Multiple-Center-of-Projection Images

Paul Rademacher Gary Bishop
University of North Carolina at Chapel Hill

ABSTRACT

In image-based rendering, images acquired from a scene are used to represent the scene itself. A number of reference images are required to fully represent even the simplest scene. This leads to a number of problems during image acquisition and subsequent reconstruction. We present the *multiple-center-of-projection image*, a single image acquired from multiple locations, which solves many of the problems of working with multiple range images.

This work develops and discusses multiple-center-of-projection images, and explains their advantages over conventional range images for image-based rendering. The contributions include greater flexibility during image acquisition and improved image reconstruction due to greater connectivity information. We discuss the acquisition and rendering of multiple-center-of-projection datasets, and the associated sampling issues. We also discuss the unique epipolar and correspondence properties of this class of image.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation – *Digitizing and scanning*, *Viewing algorithms*; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism; I.4.10 [Image Processing]: Scene Analysis
Keywords: image-based rendering, multiple-center-of-projection images

1 INTRODUCTION

In recent years, image-based rendering (IBR) has emerged as a powerful alternative to geometry-based representations of 3-D scenes. Instead of geometric primitives, the dataset in IBR is a collection of samples along viewing rays from discrete locations. Image-based methods have several advantages. They provide an alternative to laborious, error-prone geometric modeling. They can produce very realistic images when acquired from the real world, and can improve image quality when combined with geometry (e.g., texture mapping). Furthermore, the rendering time for an image-based dataset is dependent on the image sampling density, rather than the underlying spatial complexity of the scene. This can yield significant rendering speedups by replacing or augmenting traditional geometric methods [7][23][26][4].

The number and quality of viewing samples limits the quality of images reconstructed from an image-based dataset. Clearly, if we sample from every possible viewing position and along every possible viewing direction (thus sampling the entire *plenoptic function* [19][1]), then any view of the scene can be reconstructed perfectly. In practice, however, it is impossible to store or even acquire the complete plenoptic function, and so one must sample from a finite number of discrete viewing locations, thereby building a set of *reference images*. To synthesize an image from a new viewpoint, one must use data from *multiple* reference images. However, combining information from different images poses a number of difficulties that may decrease both image quality and representation efficiency. The *multiple-center-of-projection* (*MCOP*) *image* approaches these problems by combining samples from multiple viewpoints into a *single image*, which becomes the complete dataset. Figure 1 is an example MCOP image.

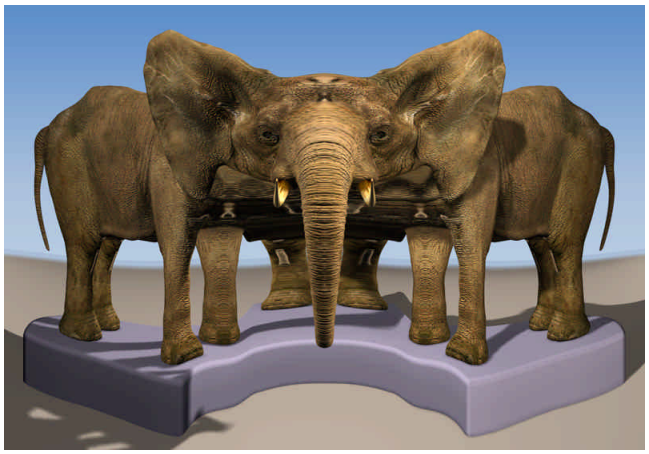

Figure 1 Example MCOP image of an elephant

The formal definition of multiple-center-of-projection images encompasses a wide range of camera configurations. This paper mainly focuses on one particular instance, based on the photographic *strip camera* [9]. This is a camera with a vertical slit directly in front of a moving strip of film (shown in Figure 2 without the lens system). As the film slides past the slit a continuous image-slice of the scene is acquired. If the camera is moved through space while the film rolls by, then different columns along the film are acquired from different vantage points. This allows the single image to capture continuous information from multiple viewpoints. The strip camera has been used extensively, e.g., in aerial photography. In this work's notion of a digital strip camera, each pixel-wide column of the image is acquired from a different center-of-projection. This single image becomes the complete dataset for IBR.

Features of multiple-center-of-projection images include:

- greater connectivity information compared with collections of standard range images, resulting in improved rendering quality,
- greater flexibility in the acquisition of image-based datasets, for example by sampling different portions of the scene at different resolutions, and
- a unique *internal* epipolar geometry which characterizes optical flow within a single image.

CB #3175 Sitterson Hall, Chapel Hill, NC, 27599-3175
rademach@cs.unc.edu, bishop@cs.unc.edu http://www.cs.unc.edu/~ibr

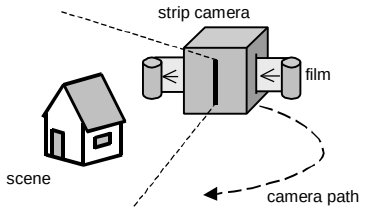

Figure 2 A *strip camera* consists of a moving strip of film behind a vertical slit.

Furthermore, MCOP images retain the desirable properties of conventional range images, such as fast incremental projection and moderate storage requirements.

In this paper we formally develop the concept of multiple-center-of-projection images, and discuss their acquisition and reprojection for image-based rendering. We describe different data structures to maintain these images, and discuss the implications of sampling during their acquisition. We also show how to perform point correspondence using a single MCOP image. We conclude by presenting several examples of MCOP images, and demonstrate their advantage over conventional range images.

2 PREVIOUS WORK

Early work in image-based rendering includes the rangeless panoramas of Chen and Williams [3][4] and Regan and Pose[22], which allow view reconstruction from a set of fixed eye locations. Plenoptic modeling [19] adds range to panoramic images, thereby allowing reprojection from arbitrary viewpoints. The concept of the plenoptic function is further explored by light slab methods [10][16], which attempt to fully sample the function within a subset of space.

Several methods exist for handling IBR range images from multiple viewpoints. *Layered depth images* [23] store multiple hits of a viewing ray in different layers of a single image, allowing, e.g., the front and back of a surface to be kept in a single data structure. The *delta tree* [6] acquires a hierarchical set of reference images on a sampling sphere around a target object, discarding redundant information when possible.

The work most closely related to this paper is the *multiperspective panorama for cel animation* [30]. This method constructs an image from multiple viewpoints for use as a backdrop in traditional cel animation. A continuous set of views along a pre-specified path can be extracted from this single backdrop. The construction of multiperspective panoramas is similar to the use of *manifold projections* in computer vision [20].

Another related work is the *extended camera for ray tracing* [11]. This method allows a ray-tracing camera to undergo arbitrary transformations as it traverses each pixel of the output image, thereby achieving a number of artistic effects.

Imaging devices similar to strip cameras have recently been explored in computer vision by Zheng and Tsuji [31] and by Hartley [13]. The former discusses their use in robot navigation, while the latter discusses the *pushbroom camera*, used in satellite imagery. One-dimensional cameras are also the basis of Cyberware scanners, which sweep a linear or circular path around an object, and systems by 3D Scanners, Ltd., which attach a 1-D scanning head to a Faro Technologies arm.

3 MULTIPLE-CENTER-OF-PROJECTION IMAGES

3.1 Definition

A multiple-center-of-projection image is an extension of a conventional image, characterized by having a *set* of cameras contributing to it, rather than only a single camera. Individual pixels or sets of pixels are acquired by different cameras, subject to certain constraints.

> A *multiple-center-of-projection image* consists of a two dimensional image and a parameterized set of cameras, meeting the following conditions:
> 1) the cameras must lie on either a continuous curve or a continuous surface
> 2) each pixel is acquired by a single camera
> 3) viewing rays vary continuously across neighboring pixels
> 4) two neighboring pixels must either correspond to the same camera or to neighboring cameras.

This definition states that the camera locations are not an unorganized set of points, but rather define a continuous curve or surface (condition 1). Condition 2 states each pixel is from a single camera, rather than a blend of samples from multiple cameras. Condition 3 imposes smoothness on the viewing rays, thereby ensuring they do not vary discontinuously. The last condition imposes an organization on the mapping of camera samples to the resulting image; it ensures we move smoothly from camera to camera as we traverse from pixel to pixel in an image.

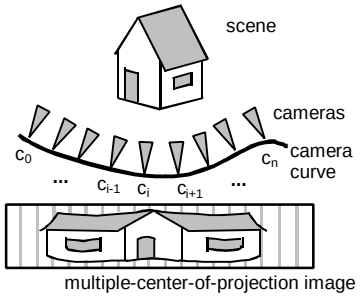

Figure 3 A multiple-center-of-projection image acquired by capturing a discrete number of image-slices along a curve. This single image (bottom) sees three sides of the house simultaneously. A similar configuration was used to create Figure 1.

Note that while the definition contains several parts, a continuous strip camera – along any continuous path – automatically satisfies the four criteria (section 4.4 discusses the sampling implications for discrete images). The remainder of this paper thus deals exclusively with the strip camera instance, unless otherwise noted.

3.2 Distinctions From Other Methods

Before delving into the details of MCOP images, we should clarify what these images *are not*. For example, what is the difference between an MCOP image and an arbitrary collection of conventional images? As will be shown in sections 4 and 5, the four constraints on an MCOP image yield advantages not found with conventional range images. These include improved image

reconstruction quality, greater flexibility during image acquisition, and unique epipolar and correspondence properties. MCOP images are *not* a data structure for maintaining collections of images.

What is the difference between an MCOP image and the multiperspective panorama [30]? Multiperspective panoramas may be considered subsets of MCOP images, since they meet the definition of section 3.1. However, a multiperspective panorama is not intended as a dataset for arbitrary view construction; it does not permit 3D reprojection, and thus can only provide views along a single predetermined path. Also, a primary goal in multiperspective panoramas is to minimize the local distortion in the bitmap – otherwise, output images will suffer from perspective errors. This will not occur with MCOP images.

What is the difference between an MCOP dataset and a polygonal mesh? MCOP images retain the same desirable characteristics that separate all image-based range datasets from polygonal datasets. Because of strong spatial coherence across neighboring pixels, the projection of image points can be computed incrementally using McMillan and Bishop's 3D warping equations [19]. Also, since the image dataset is a regular grid, the 2D pixel coordinates are implicit for all points – each pixel only needs to contain intensity and range, in contrast with intensity, *x*, *y*, and *z*. Finally, an MCOP image has projective and epipolar properties not found in polygonal meshes.

We must also distinguish between MCOP images and images with arbitrary-manifold projection surfaces. That is, we can construct a single-COP image in which the projection surface is not a plane, cylinder, or sphere, but rather an arbitrary manifold surface. Each pixel in this image is then given by an intersection of a ray from the COP through the surface. While MCOP images *do* have curved image surfaces, the two are not equivalent, since the arbitrary-manifold image can still only capture scene points visible to the single center of projection, whereas MCOP images can view from more than one location.

4 ACQUIRING MCOP IMAGES

Multiple-center-of-projection images are well-suited to applications where a useful path can be defined through or around a scene; by not tying every pixel in a reference image to the same viewpoint, they allow greater flexibility during acquisition than conventional images. For example, sampling a nearly-convex object (such as in Figure 4) results in several poorly-sampled areas, as the cameras' viewing rays approach grazing angle with the object. The MCOP image on the right, however, samples every point at a near-normal angle, thus acquiring good samples everywhere. This occurs for both quantities being measured – color and range. Other relevant sampling issues are discussed in section 4.4.

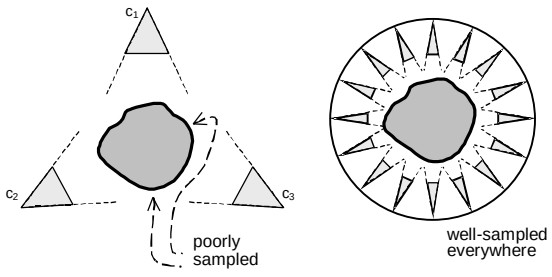

Figure 4 The three regular cameras on the left have difficulty sampling the object well. The MCOP image on the right can easily sample well the complete object.

4.1 Data Structures for MCOP Images

Although each pixel in an MCOP image may conceptually belong to a different camera, in practice we minimize storage requirements by describing the cameras parametrically in a variety of ways. At the highest level, for example, we can divide the camera curve into equally-spaced segments, then use the column index of each pixel as an implicit curve coordinate to compute the camera location. At the next level, each column of the image may explicitly store the parametric coordinate of the camera, thus allowing irregular camera spacing along the curve. Each column may instead *explicitly* contain the complete camera model, thus requiring *n* camera descriptions for an *n´m*-sized image. Or, at the extreme level, each *pixel* can explicitly store a camera model. Clearly, the most compact method should be selected given a particular application.

4.2 Synthesizing From a 3D Model

To synthesize an MCOP image from a 3D model, we first define a path through the 3D scene. This path need not be closed, nor do the viewing rays need to be normal to the curve. We then smoothly animate a camera along this curve, extracting a single-pixel-wide color image and depth map at each frame of the animation. As each slice is captured, we concatenate the color image and range map into a rectangular buffer, and store the camera information for that column (four vectors, described in 5.1) in an array. Since there is much coherence from one camera position to the next along the path, a method such as *multiple viewpoint rendering* [12] may be used to accelerate the rendering.

Figures 5 through 8 show a rendered 1000×500 MCOP image of a castle model. Details of the rendering process are given in section 8. Note that the single image captures the complete exterior of the model, which is then rendered as a single mesh. This demonstrates the increased acquisition flexibility and improved connectivity properties of the MCOP technique.

4.3 Acquiring From Real-World Scenes

We can acquire an MCOP image of the real world by constructing the digital equivalent of a strip camera. For example, we can use a 1-D CCD camera, translated along a path. One-dimensional image-strips are captured at discrete points on the path and concatenated into the image buffer. The CCD camera must be accurately tracked to prevent errors during reprojection, using for example the techniques in [28].

This method has the disadvantage of introducing a temporal element into the image, since every 1-D strip is captured at a different time. This may lead to mismatched data unless the scene is static (static scenes are a common assumption in IBR [4][19][16][10]).

Active range-finding techniques, such as laser range-finders, can be applied to MCOP images almost exactly as with regular images: simply register the laser scanner with the color camera. A Cyberware scanner, for example, is a 1-D laser range-finder registered with a 1-D linear camera. Section 6 discusses how the epipolar constraint - critical to passive range-finding methods - can be extended to MCOP images.

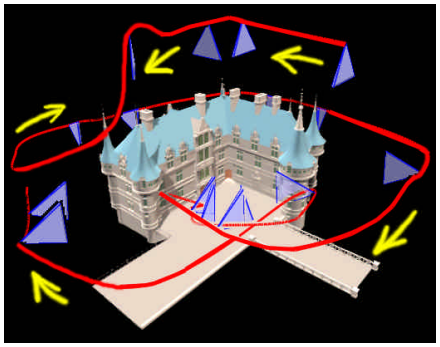

Figure 5 Castle model. The red curve is the path the camera was swept on, and the arrows indicate the direction of motion. The blue triangles are the thin frusta of each camera. Every 64th camera is shown.

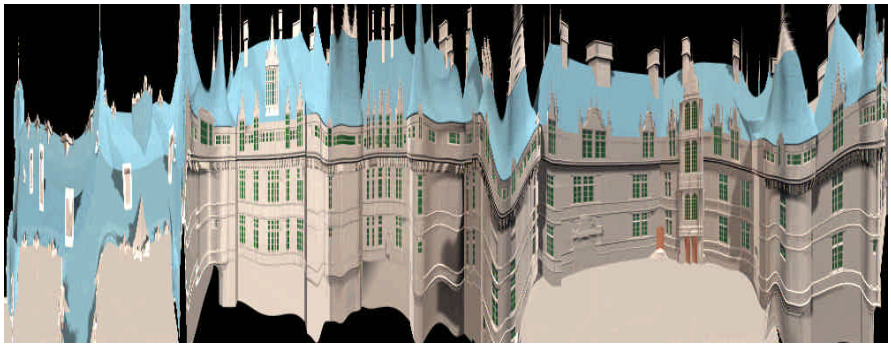

Figure 6 The resulting 1000×500 MCOP image. The first fourth of the image, on the left side, is from the camera sweeping over the roof. Note how the courtyard was sampled more finely, for added resolution.

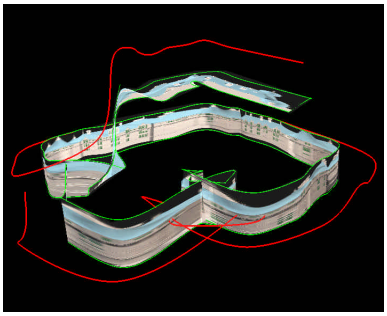

Figure 7 The projection surface (image plane) of the camera curve.

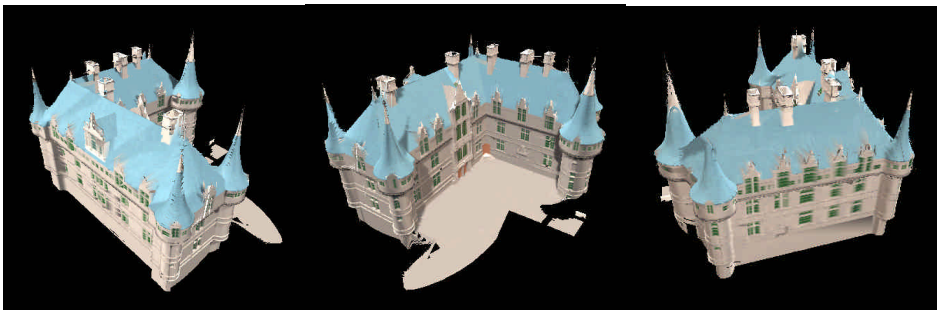

Figure 8 Three views of the castle, reconstructed solely from the single MCOP image above. This dataset captures the complete exterior of the castle.

4.4 Sampling Issues

The construction of an MCOP image is inherently a sampling process; there are two primary questions that must be asked. First, *how much* of the plenoptic function does an MCOP image capture? The light field and Lumigraph methods approach this by attempting to fully sample the function, but only over a subset of space. The MCOP method is better suited to the opposite approach - sampling the plenoptic function only partially (specifically, not capturing view-dependent lighting), but over large regions. Furthermore, MCOP range images are not bound by the *free space assumption* of light slab methods, since the range at each pixel is used to resolve visibility. Thus, MCOP images can contain both foreground and background objects (by sweeping the camera over both areas), and the reprojection viewpoint is not limited to lying between the camera path and the projection surface.

Second, *how well* does an MCOP image sample the plenoptic function? Since the functions being sampled (color and range) are not bandlimited, aliasing is inevitable. To minimize it, we must prefilter the signals (that is, perform area sampling rather than point sampling) as we must with light field rendering [16]. In that work the camera spacing is constant, and thus also the filtering kernel. However, in MCOP images the camera spacing and orientation may vary across the image. Therefore, a larger filtering kernel is required in regions of greater camera translation or rotation. However, as the filtering kernel is increased, the resolution of each sample is effectively reduced, since a greater portion of the scene is blurred into each sample. To avoid excessively-large kernels (that is, excessive blurring), the sampling rate should be increased in regions of fast camera motion.

5 REPROJECTING MCOP IMAGES

This section describes how to render a new viewpoint using an MCOP reference image. This consists of two steps: computing each pixel's reprojected location in world-space, and rendering the reprojected points using an appropriate reconstruction method. Alternatively, we may skip the reprojection into world-space, instead projecting from the 2D reference image directly into the 2D output image, as described in [18].

5.1 Reprojection Formula

Since an MCOP image conceptually contains a full camera description plus range for each pixel, the reprojection step is straightforward (in the strip camera case, we need only one camera model per column). Our implementation stores the camera information for each column i as four vectors: a center of projection C_i, a vector O_i from C_i to the image plane origin, a U_i vector defining the horizontal axis of the projection plane, and a V_i vector defining the vertical axis. Each pixel (i,j) contains *disparity* rather then depth, defined here as the distance from C_i to the image plane at a given pixel, divided by the distance from C_i to the pixel's corresponding world-space point. Thus disparity is inversely proportional to range.

Given this camera model and the disparity $\delta_{i,j}$ for a pixel (i, j), the 3-space reprojection (x, y, z) is:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{d_{i,j}} \begin{bmatrix} U_{i_x} & V_{i_x} & O_{i_x} \\ U_{i_y} & V_{i_y} & O_{i_y} \\ U_{i_z} & V_{i_z} & O_{i_z} \end{bmatrix} \begin{pmatrix} i \\ j \\ 1 \end{pmatrix} + \begin{pmatrix} C_{i_x} \\ C_{i_y} \\ C_{i_z} \end{pmatrix}$$

If the dataset is rendered in column-major order, we can reproject the pixels incrementally, since C_i, O_i, U_i, and V_i are constant for each column. Also note the (i, j) coordinates of a pixel are *implicit* (since the image is a regular grid) and do not have to be explicitly maintained.

5.2 Image Reconstruction

After calculating the reprojected coordinates of each pixel in the reference image, there are two common methods for reconstructing a *conventional* range image from a new viewpoint: splatting and meshing. Splatting consists of directly rendering each point using a variable-size reconstruction kernel (e.g., a Gaussian blob), with size dependent on distance from the eye to the point [25][29]. Meshing consists of connecting adjacent pixels with triangles, quadrilaterals, or some higher-order surface. Visibility can be determined for both methods by z-buffering.

Splatting with MCOP images is exactly as with conventional images, since each point is rendered independently: we compute each pixel's 3-space location, then render that point with an appropriate reconstruction kernel. Meshing can also be employed as with conventional images; although neighboring pixels may come from different cameras, the constraints of the MCOP definition ensure that neighboring pixels in the image represent neighboring points in space.

For proper meshing of conventional or MCOP images, discontinuities in the range image must first be detected. For example, three adjacent pixels with one belonging to a foreground object and the others belonging to the background should not be connected in the rendered image. Methods for detecting these *silhouette edges* are discussed in [17][7][27]. Such algorithms can be easily extended to the MCOP domain. For example, our silhouette-detection implementation for a *single*-COP image will not connect adjacent pixels if the surface they define is sampled by a nearly parallel ray (see Figure 9). It assumes these points probably span empty space, rather than a real surface. This method is directly extended to MCOP images by testing each triangle (from three adjacent pixels) against the rays of the two or three cameras that contribute to it: if the triangle is nearly parallel to them all, it is not rendered; otherwise, it *is* rendered. This comparison is only performed once, as a preprocess.

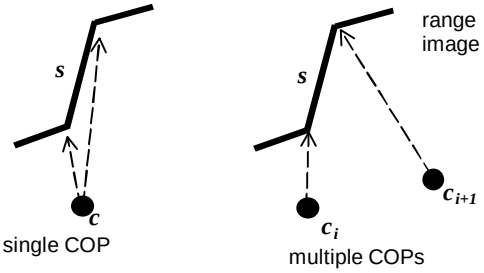

Figure 9 In the single-COP case the surface *s* is nearly parallel to the rays of camera *c*, and so is not rendered. In the MCOP case, *s* is not parallel to the ray from the second camera, and therefore *is* rendered.

MCOP images have an advantage over conventional images when rendering with meshes. In the conventional case a large number of separate images may be required to fully sample an object. One can only easily connect adjacent pixels *within each image*, but points from different images cannot be connected with a mesh unless a "zippering" preprocess [5] is performed. The amount of zippering required increases with the number of images, since each image's boundaries will tend to produce additional seams. With the MCOP method, this problem is minimized, since there is only one image in the dataset. By reducing the number of boundaries to consider, MCOP images can greatly reduce the amount of zippering necessary to fully connect a scene.

5.3 Multiple Sampling of Scene Points

An MCOP image may contain multiple samplings of the same scene point. Since these samples map to the same area in the new, reconstructed image, simple meshing or splatting methods leave only the last-rendered of the frontmost samples in the image. This, however, may not be the best sample. A better method is to blend successive samples as they are written to the image buffer, as described in [7] and [21].

6 EPIPOLAR GEOMETRY

A fundamental relationship between a pair of conventional images is the epipolar geometry they define: a pixel with unknown range in one image will map to a line in the other (planar) image [8][2]. This property characterizes the *optical flow* between the two images, and aids in solving the *correspondence problem* - if one image's pixel views a feature **x**, then to find that feature in the other image, we need only search the corresponding line that the pixel maps to. This property has led to the

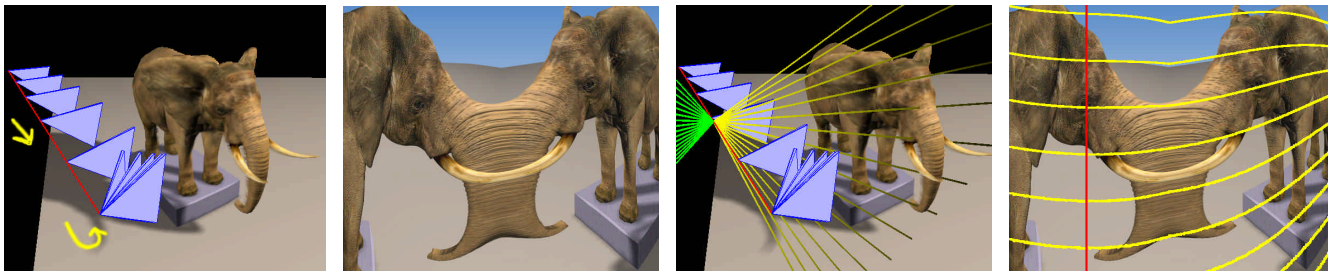

Figure 10 Internal epipolar geometry for an MCOP image. a) Elephant model. The camera first sweeps the head in one direction, then in the other direction b) the MCOP image c) rays cast by a camera, and acquired by every other camera. Positive rays are yellow, negative rays green. d) the epipolar curves induced by the rays of the leftmost eye's camera, marked in red. Note how the curve that crosses the first viewing of the eye also crosses the second viewing of the eye. Given these two corresponding points, we can find the range of the eye by triangulation.

development of occlusion-compatible rendering order for range images and layered depth images [19][23], and greatly simplifies passive range-finding methods based on correspondence.

We can similarly define an epipolar geometry between an MCOP image and a conventional image, or between two MCOP images, by merely projecting the rays of one image's pixels into the other image. However, these will usually map to curves in an MCOP image, rather than to lines.

We can also define an *internal* epipolar geometry using only a *single* MCOP image. Pick some camera $\mathbf{c_i}$ of the MCOP image, and project its rays into every other camera of the image (Figure 10). The result will be a family of curves, which characterizes the optical flow between $\mathbf{c_i}$ and the other cameras. If some feature $\mathbf{x}$ is seen multiple times in the image (e.g., the elephant's eye in the image, which is seen twice), then all sightings must lie along an epipolar curve. Thus the *epipolar constraint* holds for MCOP images, although it remains as future work whether the internal epipolar geometry will lead to useful correspondence algorithms.

7 CAMERAS ALONG A SURFACE

This paper has dealt exclusively with MCOP images constructed by placing cameras along a curve. However, the definition of multiple-center-of-projection images also allows the cameras to define a surface. This approach allows more independence among viewing locations and rays by providing an additional dimension of parameterization for the cameras. We can, for example, parameterize a 3D surface by s and t, then define a camera for m and n discrete points along s and t, respectively. The viewing rays need not be normal to the surface, but must vary smoothly across neighboring pixels.

A useful case where the cameras define a surface can be constructed as follows: consider a static rig in the shape of an arc, lined with a curved one-dimensional CCD array on the inside (Figure 12). This rig is then rotated around a target object (the CCDs need not point directly at the center of rotation). Since each CCD is considered a camera, this case constructs an MCOP image where the camera forms a surface, rather than a curve. Note that although the camera locations and orientations now differ *for every pixel* in the resulting image, they can be efficiently described by a parametric surface, and thus do not significantly increase the dataset size.

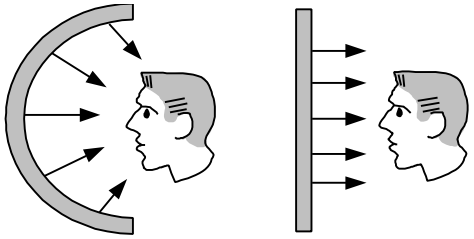

Figure 12 Cameras along an arc (left) define a surface of revolution when rotated, and may be used to capture a greater solid angle compared to conventional cylindrical scanners (right)

This configuration is very similar to Cyberware scanners, except their rig is not curved, so all of their viewing rays are parallel. A limitation of non-curved scanners is that they cannot appropriately sample the tops or bottom of many commonly scanned objects, such as the human head. The curved camera surface, however, will adequately handle this.

8 RESULTS

Figures 5 through 8 and 13 through 17 show the multiple-center-of-projection method applied to two synthetic scenes. Each MCOP image was rendered in 3D Studio MAX as described in section 4.2: we defined one curve around the model for the camera location, another curve for the camera's *lookat* point, and then rendered a single-pixel-wide image at each step along these curves. We used a custom plug-in to extract the Z-buffer and camera information (*C*, *O*, *U*, and *V* vectors) for each pixel-wide image. The color, range, and camera information were then concatenated into a rectangular color image, a rectangular disparity map, and an array of camera descriptors.

The castle model demonstrates the MCOP method's improved connectivity by rendering the entire image with a single mesh. The Titanic model demonstrates the method's flexible image acquisition, by sampling the flag at a higher resolution than the rest of the ship.

Figure 11 shows the results of an early experiment to acquire real-world data. A UNC HiBall optoelectronic tracker was rigidly attached to an Acuity Research AccuRange4000 laser range finder, which reflected from a rotating mirror. This rig was moved along the path shown in Figure 11b, sweeping out columns of range and reflectance at discrete points on the path. The camera swept over the scene a total of six times (note that the scene photographed is a recreation of that which the laser actually

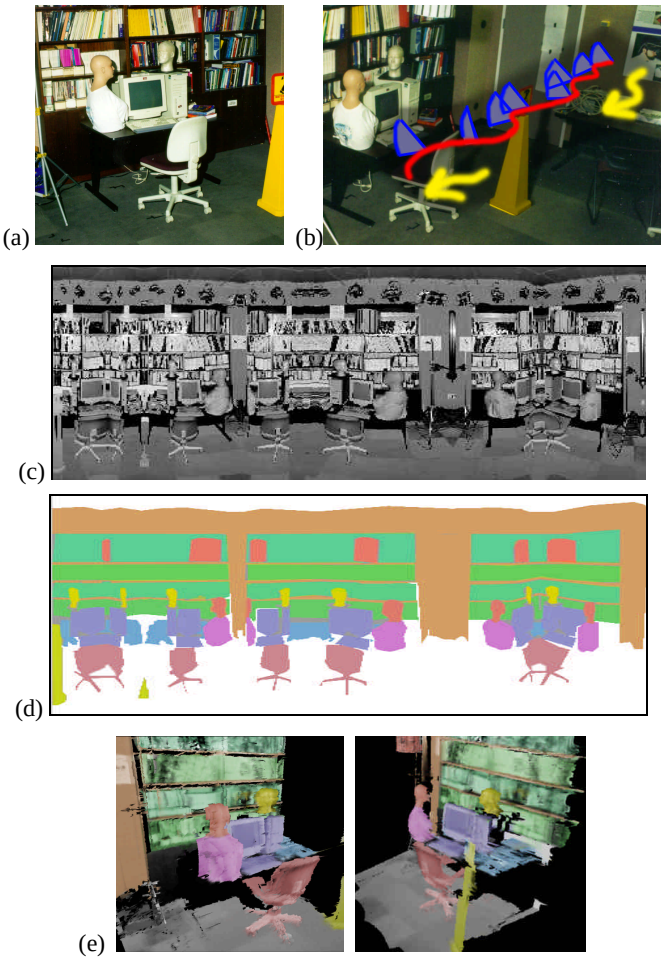

Figure 11 Results of early experiment (see section 8 for details). (a) Scene. (b) Camera path used to acquire it. (c) The grayscale acquired image. (d) Color mask applied to it. (e) Reprojected views.

scanned, since equipment was moved between the day of the experiment and the day of the photographs). The MCOP image shown in (c) is the grayscale reflectance reported by the laser. The color mask was hand-created to better distinguish the separate objects. Two reprojections of the entire dataset. The results from this early experiment were reasonable, and proved the feasibility of the MCOP method for real-world usage. The main problems encountered were static misregistration between the tracker and laser, and poor range data resulting from specular surfaces.

9 LIMITATIONS AND FUTURE WORK

The tradeoffs and limitations of MCOP images are similar to those of conventional images. In order to capture view-dependent lighting in an MCOP image, we must acquire a feature multiple times; this is analogous to sampling a point from different locations with conventional images. As with all image-based methods, the rendering quality is limited by the original image sampling; this is in contrast with geometric methods, which can represent objects as continuous entities. Also, since the number of viewpoints in an MCOP image tends to be larger than in a set of conventional images, there are more opportunities for error to be introduced by the camera tracking. Finally, while MCOP images allow greater acquisition flexibility, they do not solve the problem of finding a set of viewpoints that fully cover a scene, or sample it at some optimal quality.

Conventional range images can be rendered in occlusion-compatible order [19], a scene-independent list-priority technique that eliminates the need for z-buffering with a single reference image. This cannot be applied directly to multiple-center-of-projection images, however, due to their complex epipolar geometries. It remains as future work to classify what subsets of all possible MCOP configurations *can* be rendered in occlusion compatible order.

Another area of future work concerns the construction of complete polygonal meshes for CAD or geometric rendering [5][27], given an MCOP range image. The problem is simplified somewhat in the MCOP domain since we have connectivity information across *all* neighboring pixels, as opposed to dealing with an unorganized set of points, or multiple reference images which only contain connectivity information within themselves. Nonetheless, while MCOP images may reduce the number of seams that must be stitched, they are not altogether eliminated. Whenever a surface is viewed multiple times, the spatial connectivity across the different viewings must still be determined.

10 CONCLUSION

In this paper we have developed and discussed the multiple-center-of-projection image for image-based rendering. MCOP images alleviate many problems of image-based rendering by maintaining a *single* image, containing information from multiple viewpoints. They allow for better image reconstruction than conventional range images. They are capable of sampling different portions of a scene at different resolutions. They provide greater control over the sampling process, allowing the directions of viewing rays to vary over the image, thereby acquiring better samples than if all were bound to the same COP. They also posses a unique internal epipolar geometry that defines how multiple viewings of scene points relate to each other.

Multiple-center-of-projection images have already proven their usefulness in real-world domains, though under different names. The strip camera, for example, has existed for almost a hundred years, and has been used for such important tasks as aerial photography for the last fifty. The Cyberware scanner, another MCOP device, has proven invaluable to a wide range of computer graphics applications. This paper presents a framework by which these existing methods can be exploited further. More importantly, it extends the notion of what it means to be an "image" in image-based rendering.

ACKNOLEDGEMENTS

We would like to thank Mary Whitton, Nick England, Russ Taylor, Anselmo Lastra, Henry Fuchs, and the UNC IBR group for their assistance and insight at various stages of this work. Special thanks to Fred Brooks for his thorough critique of an early draft, and to Lars Nyland for his collaboration on the laser experiment. Thanks also to the SIGGRAPH reviewers. Photographs are by Todd Gaul. The models were provided by REM Infografica. This work is supported by DARPA ITO contract number E278, NSF MIP-9612643, DARPA ETO contract number N00019-97-C-2013, and an NSF Graduate Fellowship. Thanks also to Intel for their generous donation of equipment.

REFERENCES

[1] E. H. Adelson and J. R. Bergen. The Plenoptic Function And The Elements Of Early Vision. In *Computational Models of Visual Processing*, pp. 3-20, Edited by Michael Landy and J. Anthony Movshon, The MIT Press, Cambridge, 1991.

[2] Robert C. Bolles, H. Harlyn Baker, and David H. Marimont. Epipolar-Plane Image Analysis: An Approach To Determining Structure From Motion. In *International Journal of Computer Vision*, volume 1, page 7-55. Boston, 1987.

[3] Shenchang Eric Chen and Lance Williams. View Interpolation For Image Synthesis. In *Proceedings of SIGGRAPH 93*, pp. 279-288, New York, 1993. ACM.

[4] Shenchang Eric Chen. Quicktime VR: An Image-Based Approach To Virtual Environment Navigation. In *Proceedings of SIGGRAPH 95*, pp. 29-38, New York, 1995. ACM.

[5] Brian Curless and Marc Levoy. A Volumetric Method For Building Complex Models From Range Images. In *Proceedings of SIGGRAPH 96*, pp. 303-312. New York, 1996. ACM.

[6] William J. Dally, Leonard McMillan, Gary Bishop, and Henry Fuchs. The Delta Tree: An Object-Centered Approach To Image-Based Rendering. MIT AI Lab Technical Memo 1604, May 1996.

[7] Lucia Darsa, Bruno Costa Silva, and Amitabh Varshney. Navigating Static Environments Using Image-Space Simplification And Morphing. In *Proceedings of 1997 Symposium on Interactive 3D Graphics*, pp. 25-34. April 1997.

[8] Olivier Faugeras. *Three-Dimensional Computer Vision: A Geometric Approach.* MIT Press, Cambridge, Massachusetts, 1993.

[9] Sanjib K. Ghosh. *Analytical Photogrammetry*, second edition. Pergamon, 1988.

[10] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The Lumigraph. In *Proceedings of SIGGRAPH 96*, pp. 43-54, New York, 1996. ACM.

[11] Eduard Gröller and Helwig Löffelmann. Extended Camera Specifications for Image Synthesis. In *Machine Graphics and Vision*, 3 (3), pp. 513-530. 1994.

[12] Michael Halle. Multiple Viewpoint Rendering. To appear in *Proceedings of SIGGRAPH 98*. New York, 1998. ACM

[13] Richard Hartley and Rajiv Gupta. Linear Pushbroom Cameras. In *Proceedings of Third European Conference on Computer Vision*, pp. 555-566. New York, 1994.

[14] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. Surface Reconstruction From Unorganized Points. In *Computer Graphics (SIGGRAPH 92 Conference Proceedings)*, volume 26, pp. 71-78. New York, July 1992. ACM.

[15] Stephane Lavaeu and Olivier Faugeras. 3-D Scene Representation As A Collection Of Images. INRIA Technical Report RR-2205. February 1994, INRIA.

[16] Marc Levoy and Pat Hanrahan. Light Field Rendering. In *Proceedings of SIGGRAPH 96*, pp. 31-42, New York, 1996. ACM.

[17] William R. Mark, Leonard McMillan and Gary Bishop. Post-Rendering 3D Warping. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, page 7-16, Providence, Rhode Island, April 1997.

[18] Leonard McMillan and Gary Bishop. Shape As A Perturbation To Projective Mapping, UNC Computer Science Technical Report TR95-046, University of North Carolina, April 1995.

[19] Leonard McMillan and Gary Bishop. Plenoptic Modeling: An Image-Based Rendering System. In *Proceedings of SIGGRAPH 95*, pp. 39-46, New York, 1995. ACM.

[20] Shmuel Peleg and Joshua Herman. Panoramic Mosaics by Manifold Projection. In *Proceedings of Computer Vision and Pattern Recognition*, pp. 338-343, Washington, June 1997. IEEE.

[21] Kari Pulli, Michael Cohen, Tom Duchamp, Hugues Hoppe, Linda Shapiro, Werner Stuetzle. View-Based Rendering: Visualizing Real Objects from Scanned Range and Color Data. In *Proceedings of Eighth Eurographics Workshop on Rendering*, pp. 23-34. Eurographics, June 1997.

[22] Matthew Regan and Ronald Pose. Priority Rendering with a Virtual Reality Address Recalculation Pipeline. In *Proceedings of SIGGRAPH 94*, pp. 155-162, New York, 1994. ACM.

[23] Jonathan Shade, Steven Gortler, Li-wei He, and Richard Szeliski. Layered Depth Images. To appear in *Proceedings of SIGGRAPH 98*. New York, 1998. ACM

[24] François X. Sillion, George Drettakis and Benoit Bodelet. Efficient Impostor Manipulation For Real-Time Visualization Of Urban Scenery. In *Proceedings of Eurographics 97*, pp. 207-218. Budapest, Hungary, September 1997.

[25] Edward Swan, Klaus Mueller, Torsten Moller, Naeem Shareef, Roger Crawfis, and Roni Yagel. An Anti-Aliasing Technique For Splatting. In *Proceedings of IEEE Visualization 97*, pp. 197-204, 1997

[26] Jay Torborg, James T. Kajiya. Talisman: Commodity Real-Time Graphics For The PC. In *Proceedings of SIGGRAPH 96*, pp. 353-363. New York, August 1996. ACM.

[27] Greg Turk and Marc Levoy. Zippered Polygon Meshes From Range Images. In *Proceedings of SIGGRAPH 94*, pp. 311-318. New York, July 1994. ACM.

[28] Gregory Welch and Gary Bishop. SCAAT: Incremental Tracking With Incomplete Information. In *Proceedings of SIGGRAPH 97*, pp. 333-344. Los Angeles , August 1997. ACM.

[29] Lee Westover. Footprint Evaluation For Volume Rendering. In *Computer Graphics (SIGGRAPH 90 Conference Proceedings)*, volume 24, pp. 367-376. New York, August 1990. ACM.

[30] Daniel N. Wood, Adam Finkelstein, John F. Hughes, Craig E. Thayer, and David H. Salesin. Multiperspective Panoramas For Cel Animation. In *Proceedings of SIGGRAPH 97*, pp. 243-250. New York, 1997. ACM.

[31] Jiang Yu Zheng and Saburo Tsuji. Panoramic Representation for Route Recognition by a Mobile Robot. In *International Journal of Computer Vision* 9 (1): 55-76. Netherlands, 1992. Kluwer.

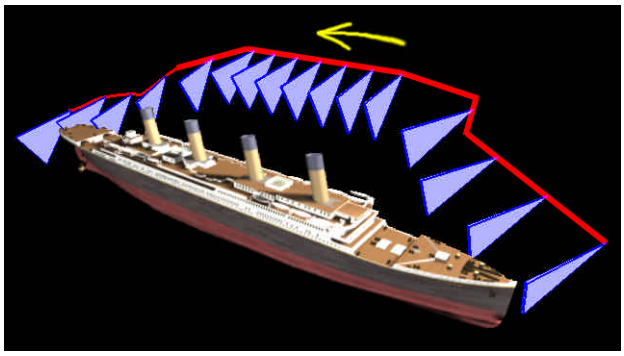

Figure 13 Titanic. We sample the fore and aft more closely than the midsection, for better reconstruction at those areas. The camera moves very close at the end of the path, to capture a small flag at the rear of the ship.

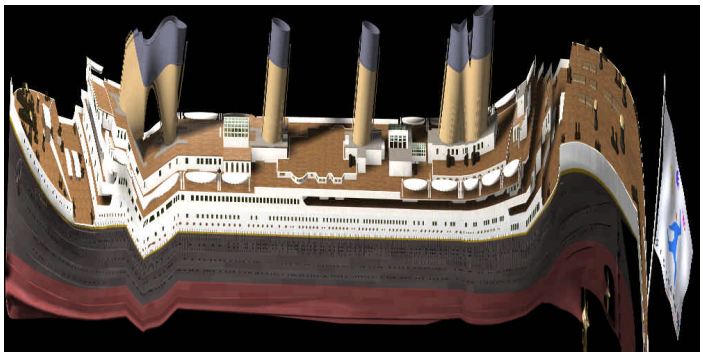

Figure 14 The resulting 1000×500 MCOP image. Because the fore and aft of the ship were sampled at a higher resolution, they occupy a larger portion of the image. The right side of the image shows the finely-sampled flag.

Figure 15 The image on the left is constructed by splatting every fourth column in the MCOP dataset. This shows the image-slices acquired by each camera on the curve. On the right is the dataset rendered by connecting adjacent points with triangles.

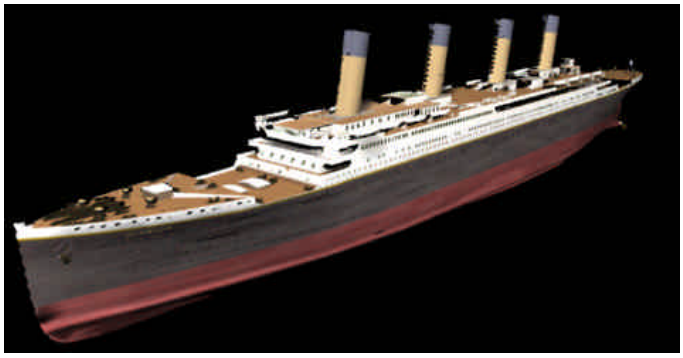

Figure 16 A full view of the reprojected dataset. The back sides of the funnels are not seen, because they were not sampled by the cameras (see Figure 13).

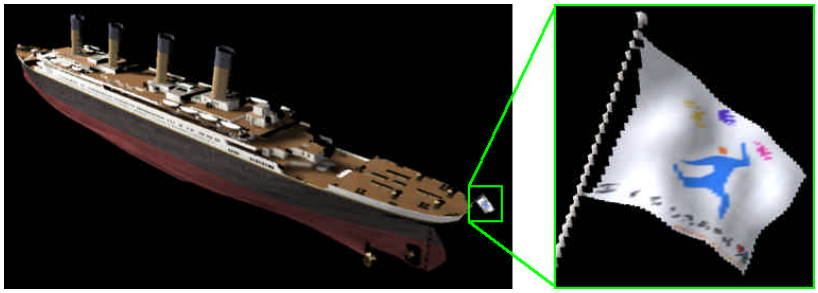

Figure 17 The rear of the ship. The flag was sampled at the highest resolution of the image (Figure 14), allowing an extreme close-up of it.