Spherical Light Field Environment Capture for Virtual Reality using a Motorized Pan/Tilt Head and Offset Camera

Paul Debevec*† Greg Downing‡ Mark Bolas† Hsuen-Yueh Peng* Jules Urbach*
*OTOY, Inc. †USC Institute for Creative Technologies ‡xRez, Inc.

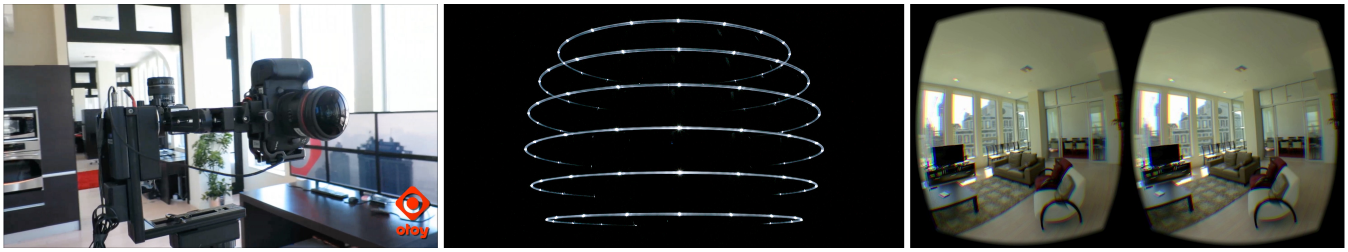

Figure 1: *(Left) Spherical light field capture system with an offset and counterweighted fisheye camera on a programmable pan/tilt head (Center) A long exposure photo of a subsampling of the camera positions. (Right) The light field displayed in a head-tracked VR Headset.*

Introduction Todays most compelling virtual reality experiences shift the users viewpoint within the virtual environment based on input from a head-tracking system, giving a compelling sense of motion parallax. While this is straightforward for computer generated scenes, photographic VR content generally does not provide motion parallax in response to head motion. Even 360° stereo panoramas, which offer separated left and right views, fail to allow the vantage point to change in response to head motion.

Background Light Field Capture and Rendering [Levoy and Hanrahan 1996] records an image-based dataset of a scene using renderings or photographs from a 2D array of different points of view, and synthesizes new points of view by out of the pixels of the images in an array. Most often, the points of view are arrayed in a plane, and can be recorded using a planar motion gantry. Notably, since radiance remains constant along a ray, the views which can be synthesized are not restricted to be within the plane of the original viewpoints, and can be generated both in front of and behind the original capture plane. However, for VR, it would be desirable to have a wide field of view looking in any direction from the range of viewpoints to which the user can move their head.

A standard way to capture a photographic light field is to use a planar light field gantry, where the camera is translated horizontally and vertically to cover a regular grid of viewpoints, but this does not provide the panoramic coverage desirable for VR. Concentric Mosaics [Shum and He 1999] record panoramic but horizontal-parallax-only light fields by rotating a video camera on a turntable with the lens offset from the center of rotation, but this misses vertical parallax and requires depth correction to move the view forward and backwards.

Apparatus We record outward-looking spherical light fields using a Rodeon motorized pan/tilt head, a Canon 5D Mark III DSLR camera with an 8mm fisheye lens, and an additional mounting rail to place the camera 35mm forward from the center of rotation as in 1(Left). To avoid straining the motors, we place a counterweight (in this case, another similar camera) on the opposite end of the rail.

SIGGRAPH 2015 Posters, August 09 – 13, 2015, Los Angeles, CA.
ACM 978-1-4503-3632-1/15/08.
http://dx.doi.org/10.1145/2787626.2787648

Aquisition We programed the pan/tilt head to shoot 72 images horizontally by 24 images vertically, spaced 5° centered around the equator, resulting in a set of 1728 images. Since the fisheye covers a hemisphere of 180°, our light field provides the data necessary to render views of the scene from any position within the 35cm radius sphere outlined by the cameras. In fact, the imagery is sufficient for rendering a 90-degree outward facing view up to nearly 50cm away from the center, with additional loss of FOV at greater distances.

Processing We process the spherical light field so that it can be viewed interactively in an HMD. The images are resized to the resolution appropriate for the HMD's pixel count and field of view, and then the image dataset is compressed using the OTOY's "ORBX" codec for approximately 1000:1 compression, taking advantage of the multiple dimensions of redundancy in the data.

Rendering The light fields are viewed using OTOY's GPU-accelerated viewer in an Oculus Rift DK2 HMD. The DK2's head tracking is used to interactively render two views of the light field according to the user's eye positions, achieving 75Hz updates. The user can move their head forward, back, up, down, left, and right in the light field within and beyond the volume recorded. The light field can be interactively refocussed at different distances, making either the furniture of the room or the buildings through the windows come into sharp focus. Out of focus regions exhibit some ghosting due to the limited 3cm resolution of the camera spacing.

Future Work Providing scene depth to the light field viewer, either through panoramic laser scanning, photogrammetry, and/or stereo correspodence would allow the light field to be refocussed accurately over the entire field of view, eliminating ghosting artifacts. Building a spherical array of outward-facing video cameras to shoot video light fields is also clearly of interest, as would tackling the significant issue of compression and decompression if one wishes to play back or broadcast a spherical video light field.

References

LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, ACM, New York, NY, USA, SIGGRAPH '96, 31–42.

SHUM, H.-Y., AND HE, L.-W. 1999. Rendering with concentric mosaics. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, SIGGRAPH '99, 299–306.