Axial-Cones: Modeling Spherical Catadioptric Cameras for Wide-Angle Light Field Rendering

Yuichi Taguchi Amit Agrawal Ashok Veeraraghavan Srikumar Ramalingam Ramesh Raskar
Mitsubishi Electric Research Labs (MERL) MIT Media Lab

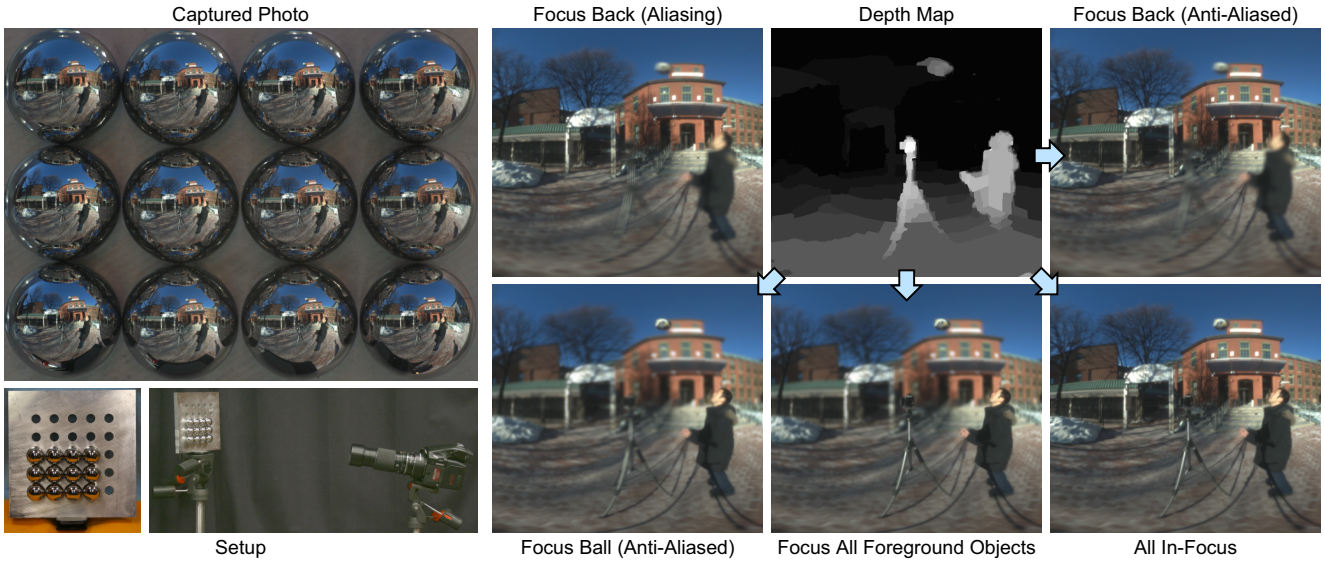

Figure 1: *Our axial-cone modeling enables wide-FOV digital refocusing and dense depth estimation using an array of spherical mirrors. We generate* $150^\circ \times 150^\circ$ *FOV refocused images and a dense depth map on a challenging outdoor scene. The estimated depth map is used for aliasing removal, surface-dependent refocusing, and all-in-focus rendering. The results are depicted using Mercator projection.*

Abstract

Catadioptric imaging systems are commonly used for wide-angle imaging, but lead to multi-perspective images which do not allow algorithms designed for perspective cameras to be used. Efficient use of such systems requires accurate geometric ray modeling as well as fast algorithms. We present accurate geometric modeling of the multi-perspective photo captured with a spherical catadioptric imaging system using *axial-cone cameras*: multiple perspective cameras lying on an axis each with a different viewpoint and a different cone of rays. This modeling avoids geometric approximations and allows several algorithms developed for perspective cameras to be applied to multi-perspective catadioptric cameras.

We demonstrate axial-cone modeling in the context of rendering wide-angle light fields, captured using a spherical mirror array. We present several applications such as spherical distortion correction, digital refocusing for artistic depth of field effects in wide-angle scenes, and wide-angle dense depth estimation. Our GPU implementation using axial-cone modeling achieves up to three orders of magnitude speed up over ray tracing for these applications.

CR Categories: I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Imaging Geometry

Keywords: Catadioptric Imaging, Computational Photography, Wide-Angle Light Fields, Digital Refocusing, Spherical Mirrors

1 Introduction

Catadioptric imaging systems (CIS) consisting of mirrors (cata) and lenses (dioptric) are widely used to increase the field of view (FOV) in a single photo. Unlike perspective cameras, such imaging systems do not have an effective single viewpoint (central model), except for few special configurations as shown in [Baker and Nayar 1999]. These configurations are: a perspective camera placed on the focus of a hyperbolic/elliptical mirror, and an orthographic camera aligned with the axis of a parabolic mirror. Common configurations such as a perspective camera looking at a spherical mirror or an array of mirrors lead to a multi-perspective image (non-central model). Thus, the well-known perspective projection model cannot be used to model the geometric structure of the captured rays.

To efficiently use such imaging systems, it is important to have an accurate geometric model amenable to fast algorithms. Previous approaches have faced the dilemma of accuracy versus speed. To reduce the complexity, a non-central CIS is often approximated as a central model, which leads to incorrect mapping of rays and skewed 3D estimation as shown in [Micusik and Pajdla 2004]. Recently, general linear cameras (GLC) [Yu and McMillan 2004; Ding et al. 2009] have been used for locally approximating a non-central CIS with an affine model that allows efficient mapping of rays, but it introduces approximation.

In this paper, we propose efficient and accurate geometric modeling of the rays captured with a perspective camera looking at a rotationally symmetric mirror using *axial-cone cameras*: multiple perspective cameras lying on the mirror axis each with a different viewpoint and a different cone of rays. The key idea is to model the non-central CIS as a combination of multiple perspective cameras. This allows several algorithms developed for perspective cameras to be directly applied to non-central catadioptric cameras. Axial-cone modeling avoids central/GLC approximations.

We demonstrate axial-cone modeling in the context of rendering wide-angle light fields, captured using a spherical mirror array. Light fields captured using perspective cameras have been extensively studied and fast rendering algorithms have been developed for them. In general, algorithms developed for perspective cameras cannot be used for non-perspective catadioptric cameras. We demonstrate that axial-cone modeling directly allows the use of the algorithms based on projective texture mapping for fast GPU implementation. We achieve up to three orders of magnitude speed up over explicit ray tracing for applications such as digital refocusing and dense depth estimation in wide-angle scenes.

Contributions: Our paper makes the following contributions:

- We propose axial-cone representation for accurate geometrical ray modeling of non-central CIS with rotationally symmetric mirror and avoid approximations common in other approaches.
- We demonstrate that axial-cone modeling allows several algorithms developed for perspective cameras to be directly used for non-central catadioptric cameras.
- We show several applications such as novel depth of field effects and dense depth estimation for wide-angle imaging.
- We present an efficient GPU implementation using projective texture mapping, achieving three orders of magnitude speed up over explicit ray tracing.

1.1 Related Work

Catadioptric Imaging Systems: Baker and Nayar [1999] presented the complete class of single-viewpoint CIS with detailed solutions and degenerate cases. A spherical mirror with a perspective camera leads to a non-single viewpoint or multi-perspective image when the camera is placed outside the sphere. Nalwa [1996] presented a single-viewpoint $360^\circ \times 50^\circ$ FOV camera using four planar mirrors in a pyramidal configuration with four cameras. Levoy et al. [2004] used an array of planar mirrors to capture a light field. Depending on the mirror shape and the relative camera-mirror pose in a CIS, the geometry of the captured rays could vary between central (perspective projection), axial [Ramalingam et al. 2006] (all projection rays intersect a line in space), or non-central [Swaminathan et al. 2001; Micusik and Pajdla 2004] models. Radial CIS [Kuthirummal and Nayar 2006] uses a camera looking through a single hollow rotationally symmetric mirror polished on the inside to perform 3D reconstruction, generation of texture maps, and computation of the BRDF parameters. Their system has a single mirror resulting in a circular locus of virtual viewpoints. In contrast, we use multiple spherical mirrors polished on the outside for wide-FOV light fields, modeled as axial locus of virtual viewpoints for each mirror. Note that this provides wide FOV in both directions, unlike horizontal panoramas.

Forward/Back Projection: In general, non-central CIS do not have a closed-form solution for the mapping of 3D points in the scene to the pixels (forward projection). This requires iterative non-linear optimization for *each* 3D point depending on the mirror shape [Micusik and Pajdla 2004; Ding et al. 2009]. Thus, explicit ray tracing for forward projection is orders of magnitude slower for non-central CIS compared to perspective imaging. Agrawal et al. [2010] have recently presented analytical solutions of forward projection for axial non-central CIS using quadric mirrors, but this still performs forward projection for each 3D point independently. The inverse mapping from pixels to rays in the scene (back projection), in contrast, can be easily computed for typical CIS given the calibration. In the most general case, it can be obtained using generic camera calibration [Grossberg and Nayar 2001; Sturm and Ramalingam 2004], where cameras are modeled by attributing an individual projection ray to each pixel. Since such a model performs back projection pixel-by-pixel, it is not suitable for rendering because of possible holes in the rendered image. Our axial-cone modeling allows efficient back projection using GPU texture mapping, which automatically performs interpolation between pixels.

Spherical mirror arrays have been used for capturing incident light field for relighting [Unger et al. 2003], triangulation of a point light source [Lensch et al. 2003], near object detection for surveillance [Kojima et al. 2005], and 3D reconstruction [Lanman et al. 2006; Ding et al. 2009]. In [Unger et al. 2003], each sphere image was re-mapped into latitude/longitude format and ray tracing was used to compute intersection with the light field plane. Lensch et al. [2003] computed the position of a point light source by back-projecting corresponding rays from multiple spheres, which is a pixel-by-pixel operation. Lanman et al. [2006] manually specified corresponding points in each sphere image and reconstructed them as a simple 3D mesh model. Their manual approach cannot be extended to natural scenes easily. Sparse 3D reconstruction have been shown using spherical mirrors [Micusik and Pajdla 2004; Kanbara et al. 2006]. To the best of our knowledge, the state-of-the-art algorithm for *dense* 3D reconstruction is by Ding et al. [2009]. However, even their approach uses GLC approximation: The captured photo of multiple spheres is tessellated into triangles and each of them is approximated using a different GLC for fast forward projection. In contrast, we present a geometrically exact model using axial-cone cameras without approximations and avoid ray tracing/forward projection. We show how axial-cone modeling can utilize efficient dense reconstruction algorithms such as plane sweeping [Collins 1996; Yang et al. 2002]. Our modeling of spherical mirrors is related to radially symmetric distortion models proposed for camera calibration [Tardif et al. 2009].

Light Fields: Levoy and Hanrahan [1996] and Gortler et al. [1996] proposed the concept of light field using two-plane parameterization. Other parameterizations such as two-sphere and sphere-plane have been proposed for sampling light fields over a sphere [Camahort et al. 1998]. Light fields captured using camera arrays have been used for synthetic aperture photography and digital refocusing [Levoy et al. 2004; Vaish et al. 2004]. Previous hand-held light field cameras trade spatial resolution for angular resolution. These approaches either insert a micro-lens array [Ng et al. 2005] or mask [Veeraraghavan et al. 2007] close to the sensor, or use an array of lens and prisms in front of the main lens [Georgiev et al. 2006]. However, these approaches offer a limited FOV, which is determined by the main lens. Using our axial-cone modeling, we show how wide-FOV light fields can be processed by capturing a single photo of a spherical mirror array, which also trades spatial resolution for angular resolution. Taguchi et al. [2010] captured multiple images as an axial light field by moving a camera along the axis of a rotationally symmetric mirror. To generate a wide-FOV perspective image at a virtual viewpoint inside the mirror, they mapped a cone of rays in each real camera to a unique cone of rays in the single virtual camera. Our axial-cone modeling is opposite: We map cones of rays in a single real camera to multiple cones each with a different viewpoint and a different angle, resulting in multiple virtual cameras. This enables wide-FOV light field applications by capturing a single image of multiple spherical mirrors.

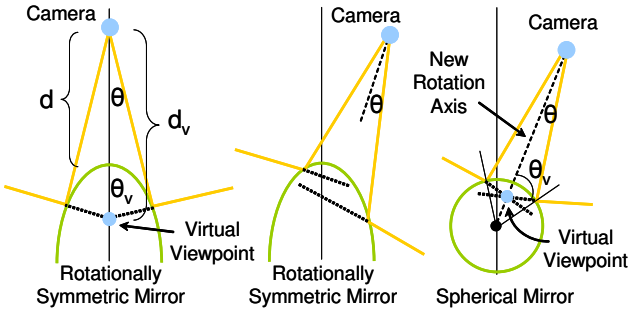

Figure 2: *For rotationally symmetric mirrors, axial-cone modeling is not applicable unless the camera lies on the mirror axis. Spherical mirror is an exception since mirror axis can be defined to contain the camera.*

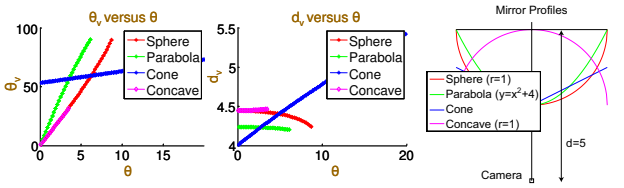

Figure 3: *Plots of θ_v and d_v versus θ for different mirror profiles shown on right. Note that the plot for concave sphere overlaps closely with that of the sphere.*

Light fields are well-studied for perspective cameras and existing techniques offer efficient operations only for the perspective case, such as texture mapping for refocusing [Isaksen et al. 2000] and plane sweep based stereo [Collins 1996; Yang et al. 2002]. Our axial-cone modeling enables utilizing these efficient algorithms for non-perspective catadioptric cameras.

2 Axial-Cone Modeling

Let us consider commonly occurring **rotationally symmetric mirrors**, such as conical and quadric (spherical, parabolic, hyperbolic, and elliptical) mirrors. We now describe axial-cone modeling for a CIS having a perspective camera placed on the rotation axis of such a mirror.

Consider a pinhole camera placed on the axis of a rotationally symmetric mirror, as shown in Figure 2 (left). On a planar slice containing the axis, consider two rays at an angle θ from the vertical passing through the pinhole. Since the mirror is rotationally symmetric, after reflection the rays will intersect at a virtual viewpoint inside the mirror at a distance d_v on the mirror axis, subtending an angle θ_v. Since the intersection lies on the mirror axis, for any out of plane rotation of the mirror about its axis in 3D, d_v and θ_v will remain fixed for a given θ. Thus, the locus of virtual viewpoints for a given θ is a single 3D point on the mirror axis. Hence a cone of rays[1] corresponding to an angle θ in the captured photo can be mapped to a cone of rays corresponding to an angle θ_v for a virtual perspective camera placed at d_v. In the captured photo, each cone of rays corresponds to a circle of pixels if the optical axis is aligned with the mirror axis. Thus, each circle of pixels in the captured photo with radius corresponding to θ can be mapped to a circle of pixels in the virtual perspective camera placed at d_v with radius corresponding to θ_v. If the optical axis of the camera is not aligned with the mirror axis, the resulting photo is related by a homography which can be applied as pre-processing.

[1]For simplicity, we refer to rays that lie on the surface of a cone as a 'cone of rays'.

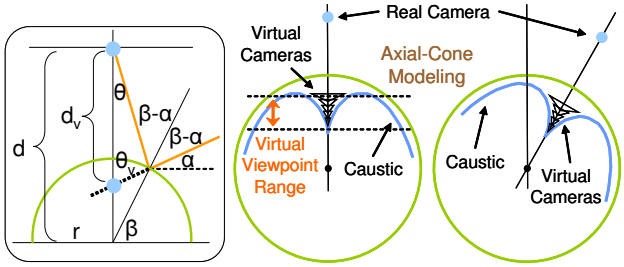

Figure 4: *Axial-cone modeling for a spherical mirror. (Left) Determining parameters of axial-cone cameras. (Middle and Right) Pictorial representation of virtual cameras for the two cases.*

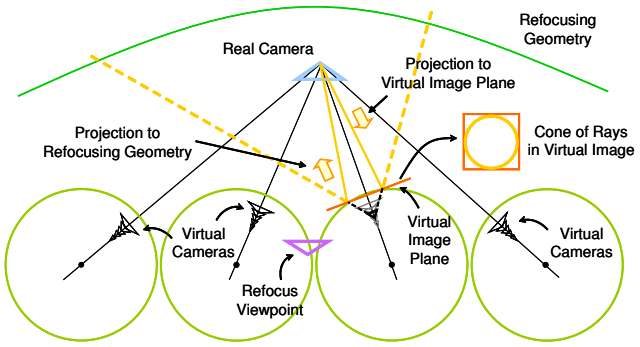

Figure 5: *Modeling of a spherical mirror array using axial-cone cameras. Refocusing is done via texture mapping all the virtual camera images to a specified refocusing geometry using a refocus viewpoint.*

A CIS will have an effective **single viewpoint** if d_v is same for *all* θ. This happens if a perspective camera is placed on the focus of a hyperbolic/elliptical mirror or an orthographic camera is used with a parabolic mirror [Baker and Nayar 1999]. For other configurations, there is no effective single viewpoint since d_v depends on θ. Since all virtual viewpoints lie on the mirror axis, we refer to the set of virtual perspective cameras as **axial-cone cameras**. Note that each virtual camera has only a cone of rays and thus a circle of pixels, not a full 2D image. Figure 3 shows θ_v and d_v versus θ for common rotationally symmetric mirrors including cone, sphere, parabola, and concave sphere (polished on inside), assuming the same distance between the pinhole camera and the apex of the mirror. The plot of θ_v versus θ shows the effective increase in field of view, while that of d_v versus θ shows the deviation from a single viewpoint system using the corresponding mirror. A concave mirror does not increase the FOV beyond 40° due to inter-reflections.

Off-Axis Case: Consider the off-axis case shown in Figure 2 (middle). Two rays subtending an angle θ from the camera's optical axis, after reflection from the mirror, do not intersect on the mirror axis. Thus, the locus of viewpoints for a given θ is not a single point as in the on-axis case. A practical consequence is that for each pixel in the captured photo, one would need to define a new virtual perspective camera. This modeling is similar to attributing an individual projection ray to each pixel as in generic camera calibration and thus does not provide any computational benefits.

2.1 Spherical Mirror

A spherical mirror represents a special case in the above analysis. Since the sphere is rotationally symmetric in *all* directions, the line joining the center of the sphere to the camera is an axis of symmetry. Thus, virtual cameras can be defined along this new axis as shown in Figure 2 (right) for *any* camera placement.

Axial-Cone Parameters: Consider Figure 4 (left), where the camera is placed at the origin and a spherical mirror of radius r is placed at a distance d. Let d_v be the distance of the virtual camera for the ray subtending an angle θ. Then, $\tan\theta = \frac{r\cos\beta}{d - r\sin\beta}$, which gives

$$\sin\beta = \left(d\sin^2\theta + \cos\theta\sqrt{r^2 - d^2\sin^2\theta}\,\right)/r. \quad (1)$$

Using the law of reflection, we have $\alpha = 2\beta - \theta - \pi/2$. Thus,

$$d_v = d - r\sin\beta - r\cos\beta/\tan(2\beta - \theta), \quad (2)$$
$$\theta_v = \pi/2 - \alpha = \pi - 2\beta + \theta. \quad (3)$$

For a given θ, β can be obtained using (1), and then the viewpoint d_v and angle θ_v can be obtained using (2) and (3). Figure 4 pictorially shows the axial-cone modeling. Note that viewpoint close to the mirror surface has a larger cone and vice-versa and that the scene geometry is *not* required for axial-cone modeling.

This modeling can be easily extended to the case of multiple spherical mirrors, since a sphere is rotationally symmetric in all directions. Figure 5 shows how the captured photo of multiple spherical mirrors is modeled. Each mirror results in axial-cone cameras on the line joining the mirror center and center of projection of the camera. Note that in principle, axial-cone modeling allows any random arrangement of spheres, since each sphere image is modeled separately using axial-cone cameras.

3 Rendering of Wide-Angle Light Fields

Now we demonstrate axial-cone modeling in the context of rendering wide-angle light fields, captured using a spherical CIS. As explained in the last section, axial-cone modeling applies to any random arrangement of spheres. Since each sphere image is modeled separately using axial-cone cameras, even the size of the spheres need not be the same. Wide-angle light fields can be captured in a variety of ways using a spherical CIS. For example, one can use a single perspective camera and capture multiple images by moving a single spherical mirror. Another way would be to rigidly attach the camera and mirror and capture multiple images by moving both of them. A *single-shot* approach would be to use an array of spheres, which we use in this paper.

3.1 Digital Refocusing Implementation

We first describe digital refocusing using axial-cone modeling from a single photo of a sphere array. For traditional light fields, refocusing can be performed by back-projecting the captured pixels to the desired scene geometry (referred to as *refocusing geometry*). Typically, a planar geometry is used and refocusing effects are produced by changing the depth of the plane [Isaksen et al. 2000]. Axial-cone modeling utilizes this key idea to project cones of rays in all the axial-cone cameras to the refocusing geometry. Note that a planar geometry is not the only choice for wide-angle light fields; one can use a variety of geometries such as spherical geometry and surfaces based on depths of objects in the scene.

The image of each sphere in the captured photo is referred to as *sphere image*. Figure 5 describes the refocusing procedure. For each virtual camera, we define a virtual image plane passing through its mirror reflection points. We then perform projective texture mapping of each sphere image to all its virtual cameras to account for the homography between the captured photo and the virtual image plane. Theoretically each virtual camera has only a circle of pixels. In practice, we uniformly sample the virtual viewpoint range using N virtual cameras. Let $\theta_v(i)$ denote the cone-angle of the i^{th} virtual camera. Each virtual camera then has a circular band of pixels corresponding to the cone-angles in the range

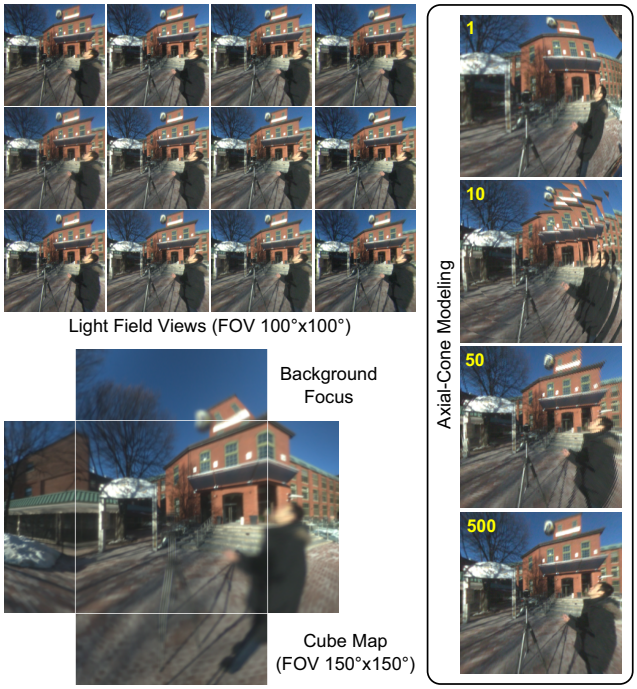

Figure 6: *(Top) Light field views obtained by projecting each sphere image on the background plane. (Right) Single light field view rendered using different numbers of axial-cone cameras. Sufficient number of the axial-cone cameras enables correct projection without distortion. (Bottom) Rendered cube map at the refocus viewpoint.*

$[\frac{\theta_v(i)+\theta_v(i-1)}{2}, \frac{\theta_v(i+1)+\theta_v(i)}{2}]$ and the extra pixels are removed by GPU fragment shader. Each virtual camera image is then projected to the refocusing geometry using a refocus viewpoint. Note that the scene geometry need not be planar since projective texture mapping can correctly handle the projection from the virtual camera. The projection of all virtual cameras for a single geometry results in a wide-angle light field 'view', rendered from the refocus viewpoint. The light field views for all sphere images are averaged to produce the refocusing result.

Discussions: The GPU implementation only requires back projection, because texture mapping automatically performs interpolation between pixels. Pixel-by-pixel back projection (on CPU) may produce holes in the rendered image; this is why one typically needs forward projection, which guarantees to fill all the pixels in the rendered image by tracing rays from scene points. In addition, our approach processes a cone of rays, instead of a single ray, in a single rendering path. Thus, it is much faster than forward projection based approach, which is a pixel-by-pixel operation.

Note that true depths of scene points are not required for the back projection to a refocusing geometry; this is similar to traditional light field refocusing, where the image for each viewpoint is projected to a refocusing geometry without knowing depths. For scene points lying on the refocusing geometry, the back-projected rays are mapped to geometrically correct positions and appear at the same pixel in all light field views. Thus, such scene points will be sharp in the refocused image. On the other hand, scene points not lying on the refocusing geometry will have parallax, similarly to the traditional light field case. The only difference is that each view corresponds to several axial-cone cameras having different viewpoints, instead of a single perspective camera in traditional light fields.

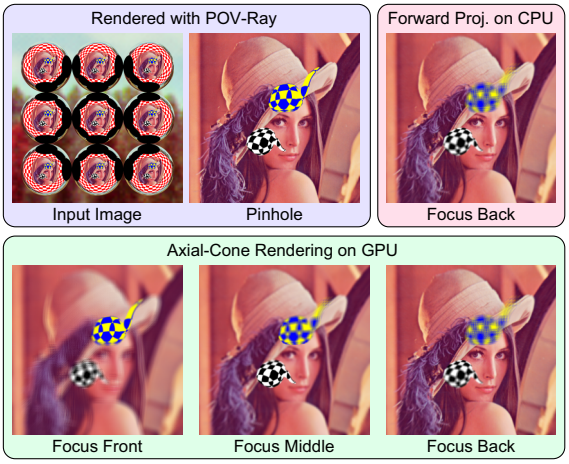

Figure 7: *Simulation results comparing refocused images rendered using axial-cone modeling and CPU-based forward projection, with the pinhole image rendered from the refocus viewpoint. The refocused images depict* 120° *FOV perspective rendering.*

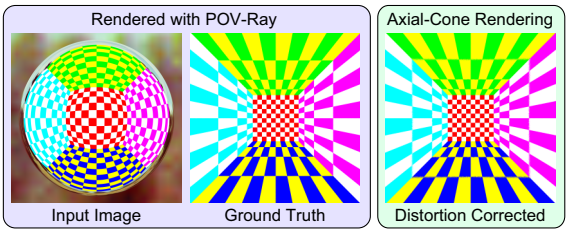

Figure 8: *Spherical distortion correction from an image of a single sphere using axial-cone modeling. The ground truth image is rendered from the same viewpoint with the same FOV (*150°*).*

3.2 Refocusing Results

Figure 6 (top) shows the 4×3 views obtained from the outdoor scene shown in Figure 1. Note that all straight lines in the scene are mapped to straight lines in the views. A perspective refocused image can be obtained by averaging the views. For narrow FOV, a single perspective refocusing output is sufficient. For wide FOV, however, such a rendering does not sample light rays uniformly across the FOV. We therefore generate a cube map at the refocus viewpoint by using 5 perspective cameras each having different directions and 90° FOV, as shown in Figure 6 (bottom). From the cube map, global projections such as Mercator projection can be applied to generate the final refocused image, as shown in Figure 1. Figure 6 (right) depicts the circular artifacts in a view when N is small. Note that straight lines in the scene are distorted if we use a single virtual camera ($N = 1$).

POV-Ray Simulations: Figure 7 verifies our axial-cone modeling implementation using an image of a 3×3 spherical mirror array rendered in POV-Ray. We use a single plane as the refocusing geometry and generate refocusing results using CPU based ray tracing (forward projection) and our GPU based axial-cone modeling. For comparison, we also show the ground truth pinhole camera image generated from the refocus viewpoint with the same FOV. Note that when focusing on the background textured plane, the texture in the refocused image has the same sharpness as the pinhole rendering (ground truth), because the rays observed from the refocus viewpoint are geometrically correct when the refocusing geometry matches the scene geometry. Also notice that the axial-cone rendering matches with the forward projection based CPU rendering. However, due to the resolution property of spherical mirrors, the resolution of the refocused image in periphery is lower than the pinhole rendering, both for CPU based forward projection and axial-cone rendering.

Performance: On a standard 2.66 GHz PC with an NVIDIA GeForce 8800 GTX graphics card, our GPU implementation took 2.2 seconds to render a 800×800 pixel refocused image from 12 sphere images using 500 axial-cone cameras for each sphere. Rendering the same resolution image on the CPU using iterative forward projection [Micusik and Pajdla 2004] took 6000 seconds. We therefore achieved $\sim$ 3000 times faster refocusing operation. GLC approximation [Ding et al. 2009] and analytical forward projection [Agrawal et al. 2010] will provide speed up for rendering, but they still require computation for each ray. In contrast, our approach does not involve approximation and processes a cone of rays in a single rendering pass, leading to a faster implementation.

3.3 Dense Depth Estimation

While sparse 3D reconstruction using spherical arrays is common, dense depth estimation is a challenging and largely unsolved problem. The state-of-the-art algorithm by Ding et al. [2009] uses a GLC approximation for forward projection. Using axial-cone modeling, we can directly apply plane sweeping [Collins 1996] and its variants commonly used for traditional light fields [Ng et al. 2005]. These methods first generate multiple refocused images at different depths (focal stack) and for each pixel select the best-focused image with some smoothness assumption to produce a depth map and all-in-focus image. Similarly, we generate light field views projected on spherical scene geometry at multiple radii. We use the variance across the views as the color consistency cost and estimate the depth map using a graph-cut framework [Boykov et al. 2001]. Using the estimated depth map, the refocused images can be combined to obtain an all-in-focus image as shown in Figure 1.

4 Applications

We show several novel results and effects using axial-cone modeling. For displaying wide-angle images, perspective projection distorts peripheral regions. To obtain visually pleasing images, several global projections have been proposed, each preserving some property but invariably leading to distortions. We use Mercator projection to display our results. Note that the distortions in displaying wide-angle images are unrelated to axial-cone modeling or the spherical array characteristics.

Spherical Distortion Correction: Our GPU implementation performs efficient spherical distortion correction by projecting all axial-cone camera images for a single sphere onto a plane. Figure 8 shows an example, where a spherical mirror is surrounded by five checkerboards. Distorted (curved) lines in the input image are corrected as straight lines in the result. Each light field view in Figure 6 also corresponds to a distortion corrected image. Note that this projection does not require depths, but the resulting image is pseudo-perspective, since the parallax is correct only for the scene points lying on the plane on which the projection is done.

Surface-Dependent Refocusing & Aliasing Removal: Our axial-cone modeling allows dense depth map computation, which can be used for surface-dependent refocusing and aliasing removal in refocused images. In scenes containing multiple objects at different depths, all foreground objects cannot be simultaneously brought into focus using a common spherical or planar refocusing geometry. We achieve this goal by using a modified depth map as the spatially depth varying refocusing geometry. We keep the depth values corresponding to the foreground objects unmodified and replace the background depth value with the nearest foreground depth value.

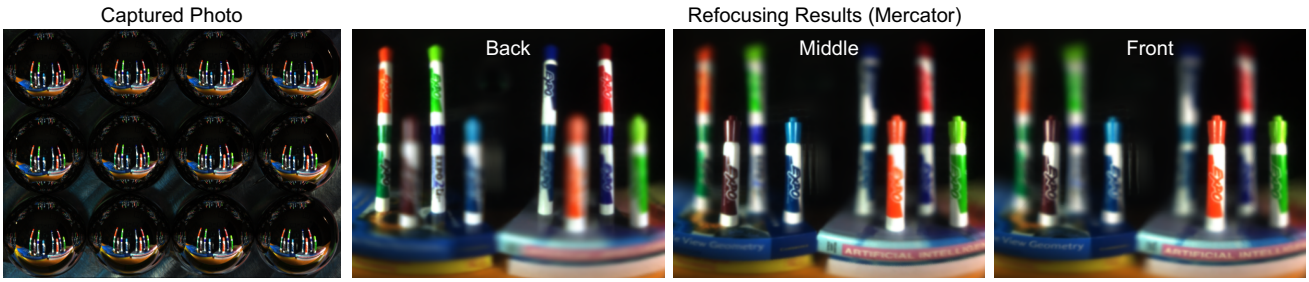

Figure 10: *Wide-angle digital refocusing on a tabletop scene. The refocusing results are depicted using Mercator projection.*

Figure 9: *(Left) Depth map estimated from light field can be used to bring all foreground objects (at different depths) in focus. (Right) Captured photo and narrow FOV refocusing results.*

Figure 9 shows an example. Note that all-in-focus image corresponds to using the depth map itself as the refocusing geometry.

Because of sparse spatial sampling, large depth variations in wide-angle scenes can cause aliasing artifacts in refocused images. This is similar to aliasing artifacts in light fields captured using camera arrays due to sparse spatial sampling. Techniques such as aperture filtering would require computation of pre-filter using forward projection. Instead, we apply a depth-dependent blur on refocused images for faster processing. We use a spatially varying Gaussian filter with $\sigma = \alpha|d_p|$, where d_p is the disparity difference between the refocusing geometry and the pixel depth, and α controls the amount of blur. Figure 1 shows that visually pleasing out-of-focus blur can be obtained by using such smoothing. However, errors in depth map could cause some artifacts at sharp discontinuities.

Figure 1 shows several wide-FOV rendering results using the estimated depth map for a challenging outdoor scene, where a person is trying to catch a ball. Figure 9 shows refocusing results on three people sitting around a table. Using a modified depth map as the refocusing geometry, we bring all three people in focus, while keeping the background out of focus. Figure 9 also shows three narrow

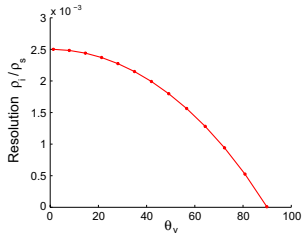

Figure 11: *Decrease in tangential resolution for a spherical mirror with half-FOV angle θ_v.*

FOV rendered images focused on all three people within the entire FOV. Figure 10 shows three refocused images on a tabletop scene.

5 Implementation and Analysis

We use an array of stainless steel balls with high sphericity glued to a planar platform as the spherical mirror array. We employ a 22 MP Mamiya 645AFD camera with 210 mm lens and a 12 MP Canon Rebel XSI with 300 mm lens. The camera was placed ~ 850 mm from the array. Each sphere image has resolution of ~ 1 MP. By using spheres of different size, our design can be easily scaled to different scene scales. We use balls of 0.25" radii for small tabletop scenes (Figure 10) and 0.5" radii for room-size/outdoor scenes.

5.1 Calibration

The internal camera parameters are estimated off-line. Since the physical sphere radius is known, we only need to determine the position of sphere centers with respect to the camera. The sphere centers are computed from the captured photo itself (*in-photo* calibration). We first estimate the sphere centers independently by clicking few points (minimum of 3) on each sphere boundary in the image. These pixels correspond to the rays that are tangential to the sphere in 3D. The sphere center lies at a distance of $\frac{r}{\sin\alpha}$ along the central ray, which makes the same angle α with all the tangent rays. The sphere centers are then refined jointly by finding corresponding SIFT features across sphere images and using a bundle adjustment approach [Micusik and Pajdla 2004; Agrawal et al. 2010].

5.2 Analysis

Spatial Resolution: The resolution characteristic for a spherical mirror can be analyzed similar to [Kuthirummal and Nayar 2006]. Let z be the distance of a scene plane from the camera and u be the image plane-lens distance. Then a circle of pixels of radius $\rho_i = u\tan\theta$ is projected to a circle of radius $\rho_s = (d_v - z)\tan\theta_v$ on the scene plane. Thus, the tangential resolution can be obtained as

$$\rho_i/\rho_s = u\tan\theta/((d_v - z)\tan\theta_v). \quad (4)$$

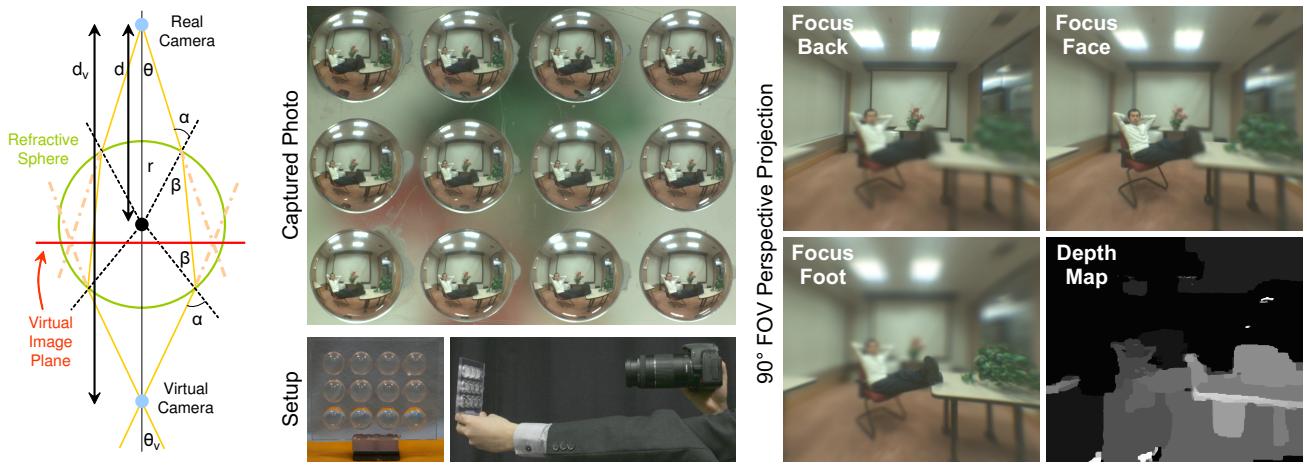

Figure 12: *(Left) Axial-cone modeling equally applies to a refractive sphere, with a different set of virtual viewpoints and associated cone angles compared to the spherical mirror. Note that the real and the virtual viewpoints lie on the opposite sides of the sphere. (Center) Single photo captured by looking through an array of refractive acrylic balls. (Right) Refocused images and the estimated depth map.*

Figure 11 plots the decrease in resolution with θ_v for a scene plane 1 m from the sphere. The resolution decreases by a factor of two at 120° FOV.

Defocus Blur: In a spherical mirror based CIS, the infinite scene depth range is compressed within a small volume inside the mirror, called the caustic volume [Swaminathan 2007]. Defocus blur in the captured photo can be avoided for the scene if the caustic volume is within the depth-of-field of the camera. We use a small $f/11$ aperture to achieve this.

Failure Cases/Artifacts: Similar to any other catadioptric system and conventional lenses, surface finish and dust on the mirror surface reduce image quality. Errors in calibration cause non-zero parallax across views even at the refocusing geometry, which reduces the sharpness of in-focus regions in the refocused images. Calibration could be improved by capturing extra images of calibration targets, although it increases the setup time compared to our simple in-photo calibration approach. Objects moving faster than the shutter speed result in motion blur (ball in Figure 1). Errors in depth estimation can cause artifacts around object boundaries in surface-dependent refocusing and aliasing removal results.

6 Axial-Cones for Refractive Spheres

Similar to spherical mirrors, axial-cone modeling can also be used to model *refractive* spheres as shown in Figure 12 (left). We assume the refractive sphere to be of constant refractive index μ and ignore specular/internal reflections from its surface and inside. Due to symmetry, rays at an angle θ from the camera intersect on the axis joining the optical center and the sphere center. Thus, similar to a spherical mirror, the cone of rays at an angle θ can be modeled as cone of rays at d_v corresponding to θ_v, where

$$d_v = d(1 + \sin\theta / \sin\theta_v), \quad (5)$$

$$\theta_v = 2\arcsin(d\sin\theta/r) - 2\arcsin(d\sin\theta/(r\mu)) - \theta. \quad (6)$$

Figure 12 (right) shows an example looking through 4×3 array of refractive acrylic balls. Notice the sharpness on the table and on the person's face in refocused images. All curved lines in the captured photo become straight in the refocused images, showing that the axial-cone cameras correctly model the refractive geometry. However, compared to a spherical mirror, refractive sphere offers a smaller FOV and leads to more defocus blur in the captured photo.

7 Discussions

Axial-Cones vs. Forward Projection: Axial-cone modeling has the benefit of utilizing fast algorithms developed for perspective cameras and thus it can avoid forward projection for several applications as shown. However, forward projection may be required for other applications such as structure-from-motion and to compute the accurate ray-space pre-filter. We envision that our paper will generate interest in hybrid algorithms, utilizing both axial-cone based back projection and forward projection.

Mirrors for Light Field Capture: Planar mirror array has been used for single-shot light field capture [Levoy et al. 2004], but it does not increase the effective FOV. Using a spherical mirror array to capture wide-FOV light field has its own tradeoffs. Since a sphere is rotationally symmetric in all directions, the camera can be placed freely. This is a big advantage over typical central CIS that require precise placement of the camera with respect to the mirror [Nene and Nayar 1998]. Spherical mirrors are easy to manufacture, are low cost, and minimize calibration parameters. As shown, *in-photo* calibration can be achieved without the need for extra calibration images, which is useful in practice. A single-shot capture allows handling dynamic scenes without any synchronization. This is similar in spirit to previous hand-held light field cameras, but our approach does not require camera modifications as in [Ng et al. 2005; Veeraraghavan et al. 2007]. On the other hand, a spherical mirror array shares the limitations of typical mirror array based systems such as the loss of resolution due to inter-reflections and gaps between the mirrors. A spherical mirror provides lower resolution in periphery compared to the other quadric (parabolic, hyperbolic, and elliptical) mirrors [Heidrich 1999; Swaminathan et al. 2001]. This can be overcome by rigidly moving a single camera-mirror pair to capture images from multiple viewpoints. Since our axial-cone modeling applies to any rotationally symmetric mirror, the other quadric mirrors can also be used in the above setting to obtain better resolution in periphery.

Conclusions: We presented the axial-cone representation to model a catadioptric system having a rotationally symmetric mirror. Our modeling avoids explicit ray tracing and approximations, and allows algorithms developed for perspective cameras to be directly applied to such non-perspective catadioptric cameras. We demonstrated axial-cone modeling for rendering wide-angle light fields captured using a spherical mirror array. We showed a fast GPU implementation for several applications such as spherical distortion

correction, digital refocusing, and dense depth estimation, achieving up to three orders of magnitude speed up over conventional ray tracing. Axial-cone modeling could prove useful for other applications such as modeling fish-eye lenses, catadioptric projectors, and rendering rain. Sampling and interpolation of mirror based light fields, both in ray space and frequency domain, also remain an interesting future work area. We hope that our paper will stimulate further research in capture, analysis, and rendering of wide-angle light fields using mirrors and refractive elements, as well as inspire new software tools for wide-angle photography.

Acknowledgments

We thank the anonymous reviewers for their valuable suggestions. We also thank Jay Thornton, Keisuke Kojima, Tim K. Marks, and John Barnwell, along with Haruhisa Okuda and Kazuhiko Sumi, Mitsubishi Electric, Japan, for their help and support.

References

AGRAWAL, A., TAGUCHI, Y., AND RAMALINGAM, S. 2010. Analytical forward projection for axial non-central dioptric and catadioptric cameras. In *ECCV*.

BAKER, S., AND NAYAR, S. K. 1999. A theory of single-viewpoint catadioptric image formation. *IJCV 35*, 2, 175–196.

BOYKOV, Y., VEKSLER, O., AND ZABIH, R. 2001. Fast approximate energy minimization via graph cuts. *PAMI 23*, 11, 1222–1239.

CAMAHORT, E., LERIOS, A., AND FUSSELL, D. 1998. Uniformly sampled light fields. In *Proc. 9th Eurographics Workshop on Rendering*, 117–130.

COLLINS, R. 1996. A space-sweep approach to true multi-image matching. In *CVPR*, 358–363.

DING, Y., YU, J., AND STURM, P. 2009. Multi-perspective stereo matching and volumetric reconstruction. In *ICCV*.

GEORGIEV, T., ZHENG, C., NAYAR, S., CURLESS, B., SALASIN, D., AND INTWALA, C. 2006. Spatio-angular resolution tradeoffs in integral photography. In *EGSR*, 263–272.

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

GROSSBERG, M., AND NAYAR, S. 2001. A general imaging model and a method for finding its parameters. In *ICCV*, vol. 2, 108–115.

HEIDRICH, W. 1999. *High-Quality Shading and Lighting for Hardware-Accelerated Rendering*. PhD thesis, University of Erlangen.

ISAKSEN, A., MCMILLAN, L., AND GORTLER, S. 2000. Dynamically reparameterized light fields. In *SIGGRAPH*, 297–306.

KANBARA, M., UKITA, N., KIDODE, M., AND YOKOYA, N. 2006. 3D scene reconstruction from reflection images in a spherical mirror. In *ICPR*, vol. 4, 874–879.

KOJIMA, Y., SAGAWA, R., ECHIGO, T., AND YAGI, Y. 2005. Calibration and performance evaluation of omnidirectional sensor with compound spherical mirrors. In *OMNIVIS*.

KUTHIRUMMAL, S., AND NAYAR, S. 2006. Multiview radial catadioptric imaging for scene capture. *ACM Trans. Graph. 25*, 3, 916–923.

LANMAN, D., CRISPELL, D., WACHS, M., AND TAUBIN, G. 2006. Spherical catadioptric arrays: Construction, multi-view geometry, and calibration. In *3DPVT*, 81–88.

LENSCH, H. P. A., KAUTZ, J., GOESELE, M., HEIDRICH, W., AND SEIDEL, H.-P. 2003. Image-based reconstruction of spatial appearance and geometric detail. *ACM Trans. Graph. 22*, 2, 234–257.

LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *SIGGRAPH*, 31–42.

LEVOY, M., CHEN, B., VAISH, V., HOROWITZ, M., MCDOWALL, I., AND BOLAS, M. 2004. Synthetic aperture confocal imaging. *ACM Trans. Graph. 23*, 3, 825–834.

MICUSIK, B., AND PAJDLA, T. 2004. Autocalibration and 3D reconstruction with non-central catadioptric cameras. In *CVPR*.

NALWA, V. 1996. A true omnidirectional viewer. Tech. rep., Bell Laboratories.

NENE, S. A., AND NAYAR, S. K. 1998. Stereo with mirrors. In *ICCV*, 1087–1094.

NG, R., LEVOY, M., BRDIF, M., DUVAL, G., HOROWITZ, M., AND HANRAHAN, P. 2005. Light field photography with a hand-held plenoptic camera. Tech. rep., Stanford Univ.

RAMALINGAM, S., STURM, P., AND LODHA, S. 2006. Theory and calibration algorithms for axial cameras. In *ACCV*.

STURM, P., AND RAMALINGAM, S. 2004. A generic concept for camera calibration. In *ECCV*, vol. 2, 1–13.

SWAMINATHAN, R., GROSSBERG, M., AND NAYAR, S. 2001. Caustics of catadioptric cameras. In *ICCV*, vol. 2, 2–9.

SWAMINATHAN, R. 2007. Focus in catadioptric imaging systems. In *ICCV*, 1–7.

TAGUCHI, Y., AGRAWAL, A., RAMALINGAM, S., AND VEERARAGHAVAN, A. 2010. Axial light field for curved mirrors: Reflect your perspective, widen your view. In *CVPR*.

TARDIF, J.-P., STURM, P., TRUDEAU, M., AND ROY, S. 2009. Calibration of cameras with radially symmetric distortion. *PAMI 31*, 9, 1552–1566.

UNGER, J., WENGER, A., HAWKINS, T., GARDNER, A., AND DEBEVEC, P. 2003. Capturing and rendering with incident light fields. In *EGSR*, 141–149.

VAISH, V., WILBURN, B., JOSHI, N., AND LEVOY, M. 2004. Using plane + parallax for calibrating dense camera arrays. In *CVPR*, 2–9.

VEERARAGHAVAN, A., RASKAR, R., AGRAWAL, A., MOHAN, A., AND TUMBLIN, J. 2007. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph. 26*, 3, 69:1–69:12.

YANG, R., WELCH, G., AND BISHOP, G. 2002. Real-time consensus-based scene reconstruction using commodity graphics hardware. In *Pacific Graphics*, 225–234.

YU, J., AND MCMILLAN, L. 2004. General linear cameras. In *ECCV*, vol. 2, 14–27.