Wavefront Coding and Recent Research

Fredo Durand
MIT EECS 6.815/6.865
Pset 11

• Use &, not and
Changes

• Pset 11
  – due Friday 9pm

• Extra Office hours
  – Fredo 5–6pm, Thursday

• Last pset due next Wednesday
  – pretty short, video magnification
Today

• Except for the first topic (wavefront coding), we’ll see research done in my group.

• The corresponding papers are available at http://people.csail.mit.edu/fredo/
Computational Imaging at MIT

- Ted Adelson (BCS)
- George Barbastathis (Mechanical Engineering)
- Fredo Durand (EECS)
- Bill Freeman (EECS)
- Berthold Horn (EECS)
- Wojciech Matusik (EECS)
- Ramesh Raskar (Medialab)
Wavefront coding
Is depth of field a blur?

From Macro Photography
Is depth of field a blur?

• Depth of field is NOT a convolution of the image
• The circle of confusion varies with depth
• There are interesting occlusion effects
• (If you really want a convolution, there is one, but in the light field...)

From Macro Photography
Hardware depth of field solution

http://www.cdm-optics.com/site/publications.php
Hardware depth of field solution

- Stop down! (use smaller aperture)
- problem: noise

http://www.cdm-optics.com/site/publications.php
Wavefront coding

• CDM-Optics, U of Colorado, Boulder
• Improve depth of field using weird optics & deconvolution
• http://www.cdm-optics.com/site/publications.php

Wavefront coding

• Idea: deconvolution to deblur out of focus regions
• Problem 1: depth of field blur is not shift-invariant
  – Depends on depth
    ➔ If depth of field is not a convolution,
      it's harder to use deconvolution ;-(
• Problem 2: Depth of field blur "kills information"
  – Fourier transform of blurring kernel has low frequency response
Wavefront coding
Wavefront coding

- **Idea: deconvolution to deblur out of focus regions**
  - Problem 1: depth of field blur is not shift-invariant
  - Problem 2: Depth of field blur "kills information"

- **Solution: change optical system so that**
  - Rays don't converge anymore
  - Image blur is the same for all depth
  - Blur spectrum is higher
Wavefront coding
Wavefront coding

• Idea: deconvolution to deblur out of focus regions
  – Problem 1: depth of field blur is not shift-invariant
  – Problem 2: Depth of field blur "kills information"

• Solution: change optical system so that
  – Rays don't converge anymore
  – Image blur is the same for all depth
  – Blur spectrum is higher

• How it's done
  – Phase plate (cubic lens \( z = y^3 + x^3 \))
  – Will do things similar to spherical aberrations
Image of a point source

• aka Point Spread Function (PSF), same as kernel

Normal lens

In focus

Out of focus

Wavefront coding

(A)

(B)

(C)

(D)
Ray version
Image of a point source

- aka Point Spread Function (PSF), same as kernel

Normal lens

In focus

Out of focus

(A)

(B)

Wavefront coding

(C)

(D)

less density

big density
Frequency response (MTF)

- Magnitude of the Fourier transform of the PSF
- Higher is better (noise less amplified by deconvolution)

Fig. 5: MTFs corresponding with the PSFs of Fig. 3 for a conventional image in and out of focus and a coded image for the same misfocus values.
Results

Traditional Optical System Image

Intermediate Extended Depth of Field Image

Stopped Down Traditional System Image

Final Wavefront Coded™ Image
Frequency response (MTF)

- Magnitude of the Fourier transform of the PSF
- Higher is better (noise less amplified by deconvolution)

Fig. 5: MTFs corresponding with the PSFs of Fig. 3 for a conventional image in and out of focus and a coded image for the same misfocus values.
Results
Lattice focal lens
Levin et al. 09: assembly of subsquares with different focal powers

- each element focuses on a different depth

New solution: The lattice-focal lens

toy lattice-focal lens with 4 elements

\[
E \left[ \left( \frac{\phi_s(\omega_{x,y})}{|\omega_{x,y}|} \right)^2 \right] \approx \frac{A^{8/3}}{S^{4/3} \Omega^{1/3}} |\omega_{x,y}|
\]
Proof of concept

- 12 subsquares cut from plano-convex spherical lenses
- Attached to main lens extra focal power needed very low
- Modest DOF extension with only 12 subsquares
Depth estimation

- Defocus kernels vary with depth

- Depth estimation as for the coded aperture camera [Levin et al. 07]
Standard lens reference
Lattice-focal lens
Lattice focal lens

While some defocus can be removed computationally using deconvolution, this comes at the expense of defocus outside the focal plane. Recent advances in computational imaging have shown that it is desirable to avoid spending power in the other regions of the frequency domain whose energy is low, noise is amplified and image reconstruction is degraded. To capture scenes with a given depth range, coarse depth estimation is needed. We have constructed a prototype and our analysis leads to the development of the lattice-focal lens—a wide-aperture lens that preserves high frequencies, we achieve a good restoration over the full depth range.

At the center of a lattice-focal lens is the point spread function (PSF). The defocus PSF is the Fourier transform (FT) of the optical transfer function (OTF). In the frequency domain, convolution is a multiplication, and deconvolution is division. In particular, deconvolution quality is a tight function of the defocus kernel. Recent advances in computational imaging have shown that it is desirable to avoid spending power in the other regions of the frequency domain whose energy is low, noise is amplified and image reconstruction is degraded.

Keywords: Computational camera, depth of field, Light field, 4D Frequency Analysis of Computational Cameras for Depth of Field Extension

The defocus kernel of this lens is designed to preserve high frequencies, we achieve a good restoration over the full depth range, but this comes at the expense of defocus outside the focal plane. This both requires and encourages the modification of the acquisition process to extend the DOF through deconvolution. Because deconvolution quality is a tight function of the defocus kernel, designs with high frequency power spectrum of the defocus kernel are preferred. Recent advances in computational imaging have shown that it is desirable to avoid spending power in the other regions of the frequency domain whose energy is low, noise is amplified and image reconstruction is degraded.
Application: Refocusing from single captured image
Application: Refocusing from single captured image
Application: Refocusing from single captured image
Depth of field analysis

• How do different cameras compare?
• What is the best that can be done?
Application: Refocusing from single captured image
Depth of field analysis

• How do different cameras compare?
• What is the best that can be done?
Our new theoretical analysis

- http://www.wisdom.weizmann.ac.il/~levina/papers/lattice/
- In the 4D light field
  - Fourier analysis
- Shows that only a 3D subset of the 4D spectrum is useful (dimensionality gap)
- Inspires new lens design: lattice-focal lens

Previous designs spend energy outside the useful subset
Comparison of different cameras
Comparing image reconstruction (simulation)

Object at in-focus depth

Object at extreme depth
Comparing image reconstruction (simulation)

- **standard lens**
- **coded aperture**
- **focus sweep**
- **wavefront coding**
- **lattice-focal lens**

**Object at in-focus depth**

**Object at extreme depth**
Comparing image reconstruction (simulation)

Object at in-focus depth

Object at extreme depth
In cell phones

- e.g. Nokia E7
- Also exploits chromatic aberrations to focus R, G & B at different depths
- [http://www.allaboutsymbian.com/features/item/12789_EDoF_versus_Auto-focus_Underst.php](http://www.allaboutsymbian.com/features/item/12789_EDoF_versus_Auto-focus_Underst.php)
- [http://www.tessera.com/Pages/tessera.aspx](http://www.tessera.com/Pages/tessera.aspx)
Motion-Invariant Photography
Our counter intuitive solution:

To reduce motion blur, increase it!

- move camera as picture is taken
Our counter intuitive solution:

To reduce motion blur, increase it!

- move camera as picture is taken

• Makes blur invariant to motion- can be removed with spatially uniform deconvolution

  - kernel is known (no need to estimate motion)
  - kernel identical over the image (no need to segment)
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.
Time 2
Static camera
Parabolic motion camera
Comparing camera reconstruction

Note: synthetic rendering, exact PSF is known
For each spatial frequency, max budget across velocity range

\[ |K(\omega_x, \omega_t)|^2 \leq \frac{1}{|\omega_x|} \]
Cameras and information preservation

**Static**
- Constant horizontally
- Spends frequency “budget” outside wedge
- Handles 2D motion

**Flutter shutter**
- Near optimal “budget” usage at all frequencies

**Parabolic**
- Bounded “budget” per column \( \omega_x \)

**Upper bound**
Ideally move sensor (requires same hardware as existing stabilization systems)
Linear rail

Static camera input-
Unknown and variable blur
Linear rail

Static camera input-
Unknown and variable blur

Our parabolic input-
is invariant to velocity

Blur
Linear rail

Static camera input - Unknown and variable blur

Our output after deblurring - NON-BLIND deconvolution
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
In General: Coded Imaging
Philosophy: Image capture

- A sensor placed alone in the middle of the visual world does not record an image.
Image capture

- Pinhole allows you to select light rays
Image formation: optics

- Optics forms an image: selects and integrates light rays
Image formation: computation

- The combination of optics & computation forms the image: selects and combines rays
Computational imaging goals

- Better capture information
- Form image as a post-process
Better capture information

- Same as communication theory: optics encodes, computation decodes
- Code seeks to minimize distortion
Form images as a post-process

- The computational part of formation can be done later and multiple times
- e.g., enable refocusing

The computational part of formation can be done later and multiple times
- e.g., enable refocusing
Learning
Photo
Adjustments
Our Dataset

• 5000 photos in RAW format
• 5 students retouched them by hand
  – Trained at the Visual School of Art in New York
  – Paid for their work (5000 photos in 2 months)

Image Descriptor

• Ideally, correlates with the adjustment.

• We tried many descriptors:
  – Luminance, color, and gradient distributions
  – Global and local (3x3, 5x5, center / surround...)
  – Scene descriptor (GIST)
  – Face detector, face-related features

• The winner is: global luminance distribution + faces
Sample result (representative performance)
• Now in Adobe Photoshop for Auto Brightness/contrast
Local Laplacian Filtering
Background on Laplacian Pyramids

- Difference between adjacent Gaussian levels

level 0

level 1

level 2

(level 3 (residual))
Possible Nonlinearities

• Detail manipulation

![Graph and images showing detail manipulation](image-url)
Speckle Imaging
Computational Re-Photography

- Given reference (old) photograph
- Take new photo at the exact same viewpoint

- Our method: the camera guides the user
  - camera tethered to a laptop
  - pose estimation
  - visualization
Results
Results
Results
Results after Style Transfer
Results after Style Transfer
Results after Style Transfer
Results after Style Transfer
Results after Style Transfer
Results after Style Transfer