The Making of Welcome to Light Fields VR

Ryan S. Overbeck
Google Inc.
rover@google.com

Daniel Erickson
Google Inc.
danerickson@google.com

Daniel Evangelakos
Google Inc.
devangelakos@google.com

Paul Debevec
Google Inc.
debevec@google.com

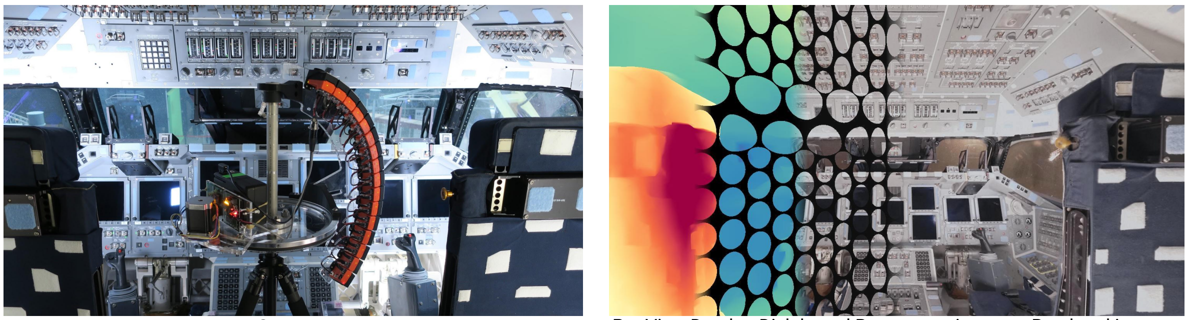

Figure 1: We introduce the systems for capturing and rendering high quality panoramic light field still images used to create the *Welcome to Light Fields* experience for virtual reality. (Left) Our rotating light field camera array is built from a modified GoPro Odyssey. (Right) Our light field renderer uses a novel reconstruction algorithm which projects a per-view geometry (the heatmap to the left) through a disk-based reconstruction basis (center-left) to achieve high quality view interpolation (right).

ABSTRACT

Light fields can provide transportive immersive experiences with a level of realism unsurpassed by any other imaging technology. Within a limited viewing volume, light fields accurately reproduce stereo parallax, motion parallax, reflections, refractions, and volumetric effects for real-world scenes. While light fields have been explored in computer graphics since the mid-90's [Gortler et al. 1996; Levoy and Hanrahan 1996], practical systems for recording, processing, and delivering high quality light field experiences have remained out of reach.

In this talk, we describe the hardware and software developed to produce the *Welcome to Light Fields* virtual reality experience, the first immersive light field experience that can be downloaded and viewed on VR-ready computers with HTC Vive, Oculus Rift, and Windows Mixed Reality HMDs. This piece enables the user to step into a collection of panoramic light field still photographs and be guided through a number of light field environments, including the flight deck of Space Shuttle Discovery. *Welcome to Light Fields* is a free VR app. on the Steam Store that was launched in March 2018 [Google VR 2018].

This piece is the culmination of a multi-year effort to develop consumer-ready light field technology. We built new camera rigs for light field capture and new algorithms for light field rendering. In this talk, we delve into the details from capture to rendering and the steps in between.

We emphasize that our system generates stereoscopic views of a light field for use in current generation consumer VR HMDs. As future work, it could be of interest to display such datasets on near eye light field displays [Lanman and Luebke 2013] to render effects of visual accommodation.

SIGGRAPH '18 Talks, August 12-16, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5820-0/18/08.
https://doi.org/10.1145/3214745.3214811

CCS CONCEPTS

• **Computing methodologies → Image-based rendering**; **Virtual reality**;

KEYWORDS

virtual reality, light fields

ACM Reference Format:
Ryan S. Overbeck, Daniel Erickson, Daniel Evangelakos, and Paul Debevec. 2018. The Making of Welcome to Light Fields VR. In *Proceedings of SIGGRAPH '18 Talks.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3214745.3214811

1 LIGHT FIELD CAPTURE

The advent of consumer VR headsets with positional tracking has created a platform to envelop the user with photographically-acquired VR experiences. However, most such VR experiences offer at most omnidirectional stereo rendering, allowing a user to see stereo in all directions (though not when looking up and down, and only with the head held level) but not to move around in the scene.

To address this, we developed an inside-out spherical light field capture system designed to be relatively easy to use and efficient to operate. Our approach was to repurpose the sixteen synchronized GoPro Hero4 action sports cameras from a GoPro Odyssey omnistereo rig [Anderson et al. 2016] to create a sweeping linear camera array as in Figure 1 (left).

We programmed a belt-driven motor system to rotate the camera arc in 30, 60, or 120 seconds, and to rotate either continuously or to start and stop (if taking 60 seconds or longer) at 72 positions around the circle. We used the fastest 30 second rotation time for brightly lit conditions when motion blur would not be a problem and when capturing people. The slower rotation times and start-stop mode were used to avoid motion blur altogether.

By extracting evenly-spaced video frames from the 16 video files ($2704 \times 2028 pixels$, 30fps), we obtain a dataset of 72 longitudinal and 16 latitudinal samples, or 6 Gigapixels. We subsample the more finely spaced images near the poles in a post-process to get a more even distribution. When subjects are close and will produce greater parallax, we can double the number of latitudinal samples to 32 by rotating the camera array twice. For this configuration, the top of the arc is attached to a string wound about a thin vertical rod, and by unwinding the rig naturally pivots down by half a camera position each rotation.

The size of the rig represents a trade-off between where it is able to fit and the size of the light field viewing volume. Our current rig sweeps the camera centers along the surface of a sphere which is 70cm in diameter, which, given the GoPro's 110° horizontal field of view, yields a light field viewing volume 60cm in diameter. We found that this was a sufficient volume to provide satisfying parallax for a seated VR experience, and it also allowed us to fit the rig into tighter spaces, such as the airlock of the space shuttle.

2 LIGHT FIELD RENDERING

After capturing a light field dataset, there remain significant challenges to delivering a compelling and immersive light field experience to a user. Specifically, we identify three main challenges to producing a light field rendering system that can run on a lower end VR Ready PC:

- **Size** - the data must be small enough to be worth downloading and storing.
- **Quality** - the reconstructed light field must be high enough quality to be believable.
- **Speed** - the renderer must deliver stereo views at 90 Hz.

This portion of our talk will focus on several fundamental algorithms that taken together help overcome the challenges listed above. We developed a novel light field compression format that both reduces our data to a reasonable size and decodes in real-time (see Section 2.1). In order to reduce the time spent decoding data and make sure the the light field data fits in a modest GPU RAM budget, we use a tile-based streaming and caching architecture (see Section 2.2). We also use a new approach to light field reconstruction that uses a disk-based reconstruction field with a multi-view geometry that achieves higher quality than previous real-time approaches with fewer images and leads to an efficient implementation on modern GPUs (see Section 2.3). Our talk will also include several other improvements that help us render high quality lightfields at 90 Hz.

2.1 Compression

Our rendering algorithm requires 1,000–2,000 images to capture a spherical light field that can be rendered with high quality. This amounts to several GBs of data per light field, which would be painful to download over a standard 10–20 Mbit/s connection and hard to load into a modest GPU RAM budget.

Fortunately, the images in a light field exhibit significant coherence which can be leveraged for compression. The obvious approach is to apply standard video compression and map the disparity between cameras to the video's time dimension. Unfortunately, a naive application of video compression would lead to a format that wouldn't be practical for real-time on-demand decoding.

We use a novel light field compression format built upon a modified version of the VP9 video codec [Mukherjee et al. 2015]. By building upon existing video coding technology, we can take advantage of future enhancements to the VP9 codec, and we pave the way to a future hardware implementation of our light field decoder.

Our light field codec compresses our light fields down to 70MB–210MB while maintaining nearly loss-less image quality. At this size, we can download a light field over a standard 10–20Mbit/s connection in a matter of minutes, and over a 100 Mbit/s connection in a matter of seconds.

2.2 Tile Streaming and Caching

In order to both minimize the time spent decoding and fit the decoded light field data into bounded GPU memory, we decode individual light field tiles and stream those to the GPU. This allows us to easily fit within a modest GPU RAM budget which is especially helpful when sharing GPU memory with other assets or for rendering multiple light fields at the same time.

2.3 Disk-based Reconstruction with Multi-view Geometry

Previous real-time light field rendering algorithms struggle to achieve high enough quality to make a believable and immersive experience. We introduce a novel real-time light field model for high quality light field reconstruction that runs efficiently on modern GPUs. As shown in Figure 1 (right), our model consists of a collection of disk-shaped windows that are blended together where the windows overlap. Each window corresponds to one of the camera views that captured the light field. Inside each window, we render a mesh with scaleable geometric complexity that is projection mapped with a light field image. Each mesh is a tessellated version of a depthmap generated using multi-view stereo algorithms. Using this model, we are able to achieve high quality rendering with fewer images than previous real-time approaches, which greatly alleviates the burden on the light field capture hardware and helps compress the data down to a more manageable size.

REFERENCES

Robert Anderson, David Gallup, Jonathan T. Barron, Janne Kontkanen, Noah Snavely, Carlos Hernandez Esteban, Sameer Agarwal, and Steven M. Seitz. 2016. Jump: Virtual Reality Video.

Google VR. 2018. Welcome to Light Fields on Steam. (March 2018). https://store.steampowered.com/app/771310/Welcome_to_Light_Fields/

Steven J Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F Cohen. 1996. The lumigraph. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 43–54.

Douglas Lanman and David Luebke. 2013. Near-eye light field displays. *ACM Transactions on Graphics (TOG)* 32, 6 (2013), 220.

Marc Levoy and Pat Hanrahan. 1996. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 31–42.

Debargha Mukherjee, Jingning Han, Jim Bankoski, Ronald Bultje, Adrian Grange, John Koleszar, Paul Wilkins, and Yaowu Xu. 2015. A technical overview of VP9–The latest open-source video codec. *SMPTE Motion Imaging Journal* 124, 1 (2015), 44–54.