

Light Field Research

Carson Vogt

PhD

Heriot-Watt University

31 August 2015

Abstract

In this review, the light field research area is discussed as well as potential future directions. The content includes an introduction to the topic and its brief history, a literature review on it and other relevant areas to the problem, work that has been done to date outside of the literature review, and proposed research including future steps.

List of Figures

1	The dense camera array used at Stanford to capture a number of light fields.	4
2	Another collection method, the gantry effectively mimics the camera array so long	4
3	The plenoptic function sampled at two points, showing all pencils of light rays which pass through those points [1].	5
4	The plenoptic function at two points, showing all pencils of light rays [1]	7
5	The dense camera array used at Stanford to capture a number of light fields.	9
6	Another collection method, the gantry effectively mimics the camera array so long as the scene is not altered between gantry locations.	9
7	the lytro immerge, designed to capture an entire scene at once, with an array of cameras built into the structure.	10
8	On the right is the camera rig behind the Google Jump VR video system [2]	10
9	This mosaic is a compilation of all the images taken during one light field acquisition session with the baxter robot.	12
10	This mosaic is a compilation of all the images taken during one light field acquisition session with the baxter robot.	13
11	This mosaic is a compilation of all the images taken during one light field acquisition session with the baxter robot.	13
12	A depiction of the second renderer based on [10] and showing the plane of focus F and the rays coming from it that would be considered in focus.	14
13	A screenshot of the light field interface in action, with the subaperture views shown in blue, the aperture shown as the red circle, and the resulting image shown in the larger window.	15
14	The dense camera array used at Stanford to capture a number of light fields.	15
15	Another collection method, the gantry effectively mimics the camera array so long	15

Contents

Introduction

Light field rendering is a fascinating image-based rendering (IBR) technique whose properties have only begun to be explored. With the seminal paper on the topic published in 1996 [17], research and applications have been rather limited, with a majority of research being directed towards virtual reality (VR) in the entertainment industry. There are a number of reasons for the limited research directions and abundance which will be covered here as well as the positive aspects and advantages of light field IBR.

The primary technique from which light field representations are derived is image-based rendering. It is described by McMillan and *** as the method of creating a "continuous representation" of the scene or subject being discretely sampled [21].

The sampling that makes up IBR comes from the plenoptic function, a seven-dimensional function that describes the sum of light rays at any point in space.

Current light field technology is built on the parameterization put forward by Levoy in 1996. If we go back just five years to the original paper regarding the plenoptic function, we see a critical element missing from the function: time. There are a number of major limitations associated with light fields. For one, the method for data collection is often bulky or slow and extremely limited in its capabilities (cite the papers). Up until recently, it has required a static scene, but advances in light field cameras and processing power have seen small advancements in light field video (cite). Regardless, this still leaves the user with the first problem. Another issue is the lack of mobility within a scene and capture. With current rendering techniques, the parameterization is adhered to exactly. That is to say, if image capture is done two dimensionally, it must remain so.

Light fields have a number of intriguing qualities that have yet to be exploited. Such abilities include the ability to recreate a background image in the presence of an occluder as scene in figure 1.

We aim to not only improve on the work done to date, but to develop a system that can take time into account for a dynamic light field experience.

Light fields are a rising method for rendering scenes. With the rising popularity of virtual reality, light field technology has the potential to revolutionize the way in which viewers interact with subjects of interest, like sports matches.

WE want to create a system that can capture a light field and effectively turn it back into a 5D problem, where time is one of the specified parameters

"There are many potential formats which could be used for VR video. Light-fields [Levoy and Hanrahan 1996] provide the greatest level of immersion if they can be captured for a suitable volume"



Figure 1: The dense camera array used at Stanford to capture a number of light fields.



Figure 2: Another collection method, the gantry effectively mimics the camera array so long

"...a practical solution for this has not yet been demonstrated"

Literature Review

The light field is, at its base, a derivation of the plenoptic function (the root *plenus* meaning complete, and optic). The plenoptic function, the concept for which was set forward by Adelson and Bergen in 1991, is a method of mathematically describing the elements of vision. The paper takes on the daunting task of defining what can be seen across 3D space, where every point (or pencil of rays) contains the information for all of the rays passing through it. This is illustrated well in figure 3. The concept of all light passing through a point is rooted in the

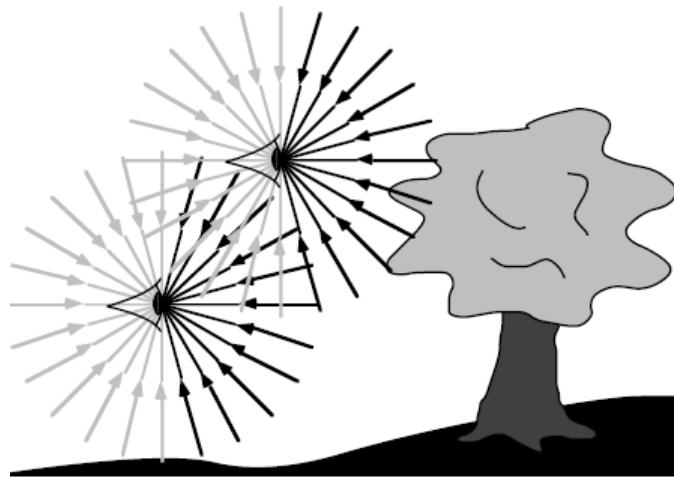


Figure 3: The plenoptic function sampled at two points, showing all pencils of light rays which pass through those points [1].

early observations of Leonardo da Vinci, who observes the effect of a hole in a wall, which reveals an image of the scene in front of the wall, upside down on the back wall [1]. This is the concept behind the pinhole camera, which is a key assumption applied to cameras in much of the work done with light fields. The next section reviews the derivation of the 4D light field from the 7D plenoptic function.

Early Research and Derivation from Plenoptic Function

Having a basic idea of what constitutes light in space, Adelson and Bergen set forth to mathematically represent this plenoptic function. The function itself

is given as

$$P = P(\theta, \phi, \lambda, t, V_x, V_y, V_z) \quad (1)$$

where θ and ϕ are the viewing angles from a point in space, λ ***mention radiance*** is wavelength of the incoming ray, t is the time parameter at which the ray is observed, and V_x, V_y, V_z define the location in three dimensional space [1]. This defines the incoming rays at every point in 3D space for all time, which would theoretically allow for the reconstruction of any view. Measuring the plenoptic function comes with its own set of difficulties though.

In [21] and [8], the 7D plenoptic function is reduced to five dimensions,

$$p = P(\theta, \phi, V_x, V_y, V_z) \quad (2)$$

where t and λ are omitted as the time of sampling is considered to be constant (therefore requiring the scene remain static) and the radiance is*** (Look in Levoy06a).

In Levoy and Hanrahan's seminal paper *Light Field Rendering* [17], the authors note that in space free of occlusions, the function can be further reduced to a four dimensional function as the ray of light reaching the pixel will not be altered in free space. In this environment, the only requirement is to know the direction of a ray, which is a function of four points shown in the following equation:

$$r = L(u, v, s, t) \quad (3)$$

where r is the individual ray passing through the planes. This parameterization requires a ray to pass through two known planes, defined as the s, t and u, v planes. This is still a standard parameterization, used in [10], [29], and [24], among others. In one case, a three dimensional parameterization is used as acquisition is completed by using a 1D row of cameras [15]. Levoy and Hanrahan introduce the two plane parameterization as the light field [17].

While [17] introduces the popular two plane parameterization, McMillan and Bishop define a cylindrical parametrization a year earlier. Having encountered difficulties storing a spherical parameterization on a computer, the authors note that with a cylinder, it can effectively be unrolled into a plane, thus making it easy to store and access [21].

For non-synthetic rendering systems.

Rendering may be viewed as a reconstruction of the light field forming an estimate of the perceived image around the objects (Kriging Paper)

In *Light Fields and Computational Imaging*, Levoy outlines the available parameterizations as of 2006, shown in figure 4

With Isaksen *et al's Dynamically Reparameterized Light Fields*
LIGHT FIELD EFFECTS Learning Based View Synthesis

Light Field Collection

While light field collection has gone relatively unchanged through the years, there are some interested new approaches to collection, but they are still under-developed at this stage.

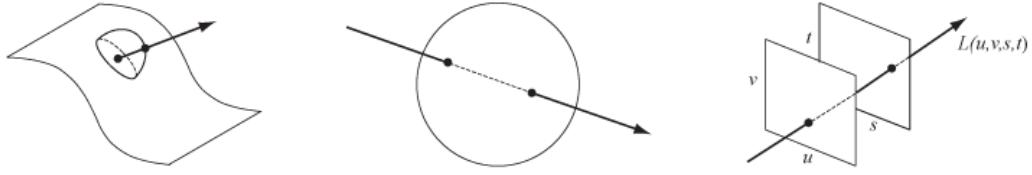


Figure 4: The plenoptic function at two points, showing all pencils of light rays [1]

The primary method by which light fields are captured is a brute force method relying on a large array of cameras [23]**** or a robotic gantry set to move to specific locations, thereby mimicing the camera array. These can be done as the parameterization used ignores the time component of the 7D plenoptic function. While a number of groups have shown interesting use cases for the camera array (some of which will be covered in later sections) such as [23]*****¹, it is not a practical method for capturing difficult scenes.

Utilizing the growing capabilities of image sensors, Ng shows in his thesis a way by which a light field can be captured by a single camera. Rather than producing a single image, a microlens array is inserted into the camera, after the lens and before the image sensor (see figure). This enables the user to create such effects as refocusing and some minimal movement around the image for a 3D effect after the image has been captured [23]. The light field camera still suffers from a lack of **spatial and angular resolution, leaving room for research to be done combining the light field camera and novel collection techniques.

Another unique data collection method was demonstrated in Davis *et al's* paper, *Unstructured Light Fields* from 2012 [5]. In it, Davis

Minimizing the amount required to collect(cite review and actual article) Move on to the more recent papers of Sparse lightfields and learning based view synthesis

Unstructured light fields paper also shows a method for collection utilizing a single camera and a SLAM algorithm known as PTAM(cite PTAM paper). Of course, this research is now a bit aged and can be improved in a number of ways, potentially expanding it significantly with the addition of a more robust and wide-ranging(?) SLAM method such as ORB-SLAM2(cite)

[23] [5] [24]

While light field rendering is the most immersive image-based rendering technique, one of the drawbacks of the light field is that the number of images required to create a high resolution and immersive scene can be prohibitively large [2]. This leads to the natural question of the amount of sampling required for a complete or usable light field. There's surprisingly limited number of papers on the subject, which likely has to do with the methods for data collection, and thus it is one of the less-studied areas within the light field field.

Images Required

[18] cites several papers which mention the requirements for obtaining a full light field using pre-2006 techniques.

Chai *et al* sets about to describe a formula which described the distance between sampling cameras [4].

[10] mentions it how?

In Gortler***'s paper, *The Lumigraph*, it is noted that fewer images are required if the collector has some knowledge of the geometry of the scene [7]. This concept is also made use of in more recent papers, such as [15]

One of the less studied areas is the amount of data required for a quality light field. In [18], Levoy briefly covers the amount of data required for an effective and high quality light field (Insert a quote here).

The subject had been previously covered first in [7] then in [10] and [4] In [26], by analyzing the sparsity in the continuous Fourier domain, show that if light fields are collected to optimize for this, the number of images required to create a light field may be drastically reduced,

View Creation

The most heavily studied area with light fields is on the methods used for creating new images given missing, sparsely sampled light fields or ***

In *Light Field Rendering*, Levoy and Hanrahan outline what would become the baseline for rendering views from a large set of collected images using the parameterization described in the previous section. [14] Viewpoint dependent stereoscopic display [15] scene reconstruction

In *Light Field Reconstruction Using Sparsity in the Continuous Fourier Domain*, the authors show that light fields are sparse in the continuous Fourier domain [26]

In one of the most recent papers on the subject, Kalantari *et al*, rather than examine light fields in the Fourier domain, they train convolutional neural networks on new view synthesis within light field cameras. Their results suggest that their method, at least on the small scale of a light field camera, produces superior images compared with the previous state of the art [13]. Potential future work might be applying this to light fields with lower spatial and angular resolution that might have been collected by robots.

[9] Principles of light field imaging

Application of Light Fields

Graphics and virtual reality are the most obvious applications for light fields and have effectively spurred on the field since the middle of the 1990s. Light fields are ve and realistic.

Riding on the work of Ng and the light field camera, Levoy *et al* struck again with a paper on light field microscopy.[19]

In [29] Vaish *et al* utilize a large camera array to Demonstrate a method by which occlusions can effectively be removed if the light field is sampled densely enough, comparing several different techniques.



Figure 5: The dense camera array used at Stanford to capture a number of light fields.

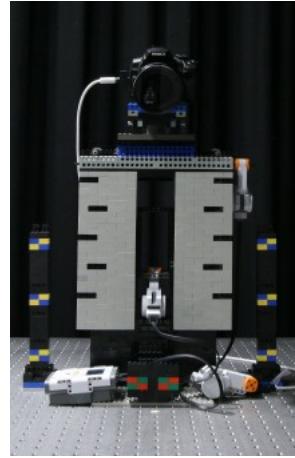


Figure 6: Another collection method, the gantry effectively mimics the camera array so long as the scene is not altered between gantry locations.

Some research has been done in light field videography, but it has been largely untouched. Wilburn *et al*, utilizing a dense camera array, create a method for creating a high framerate system. However, the large camera array comes with its own set of problems in the form of.. It is not until 2016 that Google in a paper by Anderson *et al* demonstrated effective light field video, bringing the generally left out time dimension back into the parameterization [2].

It was Davis *et al* who attempted to bring light fields to a consumer level outside of the light field camera concept put forward by Ng. While the use cases are limited, they went as far as developing an app that allowed users to capture their own light field and extract subjects from it as 3D objects [5]. Using a spherical parameterization, Davis *et al* construct a light field by guiding a user with a graphical user interface that instructs the user on where to hold the camera for data collection.

Working at about the same time as Davis *et al*, Jachnik *et al* create a very similar system. While there are differences in user interface, and algorithmic difference, the concept is quite similar. However, after capturing a light field, the authors attempt to create a crude environment map, and from that the locations of light sources. The results are unique in that for augmented reality applications, shadowing can be implemented on an object that has been placed into a scene, creating a greater sense of realism [11].

Future Collection Methods

With a few exceptions, the direction of light field capture has largely been constant with researchers relying on the large camera array depicted in figure*** While there have been other methods, such as those of Ng, Davis *et al*, and Jachnik *et al*, they are either limited in their spatial and angular representations,

or require significant human interaction to actually create the light field.

[24] demonstrated the ability to capture a light field with a seven degree of freedom robot arm on a Baxter robot. While there is no doubt that this is a definite move from traditional collection methods, there is a vast amount of research yet to be done in this area with wide ranging applications that could bring light fields to industries outside gaming or virtual reality as attempted in [19].

The current trend for light field cameras seems to be going in the direction of multi-camera mobile apparatus, as shown in figure 7. The use cases for these capture methods are limited .



Figure 7: the lytro immerge, designed to capture an entire scene at once, with an array of cameras built into the structure.



Figure 8: On the right is the camera rig behind the Google Jump VR video system [2]

As noted in the introduction, part of the goal of this project is to be able to collect light fields via cameras mounted on one or multiple multirotors, such as the one pictured in ***figure. The amount of research that has gone into multirotor navigation and use as a tool is substantial. A number of those papers were reviewed which could pertain to their use as a light field collection devices, covering topics such as trajectory, localization, and data collection.

In the first, the previously mentioned PTAM system is integrated into a new navigation system, allowing for fairly accurate localization of the multirotor [6].

In another paper, the authors*** were interested in aircraft placement for optimum lighting effects in a scene... [28]

The final paper with In another, a system was developed to define the trajectory of a multirotor... [25]

Advent of the Light Field Camera

This section goes through, in detail, a paper of particular interest written by Ng et al entitled *Light Field Photography with a Hand-held Plenoptic Camera*. In it, the authors outline a method for collecting light fields with a single, modified camera rather than with the bulky and difficult camera array which is also faster than the single camera method employed in papers such as [5].

Progress to Date

A considerable amount of work has been done so far in terms of learning, constructing light field viewers, and scratching at the limit of current state of the art techniques and technologies. This sections covers some of the hurdles in deducing what platforms to use, such as quantifying the accuracy of the systems involved, as well as collection, and finally on to intial rendering of acquired light fields.

Comparison of Platforms

The method for collecting light fields is highly variable and a number of those methods have been explored through experimentation, which is detailed here. In an attempt to acquire light fields with precise position values, a comparison was done between what was at the most available and standard SLAM software, and the baxter robot by Rethink Robotics. This does not account for all of the methods that can or will be used to acquire light fields.

The first method tested was the baxter 7 DOF arm, using the built in camera. A laser was grasped by the robot and directed at a black frame of known dimensions on a white background. The arm was programmed to move in a grid pattern, stopping periodically so that the camera below the laser could take a picture of the scene. The pictures were then analyzed for accuracy against pre-determined points that the laser should have aligned with. The results of these tests can be seen in figures 9 and 10.

SLAM

In the literature review section of this report, of interest is SLAM techniques. This will likely be key in determining the method by which

Collected Light Fields

System Built

The following section describes light field renderers that have been successfully constructed to date. Three different systems have been built and iterated upon, and utilize three distinct methods for storing rays and querying a database for

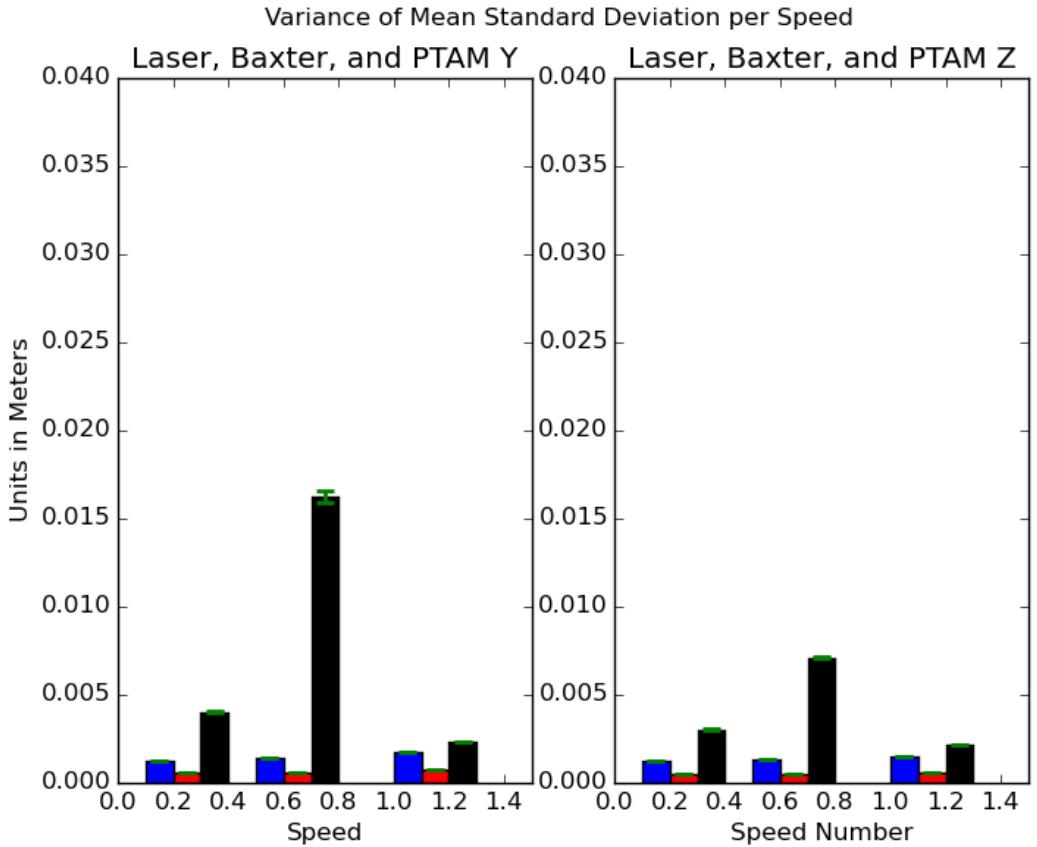


Figure 9: This mosaic is a compilation of all the images taken during one light field acquisition session with the baxter robot.

the appropriate ray parameters. The first system built takes its form from the seminal paper by Levoy and *** [17], the second is based on the work by Isaksen *** [10], and the third is a combination of the work by LEvoy *** and the renderer created by the New Stanford Light Field Archive [31] **cite Stanford archive.

also assumes that the pixel is a sum of the rays at that point. The first renderer utilizes the common, and original two-plane parameterization set forward in [17]. As with all parameterizations, a pinhole model is assumed for the camera being used. A number of known values are required for this particular parameterization to work, such as distance of the subject from the camera. With this knowledge, the *ST* and *UV* planes may be appropriately placed such that light rays from the subject pass through both before entering the camera. Aside from the distance to the subject, the internal physical camera parameters must be known, such as the size of the image sensor, the camera flange focal length, the position of the camera in a global reference frame, and the number of pixels on the image sensor. Knowing these parameters allows for the storage of two points which are defined by the two planes $P = (S, T)$, $Q = (U, V)$, and together give the direction of the ray along with its recorded intensity value. The new ray database can then be queried at any time for a new camera position. With the new position, a k-Nearest Neighbors (KNN) algorithm is used to compare ray directions with

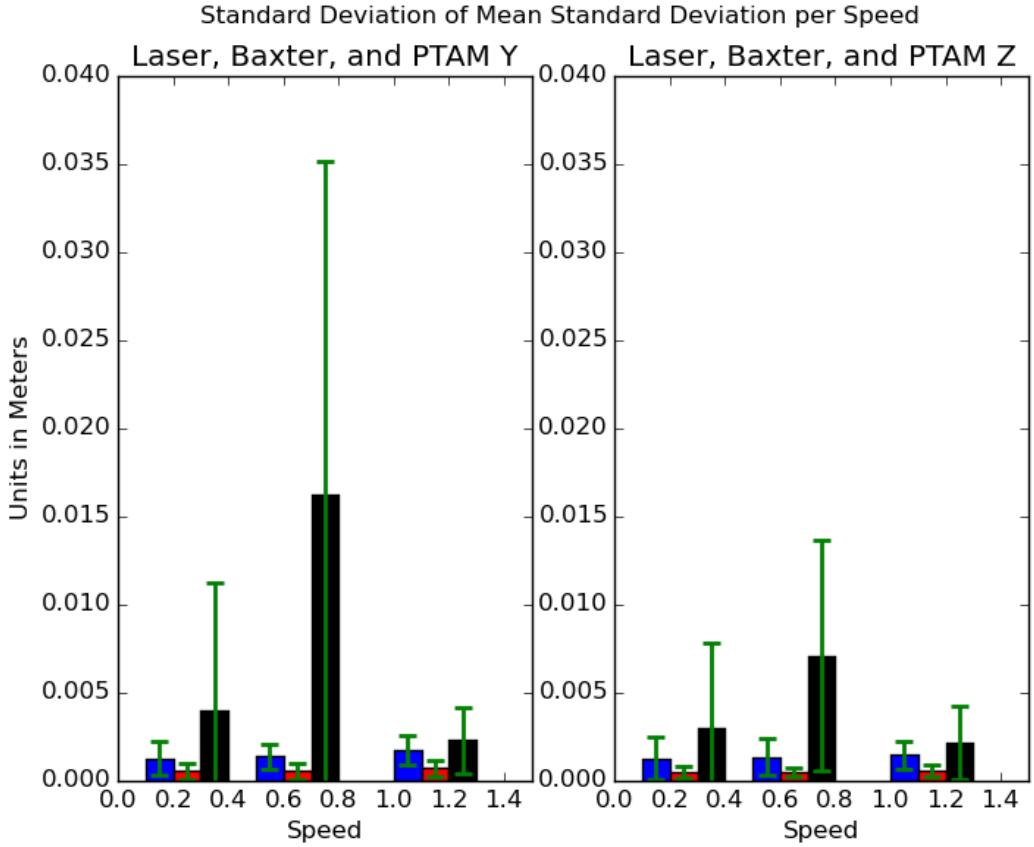


Figure 10: This mosaic is a compilation of all the images taken during one light field acquisition session with the baxter robot.

stored directions and an average is taken of intensity values for the nearest rays to create a new image.

The second renderer built is based on [10]. It is similar to the previous method but moves the planes to the internals of the camera where there are still no occlusions. Regardless of the plane locations, rays from the scene are still calculated and stored the same way, but there is the addition of a mobile "focus plane". This focus plane allows for focus at different points in the scene via finding the rays coming from the object at the plane position as illustrated in figure 12.

A third renderer was built based on the online viewer found on the New Stanford Light Field Archive website [31]. After downloading the ActionScript/MXMLsource code and reverse engineering it, it was rewritten in C++ and OpenCV. It is pictured in figure 13. It is a much simplified version of the first renderer that was programmed, but is considerably faster in its run time. In this case, the images are required to be taken specifically from a single plane as any deviation would lead to ghosting. This is caused by the fact that the current rendering technique that this is based on

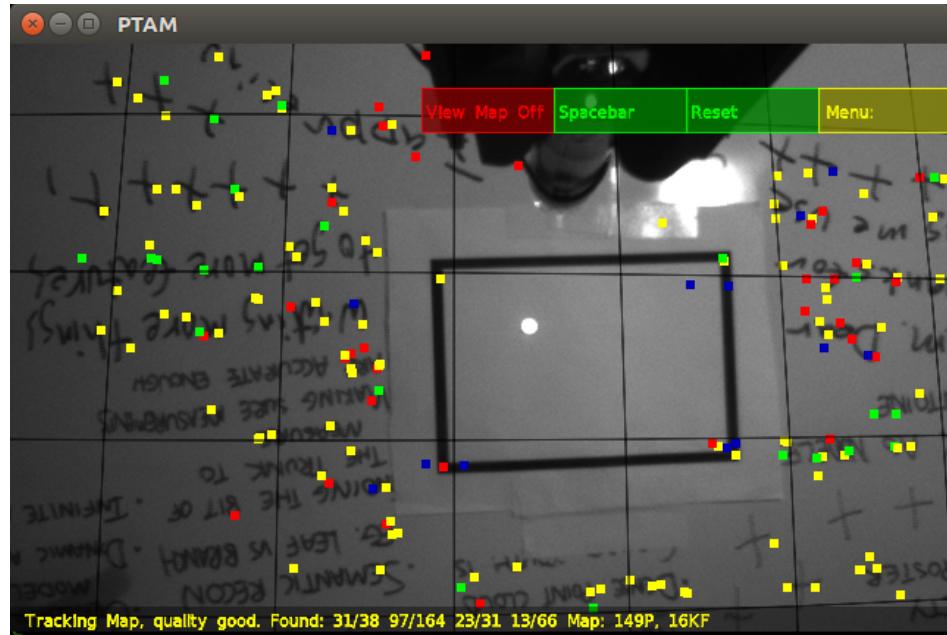


Figure 11: This mosaic is a compilation of all the images taken during one light field acquisition session with the baxter robot.



Figure 12: This mosaic is a compilation of all the images taken during one light field acquisition session with the baxter robot.

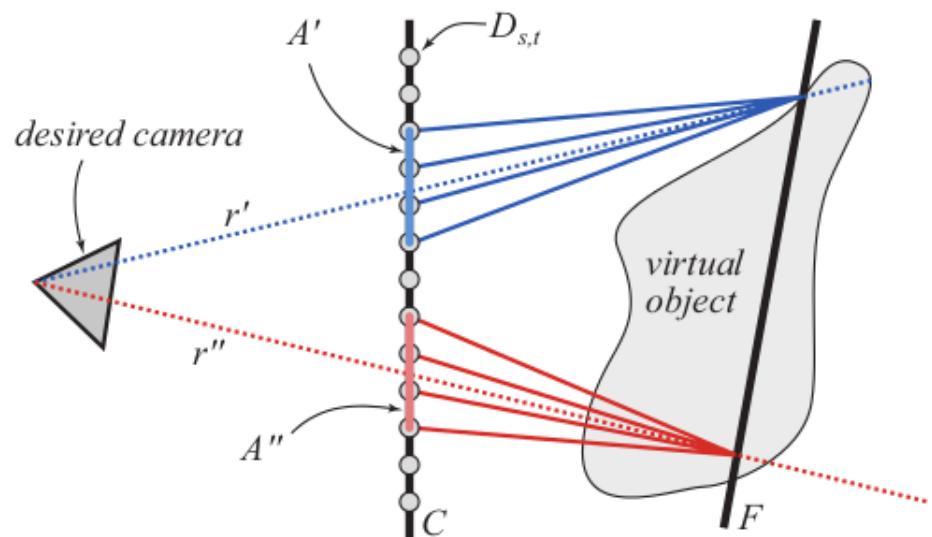


Figure 13: A depiction of the second renderer based on [10] and showing the plane of focus F and the rays coming from it that would be considered in focus.

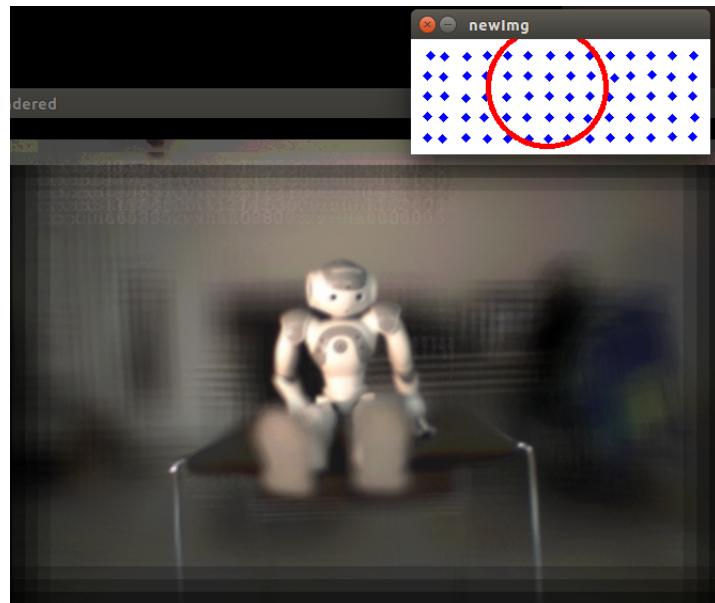


Figure 14: A screenshot of the light field interface in action, with the subaperture views shown in blue, the aperture shown as the red circle, and the resulting image shown in the larger window.

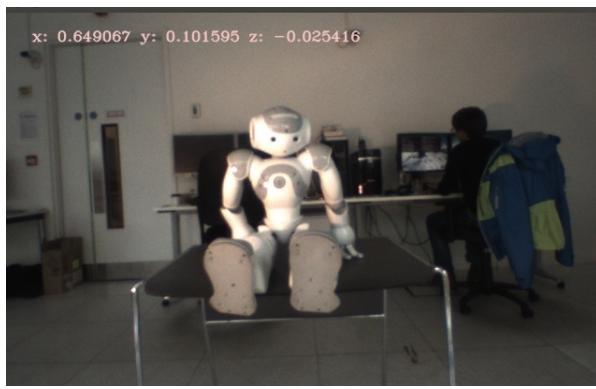


Figure 15: The dense camera array used at Stanford to capture a number of light fields.

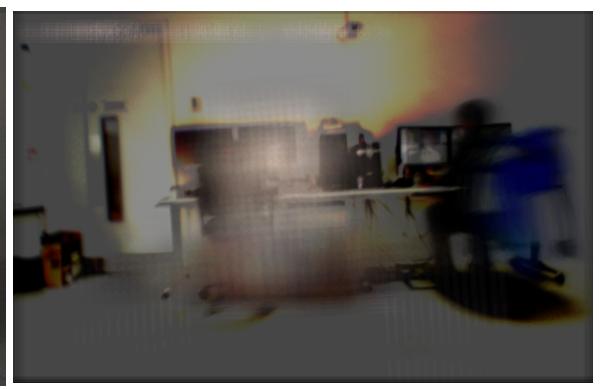


Figure 16: Another collection method, the gantry effectively mimics the camera array so long

Proposed Research

steps

Looking at the work utilizing SLAM in unstructured light fields and improving upon it by initially using the superior ORB-SLAM technique, then further by applying the methodology to

PR2 image

Show ardrone image

While there are existing drones that can easily be controlled from the computer and capture data, there is clear lag high risk of data loss due to the transmission method (wifi). Instead, I would propose to make some simple modifications that would allow data capture and control in real time. Drone Image + hardware

First Step: Unstructured Light Fields with Robots

First to replicate the experiment done with Davis, then with PR2 and Baxter, then on to drones.

Second Step: Capturing Light Fields in Time

The ability to allow a user to move through space and time could have huge implications in the realm of science and industry, allowing scientists to monitor experiments more closely than ever, and industry to monitor apparatus more in depth and with greater detail than before.

Bibliography

- [1] Adelson, Edward H. (1991). *The Plenoptic Function and Elements of Early Vision*
- [2] (2016). *Jump: Virtual Reality Video*
- [3] (1987). *Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion*
- [4] (2000). *Plenoptic Sampling*
- [5] (2012). *Unstructured Light Fields*
- [6] (2012). *Accurate Figure Flying with a Quadrocopter Using Onboard Visual and Intertial Sensing*
- [7] (1996). *The Lumigraph*
- [8] (2014). *Light field modelling and interpolation using Kriging techniques*
- [9] (2016). *Principles of Light Field Imaging*
- [10] (2001?). *Dynamically Reparameterized Light Fields*
- [11] (2013). *Real-Time Surface Light-field Capture for Augmentation of Planar Specular Surfaces*
- [12] (2015). *An Interactive Tool for Designing Quadrotor Camera Shots*
- [13] (2016). *Learning-Based View Synthesis for Light Field Cameras*
- [14] (1995). *A Viewpoint Dependent Stereoscopic Display Using Interpolation of Multi-Viewpoint Images*
- [15] (2013). *Scene Reconstruction from High Spatio-Angular Resolution Light Fields*
- [16] (2007). *Parallel Tracking and Mapping for Small AR Workspaces*
- [17] (1996). *Light Field Rendering*
- [18] (2006). *Light Fields and Computation Imaging*
- [19] (2006). *Light Field Microscopy*

- [20] (2009). *Moving Gradients: A Path-Based Method for Plausible Image Interpolation*
- [21] (1995). *Plenoptic Modelling: An Image-Based Rendering System*
- [22] (2015). *ORB-SLAM: a Versatile and Accurate Monocular SLAM System*
- [23] (2006). *Digital Light Field Photography*
- [24] (2016). *Time-Lapse Light Field Photography With a 7 DoF Arm*
- [25] (2016). *Generating Dynamically Feasible Drone Trajectories for Quadrotor Cameras*
- [26] (2014). *Light Field Reconstruction in the Continuous Fourier Domain*
- [27] (2001). *A Review of Image-based Rendering Techniques*
- [28] (2014). *Computational Rim Illumination with Aerial Robots*
- [29] (2006). *Reconstructing Occluded Surfaces using Synthetic Apertures: Stereo, Focus, and Robust Measures*
- [30] (2005). *High-Speed Videography Using a Dense Camera Array*
- [31] stanford new light field archive <http://lightfield.stanford.edu/>