Light Field Imaging
Plenoptic Function, Light Field Cameras, 3D Displays

Stanford Lightfield Dataset

*slides adapted from Yannis Gkioulekas, Gordon Wetzstein, Fredo Durand, Marc Levoy
Course Projects & Proposal

• **Due Wednesday**: short project proposal (part of your project grade!) = 1-2 pages
  • motivation
  • related work
  • project overview
  • milestones, timeline & goals
  • at least 3 scientific references
  • we may ask you to revise the proposal, will assign a mentor to your team
Poster Session

- Thursday Dec 7th 2-4pm
  - Bahen Atrium
  - you MUST attend in person or make alternative arrangements with me (only in extenuating circumstances)
- You are responsible for making sure that your poster is printed on time!
  - use the “same-day” or ”next-day” ordering option from https://utposter.com/
  - use the offer code on Quercus (“pages”) so you don’t have to pay
  - Pick up poster from 339 Bloor St W
  - bring it to the poster session and put it up before poster session begin
- You have from November 18 until December 6 (the day before the poster session) to print your poster using this service.
Use the offer code on Quercus
Machine learning for determining protein structure and dynamics from cryo-EM images

Ellen Zhong  
Assistant Professor, Princeton University

Wednesday Nov 22, 3-4 PM  
DGP Seminar Room (Bahen 5187)
Computational Imaging for Enabling Vision Beyond Human Perception

Mark Sheinin
Incoming Assistant Professor,
Weizmann Institute of Science

Monday Nov 27, 2-4 PM (lecture 11)
Next Generation Black-Hole Imaging

Aviad Levis
Incoming Assistant Professor,
University of Toronto

Monday Dec 4, 2-4 PM (lecture 12)
Overview of today’s lecture

• Quick reminder: pinhole vs lens cameras.
• Focal stacks
• Lightfield
• Measuring lightfields
• Plenoptic camera
• Images from lightfields
Pinhole vs lens cameras
Pinhole camera
Pinhole camera

- Everything is in focus.
- Very light inefficient.
Lens camera

focus plane
Lens camera

- Only one plane is in focus.
- Very light efficient.
Lens camera

- Only one plane is in focus.
- Very light efficient.

How can we get an all in-focus image?
Focal stack
Focal stack

- Capture images focused at multiple planes.
- Merge them into a single all in-focus image.

Analogous to what we did in HDR
- Focal stack instead of exposure stack.
Focal stack imaging
Focal stack imaging

1. Capture a focal stack

2. Merge into an all in-focus image
Focal stack imaging

1. Capture a focal stack

2. Merge into an all in-focus image
How do you capture a focal stack?

Which of these parameters would you change (and how)?

- lens-sensor distance $D'$
- lens-object distance $D$
- lens focal length $f$
- lens aperture $f/#$
How do you capture a focal stack?

Which of these parameters would you change (and how)?

- lens-object distance $D$
- lens focal length $f$
- lens aperture $f/#$
- rotate lens focus ring (not zoom!)
- lens-sensor distance $D'$
How do you capture a focal stack?

Which of these parameters would you change (and how)?

- lens-sensor distance $D'$
- lens-object distance $D$
- lens focal length $f$
- lens aperture $f/#$
Capturing a focal stack

In-focus plane in each stack image
Focal stack imaging

1. Capture a focal stack

2. Merge into an all in-focus image
How do you merge a focal stack?
How do you merge a focal stack?

1. Align images
2. Assign per-pixel weights representing “in-focus”-ness
3. Compute image average
How do you merge a focal stack?

1. Align images
2. Assign per-pixel weights representing “in-focus”-ness
3. Compute image average
Image alignment

Why do we need to align the images?
Why do we need to align the images?

• When we change focus distance, we also change field of view (magnification).
• Also, scene may not be static (but we will be ignoring this for now).
Image alignment
Image alignment
Why do we need to align the images?
• When we change focus distance, we also change field of view (magnification).
• Also, scene may not be static (but we will be ignoring this for now).

• Assume we know $f$ and all $D'$ values.
• How do can we align the images?
Why do we need to align the images?

- When we change focus distance, we also change field of view (magnification).
- Also, scene may not be static (but we will be ignoring this for now).

\[
m = \frac{f}{D' - f} \quad \frac{1}{D'} + \frac{1}{D} = \frac{1}{f}
\]

Assume we know \( f \) and all \( D' \) values.

- How do can we align the images?
How can we avoid having to do alignment?
Use a telecentric lens

Place a pinhole at focal length, so that only rays parallel to primary ray pass through.

Magnification independent of object depth.

Magnification depends only on sensor-lens distance $S'$. 

[Watanabe and Nayar, PAMI 1997]
How do you merge a focal stack?

1. Align images
2. Assign per-pixel weights representing “in-focus”-ness
3. Compute image average
Weight assignment

How do we measure how much “in-focus” each pixel is?
Weight assignment

How do we measure how much “in-focus” each pixel is?
• Measure local sharpness.

How do we measure local sharpness?
Weight assignment

How do we measure how much “in-focus” each pixel is?
• Measure local sharpness.

run Laplacian operator
do some Gaussian blurring (why?)
Weight assignment

How do we measure how much “in-focus” each pixel is?
• Measure local sharpness.

run Laplacian operator

do some Gaussian blurring (so that nearby pixels have similar weights)

Just one example, many alternatives possible.
How do you merge a focal stack?

1. Align images
2. Assign per-pixel weights representing “in-focus”-ness
3. Compute image average
Focal stack merging

and divide by sum of weights
Some results

element image from stack

all in-focus image
Another example

Useful in macrophotography, where depths of field are very shallow
Another example

middle image from stack

all in-focus image
Another look at the mixing weights

What do the mixing weights look like?
Another look at the mixing weights

Depth from focus = determining sharpest pixel in focal stack
Depth from focus on a mobile phone

Use focal stack from autofocus

[Suwajanakorn et al., CVPR 2015]
Another look at the mixing weights

What is a problem of these depth maps?
Another look at the mixing weights

What is a problem of these depth maps?

- Blurry because we need to process entire neighborhoods to compute sharpness.
- Can we use any extra information to get per-pixel depth?
Confocal stereo
Confocal stereo

- Capture a 2D stack by varying both focus and aperture
- Analyze each pixel’s focus-aperture image to find true depth

[Hassinof and Kutulakos, ECCV 2006]
Aperture-Focus Image of pixel p

( aperture $\alpha_i$, focus $f_j$ )
AFI model fitting

Key idea: When a point is in focus, its color and intensity remain the same for all apertures.

• This property is called confocal constancy.
• We can find correct depth by checking intensity variation across apertures for each focus setting.
Depth inference techniques

Confocal stereo:
• Requires focus-aperture stack.
• Gives per-pixel depth estimates.

Depth from focus:
• Requires focus stack.
• Gives per-patch depth estimates.

What is a downside of these two approaches?
Reminder: Circle of confusion

Size of circle of confusion depends on distance from in-focus plane.

\[
\frac{y}{D/2} = \frac{|O - S|}{O}
\]

\[
\frac{y}{c/2} = \frac{1}{m}
\]

\[
c = mD \frac{|O - S|}{O}
\]
Depth from defocus

Use as few as two images (two-frame depth from defocus):
• Assume circle of confusion can be modeled as (typically Gaussian) blur kernel with varying $\sigma$.
• Use pair of images to estimate how blur size at each pixel changes from one depth to another.
• Relate this blur size to depth.
Requires very elaborate modeling and priors to be robust.

[Tang et al., CVPR 2017]
Depth inference techniques

Confocal stereo:
• Requires focus-aperture stack.
• Gives per-pixel depth estimates.

Depth from focus:
• Requires focus stack.
• Gives per-patch depth estimates.

Depth from defocus:
• Requires only two images.
• Gives per-patch depth estimates.

What is a downside of these three approaches?
Focal flow

Use a dense focal stack (e.g., from autofocus):

- Assume circle of confusion can be modeled as (typically Gaussian) blur kernel with varying $\sigma$.
- Estimate optical flow between successive frames in focal stack.
- Relate optical flow change to depth.

Requires very little computation (no convolutions) but only works on textured patches (why?).

\[
\begin{bmatrix}
I_x & I_y & (xI_x + yI_y) & (I_{xx} + I_{yy})
\end{bmatrix} \begin{bmatrix}
\bar{u}_4 + I_t
\end{bmatrix} = 0
\]

[Alexander et al., ECCV 2016; Guo et al., ICCV 2017]
Use a dense focal stack (e.g., from autofocus):
• Assume circle of confusion can be modeled as (typically Gaussian) blur kernel with varying $\sigma$.
• Estimate optical flow between successive frames in focal stack.
• Relate optical flow change to depth.
Requires very little computation (no convolutions) but only works on textured patches (why?).

Biologically-inspired from jumping spiders!

[Alexander et al., ECCV 2016; Guo et al., ICCV 2017, PNAS 2019]
Focal flow

Use a dense focal stack (e.g., from autofocus):
- Assume circle of confusion can be modeled as (typically Gaussian) blur kernel with varying $\sigma$.
- Estimate optical flow between successive frames in focal stack.
- Relate optical flow change to depth.

Requires very little computation (no convolutions) but only works on textured patches (why?).

[Alexander et al., ECCV 2016; Guo et al., ICCV 2017, PNAS 2019]
Depth inference techniques

Confocal stereo:
• Requires focus-aperture stack.
• Gives per-pixel depth estimates.

Depth from focus:
• Requires focus stack.
• Gives per-patch depth estimates.

Focal flow:
• Requires dense focus stack.
• Efficient per-patch depth estimates.

Depth from defocus:
• Requires only two images.
• Gives per-patch depth estimates.
Lightfield
A lens measures all rays radiated from the object (up to aperture size).
Measuring rays

We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
Measuring rays

We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
Measuring rays

We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
Measuring rays

We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
We can capture the same set of rays by using a pinhole camera from multiple viewpoints. How would you merge these images into a lens-based, defocused image?
What is the dimension of the lightfield?

Parameterize every ray based on its intersections with two planes.
Lightfield: all rays in a scene

Reference plane \((s, t)\)  Aperture plane \((u, v)\)

4-dimensional function \(L(u, v, s, t)\)

Parameterize every ray based on its intersections with two planes.

- How else can we parameterize the lightfield?
Parameterize every ray based on its intersections with two planes.
Lightfield slices

What does $L(u = u_o, v = v_o, s, t)$ look like?

4-dimensional function $L(u, v, s, t)$
(conjugate of scene-based function)
Lightfield slices

Sensor plane \((s, t)\)

Aperture plane \((u, v)\)

What does \(L(u = u_0, v = v_0, s, t)\) look like?

A pinhole image from a certain viewpoint

4-dimensional function \(L(u, v, s, t)\)

(conjugate of scene-based function)
Lightfield slices

reference plane \((s, t)\) \hspace{2cm} aperture plane \((u, v)\) \hspace{2cm} sensor plane \((s, t)\)

aperture coordinates \(u = u_0, v = v_0\)

Lightfield slice \(L(u = u_0, v = v_0, s, t)\)
Lightfield slices

What does $L(u = u_0, v = v_0, s, t)$ look like?
A pinhole image from a certain viewpoint

What does $L(u, v, s = s_0, t = t_0)$ look like?

4-dimensional function $L(u, v, s, t)$
(conjugate of scene-based function)
What does $L(u = u_0, v = v_0, s, t)$ look like?

• a pinhole image from a certain viewpoint

What does $L(u, v, s = s_0, t = t_0)$ look like?

• radiance emitted by a certain (in-focus) point at various directions
Lightfield slices

reference plane \((s, t)\)    aperture plane \((u, v)\)    sensor plane \((s, t)\)

Reference/sens or coordinates

\(s = s_0, t = t_0\)

Lightfield slice \(L(u, v, s = s_0, t = t_0)\)
Lightfield slices

What does $L(u = u_o, v = v_o, s, t)$ look like?
- a pinhole image from a certain viewpoint

What does $L(u, v, s = s_o, t = t_o)$ look like?
- radiance emitted by a certain (in-focus) point at various directions

4-dimensional function $L(u, v, s, t)$
(conjugate of scene-based function)
Lightfield visualization

$L(u, v, s = s_o, t = t_o)$ is the radiance emitted by a certain (in-focus) point at various directions.

$L(u = u_o, v = v_o, s, t)$ is a pinhole image from a certain viewpoint.

Demo: [http://lightfield.stanford.edu/lfs.html](http://lightfield.stanford.edu/lfs.html)
Same LF, Resort Pixels
4D Light Field
This slide has a 16:9 media window
viewer moves right
global 4D Light Field
viewer moves down
1D Scanline

- View moves left
- View moves right
- Center view
- Left view
- Right view
Object Depth

local slope corresponds to depth!

View moves left
Occlusions
Specular Reflections

view moves left
How can you capture the lightfield of a scene?
Measuring lightfields
How to capture a lightfield?

How can you do this?
Option 1: use multiple cameras

Stanford camera array

(“synthetic aperture”)

[Willburn et al., SIGGRAPH 2005]
Option 1: use multiple cameras

Stanford camera array

What kind of lens would you use for this?

(“synthetic aperture”)

[Willburn et al., SIGGRAPH 2005]
Compound Eye vs Light Field

Image: National Geographic

Mosquito eye, image: Raija Peura, University of Oulu
Option 2: take multiple images with one camera

Single camera mounted on LEGO motor.
Sequential

- early explorations in the context of image based rendering
- Digital Michaelangelo Project: Michaelangelo’s’s statue of Night

[Levoy et al. 2000]
Sequential

- programmable aperture photography

[Liang et al. 2008]
Plenoptic camera
Option 3: use a plenoptic camera

plenoptic = plenus (Latin for “full”) + optic (Greek for “seeing”, in this case)
Making a plenoptic camera

reference plane \((s, t)\)
aperture plane \((u, v)\)
sensor plane \((s, t)\)

Lightfield slice \(L(u, v, s = s_o, t = t_o)\)

reference/sensor coordinates \(s = s_o, t = t_o\)
Making a plenoptic camera

reference plane \((s, t)\)  
aperture plane \((u, v)\)  
sensor plane \((s, t)\)

reference/sensor coordinates  
\(s = s_0, t = t_0\)

Lightfield slice \(L(u, v, s = s_0, t = t_0)\)

each pixel measures ray  
\(L(u, v, s = s_0, t = t_0)\) for different \((u, v)\)
Making a plenoptic camera

reference plane \((s, t)\)  
aperture plane \((u, v)\)  
sensor plane \((s, t)\)

Each pixel measures a ray \(L(u, v, s, t)\)

Each pinhole corresponds to a slice \(L(u, v, s = s_o, t = t_o)\)

Lightfield \(L(u, v, s, t)\)  
How can we make this more light efficient?
Making a plenoptic camera

reference plane \((s, t)\)  
aperture plane \((u, v)\)  
sensor plane \((s, t)\)

Each pinhole corresponds to a slice \(L(u, v, s = s_o, t = t_o)\)

Lightfield \(L(u, v, s, t)\)

How can we make this more light efficient?  
- Replace pinholes with lenslets
History of the plenoptic camera

First conceptualized by Gabriel Lippmann, who called it integral photography.

- Original article appeared in French journal three years earlier, what is shown here is an American re-print.
History of the plenoptic camera

First conceptualized by Gabriel Lippmann, who called it integral photography.
Reappeared under different forms and names throughout the century.

- The left paper is from 1930, the right one from 1970.
History of the plenoptic camera

Single Lens Stereo with a Plenoptic Camera

Edward H. Adelson and John Y.A. Wang

Abstract—Ordinary cameras gather light across the area of their lens aperture, and the light striking a given subregion of the aperture is structured somewhat differently than the light striking an adjacent subregion. By analyzing this optical structure, one can infer the depths of objects in the scene, i.e., one can achieve “single lens stereo.” We describe a novel camera for performing this analysis. It incorporates a single main lens along with a lenticular array placed at the sensor plane. The resulting “plenoptic camera” provides information about how the scene would look when viewed from a continuum of possible viewpoints bounded by the main lens aperture. Deriving depth information is simpler than in a binocular stereo system because the correspondence problem is minimized. The camera extracts information about both horizontal and vertical parallax, which improves the reliability of the depth estimates.

I. INTRODUCTION

"EVERY BODY in the light and shade fills the surrounding air with infinite images of itself; and these, by infinite pyramids diffused in the air, represent this body throughout space and on every side.” Leonardo da Vinci [11] was

Fig. 1. Diagram from Leonardo’s notebooks illustrating the fact that the light rays leaving an object’s surface may be considered to form a collection of cones (which Leonardo calls “pyramids”), each cone constituting an image that would be seen by a pinhole camera at a given location.
History of the plenoptic camera

Figure from the 1992 paper, which shows the version of the plenoptic camera with pinholes instead of lenslets on the image plane.

Fig. 5. Array of miniature pinhole cameras placed at the image plane can be used to analyze the structure of the light striking each macropixel.
History of the plenoptic camera

Light Field Photography with a Hand-held Plenoptic Camera

Ren Ng*  Marc Levoy*  Mathieu Brédif*  Gene Duval†  Mark Horowitz*  Pat Hanrahan*
*Stanford University  †Duval Design

Abstract

This paper presents a camera that samples the 4D light field on its sensor in a single photographic exposure. This is achieved by inserting a microlens array between the sensor and main lens, creating a plenoptic camera. Each microlens measures not just the total amount of light deposited at that location, but how much light arrives along each ray. By re-sorting the measured rays of light to where they would have terminated in slightly different, synthetic cameras, we can compute sharp photographs focused at different depths. We show that a linear increase in the resolution of images under each microlens results in a linear increase in the sharpness of the refocused photographs. This property allows us to extend the depth of field of the camera without reducing the aperture, enabling shorter exposures and lower image noise. Especially in the macrophotography regime, we demonstrate that we can also compute synthetic photographs from a range of different viewpoints. These capabilities argue for a different strategy in designing photographic imaging systems.

To the photographer, the plenoptic camera operates exactly like an ordinary hand-held camera. We have used our prototype to take hundreds of light field photographs, and we present examples of portraits, high-speed action and macro close-ups.

Keywords: Digital photography, light field, microlens array, synthetic photography, refocusing.

Externally, our hand-held light field camera looks and operates exactly like a conventional camera: the viewfinder, focusing mechanism, length of exposure, etc. are identical. Internally, we augment the 2D photosensor by placing a microlens array in front of it, as proposed by Adelson and Wang [1992] in their work on the “plenoptic camera” (They did not build this device, but prototyped a non-portable version containing a relay lens.) Each microlens forms a tiny sharp image of the lens aperture, measuring the directional distribution of light at that microlens.

This paper explains the optical recipe of this camera in detail, and develops its theory of operation. We describe an implementation using a medium format digital camera and microlens array. Using this prototype, we have performed resolution experiments that corroborate the limits of refocusing predicted by the theory. Finally, we demonstrate examples of refocusing and view-point manipulation involving close-up macro subjects, human portraits, and high-speed action.

2 Related Work

The optical design of our camera is very similar to that of Adelson and Wang’s plenoptic camera [1992]. Compared to Adelson and Wang, our prototype contains two fewer lenses, which significantly shortens the optical path, resulting in a portable camera. These differences are explained in more detail Section 3.1 once sufficient technical background has been introduced. The other main differ-
Prototype plenoptic camera

- predecessor of Lytro
- resolution: 292x292px, 14x14 light field views

Kodak 16-megapixel sensor

125μ square-sided microlenses

[Ng et al., Stanford Technical Report 2005]
Commercial plenoptic camera

**Lens**
The Lytro Light Field Camera starts with an 8X optical zoom, f/2 aperture lens. The aperture is constant across the zoom range allowing for unheard of light capture.

**Light Field Engine 1.0**
The Light Field Engine replaces the supercomputer from the lab and processes the light ray data captured by the sensor.

The Light Field Engine travels with every living picture as it is shared, letting you refocus pictures right on the camera, on your desktop and online.

**Light Field Sensor**
From a roomful of cameras to a micro-lens array spacially adhered to a standard sensor, the Lytro’s Light Field Sensor captures 11 million light rays.
Commercial plenoptic camera

Lens
The Lytro Light Field Camera starts with an 8X optical zoom, f/2 aperture lens. The aperture is constant across the zoom range allowing for unheard of light capture.

Light Field Engine 1.0
The Light Field Engine replaces the supercomputer from the lab and processes the light ray data captured by the sensor.

Light Field Sensor
From a roomful of cameras to a micro-lens array specifically adhered to a standard sensor, the Lytro’s Light Field Sensor captures 11 million light rays.
Commercial plenoptic camera

newer version with higher resolution

Lytro Illium
Lytro Cinema Camera

- 755 MP Camera
- 1 s of footage needed 400 GB
- Needed to be moved by crane

- Commercially unsuccessful
- Lytro eventually bought out by Google
Industrial plenoptic cameras

Plenoptic cameras have become quite popular in lab and industrial settings.

- Much higher resolution, both spatial and angular, than commercial cameras.
- Support interchangeable lenses.
- Can do video.
- Very expensive.
Shack-Hartmann wavefront sensors

• Completely different use: measuring how close a wavefront is to being planar
• Exactly the same optics as a plenoptic camera.
• Common instrument in optics labs and medical imaging.
Given that the plenoptic camera design was known since 1908, why did it take a century for such a camera to be made commercially available?
Question

Given that the plenoptic camera design was known since 1908, why did it take a century for such a camera to be made commercially available?

- Difficult to manufacture good lenslet arrays at high resolution (unless attached to sensors).
- Digital sensors did not have sufficiently high resolution.
- No known good use for them besides depth sensing.
  - The introduction of the handheld plenoptic camera was strongly influenced by the 1996 lightfield papers, which demonstrated all of the useful photographic operations one could replicate by having access to the entire lightfield.
- And of course, nobody thought of commercializing them.
Images from lightfields
A plenoptic “image”

What are these circles?
A plenoptic camera

reference plane \((s, t)\)  
aperture plane \((u, v)\)  
sensor plane \((s, t)\)

Lightfield \(L(u, v, s, t)\)

Each lenslet corresponds to a slice \(L(u, v, s = s_o, t = t_o)\)
The plenoptic image

Which coordinates do I change when I move from one circle to another?

Which coordinates do I change when I move within each circle?
The plenoptic image

Which coordinates do I change when I move from one circle to another?
• I change s, t (sensor plane) coordinates.

Which coordinates do I change when I move within each circle?
• I change u, v (aperture plane) coordinates.
Simulate different viewpoints?
Simulate different viewpoints?

• Pick same pixel within each lenslet view
Changing the viewpoint

Viewpoint change is limited by the aperture of each of the lenslets.
How do I…

Simulate different viewpoints?
• Pick same pixel within each lenslet view

Simulate different aperture sizes?
Simulate different viewpoints?
• Pick same pixel within each lenslet view

Simulate different aperture sizes?
• Sum more than one pixels within each lenslet view
How do I...

Simulate different viewpoints?
- Pick same pixel within each lenslet view

Simulate different aperture sizes?
- Sum more than one pixels within each lenslet view

Simulate lens at current focus setting?
How do I...

Simulate different viewpoints?
• Pick same pixel within each lenslet view

Simulate different aperture sizes?
• Sum more than one pixels within each lenslet view

Simulate lens at current focus setting?
• Same as above. Sum all pixels for max aperture setting.
How do I... Change the focus setting?
We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
Measuring rays

We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
Measuring rays

We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
Measuring rays

We can capture the same set of rays by using a pinhole camera from multiple viewpoints.
We can capture the same set of rays by using a pinhole camera from multiple viewpoints
• How would you merge these images into a lens, defocused image?
Measuring rays

Sum all pixels in each lenslet view.
Form lens image

reference plane (s, t)  aperture plane (u, v)  sensor plane (s, t)

How do I refocus?

Sum all pixels in each lenslet view.
Form lens image

reference plane (s, t) aperture plane (u, v) sensor plane (s, t)

How do I refocus?
• Need to move sensor plane to a different location.
Understanding Refocus

- consider light field inside camera
- synthesize image on sensor \( i_{d=0}(x) = \int_\Omega l(x, \nu) d\nu \)

\[
i_d(x) = \int_\Omega l(x + d\nu, \nu) d\nu
\]
Understanding Refocus - Parameterization

[Ng 2006]
Understanding Refocus – Integration

[Ng 2006]
Understanding Refocus – Shift+Add

- transport in free space + integrate (shift + add)
- equal to projection of light field

\[
i_0(x) = \int_{\Omega} l(x,\nu) d\nu
\]
\[
i_{d_1}(x) = \int_{\Omega} l(x + d_1\nu,\nu) d\nu
\]
\[
i_{d_2}(x) = \int_{\Omega} l(x + d_2\nu,\nu) d\nu
\]
Understanding Refocus – Shift+Add

- transport in free space + integrate (shift + add)
- equal to projection of light field

[Ng 2006]
Understanding Refocus – Fourier Slicing

- Fourier slice theorem: projection in primal is slicing in Fourier space

\[ i_d(x) = \int_{\Omega} l(x + dv, v) dv \quad \leftrightarrow \quad \hat{i}_d(f_x) = \hat{l}(f_x, f_v - df_x) \]
Equations:

\[ u_5 = -2 \]
\[ u_4 = -1 \]
\[ u_3 = 0 \]
\[ u_2 = 1 \]
\[ u_1 = 2 \]
at least 3 objects: front, center, and far
slide phone and take photos
keep exposure & focus as steady as possible
avoid phone tilt (e.g. attach it to a bigger object)
mark equidistant locations
focal plane
focal plane

- reference (rectified light field): shift so that center object aligns!
\begin{align*}
    u_5 &= -2 \\
    u_4 &= -1 \\
    u_3 &= 0 \\
    u_2 &= 1 \\
    u_1 &= 2
\end{align*}

\begin{align*}
    i_0(x) &= \int_{\Omega} l(x, u) du \\
    i_{d_1}(x) &= \int_{\Omega} l(x + d_1 u, u) du \\
    i_{d_2}(x) &= \int_{\Omega} l(x + d_2 u, u) du
\end{align*}

- reference (rectified light field): shift so that center object aligns!
average = synthetic aperture, focus on center
• **shift+add for refocus!**
focal plane

- shift+add for refocus!
• shift+add for refocus!
\[ u_5 = -2 \quad u_4 = -1 \quad u_3 = 0 \quad u_2 = 1 \quad u_1 = 2 \]

\[
\begin{align*}
i_0(x) &= \int_{\Omega} l(x,u) \, du \\
i_{d_1}(x) &= \int_{\Omega} l(x + d_1 u, u) \, du \\
i_{d_2}(x) &= \int_{\Omega} l(x + d_2 u, u) \, du
\end{align*}
\]

- reference (rectified light field): shift so that front object aligns!
average = synthetic aperture, focus on front
focal plane

- shift+add for refocus!
• shift so that back object aligns!
$u_5 = -2 \quad u_4 = -1 \quad u_3 = 0 \quad u_2 = 1 \quad u_1 = 2$

\[ i_0(x) = \int_\Omega l(x,u)du \]
\[ i_{d1}(x) = \int_\Omega l(x + d_1u,u)du \]
\[ i_{d2}(x) = \int_\Omega l(x + d_2u,u)du \]

• reference (rectified light field): shift so that center object aligns!
average = synthetic aperture, focus on rear
Refocusing example
Refocusing example
Refocusing example
Light Fields with Camera Arrays

- Stanford Multi-Camera Array

Three ways to measure a lightfield

1) Use a plenoptic camera
2) Use a camera array
3) Use one camera multiple times

What are the pros and cons of each?
Three ways to measure a lightfield

1) Use a plenoptic camera
   pro: inexpensive, fast
   con: low-resolution

2) Use a camera array
   pro: high-resolution, fast, large baseline
   con: expensive, difficult to calibrate, bulky

3) Use one camera multiple times
   pro: inexpensive, high-res
   con: slow, no dynamic scenes
Overview of Other Light Field Applications
Depth Extraction

- analyze epipolar lines: slope = depth (a little more complicated…)

[Wanner & Goldluecke 2012]
Depth Extraction

[Wanner & Goldluecke 2012]

center view  depth from stereo  depth from light field
Computer Vision Applications

- 4D feature detection
- segmentation
- tracking
- object removal and inpainting
- augmented reality
- image and video understanding
- usually analysis of 4D structure …
Glare Reduction

- glare comes from inter-reflection in lens elements

(original photo)

[Raskar et al. 2008]
Glare Reduction

- glare appears as high-frequency “outliers” (highly view-dependent effect) in 4D light field
- apply 4D filtering to reduce

[Raskar et al. 2008]
Glare Reduction

- glare comes from inter-reflection in lens elements

[Raskar et al. 2008]
Aberration Correction

- digitally correct for optical aberrations by distorting the light field

[Ng and Hanrahan 2006]
Light Field Microscopy

• can do refocus, but more interesting: instantaneous 3D volume (for fluorescence)!
• diffraction becomes an issue

[Levoy et al. 2006]
Next: Time-of-flight imaging

- Lidar
- Indirect TOF
- NLOS Imaging
- Imaging through scattering media
References

Basic reading:
• Szeliski textbook, Section 12.1.3, 13.3.

Additional reading:
  One of the (relatively) early papers on depth from defocus.
• Suwajanakorn et al., “Depth from Focus with Your Mobile Phone,” CVPR 2015.
  Implementing depth from defocus on a mobile phone using the autofocus focal stack.
  The paper on high resolution depth from a focus and aperture stack.
  Continuously change focus within one exposure, without stopping to capture a stack.
• Levoy and Hanrahan, “Light Field Rendering,” SIGGRAPH 1996.
  The two papers introducing the light field.
  The paper (re)-introducing the plenoptic camera to computer vision and graphics.
• Ng et al., “Light field photography with a hand-held plenoptic camera,” Stanford TR 2005.
  The paper (re)-(re)-introducing the plenoptic camera, and the precursor to Lytro.
• Ng, “Fourier Slice Photography,” SIGGRAPH 2005.
  The paper on frequency-space analysis of refocusing and lightfield measurements.
  The camera array paper.
  Make a lightfield camera from a pinhole array (and many other interesting stuff about lightfield cameras).