Computational Photography

Bill Freeman,

Computer Science and Artificial Intelligence Lab (CSAIL)  
MIT

January 24, 2012
Cameras: optics plus sensors plus computation

Early photography
Cameras: optics plus sensors plus computation

Early photography

Contemporary photography

a good fit with the toolkit of workshop attendees: great image priors
electronic camera

http://www.zimfamilycockerers.com/CanonDigitalRebelXS-Side.JPEG
exploded view of camera

exploded view of camera

Much of what I’ll present is not my work (just “a few of my favorite things”)  
Apologies if I didn’t include any particular project.

large depth-of-field imaging

seminal work

Applied Optics, 1995

Extended depth of field through wave-front coding

Edward R. Dowski, Jr., and W. Thomas Cathey

We designed an optical–digital system that delivers near-diffraction-limited imaging performance with a large depth of field. This system is the standard incoherent optical system modified by a phase mask with digital processing of the resulting intermediate image. The phase mask alters or codes the received incoherent wave front in such a way that the point-spread function and the optical transfer function do not change appreciably as a function of misfocus. Focus-independent digital filtering of the intermediate image is used to produce a combined optical–digital system that has a nearly diffraction limited point-spread function. This high-resolution extended depth of field is obtained through the expense of an increased dynamic range of the incoherent system. We use both the ambiguity function and the stationary-phase method to design these phase masks.

Keywords: Extended depth of field, extended depth of focus, wave-front coding.

1. Introduction
Extending the depth of field of incoherent optical systems has been an active research topic for many years. The majority of the literature on this topic has concerned methods of employing an optical power-absorbing apodizer, with possible $\pm \pi$ phase varia-

regions of zeros, digital processing can be used to restore the sampled intermediate image. Further, because the OTF is insensitive to misfocus, the same digital processing restores the image for all values of misfocus. This combined optical–digital system produces a PSF that is comparable to that of the diffrac-
Wavefront coding

http://www.wirelessweek.com/uploadedImages/WW/articles/4-TrueFocus.jpg
Controllable depth of field

Coded aperture

Behavior of lens with coded aperture

Anat Levin, Rob Fergus, Fredo Durand, William Freeman, SIGGRAPH 2007, also related work by Ashok Veeraraghavan, Ramesh Raskar, Amit Agrawal, Ankit Mohan and Jack Tumblin, SIGGRAPH 2007
Behavior of lens with coded aperture
Behavior of lens with coded aperture

- Aperture pattern
- Image of a defocused point light source

Object → Lens with coded aperture → Camera sensor

Point spread function

Focal plane
Behavior of lens with coded aperture

Aperture pattern → Image of a defocused point light source

Object → Focal plane → Lens with coded aperture → Camera sensor

Point spread function
Behavior of lens with coded aperture

Aperture pattern

Image of a defocused point light source

Object

Lens with coded aperture

Camera sensor

Focal plane

Point spread function
Behavior of lens with coded aperture

Aperture pattern

Image of a defocused point light source

Object

Lens with coded aperture

Camera sensor

Point spread function

Focal plane

Saturday, March 31, 12
Behavior of lens with coded aperture

Aperture pattern

Image of a defocused point light source

Object

Lens with coded aperture

Camera sensor

Point spread function

Focal plane
Challenges, using conventional aperture

- Hard to deconvolve even when kernel is known

![Input](image1)

![Ringing with the traditional Richardson-Lucy deconvolution algorithm](image2)
Challenges, using conventional aperture

- Hard to deconvolve even when kernel is known

- Hard to identify correct scale:
  - Larger scale
  - Correct scale
  - Smaller scale
Challenges, using conventional aperture

- Hard to deconvolve even when kernel is known

- Hard to identify correct scale:
  - Larger scale
  - Correct scale
  - Smaller scale
Challenges, using conventional aperture

- Hard to deconvolve even when kernel is known

- Hard to identify correct scale:
  - Larger scale
  - Correct scale
  - Smaller scale
Challenges, using conventional aperture

- Hard to deconvolve even when kernel is known

- Hard to identify correct scale:
  - Larger scale
  - Correct scale
  - Smaller scale

Input

Ringing with the traditional Richardson-Lucy deconvolution algorithm
Challenges, using conventional aperture

• Hard to deconvolve even when kernel is known

• Hard to identify correct scale:

  ?

  =

  ×

  Larger scale

  ?

  =

  ×

  Correct scale

  ?

  =

  ×

  Smaller scale
Challenges, using conventional aperture

- Hard to deconvolve even when kernel is known

- Hard to identify correct scale:
  - Larger scale
  - Correct scale
  - Smaller scale

Input
Ringing with the traditional Richardson-Lucy deconvolution algorithm
Intuition behind use of coded aperture

Coded aperture - reduce uncertainty in scale identification
Intuition behind use of coded aperture

Coded aperture - reduce uncertainty in scale identification

<table>
<thead>
<tr>
<th>Larger scale</th>
<th>Conventional</th>
<th>Coded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smaller scale</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Intuition behind use of coded aperture

Coded aperture - reduce uncertainty in scale identification

Conventional  Coded

Larger scale

Correct scale

Smaller scale
Intuition behind use of coded aperture

Coded aperture - reduce uncertainty in scale identification

Conventional

Coded

Larger scale

Correct scale

Smaller scale
Intuition behind use of coded aperture

Coded aperture - reduce uncertainty in scale identification

**Conventional**

Larger scale

Correct scale

Smaller scale

**Coded**
Intuition behind use of coded aperture

Coded aperture - reduce uncertainty in scale identification

<table>
<thead>
<tr>
<th>Larger scale</th>
<th>Conventional</th>
<th>Coded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smaller scale</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Coded aperture camera tasks that require image priors

- Depth estimation from defocused image data.
- Deconvolution of defocused data.
- Optimization of aperture code.
Regularizing depth estimation

Try deblurring with 10 different aperture scales

\[ x = \underset{\mathbf{x}}{\text{arg min}} \quad \left| f \otimes \mathbf{x} - y \right|^2 + \lambda \sum_i \rho(\nabla x_i) \]

Convolution error

Derivatives prior

Keep minimal error scale in each local window

Input

Local depth estimation
Regularizing depth estimation

Try deblurring with 10 different aperture scales

\[ x = \arg \min_x \left\| f \otimes x - y \right\|^2 + \lambda \sum_i \rho(\nabla x_i) \]

Convolution error
Derivatives prior

Keep minimal error scale in each local window + regularization

Input
Local depth estimation
Regularized depth
Filter search

Design constraints:

1. Binary pattern
2. Minimum hole size $\rightarrow 1\text{mm}^2$ (due to diffraction)
3. No floating parts
4. Maximum aperture $\rightarrow f/2.8$ (minimize radial distortion)

Sample binary filters:

Sample patterns and test KL score
Freq domain slices through conventional aperture and winning coded aperture

Higher frequencies are passed and zeros are not harmonically related.
Comparison – conventional aperture result

Ringing due to wrong scale estimation
Comparison - coded aperture result
Application: Digital refocusing from a single image
Application: Digital refocusing from a single image
Application: Digital refocusing from a single image
Application: Digital refocusing from a single image
Easy-to-deblur motion blur

flutter shutter

exposure
time

Blurring == Convolution

Traditional Camera: Box Filter
Blurring == Convolution

Sync Function

Traditional Camera: Box Filter
Flutter Shutter: Coded Filter

Saturday, March 31, 12
Flutter Shutter: Coded Filter

Preserves High Frequencies

Flutter Shutter: Coded Filter
Measure material properties

structured light projection

Fast Separation of Direct and Global Images

Using High Frequency Illumination

Shree K. Nayar

Gurunandan G. Krishnan

Columbia University

Michael D. Grossberg

City College of New York

Ramesh Raskar

MERL

SIGGRAPH Conference
Boston, July 2006

Support: ONR, NSF, MERL
Direct and Global Illumination

source

camera

surface

P
Direct and Global Illumination

source

camera

A

P

surface

A : Direct
Direct and Global Illumination

A : Direct
B : Interrelection

source

surface

A

B

P

camera

A : Direct
B : Interrelection
Direct and Global Illumination

A : Direct
B : Interreflection
C : Subsurface
Direct and Global Illumination

A : Direct
B : Interreflection
C : Subsurface
D : Volumetric
Direct and Global Illumination

A : Direct
B : Interreflection
C : Subsurface
D : Volumetric
E : Diffusion
Direct and Global Illumination

A : Direct
B : Interreflection
C : Subsurface
D : Volumetric
E : Diffusion
Direct and Global Components: Interreflections

source

camera

surface
Direct and Global Components: Interreflections
Direct and Global Components: Interreflections

\[ L[c, i] = L_d[c, i] + L_g[c, i] \]

radiance  direct  global
Direct and Global Components: Interreflections

\[ L[c, i] = L_d[c, i] + L_g[c, i] \]

radiance \hspace{1cm} direct \hspace{1cm} global

\[ L_g[c, i] = \sum_P A[i, j] L_s[i, j] \]

BRDF and geometry

Saturday, March 31, 12
High Frequency Illumination Pattern
High Frequency Illumination Pattern
High Frequency Illumination Pattern

Source

Camera

Surface

\[ L^+ [c, i] = L_d [c, i] + \alpha L_g [c, i] \]

Fraction of activated source elements
High Frequency Illumination Pattern

\[ L^+[c,i] = L_d[c,i] + \alpha L_g[c,i] \]

fraction of activated source elements
High Frequency Illumination Pattern

\[ L^+[c,i] = L_a[c,i] + \alpha L_g[c,i] \]

\[ L^-[c,i] = (1-\alpha) L_g[c,i] \]

fraction of activated source elements
Separation from Two Images

\[ \alpha = \frac{1}{2} \]

\[ L_d = L^+_{\text{max}} - L^-_{\text{min}}, \quad L_g = 2L^-_{\text{min}} \]

direct \quad \text{global}
Diffuse Interreflections
Specular Interreflections
Volumetric Scattering
Subsurface Scattering
Diffusion
Scene
Scene
Scene

Direct

Global
Kitchen Sink: Volumetric Scattering

[Image of a kitchen sink with a spoon and cup]

Volumetric Scattering:
Chandrasekar 50, Ishimaru 78
Kitchen Sink: Volumetric Scattering

Volumetric Scattering:
Chandrasekar 50, Ishimaru 78

Direct
Global
Peppers: Subsurface Scattering
Peppers: Subsurface Scattering

Direct

Global
Hand

Skin: Hanrahan and Krueger 93, Uchida 96, Haro 01, Jensen et al. 01, Cula and Dana 02, Igarashi et al. 05, Weyrich et al. 05
Hand

Skin: Hanrahan and Krueger 93, Uchida 96, Haro 01, Jensen et al. 01, Cula and Dana 02, Igarashi et al. 05, Weyrich et al. 05
Hand

Skin: Hanrahan and Krueger 93, Uchida 96, Haro 01, Jensen et al. 01, Cula and Dana 02, Igarashi et al. 05, Weyrich et al. 05
Combine modalities for optimal view

light projection

No-flash

Flash
Flash/no Flash

Simultaneously developed by:

PETSCHNIGG, AGRAWALA, HOPPE, SZELISKI, COHEN, AND TOYAMA. 2004. Digital photography with flash and no-flash image pairs. ACM Trans. on Graphics 2004
Motion invariant photography

move detector during exposure

Motion invariant PSF

Anat Levin, Taeg Sang Cho, Peter Sand, Fredo Durand, William Freeman,
SIGGRAPH 2008

Integration curves

Sheared integration curves

Projected PSF

Static

Image data

Velocity

Saturday, March 31, 12
Is there motion path (space-time curve) for the sensor pixels that results in a projected PSF that is invariant to the velocity of the image data?
Solution: a parabolic curve is shear invariant

\[ f(t) = t^2 \]

Sheared parabola \( (x, t) \) \[ (x - st, t) \] \[ \Rightarrow \]

\[ f_s(t) = t^2 - st \]

Shifted parabola
Solution: a parabolic curve is shear invariant

\[ f(t) = t^2 \]

\[ f_s(t) = t^2 - st \]

\[ = \left(t - \frac{s}{2}\right)^2 - \frac{s^2}{4} \]
Solution: a parabolic curve is shear invariant

\[ f(t) = t^2 \]

\[ f_s(t) = t^2 - st \]

\[ = \left( t - \frac{s}{2} \right)^2 - \frac{s^2}{4} \]

Sheared parabola \rightarrow Shifted parabola
Motion invariant PSF

Anat Levin, Taeg Sang Cho, Peter Sand, Fredo Durand, William Freeman,
SIGGRAPH 2008

integration curves

sheared integration curves

projected PSF

Static

Parabolic

Up to tail clipping, PSF is invariant to slope

Image data
Velocity
Parabolic sweep

- Start by moving very fast to the right
- Continuously slow down until stop
- Continuously accelerate to the left

Intuition: For any velocity, there is one instant where we track perfectly. We spend an equal amount of time tracing each velocity.
Static camera
Parabolic motion camera
Experimental set-up

Ideally move sensor (requires same hardware as existing stabilization systems)
Input from a static camera

Input from our parabolic camera - identical blur over both static and moving parts

Deblurred output - entire image deblurred with identical known PSF, no segmentation and no motion estimation
Record all light rays passing through lens

lens array at detector to capture 4-d lightfield

developed by John Wang, Edward Adelson, Ren Ng, and Marc Levoy

**Figure 1:** Conceptual schematic (not drawn to scale) of our camera, which is composed of a main lens, microlens array and a photosensor. The main lens focuses the subject onto the microlens array. The microlens array separates the converging rays into an image on the photosensor behind it.
Figure 8: Top: Exploded view of assembly for attaching the microlens array to the digital back. Bottom: Cross-section through assembled parts.
Light Field in a Single Exposure
Light Field in a Single Exposure
Light Field in a Single Exposure
Light Field in a Single Exposure
Light Field Inside the Camera Body

Ray carrying $L(u, v, x, y)$
Light Field Inside the Camera Body

Ray carrying $L(u, v, x, y)$
Light Field Inside the Camera Body

Ray carrying $L(u, v, x, y)$
Digital Refocusing
Digital Refocusing
Digital Refocusing

Slide by Ren Ng.
Digital Refocusing
Digital Refocusing

Slide by Ren Ng.
Digital Refocusing
Digital Refocusing
Digital Refocusing
Digital Refocusing
Digital Refocusing
Digital Refocusing
Digital Refocusing
Digital Refocusing
Digital Refocusing

Slide by Ren Ng.
Digital Refocusing

Slide by Ren Ng.
Image restoration

Natural image statistics

Characteristic distribution with heavy tails

Histogram of image gradients
Blurry images have different statistics

Histogram of image gradients
Image formation process

Blurry image = Sharp image

Blur kernel
Image formation process

Blurry image = Sharp image

Input to algorithm

Blur kernel
Image formation process

Blurry image = Sharp image

Blur kernel

Input to algorithm

Desired output
Image formation process

Blurred image = Sharp image

Input to algorithm = Desired output

Convolution operator × Blur kernel
Image formation process

- Blurry image
  - Input to algorithm

- Sharp image
  - Desired output

- Convolution operator

- Assume unknown

- Blur kernel
Multiple possible solutions
Multiple possible solutions

Sharp image

Blur kernel

Blurry image
Multiple possible solutions

\[ \text{Blurry image} = \text{Sharp image} \otimes \text{Blur kernel} \]
Multiple possible solutions

Blurry image

= Sharp image

= Blur kernel
Multiple possible solutions

\[ \text{Blurry image} = \text{Sharp image} \times \text{Blur kernel} \]
Multiple possible solutions

- Blurry image
- Sharp image
- Blur kernel

\[ \text{Blury image} \rightarrow \text{Sharpe image} \rightarrow \text{Blur kernel} \]
Three sources of information

Rob Fergus, Barun Singh, Aaron Hertzmann, Sam Roweis, and William Freeman, SIGGRAPH 2006
Three sources of information

Rob Fergus, Barun Singh, Aaron Hertzmann, Sam Roweis, and William Freeman, SIGGRAPH 2006

1. Reconstruction constraint:

Estimated sharp image \( \otimes \) Estimated blur kernel = Input blurry image
Three sources of information

Rob Fergus, Barun Singh, Aaron Hertzmann, Sam Roweis, and William Freeman, SIGGRAPH 2006

1. Reconstruction constraint:

- Estimated sharp image
- Estimated blur kernel
- Input blurry image

2. Image prior:

- Distribution of gradients
Three sources of information

1. Reconstruction constraint:

Estimated sharp image \(\otimes\) Estimated blur kernel = Input blurry image

2. Image prior:

Distribution of gradients

3. Blur prior:

Positive & Sparse
Blur kernel

Our output
Matlab’s `deconvblind`
Close-up of garland

Original

Matlab’s deconvblind

Our output
Our output
Close-up of bird

Original

Unsharp mask

Our output
Video analysis and re-rendering

Post-capture processing

We can register, then amplify, one motion *relative* to another.

Original footage courtesy of Paul Robertson, BBN.
We can register, then amplify, one motion relative to another.

Empty trunk

Full trunk (motion difference amplified)

Original footage courtesy of Paul Robertson, BBN.
Summary of some computational photography work

- lens array at detector to capture 4-d lightfield
- structured light projection, flash/no flash
- move detector during exposure
- flutter shutter
- wavefront coding
- Coded aperture
- priors on images

Where much of this work was published

- **SIGGRAPH** (premier computer graphics conference, submission deadline: late January).

- **CVPR** (Computer Vision and Pattern Recognition, premier computer vision conference, submission deadline: November).

- **ICCP** (International Conference on Computational Photography, good new conference, high quality work, interdisciplinary audience, submission deadline: November).