Mask-based Light Field Capture and Display

Douglas Lanman
MIT Media Lab

Machine Learning Meets Computational Photography
NIPS 2010 Workshops, Whistler
11 December 2010
Achieving three-dimensional (3D) photography
• Modify cameras to record perceptual depth cues
• Modify displays to reconstruct perceptual depth cues
Achieving three-dimensional (3D) photography
• Modify cameras to record perceptual depth cues
• Modify displays to reconstruct perceptual depth cues
• Achieve with compact cameras and without encumbering viewers
Representing Light Transport

- Assume geometrical (ray) optics approximation
- Parameterize radiance of 3D rays using a 4D light field $L(u,v,s,t)$

Marc Levoy and Pat Hanrahan. Light Field Rendering. 1996.
Steven Gortler et al. The Lumigraph. 1996.
Representing Occlusion

- Represent attenuation using 4D *shield field* $M(u,v,s,t)$
- Assumes no reflection, refraction, or scattering due to occluder

$$L_{received}(u,v,s,t) = M(u,v,s,t)L_{incident}(u,v,s,t)$$
Representing Occlusion

- Represent attenuation using 4D shield field $M(u,v,s,t)$
- Occluder modeled by convolution in frequency domain
Conventional vs. Light Field Cameras

Conventional photography loses 3D scene structure
- Consider two-plane light field parameterization
- Variation of incident intensity (over aperture) is averaged at sensor

Light field photography preserves 3D scene information
- Achieved with pinhole/lenslet array placed close to the sensor
- Trades decreased spatial resolution for increased angular resolution

Frederic Ives. Parallax Stereogram and Process of Making the Same. 1903.
Gabriel Lippmann. Épreuves Réversibles Donnant la Sensation du Relief. 1908.
Conventional displays are view-independent
- Consider two-plane light field parameterization
- Pixel intensity/color does not vary as a function of viewing angle

Light field displays are view-dependent (automultiscopic)
- Achieved with pinhole/lenslet array placed close to the display
- Trades *decreased* spatial resolution for *increased* angular resolution

Frederic Ives. Parallax Stereogram and Process of Making the Same. 1903.
Gabriel Lippmann. Épreuves Réversibles Donnant la Sensation du Relief. 1908.
Previous Approaches

Achieving 3D Photography

- Two alternatives for spatially-multiplexed light field capture:
  - Attenuating masks using *parallax barriers* [Ives, 1903]
  - Refracting *lenslet arrays* [Lippmann, 1908]
- Barriers cause severe attenuation $\rightarrow$ long exposures, dim displays
- Lenslets impose fixed trade-off between spatial and angular resolution
- *Masks can be optimized to increase optical efficiency*
  - Enables low-cost, large-format, and programmable 3D cameras/displays
**Proposed Approach**

**Tiled-Broadband Patterns**

**Content-Adaptive Parallax Barriers**

**Mask-based Light Field Capture and Display**

- To optimize mask-based light field capture:
  - Barriers generalized with optical heterodyning [Veeraraghavan et al., 2007]
  - *Tiled-broadband patterns* pass significantly more light than slits/pinholes

- To optimize mask-based light field display:
  - Tiled-broadband patterns cannot be used to improve efficiency
  - *Content-adaptive parallax barriers* increase brightness and refresh rate

- Applied to 3D scanning, human-computer interaction, and 3D display
Outline

• Introduction

➢ Light Field Capture
  – Optical Heterodyning and Tiled-Broadband Patterns
  – Single-shot Visual Hulls using Shield Field Cameras
  – A Thin, Depth-sensing LCD for Gestural Interaction

• Light Field Display
  – Content-Adaptive Parallax Barriers

• Conclusion
Light Field Photography

Light Field Camera Designs

- Hand-held plenoptic camera [Ng et al., 2005]
- Heterodyne light field camera [Veeraraghavan et al., 2007]

Ren Ng et al. Light Field Photography with a Hand-held Plenoptic Camera. 2005.
Light Field Analysis of Barrier Cameras

Pinhole Array

Sensor

Lenses & Apertures

Scene

\[ L_{\text{received}}(u,s) \]

\[ \mathbf{u} \rightarrow s \rightarrow L \]

\[ f_u, f_s \]

\[ \hat{L}_{\text{received}}(f_u, f_s) \]

\[ u, s \]

Band-limited light field

2D Fourier Transform
Modeling Sensors in the Frequency Domain

Heterodyne Light Field Cameras

Heterodyne Mask

Sensor

Lenses & Apertures

Scene

The modulation function is an impulse train modulated along a slanted line through the origin.

\( \hat{L}_{\text{received}}(f_u, f_s) \)

Heterodyne Light Field Cameras

Heterodyne Mask

Sensor

Lenses & Apertures

Scene

$\hat{L}_{\text{received}}(f_u f_s)$

reorder spectral components

Heterodyne Light Field Cameras

Heterodyne Mask

Sensor

Lenses & Apertures

Scene

\[ \hat{L}_{\text{received}}(f_u, f_s) \]

2D Inverse Fourier Transform

\[ L_{\text{received}}(u, s) \]
Which Heterodyne Mask to Use?

- Conditions for heterodyne light field detection
  - Mask spectrum must be a (windowed) 2D impulse train
  - Can be achieved (approximately) with a pinhole array

\[ M_{\text{pinhole}}(x,y) \]

\[ \hat{M}_{\text{pinhole}}(f_x, f_y) \]
Which Heterodyne Mask to Use?

- Conditions for heterodyne light field detection
  - Mask spectrum must be a (windowed) 2D impulse train
  - Can be achieved exactly by Sum-of-Sinusoids (SoS) 
    [Veeraraghavan et al., 2007]

\[ M_{\text{SoS}}(x,y) \]

\[ \hat{M}_{\text{SoS}}(f_x,f_y) \]
Which Heterodyne Mask to Use?

- Conditions for heterodyne light field detection
  - Both pinhole array and SoS are periodic functions
  - What other tiles lead to impulse trains?

$M_{\text{general}}(x,y)$

$\hat{M}_{\text{general}}(f_x,f_y)$
General Tiled-broadband Patterns

- Conditions for heterodyne light field detection
  - (Almost) any 2D tile can be used (tiling $\rightarrow$ impulse train)
  - Amplitude/phase of impulses given by Fourier series of tile
Specific Choice: Tiled-MURA

- Conditions for optimal heterodyne light field detection
  - Equal-amplitude impulses $\rightarrow$ tiled-broadband mask
  - One option: Modified Uniformly Redundant Array (MURA)
Benefits of Heterodyne Masks

- **Benefits and Limitations**
  - Sum-of-Sinusoids converges to $\approx 18\%$ transmission
  - Tiled-MURA near 50\% (but only for prime-valued lengths)
  - Binary vs. continuous-tone process (quantization)
Implementation

Components

• 22 megapixel Mamiya 645ZD digital camera [5344 × 4008 pixels]
• 7 × 7 tiled-MURA pattern
• Mask printed at 5,080 DPI using Heidelberg Herkules emulsion printer
Results: Mask-based Light Field Capture

Captured Image (Modulated by Tiled-MURA Mask)
Results: Mask-based Light Field Capture
Results: Sub-Aperture Images

Moving across the lens from left to right (s)

Reconstructed Light Field
Results: Digital Image Refocusing

- Images focused at different depths correspond to 2D projections
- Efficiently-computed using Fourier Projection-Slice Theorem

Ren Ng. Fourier Slice Photography. 2005.
• Introduction

➢ Light Field Capture
  – Optical Heterodyning and Tiled-Broadband Patterns
  – *Single-shot Visual Hulls using Shield Field Cameras*
  – A Thin, Depth-sensing LCD for Gestural Interaction

• Light Field Display
  – Content-Adaptive Parallax Barriers

• Conclusion
Shadowgraphy vs. Traditional Visual Hull Hardware?

- Limitations of Traditional Visual Hull
- Shadowgram-based Visual Hull
- How to Simplify Visual Hull Hardware?

Simultaneous Projection Field Capture

Naïve Solution: Pinhole Arrays

Large-Format Light Field Capture

\[(u, s)\]
Single-Shot Visual Hulls using Masks

Naïve Solution: Pinhole Arrays

Capturing Shield Fields

\[
M_{\text{occluder}}(u,s) = \left[ \frac{L_{\text{occluder}}(u,s)}{L_{\text{incident}}(u,s)} \right]
\]

\[(u,s)\]

tiled-broadband mask

diffuser
Implementation

Components

- 8.0 megapixel Canon EOS Digital Rebel XT [3456 × 2304 pixels]
- 6 × 6 array of Philips Luxeon Rebel LEDs [1.2 × 1.2 m]
- 5,080 DPI mask and a paper vellum diffuser [75 × 55 cm]
Pinhole Array Results: Test Object

Test Object: Mannequin
Tiled-MURA Results: Sensor Image

Sensor Image (0.25 s Exposure)
Tiled-MURA Results: Shadowgrams

Light Field Reconstruction
Pinhole Array Results: Visual Hull

Visual Hull Reconstruction
• Introduction

➢ Light Field Capture
  – Optical Heterodyning and Tiled-Broadband Patterns
  – Single-shot Visual Hulls using Shield Field Cameras
  – *A Thin, Depth-sensing LCD for Gestural Interaction*

• Light Field Display
  – Content-Adaptive Parallax Barriers

• Conclusion
How to Capture Touch and Gestures?

Multi-touch Interfaces
Gesture-based Interfaces

Lighting-Sensitive Displays

Optical Touch and Gesture Recognition

BiDi Screen: A Thin, Depth-Sensing LCD

Embedded photodetector arrays
Depth cameras (e.g., time-of-flight)
Ambient light sensors

Sensing multi-touch, gestures, and ambient light using LCDs

Design Goals
- Capture light field reflected or emitted by objects
- Prevent image capture from interfering with image display
- Support unencumbered interaction with thin form factor

Designing a Thin, Depth-Sensing LCD

- Ambient light source
- Angle-limiting film
- Object
- Printed mask
- Diffuser
- Camera
Designing a Thin, Depth-Sensing LCD

- Ambient light source
- Angle-limiting film
- Object
- Printed mask
- Sensor/Backlight
- Camera
- LCD
Implementation

Components

• 20.1 inch Sceptre X20WG-Nagall LCD [1680 × 1050 @ 60 fps]
• Point Grey Flea2 video cameras [1280 × 960 @ 7 fps]
Model Viewer Application
Outline

• Introduction

• Light Field Capture
  – Optical Heterodyning and Tiled-Broadband Patterns
  – Single-shot Visual Hulls using Shield Field Cameras
  – A Thin, Depth-sensing LCD for Gestural Interaction

➢ **Light Field Display**
  – *Content-Adaptive Parallax Barriers*

• Conclusion
Parallax Barrier Light Field Displays

Virtual Scene

Display

Pinhole Array

Viewers

u

s
Target 4D Light Field
Target 4D Light Field

viewer moves right

viewer moves up
Light Field Analysis of Barriers

\[ L[i, k] = f[i] \cdot g[k] \]

\[ L[i, k] = f[i] \otimes g[k] \]
Time-Multiplexing using Shifted Pinholes

\[ L[i, k] = \sum_{t=1}^{T} f_t[i] \otimes g_t[k] \]

Content-Adaptive Parallax Barriers

\[ \tilde{L} = FG \]
Content-Adaptive Parallax Barriers

\[ \tilde{L} = FG \]
Content-Adaptive Parallax Barriers

\[
\arg \min_{F,G} \frac{1}{2} \| L - FG \|^2, \text{ for } F, G \geq 0
\]
Content-Adaptive Parallax Barriers

\[
\arg \min_{F,G} \frac{1}{2} \left\| L - FG \right\|^2_W, \text{ for } F, G \geq 0
\]
Content-Adaptive Parallax Barriers

\[
\begin{align*}
\bar{L} &= F \\
\arg \min_{F, G} \frac{1}{2} \left\| L - FG \right\|_W^2, & \text{for } F, G \geq 0
\end{align*}
\]
Target 4D Light Field

viewer moves right

viewer moves up
Optimization: Iteration 1

rear mask: $f_1[i,j]$

$$F \leftarrow F \circ \frac{[(W \circ L)G^t]}{[(W \circ (FG))G^t]}$$

$$G \leftarrow G \circ \frac{[F^t(W \circ L)]}{[F^t(W \circ (FG))]}$$

front mask: $g_1[k,l]$

reconstruction (central view)

Optimization: Iteration 10

![Diagram of optimization process]

**rear mask:** $f_1[i,j]$

**front mask:** $g_1[k,l]$

\[
F \leftarrow F \circ \frac{[(W \circ L)G^t]}{[(W \circ (FG))G^t]}
\]

\[
G \leftarrow G \circ \frac{[F^t(W \circ L)]}{[F^t(W \circ (FG))]}\]


reconstruction (central view)
Optimization: Iteration 20

rear mask: $f_1[i,j]$

$$F \leftarrow F \odot \frac{[(W \odot L)G^t]}{[(W \odot (FG))G^t]}$$

$$G \leftarrow G \odot \frac{[F^t(W \odot L)]}{[F^t(W \odot (FG))]}$$

front mask: $g_1[k,l]$

reconstruction (central view)

Optimization: Iteration 30

**rear mask: $f_1[i,j]$**

**front mask: $g_1[k,l]$**

\[
F \leftarrow F \odot \frac{[(W \odot L)G^t]}{[(W \odot (FG))G^t]}
\]

\[
G \leftarrow G \odot \frac{[F^t(W \odot L)]}{[F^t(W \odot (FG))]}\]

Optimization: Iteration 40

rear mask: $f_1[i,j]$

$$F \leftarrow F \odot \frac{[(W \circ L)G^t]}{[(W \circ (FG))G^t]}$$

$$G \leftarrow G \odot \frac{[F^t(W \circ L)]}{[F^t(W \circ (FG))]}$$

front mask: $g_1[k,l]$

reconstruction (central view)

Optimization: Iteration 50

rear mask: $f_1[i,j]$  

front mask: $g_1[k,l]$  

$$F \leftarrow F \circ \frac{[(W \circ L)G^t]}{[(W \circ (FG))G^t]}$$  

$$G \leftarrow G \circ \frac{[F^t(W \circ L)]}{[F^t(W \circ (FG))]}$$  

Optimization: Iteration 60

rear mask: $f_1[i,j]$

$F \leftarrow F \odot \frac{[(W \odot L)G^t]}{[(W \odot (FG))G^t]}$

$G \leftarrow G \odot \frac{[F^t(W \odot L)]}{[F^t(W \odot (FG))]}$

front mask: $g_1[k,l]$

reconstruction (central view)

Optimization: Iteration 70

rear mask: $f_1[i,j]$

$$F \leftarrow F \circ \frac{[(W \circ L)G^t]}{[(W \circ (FG))G^t]}$$

$$G \leftarrow G \circ \frac{[F^t(W \circ L)]}{[F^t(W \circ (FG))]}$$

front mask: $g_1[k,l]$

reconstruction (central view)

Optimization: Iteration 80

rear mask: $f_1[i,j]$

front mask: $g_1[k,l]$

$$F \leftarrow F \odot \frac{[(W \odot L)G^t]}{[(W \odot (FG))G^t]}$$

$$G \leftarrow G \odot \frac{[F^t(W \odot L)]}{[F^t(W \odot (FG))]}$$

reconstruction (central view)

Optimization: Iteration 90

\[ F \leftarrow F \odot \frac{[(W \odot L)G^t]}{[(W \odot (FG))G^t]} \]
\[ G \leftarrow G \odot \frac{[F^t(W \odot L)]}{[F^t(W \odot (FG))]} \]

Content-Adaptive Rear Mask (1 of 9)
Emitted 4D Light Field
Benefits of Content-Adaptation

1) Increasing brightness:

$$\arg \min_{F,G} \frac{1}{2} \left\| \alpha L - FG \right\|_W^2, \text{ for } F, G \geq 0$$

2) Increasing refresh rate:

$$L[i, j, k, l] = \sum_{t=1}^{T} f_t[i, j] \otimes g_t[k, l], \text{ for } T < N_h N_v$$
Implementation

Components

- 22 inch ViewSonic FuHzion VX2265wm LCD [1880 × 1050 @ 120 fps]
Motion Parallax

Time-Shifted Pinhole Arrays

Content-Adaptive Parallax Barrier

viewer moves right

Light Field

viewer moves up
Increasing Brightness and Refresh Rate

1) Increasing brightness:
\[
\arg\min_{F,G} \frac{1}{2} \| \alpha L - FG \|^2_w, \text{ for } F, G \geq 0
\]

2) Increasing refresh rate:
\[
L[i, j, k, l] = \sum_{t=1}^{T} f_{t}[i, j] \otimes g_{t}[k, l], \text{ for } T < N_h N_v
\]
1) Increasing brightness:
\[ \arg \min_{F,G} \frac{1}{2} \| \alpha L - FG \|_W^2, \text{ for } F, G \geq 0 \]

2) Increasing refresh rate:
\[ L[i, j, k, l] = \sum_{t=1}^{T} f_t[i, j] \otimes g_t[k, l], \text{ for } T < N_h N_v \]
• Introduction

• Light Field Capture
  – Optical Heterodyning and Tiled-Broadband Patterns
  – Single-shot Visual Hulls using Shield Field Cameras
  – A Thin, Depth-sensing LCD for Gestural Interaction

• Light Field Display
  – Limitations of Optical Heterodyning
  – Content-Adaptive Parallax Barriers

➤ Conclusion
Contributions

• Introduced the \textit{shield field} representation of volumetric occluders
• Developed \textit{tiled-broadband patterns} for light field capture
  – Mask-based, hand-held light field photography (3x shorter exposures)
  – First shadowgram-based visual hull reconstruction of moving objects
  – First LCD to support both multi-touch and gesture-based interaction
• Developed \textit{content-adaptive parallax barriers} for light field display
  – Demonstrated limitations of optical heterodyning for 3D displays
  – Increased brightness and refresh rate of dual-layer 3D displays

\textit{Masks enable low-cost, large-format, programmable 3D cameras/displays}