only showing the fit of the computed analytical model to the red channel of the measured data, in Figure 8(d). The plots for the green and blue channels are similar. We can now render objects with the estimated BRDF, as shown in Figure 8(b). Figure 1(b) shows another example. It should be noted that our approach to sampling appearance cannot be used if the material has a very sharp specular component as then the specularity might not be captured by any of the four virtual viewpoint circles.

6 Conical Mirror

In this section, we present radial imaging systems with cones of different parameters. Having unequal radii at the ends allows for greater flexibility in selecting the size and location of the viewpoint locus and the fields of view.

6.1 Properties

As we discussed in Section 4, β is one of the parameters that defines a radial imaging system. Let us consider separately the cases of $\beta > 0$ and $\beta < 0$. For systems with $\beta > 0$, depending on the application's needs, the virtual viewpoint locus can be varied to lie in between the

---

[6]For the geometry of our prototype this was achieved by rotating the sample by 27° about the vertical axis and positioning a distant point light source at an angle of 45° with the normal to the sample in the horizontal plane.

real viewpoint of the camera and $v_d = -r\tan(\theta/2)$. There is also flexibility in terms of fields of view – the virtual viewpoint fovs can be lesser than, equal to, or greater than the effective fov of the real viewpoint of the camera. Also, the viewpoint fovs may converge or diverge. For systems with $\beta < 0$, the locus of virtual viewpoints can be varied to lie in between the camera's real viewpoint and $v_d = r\cot(\theta/2)$. Unlike configurations with $\beta > 0$, in these systems the virtual viewpoint fovs are smaller than the effective fov of the real viewpoint of the camera. Also, the viewpoint fovs always converge.

6.2 Reconstruction of 3D Objects

We now describe how to reconstruct 3D objects using a radial imaging system with $\beta > 0$ – like the one shown in Figure 2(a). Using a cylindrical mirror, as in the previous section, causes the fovs of the viewpoints of the system to converge. Consequently, such a system is suited for recovering the properties of small nearby objects. In order to realize a system that can be used for larger and more distant objects, we would like the fovs of the virtual viewpoints to 'look straight', i.e., we would like the central ray of each virtual viewpoint's fov to be parallel to the optical axis. This implies that δ – the angle made by the central ray in a virtual viewpoint's fov with the optical axis – should be zero. Examining Equations 3 and 8 tells us that for this to be true the length of the cone has to be infinite – clearly an impractical solution. Therefore, we pose the following problem: Given the fov of the camera, the radius of the near end of the cone, and the ratio γ of the effective fovs of the real and virtual viewpoints, determine the cone's half-angle β at its apex and its length l. A simple geometrical analysis yields the following solution:

$$\beta = \frac{\theta(\gamma+1)}{2(\gamma+2)}, \quad l = \frac{r\sin(2\theta/(\gamma+2))\cos(\beta)}{\sin(\theta)\sin(\theta(\gamma-1)/(2(\gamma+2)))}, \quad \gamma > 1. \quad (9)$$

The prototype we built based on the above solution is shown in Figure 9(a). The radial image lies within a 2158×2158 pixel square of the 3504×2336 pixel captured image. The effective fov of the camera's real viewpoint intercepts 1078 pixels along a radial line in the image. The fovs of the two virtual viewpoints intercept 540 pixels each. We have used this system to compute the 3D structures of faces, a problem that has attracted much interest in recent years. Commercial face scanning systems are now available, such as those from Cyberware and Eyetronics, which produce high quality face models. However, these use sophisticated hardware and are expensive.

Figures 1(c) and 10(a) show two images captured by the system in Figure 9(a). Since these images are identical in structure to those taken by the system in Section 5.2, we can create the three view images, perform stereo matching and do reconstruction as before. However, there is one small difference. In a radial slice, the effective image line (analogous to the image plane) for a virtual viewpoint is the reflection of the real image line. Since the mirrors are not orthogonal to the real image line in this case, for any two viewpoints in a slice their effective image lines would not be parallel to the line joining the two viewpoints. Therefore, before matching[7] two view images, they must be rectified.

A view of the 470×610 pixel face model reconstructed from the image in Figure 10(a) is shown in Figure 10(b). Figure 1(c) shows another example. To determine the accuracy of reconstructions produced by this system, we imaged a plane placed 40 cm from the camera's real viewpoint and computed its geometry. The captured image is shown in Figure 11(a). The rms error obtained by fitting a plane to the reconstructed points is 0.83 mm, indicating high accuracy. Figure 11(b) shows the slice of the best-fit plane and some of the reconstructed points corresponding to the vertical radial line in the captured image.

---

[7]Correspondence matches in specular regions (eyes and nose tip, identified manually) and texture-less regions are discarded. The depth at such a pixel is obtained by interpolating the depths at neighboring pixels with valid matches.
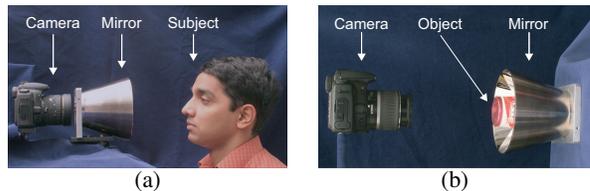

(a)   (b)

Figure 9: Radial imaging systems comprised of a cone of length 12.7 cm and radii 3.4 cm and 7.4 cm at the two ends. The half-angle at the apex of the cone is $17.48°$. Both systems use a Canon 20D camera. (a) System used for reconstructing objects such as faces. A Sigma 8mm fish-eye lens was used in this system. (b) System used to capture the complete texture and geometry of a convex object. A Canon 18-55 mm lens was used in this system.

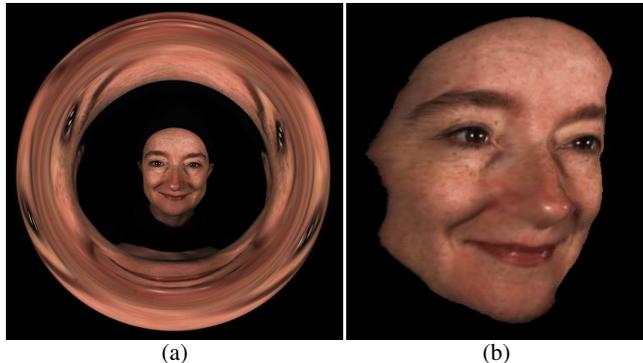(a)   (b)

Figure 10: **Recovering the Geometry of a Face.** (a) Image of a face captured by the conical mirror imaging system shown in Figure 9(a). (b) A view of the reconstructed face.

6.3 Capturing Complete Texture Maps

We now show how a radial imaging system can be used to capture, in a single image, the entire texture map of a convex object – its top and all sides (the bottom surface is not always captured). To do so, the object must be imaged from a locus of viewpoints that goes all around it. Therefore, the radius of the circular locus of virtual viewpoints should be greater than the radius of the smallest cylinder that encloses the object; the cylinder's axis being coincident with the optical axis of the camera. Since radial imaging systems with $\beta < 0$, like the one in Figure 2(e), have virtual viewpoint loci of larger radii, they are best suited for this application. While the real viewpoint of the camera captures the top view of the object, the circular locus of virtual viewpoints images the side views. Thus, the captured images have more information than the cyclographs presented in [Seitz and Kim 2002]. Figure 9(b) shows our prototype system. The radial image lies within a 2113×2113 pixel square of the 3504×2336 pixel captured image. In a radial slice, the effective fov of the camera's real viewpoint intercepts 675 pixels on the corresponding radial image line, while the virtual viewpoints each intercept 719 pixels. An image of a conical object captured by this system is shown in Figure 12(a). Figure 12(b) shows a cone texture-mapped with this image. Another example, of a cylindrical object, is shown in Figure 1(d).

6.4 Recovering Complete Object Geometry

We have shown above how the complete texture map of a convex object can be captured in a single image using a radial imaging system with $\beta < 0$. If we take two such images, with parallax, we can compute the complete 3D structure of the object. Figures 1(e) and 13(a) show two images obtained by translating a toy head along the optical axis of the system by 0.5 cm[8]. Due to this motion of the object, the epipolar lines for the two images are radial. In order to use conventional stereo matching algorithms, we need to map radial lines to

---

[8]To move the object accurately, we placed it on a linear translation stage that was oriented to move approximately parallel to the camera's optical axis.
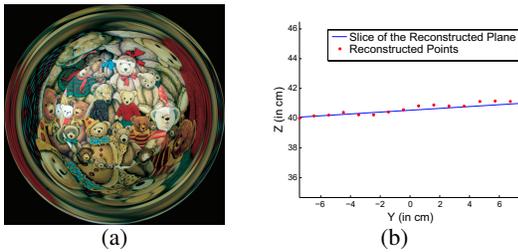
(a)   (b)

Figure 11: Determining the reconstruction accuracy of the system shown in Figure 9(a). (a) Captured image of a plane. (b) Some reconstructed points and the slice of the best-fit plane corresponding to the vertical radial line in the image. (See text for details.)
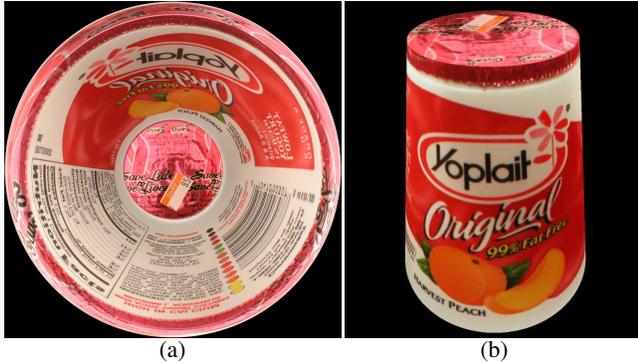

(a)   (b)

Figure 12: **Capturing the Complete Texture Map of a Convex Object.** (a) Image of a conical object captured by the system shown in Figure 9(b). (b) A cone texture-mapped with the image in (a).

horizontal lines. Therefore, we transform the captured images from Cartesian to polar coordinates – the radial coordinate maps to the horizontal axis[9]. As before, the two images are rectified. We then perform stereo matching on them and compute the 3D structure of the object. Figure 13(b) shows a view of the complete geometry of the object shown in Figure 13(a). To our knowledge, this is the first system capable of recovering the complete geometry of convex objects by capturing just two images.

7   Conclusion

In this paper, we have introduced a class of imaging systems called radial imaging systems that capture a scene from the real viewpoint of the camera as well as one or more circular loci of virtual viewpoints, instantly, within a single image. We have derived analytic expressions that describe the properties of a complete class of radial imaging systems. As we have shown, these systems can recover geometry, reflectance, and texture by capturing one or at most two images. In this work, we have focused on the use of conical mirrors for radial imaging. In future work, we would like to explore the benefits of using more complex mirror profiles. Another interesting direction is the use of multiple mirrors within a system. We believe that the use of multiple mirrors would yield even greater flexibility in terms of the imaging properties of the system, and at the same time enable us to optically fold the system to make it more compact.

Acknowledgments

References

DANA, K. J. 2001. BRDF/BTF Measurement Device. In *Proc. of ICCV*, 460–466.

---

[9]In this setup many scene features might project as radial edges in a captured image, giving rise to ambiguities in matching. The ambiguity advantage of having radial epipolar geometry is lost in this case.
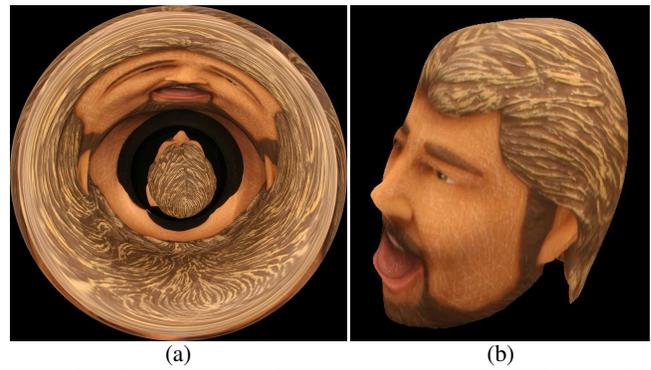

(a)   (b)

Figure 13: **Recovering the Complete Geometry of a Convex Object.** (a) Image of a toy head captured by the imaging system shown in Figure 9(b). (b) Recovered 3D model of the toy head shown in (a).

DAVIDHAZY, A. 1987. Peripheral Photography: Shooting full circle. *Industrial Photography 36*, 28–31.

EFROS, A. A., AND FREEMAN, W. T. 2001. Image Quilting for Texture Synthesis and Transfer. In *Proc. of SIGGRAPH*, 341–346.

EFROS, A. A., AND LEUNG, T. K. 1999. Texture Synthesis by Non-parametric Sampling. In *Proc. of ICCV*, 1033–1038.

GLUCKMAN, J., AND NAYAR, S. K. 1999. Planar Catadioptric Stereo: Geometry and Calibration. In *Proc. of CVPR*, 1022–1028.

GLUCKMAN, J., THOREK, K., AND NAYAR, S. K. 1998. Real time panoramic stereo. In *Proc. of Image Understanding Workshop*.

GORTLER, S., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. 1996. The Lumigraph. In *Proc. of SIGGRAPH*, 43–54.

HAN, J. Y., AND PERLIN, K. 2003. Measuring bidirectional texture reflectance with a kaleidoscope. In *Proc. of SIGGRAPH*, 741–748.

HAWKINS, T., EINARSSON, P., AND DEBEVEC, P. 2005. Acquisition of time-varying participating media. In *Proc. of SIGGRAPH*, 812–815.

KANADE, T., YOSHIDA, A., ODA, K., KANO, H., AND TANAKA, M. 1996. A Stereo Machine for Video-rate Dense Depth Mapping and its New Applications. In *Proc. of CVPR*, 196–202.

KANADE, T., RANDER, P., AND NARAYANAN, P. 1997. Virtualized Reality: Constructing Virtual Worlds from Real Scenes. In *IEEE Multimedia*, 34–47.

KWATRA, V., SCHDL, A., ESSA, I., TURK, G., AND BOBICK, A. 2003. Graphcut Textures: Image and Video Synthesis Using Graph Cuts. In *Proc. of SIGGRAPH*, 277–286.

LEVOY, M., AND HANRAHAN, P. 1996. Light Field Rendering. In *Proc. of SIGGRAPH*, 31–42.

LEVOY, M., CHEN, B., VAISH, V., HOROWITZ, M., MCDOWALL, I., AND BOLAS, M. 2004. Synthetic aperture confocal imaging. In *Proc. of SIGGRAPH*, 825–834.

LIN, S.-S., AND BAJCSY, R. 2003. High Resolution Catadioptric Omni-Directional Stereo Sensor for Robot Vision. In *Proc. of ICRA*, 1694–1699.

LIU, X., YU, Y., AND SHUM, H.-Y. 2001. Synthesizing Bidirectional Texture Functions for Real-World Surfaces. In *Proc. of SIGGRAPH*, 97–106.

NENE, S., AND NAYAR, S. K. 1998. Stereo with Mirrors. In *Proc. of ICCV*, 1087 – 1094.

PELEG, S., AND HERMAN, J. 1997. Panoramic Mosaics by Manifold Projection. In *Proc. of CVPR*, 338–343.

SCHARSTEIN, D., AND SZELISKI, R. 2002. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *IJCV 47*, 7–42.

SEITZ, S. M., AND KIM, J. 2002. The Space of All Stereo Images. *IJCV 48*, 21–38.

SHUM, H.-Y., AND HE, L.-W. 1999. Rendering with Concentric Mosaics. In *Proc. of SIGGRAPH*, 299 – 306.

SOUTHWELL, D., BASU, A., FIALA, M., AND REYDA, J. 1996. Panoramic Stereo. In *Proc. of ICPR*, 378–382.

UNGER, J., WENGER, A., HAWKINS, T., GARDNER, A., AND DEBEVEC, P. 2003. Capturing and Rendering with Incident Light Fields. In *Proc. of EGSR*, 141–149.

WARD, G. J. 1992. Measuring and Modeling Anisotropic Reflection. In *Proc. of SIGGRAPH*, 265–272.