CONCAVE SURROUND OPTICS FOR RAPID MULTI-VIEW IMAGING

A. Jones*, P. Debevec,
USC Institute for Creative Technologies
Marina del Rey, CA 90292

M. Bolas
University of Southern California
School of Cinematic Arts
Los Angeles, CA 90089

I. McDowall
Fakespace Labs
Mountain View, CA 94043

ABSTRACT

Many image-based modeling and rendering techniques involve photographing a scene from an array of different viewpoints. Usually, this is achieved by moving the camera or the subject to successive positions, or by photographing the scene with an array of cameras. In this work, we present a system of mirrors to simulate the appearance of camera movement around a scene while the physical camera remains stationary. The system thus is amenable to capturing dynamic events avoiding the need to construct and calibrate an array of cameras. We demonstrate the system with a high speed video of a dynamic scene. We show smooth camera motion rotating 360 degrees around the scene. We discuss the optical performance of our system and compare with alternate setups.

1. INTRODUCTION

When analyzing the appearance and dynamics of real-world scenes, it is often useful to photograph the scene from many viewpoints. Just as human stereo vision provides 3D information about the world, multiple camera viewpoints can provide key insight into the 3D structure and dynamics of real-world scenes.

High speed photography has been widely used over the past decades for the analysis of complex motion such as turbulent liquids, human motion, and ballistics. Combining high-speed photography with multi-view imaging may reveal much more information concerning the dynamics of such events.

Previous multi-view techniques have involved either mechanically rotating the subject or camera or using multiple cameras. However, it is often difficult to move the subject and camera at the high speeds required to obtain significantly varying viewpoints with respect to the rate of high speed photography. As a result, slow-motion photography (taken with a high speed camera) is usually photographed from a single viewpoint. The rare cases showing camera motion typically use arrays of cameras which are often hard to assemble and calibrate.

In this paper we present an optical system capable of rapidly moving the viewpoint around a scene. Our system uses a cylindrical mirror which surrounds the scene, and a smaller spinning mirror to direct the camera's viewpoint toward different positions on the cylindrical mirror. The final result is a circular array of virtual viewpoints centered on the scene. We explore the focal length properties of such a system and compare against other mirror setups.

2. BACKGROUND AND RELATED WORK

Previous systems for multi-view photography can be generally classified as multi-camera systems, motorized systems, or systems with multi-view optics. Frequently, these systems are designed to address the broader problem of reflectance capture where both lighting and viewpoint vary. In this paper we focus on viewpoint variation, though our technique could be integrated into a reflectometry system (Hawkins, 2001; Matusik 2002; Han, 2003; Tong 2005) in which the incident lighting is varied as well.

Camera arrays have been used for multi-viewpoint capture as early as the 19th century. In 1878, Eadweard Muybridge used a linear array of still cameras to capture the motion of a running horse [Muybridge, 1878; Muybridge, 1885; Solnit 2004]. However, the slightly differing viewpoints acquired were an unintended artifact of the system. More recently, various artists and technologists have used still camera arrays to create "time-slice" virtual camera moves in which the scene is shown frozen or very slowly moving as the camera viewpoint changes [Macmillan 1984; Taylor 1996]. In the mid-90's the French visual effects firm BUF Compagnie used view interpolation between two camera positions to synthesize camera motion across a frozen dynamic subject in several television commercials and music videos [Buffin 1996]. Camera arrays have been explored for purposes of multi-view video transmission [Yang 2002] and high-speed, high dynamic range, and high-resolution applications [Wilburn 2005]. While providing additional flexibility, large camera arrays are typically expensive and require significant effort to calibrate temporally, geometrically and chromatically. By contrast, our system uses a single camera and a relatively simple arrangement of mirrors to acquire multiple viewpoints without moving the camera or the scene.

The most common way to capture multiple views of a scene is to rotate the subject on a turntable, move the camera around the scene manually, or use a dolly or motion control system. For example, Kaidan (www.kaidan.com) rotation tables are frequently used to shoot Quicktime VR object movies [Chen 95]. Motion-control systems have been used to rotate cameras and samples for reflectance capture [Murray-Coleman and Smith, 1990; Dana, 1999; Dana, 2002]. Often, motorized turntables are combined with linear camera arrays [Hawkins, 2001; Matusik, 2002; Tong, 2005]. Unfortunately, motorized systems are generally too slow for dynamic scene capture, and non-rigid scenes can be undesirably affected by being moved during image capture. Our system avoids these problems by rotating only a lightweight mirror element about its axis to produce virtual motion around the scene.

Multi-view optical systems use additional optical elements such as mirrors and/or lenslets to create the appearance of many viewpoints within a single image. Such systems have few or no moving parts, and can produce useful image datasets when used with a sufficiently high-resolution camera. [Yang, 2000; Ng, 2005; Georgiev, 2006] add additional lenslets either behind or in front of the lens to capture multiple angular samples for each scene point. These techniques require trading off spatial resolution for angular resolution. In contrast to our work, the lenslets are not arranged to surround the scene, and allow for capturing the scene only from the same fixed arrangement of viewpoints. [Ward 1992] placed a reflectance sample and a fisheye camera near the center of a hemispherical mirror, so that the hemisphere of radiant light from the subject is reflected back to the lens of the camera. This allowed a point sample to be viewed from many different directions in a single photograph. Related optics were explored by [Carter, 1999; Mattison, 1998]. [Dana 2001] combined a parabolic mirror with a translation stage to capture the outgoing radiance from a point of a reflectance sample. Unlike our system, these record only a point sample of the scene in any particular image. [Levoy 2004] used a 4×4 array of small flat mirrors and a high-resolution still camera to simulate a small camera array, however, the set of views was fixed and discrete rather than continuous.

The kaleidoscope system of [Han and Perlin, 2003] uses a prismatic conical mirror placed around the scene. Interreflections between the mirrors yield a discrete sampling of views across the upper viewing hemisphere in a single image. By adjusting the taper angle of the kaleidoscope, users can trade-off between spatial and angular resolution in a single photograph. [Hawkins, 2005] uses a smooth mirrored cone around an illuminated volume of participating media to measure its phase function in a single image. [Kuthirummal and Nayar, 2006] surround the scene using mirrored cones and cylinders to produce multi-perspective views of a scene within a single photograph. However, these optical systems generally obtain only one complete image of the scene within a frame. In contrast, our system uses two mirrors, one cylindrical and one rotating flat mirror, which produce continuously variable camera motion around the scene in successive frames of a video sequence.

In this paper we use our system to capture high-speed video of a milk splash in which the camera angle rotates continuously during the event. This work is inspired by the well-known photographic work of Harold Edgerton who pioneered the use of a stroboscope flash to freeze high-speed motion [Kayafas 2001]. Due to the complexity of the photographic equipment and the non-rigid nature of his subjects, the majority of Edgerton's work shows a fixed relationship between the camera and the subject. In 1994, Tim Macmillan produced camera motion around a frozen milk drop using an array of approximately seventy macro cameras [Macmillan, 1984]. By contrast, our system produces video of a milk splash in slow motion from a continuously rotating viewpoint using a single high-speed camera instead of a camera array. While this makes our setup more straightforward in some respects, we are not able to completely freeze the motion since our images are taken in succession. Furthermore, while our system does not require the complex calibration and alignment of an array of images taken from different cameras, it exhibits a limited depth of field and some image warping in it current instantiation.

3. SYSTEM DESIGN

When using a smooth conical mirror, additional views can be generated by moving the camera off-axis. As the radial slices are no longer all focused on the same vertical line, we can reconstruct a perspective view. Instead of physically moving the primary camera, we place a tilted flat mirror between the camera and the subject. By rotating this intermediate mirror we can generate many novel perspective views of the scene.

Our system, shown in Figure 1, consists of a small scene, the camera, a flat spinning mirror, and a relatively

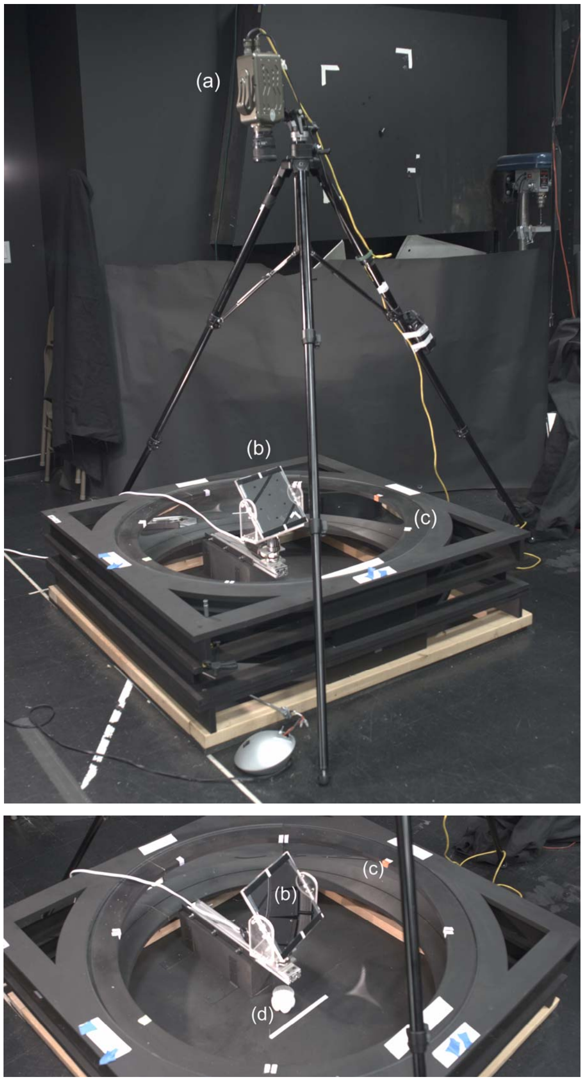

Figure 1: Photograph of apparatus including (a) the high speed camera, (b) the spinning mirror (c) the cylindrical mirror, and (d) the scene, in this case a plastic cup of milk.

large cylindrical mirror surrounding the scene. Our cylindrical mirror has a radius (r) of 53cm and a height of 20cm in height. The camera, subject, and spinning mirror are placed along the optical axis of this cylinder. The subject is placed 20cm below the center of the cylindrical mirror, and the spinning mirror is mounted on a vertical motor shaft such that its center is an equal distance above the cylindrical mirror. Finally, the camera is placed approximately 100cm above the spinning mirror, aimed looking directly down on it. By tilting the angle (α) of the spinning mirror approximately 35° from vertical, an image of the scene reflects out toward the cylindrical mirror, then back to the spinning mirror, and then up to the camera. The rotation of the mirror allows different viewpoints of the scene to be reflected toward the camera. Since the spinning mirror is lightweight and need only rotate about its center of mass, it is easily driven to rotate as much as 600rpm by an inexpensive motor. In comparison, mechanically rotating the camera around the scene at such a speed at a distance of over a meter would be far less practical and safe.

The camera used is a Vision Research Phantom v7.2 camera which can capture 800x600 color images at up to 4800 fps. As the secondary mirror is relatively small, it can rotate at fast speeds without affecting the scene or camera.

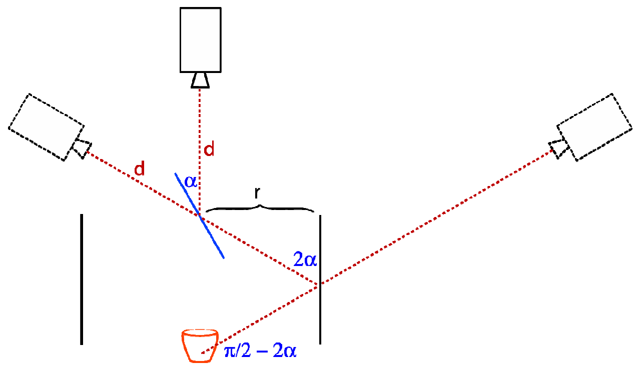

Figure 2: In our setup, rotating the central mirror produces a continuous set of virtual camera positions around the scene. The position of the virtual camera depends on the distance **d** of the camera to the spinning mirror, the angle of the tilted mirror **α**, and the radius of the cylindrical mirror **r**.

If we unfold the optical system (Figure 2), the virtual camera position follows a circular path centered at the subject. The virtual camera view has elevation angle = $\pi/2 - 2\alpha$, at a distance, $d + 2r / \sin(2\alpha)$. As the actual camera does not rotate with the mirror, the virtual camera appears to roll around its viewing axis as it orbits the scene.

4. OPTICAL PERFORMANCE

The geometrical arrangement of our system is such that the cylindrical curved mirror magnifies the image seen by the camera. The cylindrical element results in an anamorphic optical system. Consequently, there is in effect a different focal length lens combination in the horizontal and vertical directions on the image sensor. This has two major effects as seen in the camera image: first, the aspect ratio of the scene appears stretched horizontally; secondly, the unequal magnification has the effect of creating different foci in the horizontal and vertical directions.

As a result, one can focus the camera lens such that the image is focused on horizontal features of the scene and vertical edges are blurred (see Figure 3a and 4a). Alternatively, one can adjust the focus and bring vertical features of the scene into focus and compromise focus in the horizontal (Figures 3b and 4b). In between these two extremes is a focus where the blur in the horizontal and vertical directions is similar, resulting in an overall defocused image (Figure 4c).

In order to bring the entire scene into focus, we increase the ambient illumination of the scene and reduce the aperture of the camera lens to the point where, at our camera's resolution, the increased depth of field results in an image which appears to be in focus. As illustrated in Figure 5, the aperture is simply reduced such that the pencil of rays becomes small enough to create an apparently focused image. Increasing the illumination in our case is relatively straight forward as we do not rely on any form of structured illumination. Using our experimental setup, the camera lens is generally stopped down to F11 and the scene illuminated by standard theatrical lights.

An alternative solution would be to employ a corrective mirror or lens. There is a number of possible corrective optics; one is the introduction of a second curved mirror (or equivalent lens) between the spinning mirror and the camera lens, as illustrated schematically in Figure 4e. In our system however this would be quite difficult to implement - this additional optical assembly would have to rotate exactly synchronously to the existing spinning mirror.

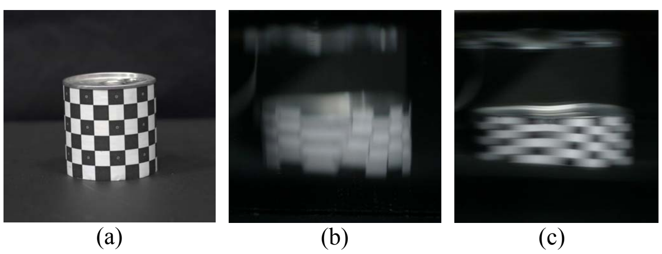

Figure 3: A cylindrical can with a checkerboard label (a) can appear both stretched and out of focus ((b) & (c)) when seen through our optical system. Photographing the scene with a wide lens aperture demonstrates the differing focal distances of vertical and horizontal image detail. With the lens focused at the distance of the spinning mirror, vertical image detail is in focus (b). With the lens focused an additional distance of $2r/\sin(2\alpha)$ beyond the mirror, horizontal detail comes into focus (c). This effect can be compensated for by using either a small aperture or corrective optics.

It is preferable, considering the implementation issues, to improve the image quality through modifying the surface profile of the existing spinning mirror shown in Figure 1b. This optimization has been performed using numerical methods. This optimized mirror profile improves the focus of the image at the camera and further distorts the aspect ratio. An example of the improvement in image quality, relative to Figure 4c, is shown in Figure 4d.

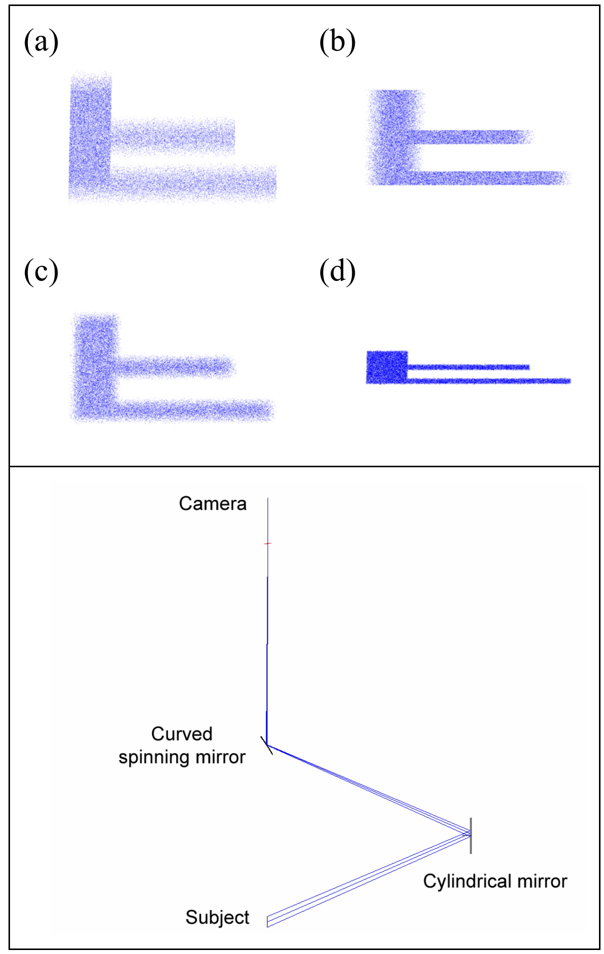

Figure 4: Comparison between original flat spinning mirror and an optimized curved spinning mirror. (top) (a)-(c) show the original mirror focused (a) horizontally (b) vertically, or (c) with the best average focus. (d) uses the optimized corrective lens shown below.

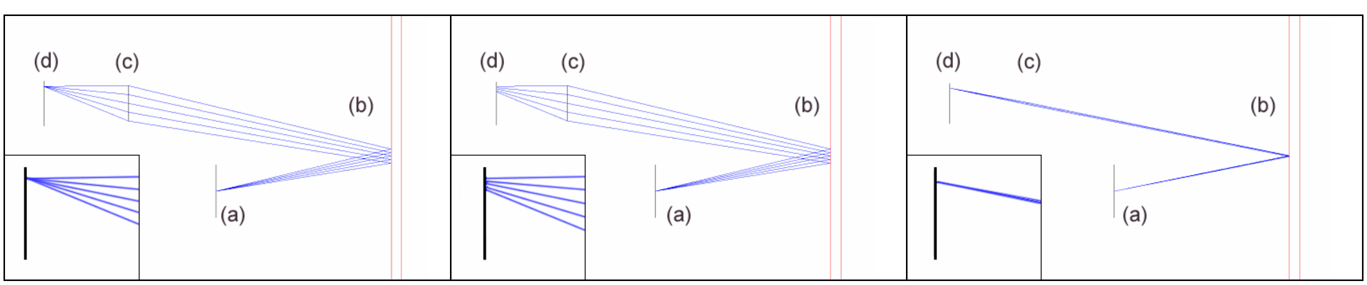

Figure 5: Schematic showing the path of vertical light rays. Rays from the scene (a) reflect off the cylindrical mirror (b), are focused by the camera lens (c) and intersect the image plane (d). At the vertical focal length, the vertical rays are focused on the image plane (left). At the horizontal focal length, the rays do not intersect the plane at a single point. (center). Using a smaller aperture, both the horizontal and vertical rays converge (right). Insets show the intersection of rays with the image plane.

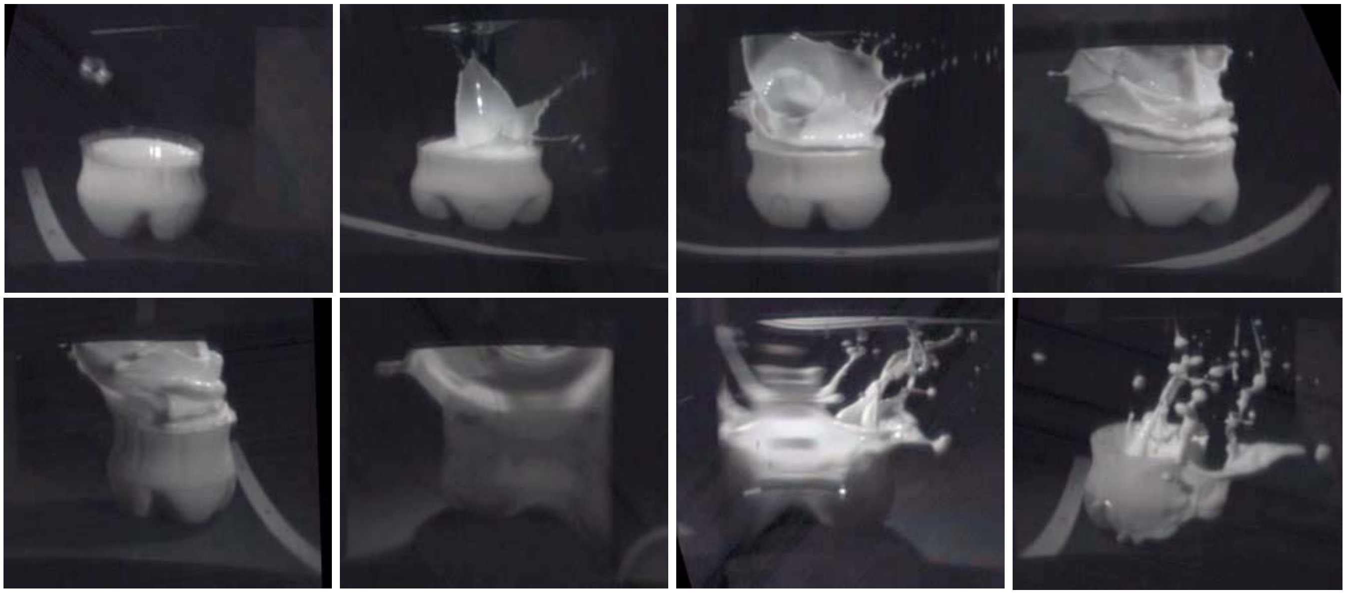

Figure 6: Rapid viewpoint change is shown in a high-speed video of splashing milk. In these still images, the viewpoint rotation is most easily seen in the strip of tape in front of the plastic cup. The original video was captured at 2000 frames per second with the viewpoint virtually rotating around the scene every 500 frames. The mirror distortion is visible in the 6^{th} image in the sequence.

5. RESULTS

To demonstrate our system, we filmed a steel hex nut thrown into a plastic cup of milk. We ran the high-speed camera at 2000 frames per second and rotated the mirror approximately four times per second, achieving slightly less than one degree of rotational motion around the scene per frame of video. The resulting video, once processed, shows slow motion photography of the resulting splash with a continuously rotating point of view (Figure 6).

Before processing, the raw video rotates both around the scene and around the virtual camera's optical axis. We digitally counter-rotate the image to remove this camera roll. This resulting video exhibits the horizontal stretching effect, which we again compensate for with digital image processing to scale the image to the correct aspect ratio. This yields the final processed video.

The final video exhibits subtle temporally-varying warping due to the imperfect shape of the cylindrical mirror, which is formed from a bent sheet of mylar. If better-machined glass optics were employed, this effect would not be present.

6. COMPARISON WITH A FACETED MIRROR APPROACH

A related alternative approach to capturing multiple views of a scene is to surround the scene with multiple planar mirrors (reminiscent of the faceted mirrors of a Praxinoscope, though pointing inward rather than outward) rather than a smooth cylindrical mirror (see Figure 7). In this alternative setup, each planar facet reflects its own view of the scene toward the camera, and any number of such facets could be placed around the scene until the field of view is reduced to less than the extent of the scene. If a particularly wide-angle lens such as a fisheye lens were used, all of these views could be photographed simultaneously in a single image, but at greatly reduced image resolution per view.

Following our principal approach, the same spinning mirror setup can be used to direct the light of each view toward the camera in sequence. Since all of the mirror elements of the system are planar, the resulting images of the scene seen by the camera have none of the optical difficulties of the cylindrical mirror setup: the image is not stretched, and the horizontal and vertical image detail becomes clear at the same focal distance (the images still appear to roll about the z axis as the mirror spins, but this is easily corrected).

However, the faceted mirror approach has three disadvantages compared to the smooth cylindrical mirror approach. First, the number of viewpoints achievable is locked to the number of mirror facets, and is limited by the minimum field of view necessary to see the subject through any one facet. In the smooth mirror approach, the viewpoint rotates continuously with as the spinning mirror rotates. Thus, any number of views around the scene can be obtained by adjusting the speed of the mirror and/or the frame rate of the camera.

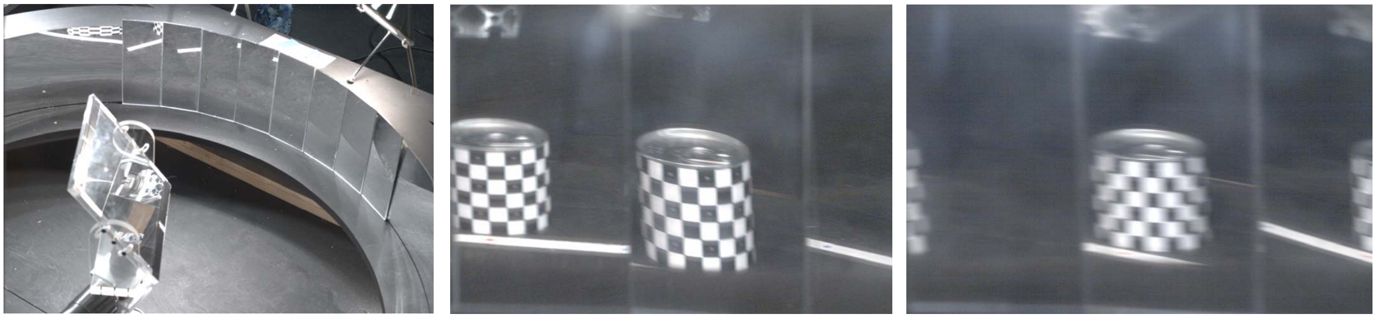

Figure 7: An alternative setup, created for purposes of comparison, of a partial cylinder made of planer facets (left). While this arrangement produces sharp images of the scene for a stationary camera (center), spinning the central mirror produces rapid translation of the image of the scene. Unless photographed at very short shutter speeds, this will cause motion blur in the resulting images (right).

Second, the faceted mirror approach requires synchronization between the camera and the spinning mirror. If the camera were to expose an image with the spinning mirror aimed between two mirror facets, the resulting image would show two fragmented views of the scene at the edges of the frame and be difficult to use. Achieving such synchronization would require an additional encoder and/or a nontrivial motion control system. In contrast, the smooth mirror approach requires no synchronization and produces a seamless image centered in the frame for any mirror position.

The third disadvantage of the faceted mirror approach is that the images reflected in the facets translate rapidly across the field of view of the camera as the mirror spins, similar to cars of a train passing by at high speed. In contrast, the smooth mirror approach produces a steady image of the scene that stays centered within the camera's field of the view. The translational motion of the successive image viewpoints requires a very short shutter speed to obtain a clear image of the scene. If the sensor integration time lasts even a small fraction of the time it takes the mirror to spin from one facet to the next, the resulting image will exhibit translational motion blur (Figure 7, right). This translational motion is significantly more extreme than the additional rolling motion inherent in either spinning mirror system. This is because the roll motion shows the scene spin about the camera axis just once per mirror rotation while the faceted system shows the scene travel across the frame every time the mirror moves from one facet to the next. Having the shutter be open for such a small percentage of the available time is counter-indicated by the exposure needs of high speed photography; typically, a large fraction of the available time between frames is required to sufficiently expose the sensor. Alternatively, a strobe lighting system could be used to freeze the image motion, but this would introduce significant additional system complexity.

For these reasons, our proposed smooth cylindrical mirror approach to multi-viewpoint imaging has significant advantages over the faceted mirror approach. In particular, if corrective optics (rather than a small aperture) is used to correct the astigmatism of the smooth cylindrical mirror system, then the ability to make efficient use of the light available makes it far superior to the faceted approach

7. FUTURE WORK

Currently, the system produces only a one-dimensional array of viewpoints around the scene. For light field acquisition applications, it could be of interest to produce views of the scene from differing inclinations (e.g. from above, straight-on, or below) as well. Ongoing work has shown that this can be done if the cylindrical mirror is replaced with an inward-pointing ellipsoidal mirror having its minor axis coincident with the center of the cylindrical mirror and with its two foci at the centers of the subject and the spinning mirror. In this manner, changing the azimuth of the spinning mirror still changes the azimuth of the virtual viewpoint while changing the inclination of the mirror now changes the viewpoint's inclination, allowing for the capture of a continuous two-dimensional array of viewpoints. Furthermore, the astigmatic nature of the system is largely eliminated as the ellipsoidal mirror is similarly curved horizontally and vertically. However, constructing an ellipsoidal mirror of sufficient imaging quality is nontrivial, and foreshortening of the flat mirror at near-vertical angles prevents achieving viewpoints from directions that approach being directly above or below the scene.

Compared to building a camera array, our apparatus is relatively inexpensive and simple to construct. The current system is built of mirror-coated mylar strips and a lightweight wooden frame. The largest error of the current system is the slight waviness of the cylindrical mirror and the need for corrective optics to compensate for the system's astigmatism, which would allow wider camera apertures to be used.

7. CONCLUSION

We have presented a single-camera technique for capturing dynamic events from multiple viewpoints. Using a small rotating mirror and a larger cylindrical mirror, our system generates many views rotating around the scene while moving neither the camera nor the scene. We believe this multi-view image acquisition process could be used in a variety of computer graphics and vision application involving 3D reconstruction, reflectance capture, and scene understanding.

ACKNOWLEDGEMENTS

The authors wish to thank Bruce Lamond, Tomas Pereira, Laurie Swanson, Larry Vladic, Vision Research, Inc., Bill Swartout, David Wertheimer, Randolph Hall, and Max Nikias for their support and assistance with this work. This work was sponsored by the University of Southern California Office of the Provost and the U.S. Army Research, Development, and Engineering Command (RDECOM). The content of the information does not necessarily reflect the position or the policy of the US Government, and no official endorsement should be inferred.

REFERENCES

Buffin, P., 1996: "Rolling Stones – Like a Rolling Stone". BUF Compagnie. *SIGGRAPH 1996 Electronic Theater*. August 1996.

Carter, R. R., and L. K. Pleskot, 1999: "Imaging scatterometer", *US Patent 5912741*.

Chen, S. E., 1995: "Quicktime VR - An Image-Based Approach to Virtual Environment Navigation", *Proceedings of SIGGRAPH 95,* pp. 29-38.

Dana, K., 2001: "BRDF/BTF Measurement Device", *International Conference on Computer Vision*, pp. 460-466

Dana K., and J. Wang, 2004: "Device for Convenient Measurement of spatially varying bidirectional reflectance", *Journal of the Optical Society of America*, pp. 1-12.

Levoy, M., Chen, B, Vaish, V., Horowitz, I., McDowall, I., Bolas, M., 2004: "Synthetic Aperture Confocal Imaging", *ACM Transactions. on Graphics (Proc. ACM SIGGRAPH 2004) 23(3)*, pp. 825-834.

Han, J. Y. and K. Perlin, 2003: "Measuring bidirectional texture reflectance with a kaleidoscope", *ACM Transactions on Graphics (Proc. ACM SIGGRAPH 2003)* 22(3), pp. 741–748.

Georgiev, T., C. Zheng,, S. Nayar, B. Curless, D. Salasin, and C. Intwala, 2006: "Spatio-angular Resolution Trade-offs in Integral Photography", *Eurographics Workshop on Rendering 2006.*

Hawkins, T., J. Cohen, and P. Debevec, 2001: "A Photometric Approach to Digitizing Cultural Artifacts", *2nd International Symposium on Virtual Reality, Archaeology, and Cultural Heritage (VAST 2001).*

Hawkins, T., P. Einarsson, P. Debevec, 2005: "Acquisition of time-varying participating media", *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2005). 24(3)*, pp. 812-815.

Kayafas G., and E. Jussim, 2001: *Stopping Time: The Photographs of Harold Edgerton*, Harry N. Abrams Inc., pp 168.

Kuthirummal, S., and S. K. Nayar, 2006: "Multiview Radial Catadioptric Imaging for Scene Capture", *ACM Transactions on Graphics (Proc. of ACM SIGGRAPH 2006) 25(3).*

Macmillan, T. 1984: "Split Milk", Film installation, London Film-Makers Co-op, London, England.

Mattison, P. R,, M. S. Dombrowski, J. Lorenz,, K. Davis, H. Mann, P. Johnson, and B. Foos. 1998: "Hand-held directional reflectometer: an angular imaging device to measure BRDF and HDR in real-time", *Proceedings of SPIE, The International Society for Optical Engineering, Scattering and Surface Roughness II*, 3426:pp240—251.

Matusik, W., P. Hanspeter, A. Ngan, P. Beardsley, R. Ziegler, and L. McMillan , 2002: "Image Based 3D Photography using Opacity Hulls", *ACM Transactions on Graphics (Proc. ACM SIGGRAPH 2002) 21(3)*, pp 427-437

Murray-Coleman. F., and A. M. Smith, 1990: "The Automated Measurement of BRDFs and their Application to Luminaire Modeling", *Journal of the Illuminating Engineering Society*, pp. 87--99

Muybridge, E., 1878: "The Horse in Motion", Photographic Series..

Muybridge, E., 1875: "Horses and Other Animals in Motion: 45 Classic Photographic Sequences". Dover Publications.

Ng, R. , M. Levoy, M. Brédif, G. Duval, M. Horowitz and P. Hanrahan, 2005: "Light Field Photography with a Hand-Held Plenoptic Camera", *Stanford University Computer Science Tech Report* CSTR 2005-02

Solnit. R. 2004: *River of Shadows: Eadweard Muybridge and the Technological Wild West,* Penguin Group, pp.320

Taylor, D, 1996: **"**Virtual Camera Movement: The Way of the Future?", *American Cinematographer*, Vol 77, No. 9, September 1996, pp.93-100.

Tong, X., J. Wang, S. Lin, B. Guo and H. Y. Shum, 2005: "Modeling and Rendering of Quasi-homogeneous Materials", *ACM Transactions on Graphics (Proc. of SIGGRAPH 2005) 24(3)*

Ward, G., 1992: "Measuring and Modeling Anisotropic Reflection", *Computer Graphics*, Vol. 26, No. 2.

Wilburn, B., N. Joshi, V. Vaish, E. V.Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, 2005: "High performance imaging using large camera arrays". *ACM Transactions on Graphics (Proc. of SIGGRAPH 2005) 24(3)*, pp765–776.

Yang, J., 2000: A Light Field Camera for Image Based Rendering, Ph.D. thesis, MIT.

Yang, J., M. Everett, C. Buehler, and L. McMillan, 2002: "A Real-Time Distributed Light Field Camera", *13th Eurographics Workshop on Rendering*. pp. 77-86.