

Fake Depth of Field, Wavefront coding

Frédo Durand

MIT EECS

De deblurring lecture first

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...)**

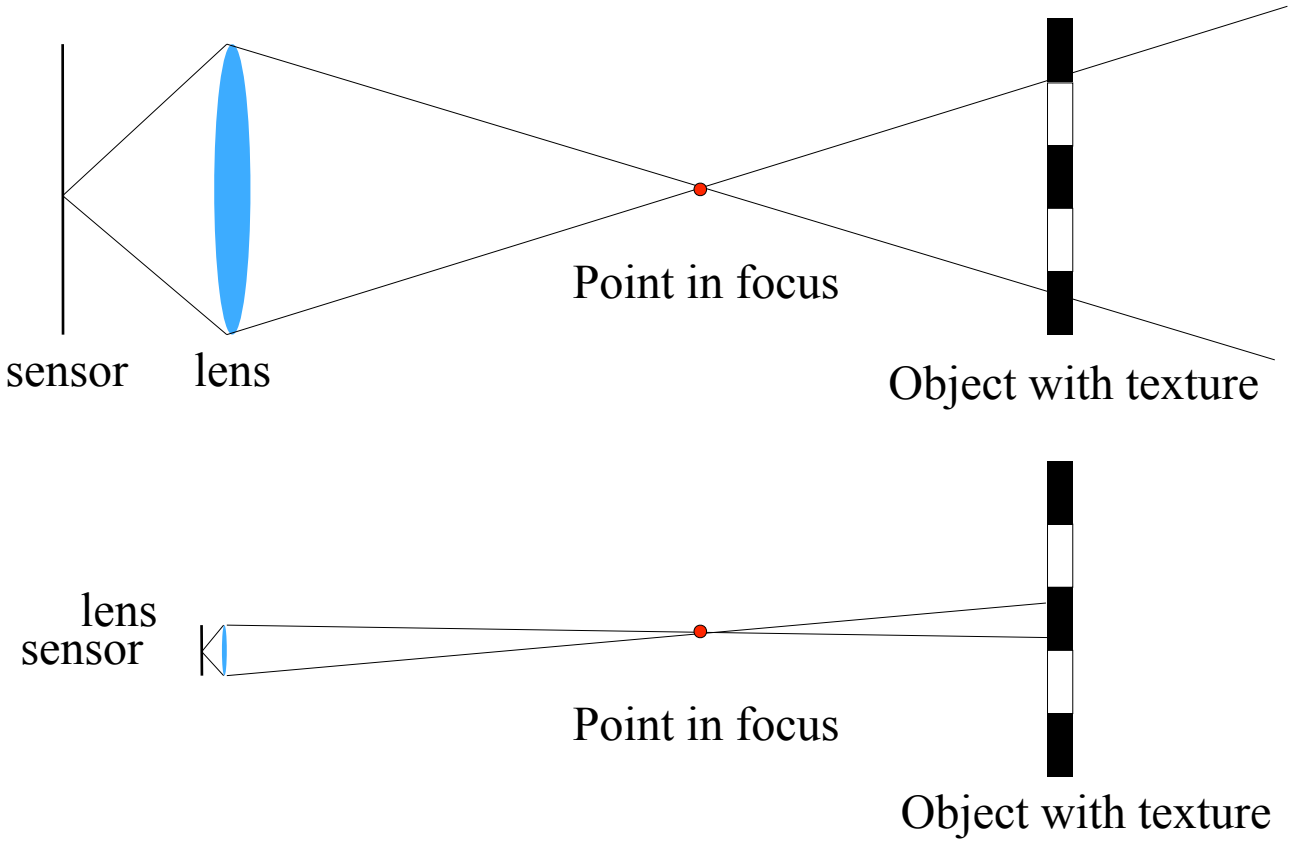


From Macro Photography

Fake depth of field

Challenge

- **Shallow depth of field hard with small sensors**



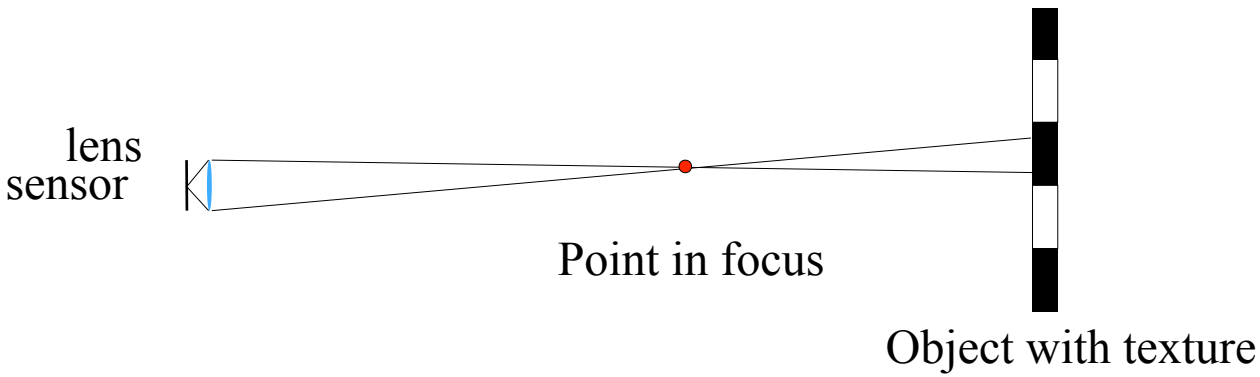
Challenge



(a) large sensor (22.2 x 14.8 mm)



(b) small sensor (7.18 x 5.32 mm)



- **https://people.csail.mit.edu/soonmin/dof/**

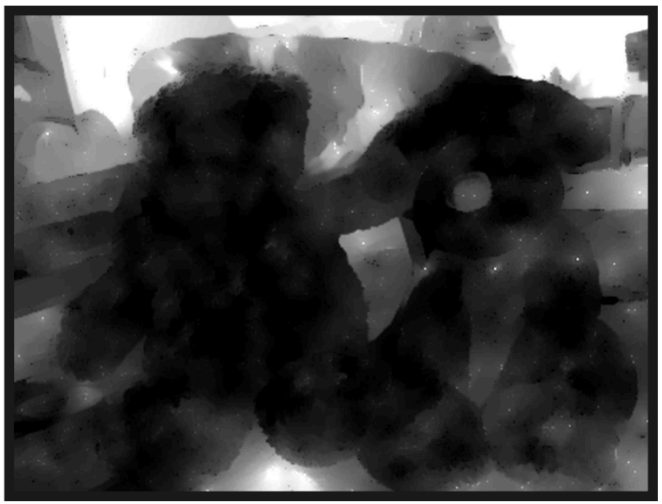
Defocus Magnification

Soonmin Bae Frédo Durand

Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology



(a) input



(b) defocus map



(c) our result with magnified defocus

Figure 1: *Our technique magnifies defocus given a single image. Our defocus map characterizes blurriness at edges. This enables shallow depth of field effects by magnifying existing defocus. The input photo was taken by a Canon PowerShot A80, a point-and-shoot camera with a sensor size of 7.18 × 5.32 mm, and a 7.8 mm lens at f/2.8.*

Portrait mode

- https://research.google/blog/portrait-mode-on-the-pixel-2-and-pixel-2-xl-smartphones/
- https://arxiv.org/pdf/1806.04171





- **Machine learning-based foreground-background segmentation**
- **Use dual pixels to get a depth map (or analyze local defocus)**
- **Predict desired per-pixel blur radius**
- **Render with non-linear blur (more later)**

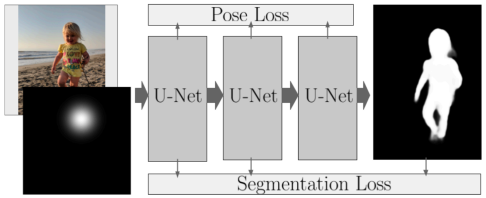
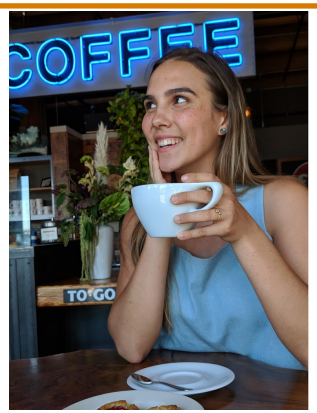
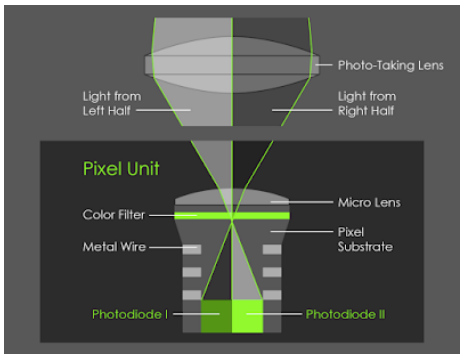


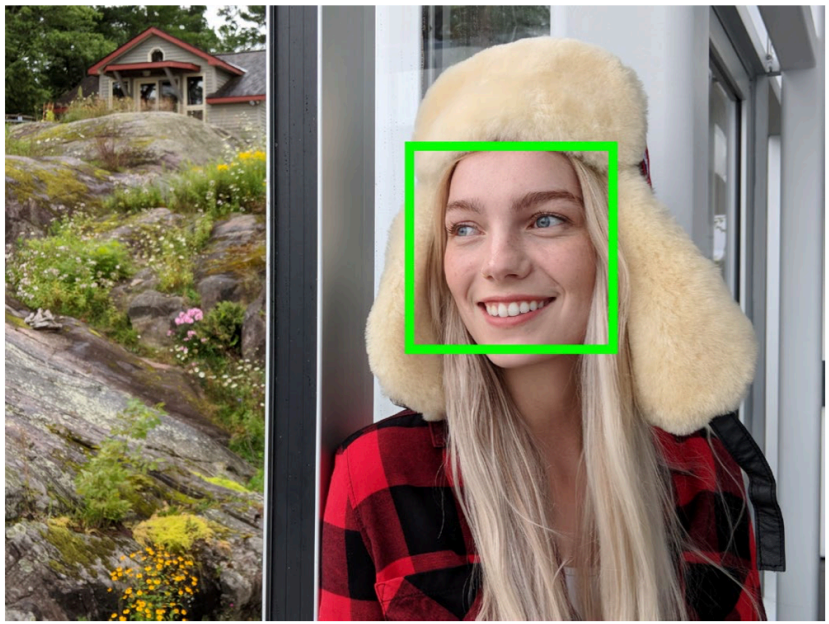
Fig. 3. **Person Segmentation Network.** RGB image (top-left) and face location (bottom-left) are the inputs to a three stage model with pose and segmentation losses after each stage.











(a) Input image with detected face



(b) Person segmentation mask



(c) Mask + disparity from DP



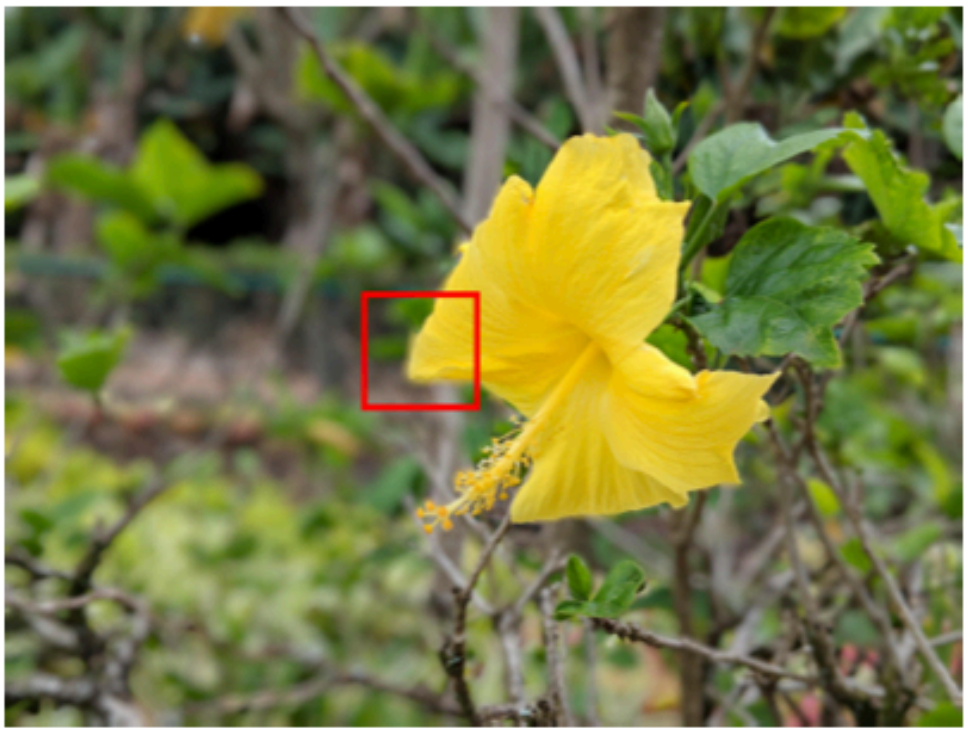
(d) Our output synthetic shallow depth-of-field image

Fig. 1. We present a system that uses a person segmentation mask (b) and a noisy depth map computed using the camera's dual-pixel (DP) auto-focus hardware (c) to produce a synthetic shallow depth-of-field image (d) with a depth-dependent blur on a mobile phone. Our system is marketed as "Portrait Mode" on several Google-branded phones.

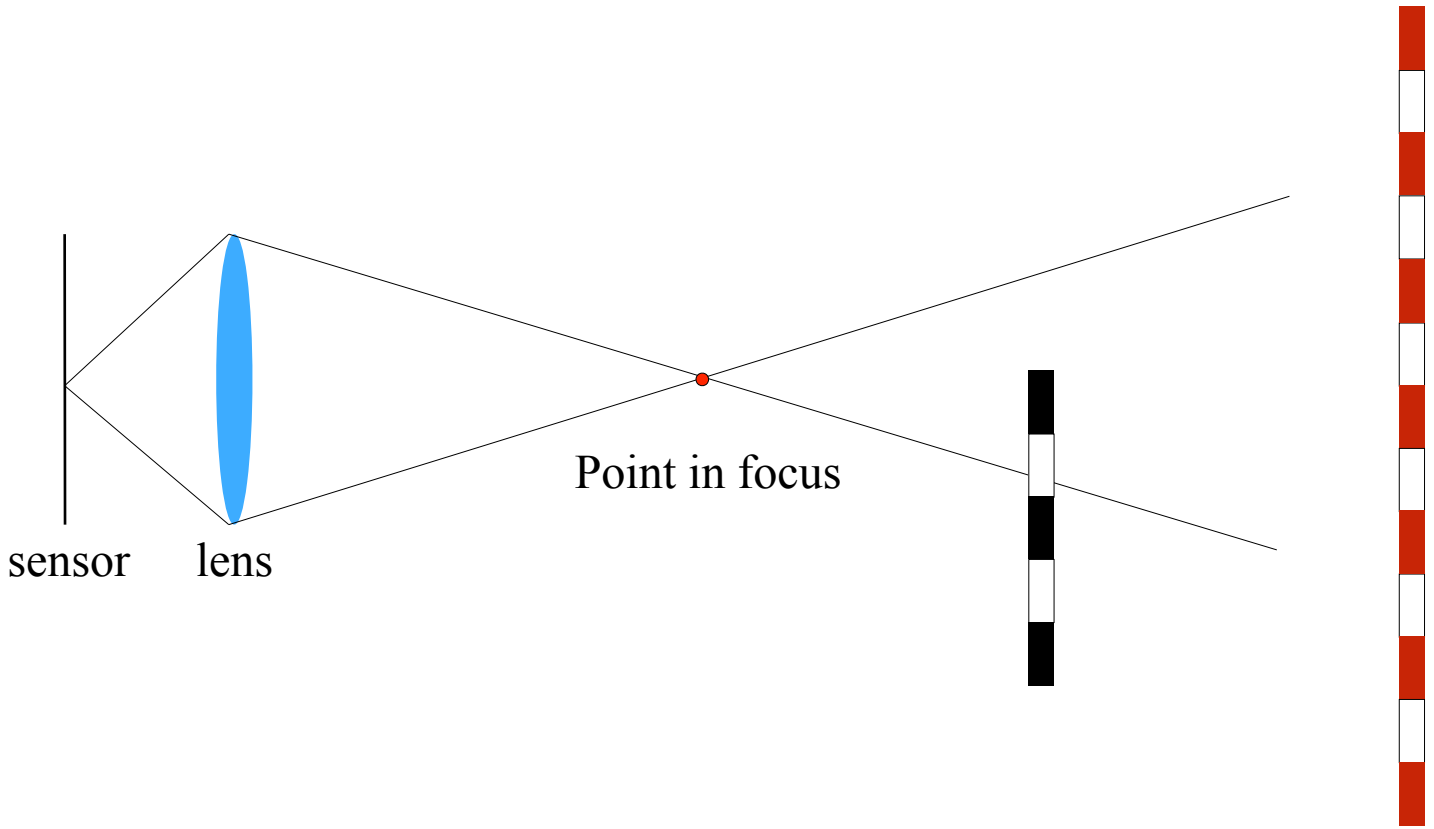
Challenge with defocus blur

- **Remember, not a convolution!**
- **Blur radius depends on input scene depth, not output pixel location**
- **Plus occlusion matters**

Naïve gather






Point in focus
sensor
lens

$$B_{scatter}(\mathbf{x}) = \sum_{\Delta \mathbf{x}} I(\mathbf{x} + \Delta \mathbf{x}) K_{\mathbf{x}+\Delta \mathbf{x}}(-\Delta \mathbf{x}).$$

Occlusion

- **Divide the image into k (=5) depth layers with opacity values**
- **Process each layer independently**
- **Composite back to front (with “over” operator)**


THC




COFFEE
TO GO


COFFEE
TO GO

Challenges

- **Masks are hard, especially around hair!**
- **Need notion of sub-pixel coverage**
-

Some more recent work

- https://arxiv.org/abs/1904.05822
- https://imaging.cs.cmu.edu/dual_pixels/

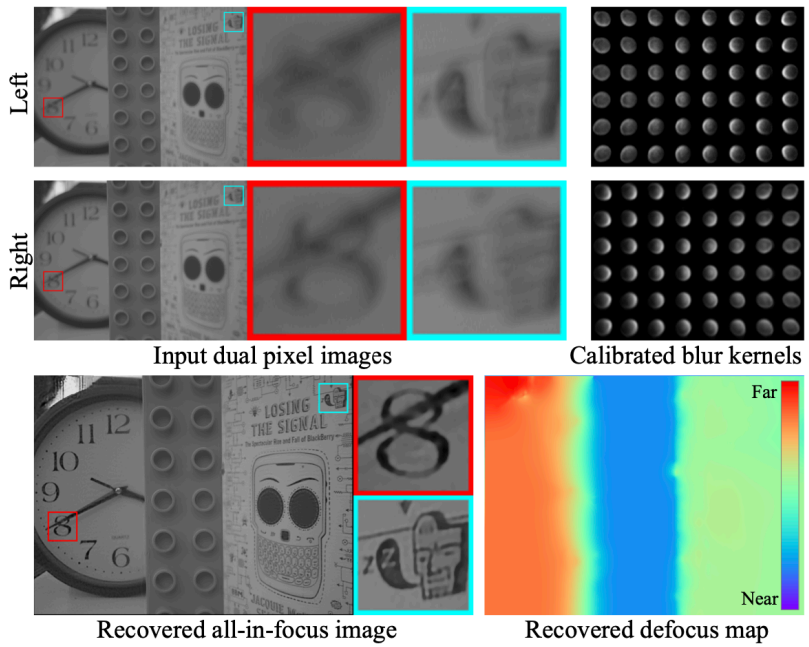


Figure 1: Given left and right dual-pixel (DP) images and corresponding spatially-varying blur kernels, our method jointly estimates an all-in-focus image and defocus map.

Wavefront coding

Hardware depth of field solution

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

Traditional Optical System Image



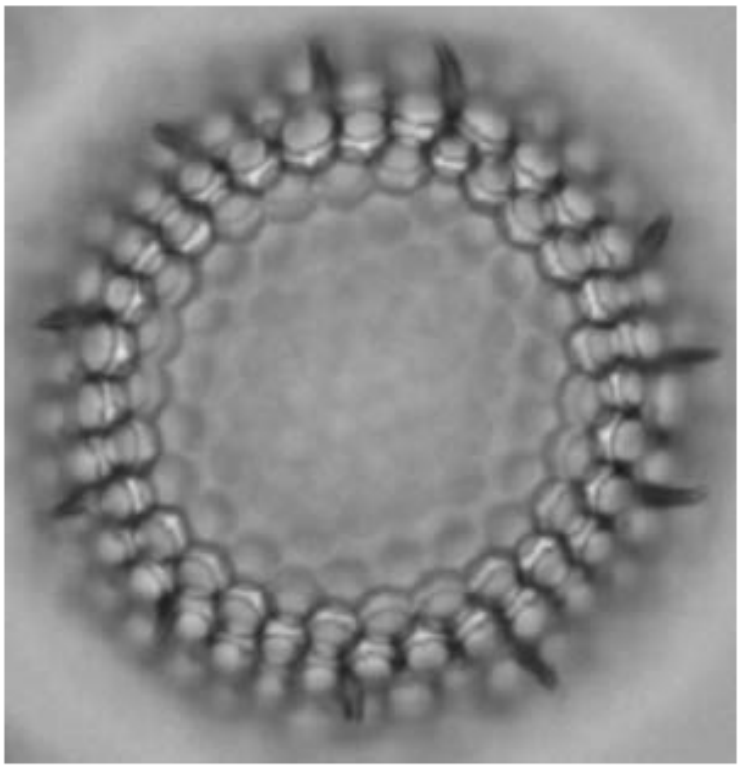
Stopped Down Traditional System Image



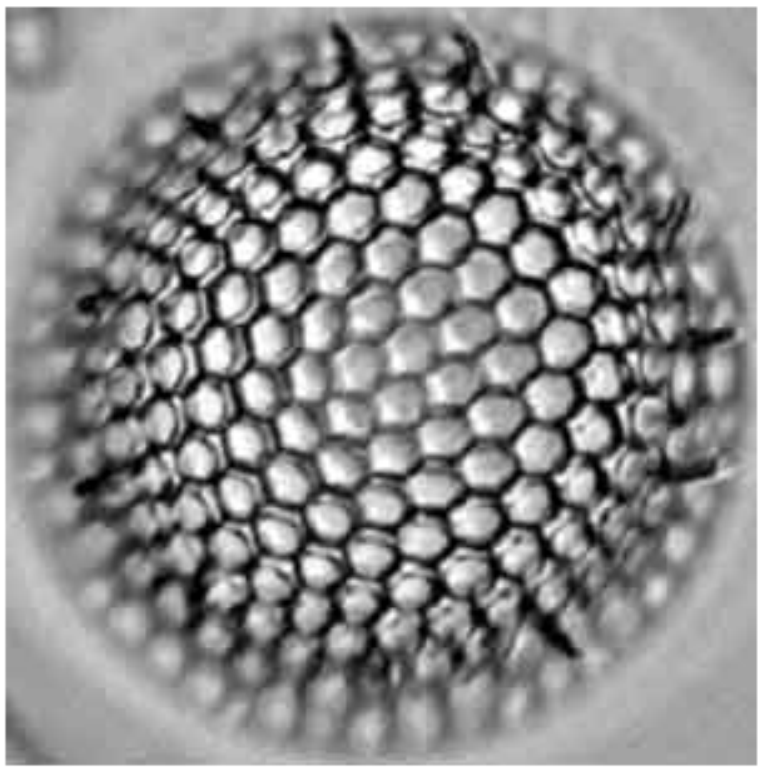
http://www.cdm-optics.com/site/publications.php

Wavefront coding

- **CDM-Optics, U of Colorado, Boulder**
- **Improve depth of field using weird optics & deconvolution**
- **http://www.cdm-optics.com/site/publications.php**
  - The worst title ever: "A New Paradigm for Imaging Systems", Cathey and Dowski, Appl. Optics, 2002



Single-cell algae imaged without wavefront coding.



Single-cell algae imaged with wavefront coding.

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

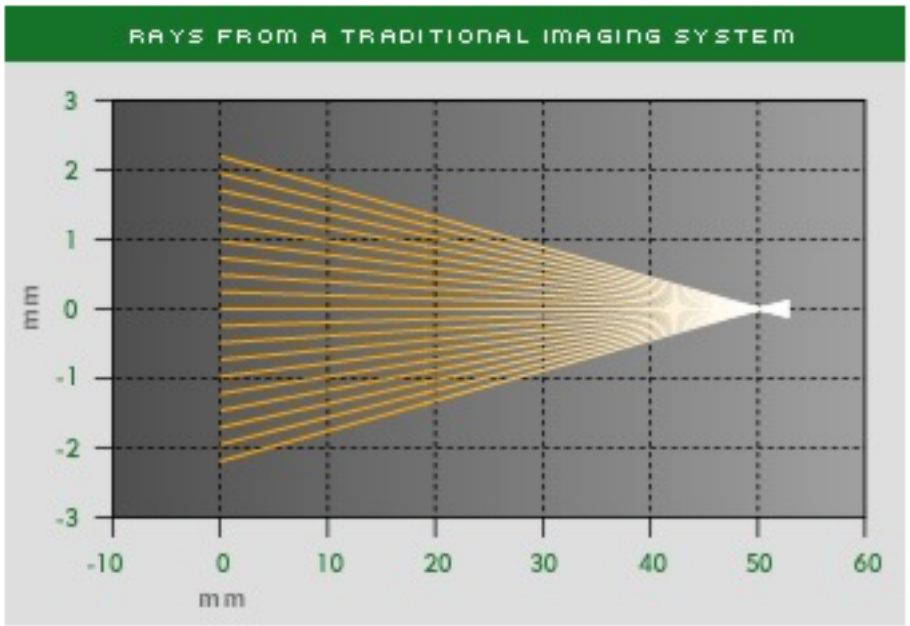
Wavefront coding

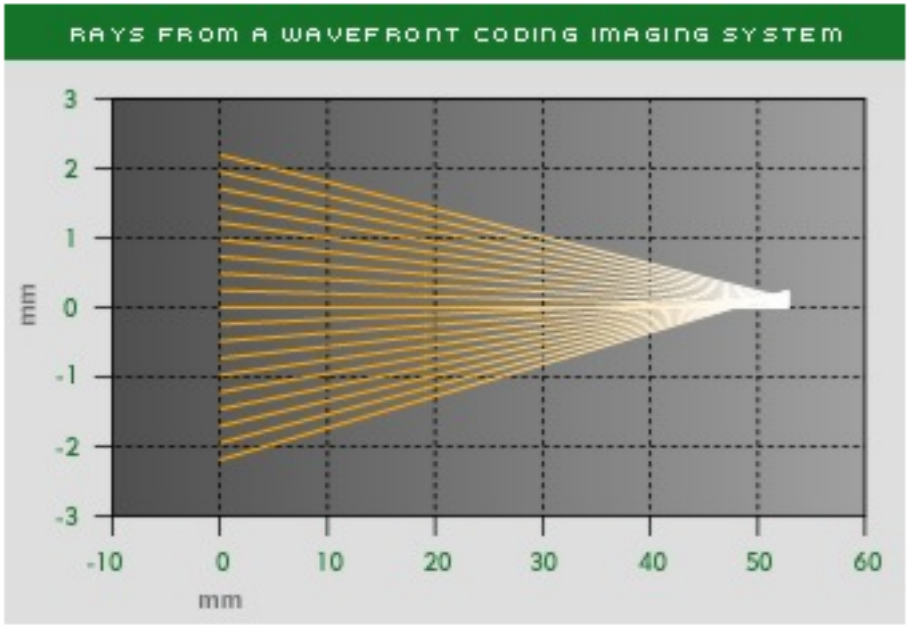
- **Idea: deconvolution to deblur out of focus regions**
  - Problem 1: depth of field blur is not shift-invariant
  - Problem 2: Depth of field blur "kills information"
- **Solution: change optical system so that**
  - Rays don't converge anymore
  - Image blur is the same for all depth
  - Blur spectrum is higher

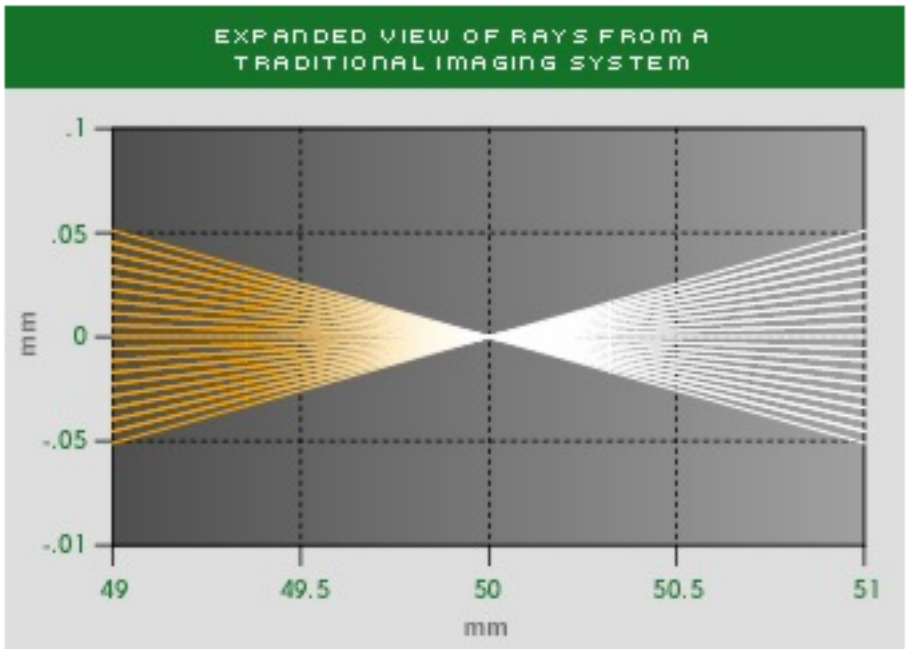
Wavefront coding

- **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

Ray version


RAYS FROM A TRADITIONAL IMAGING SYSTEM
3
2
1
0
-1
-2
-3
mm
-10
0
10
20
30
40
50
60
mm


RAYS FROM A WAVEFRONT CODING IMAGING SYSTEM
3
2
1
0
-1
-2
-3
mm
-10
0
10
20
30
40
50
60
mm


EXPANDED VIEW OF RAYS FROM A TRADITIONAL IMAGING SYSTEM
.1
.05
0
-.05
-.01
mm
49
49.5
50
50.5
51
mm

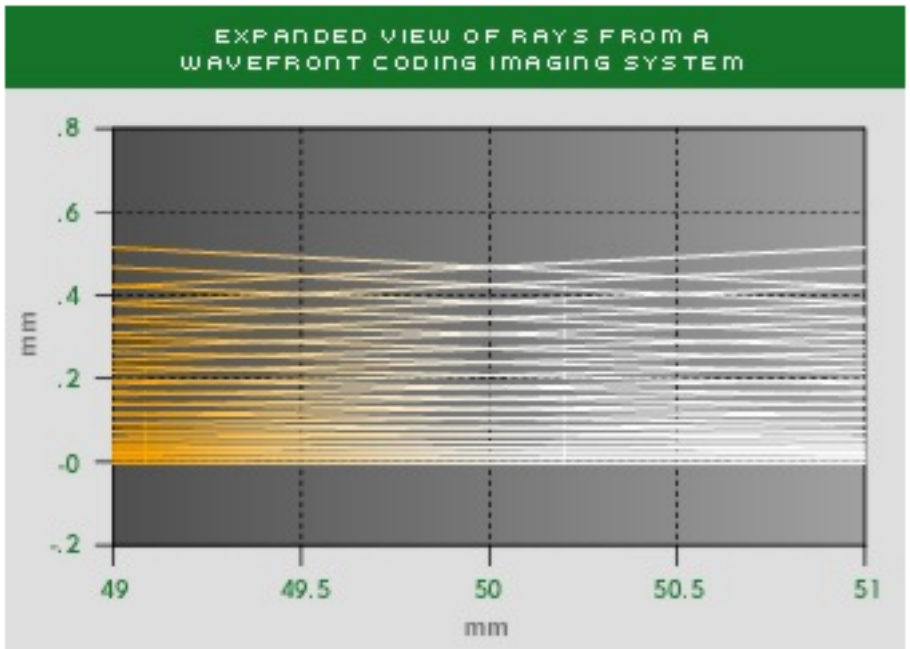
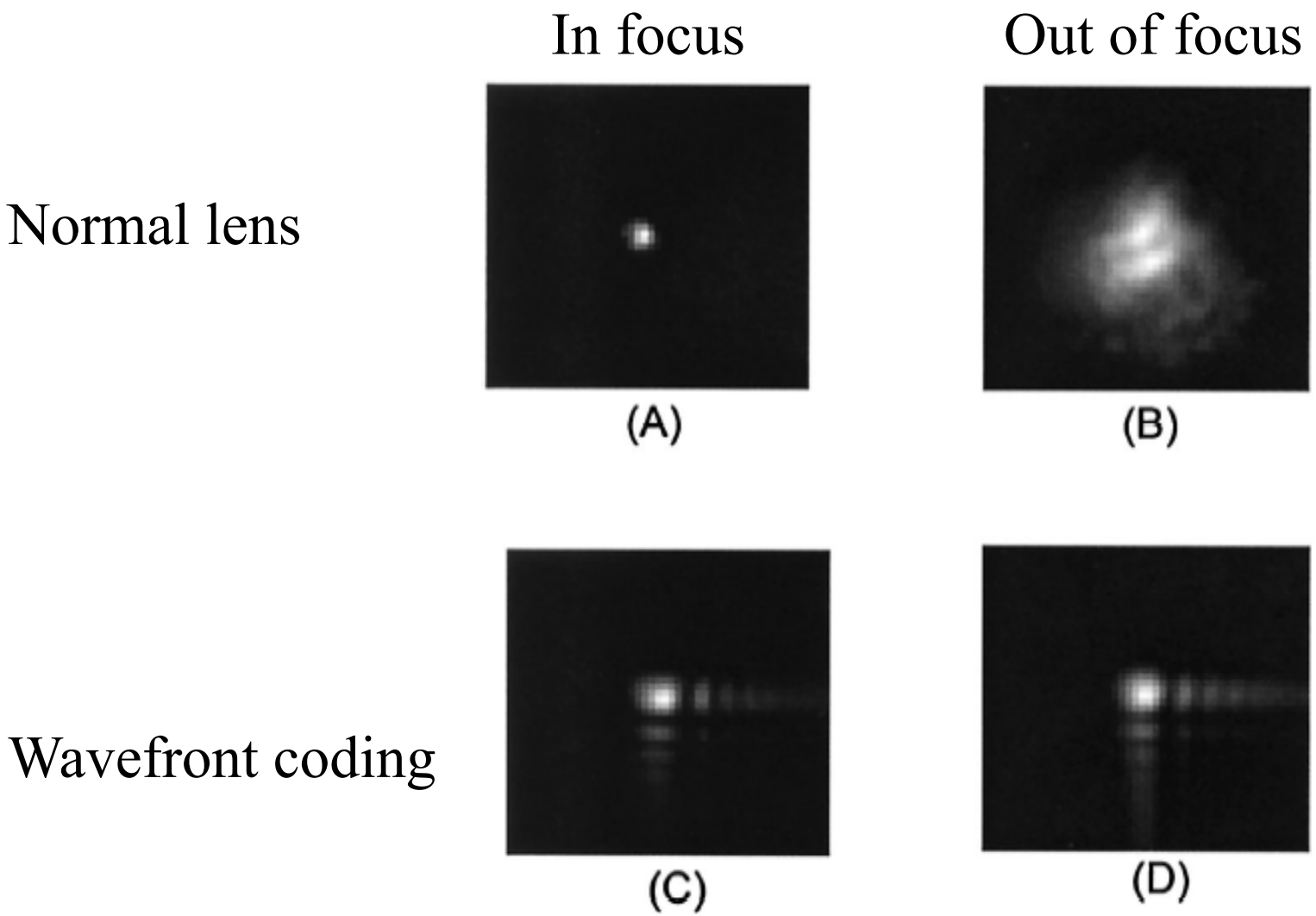

EXPANDED VIEW OF RAYS FROM A WAVEFRONT CODING IMAGING SYSTEM
.8
.6
.4
.2
-0
-.2
mm
49
49.5
50
50.5
51
mm

Image of a point source

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



Frequency response (MTF)

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

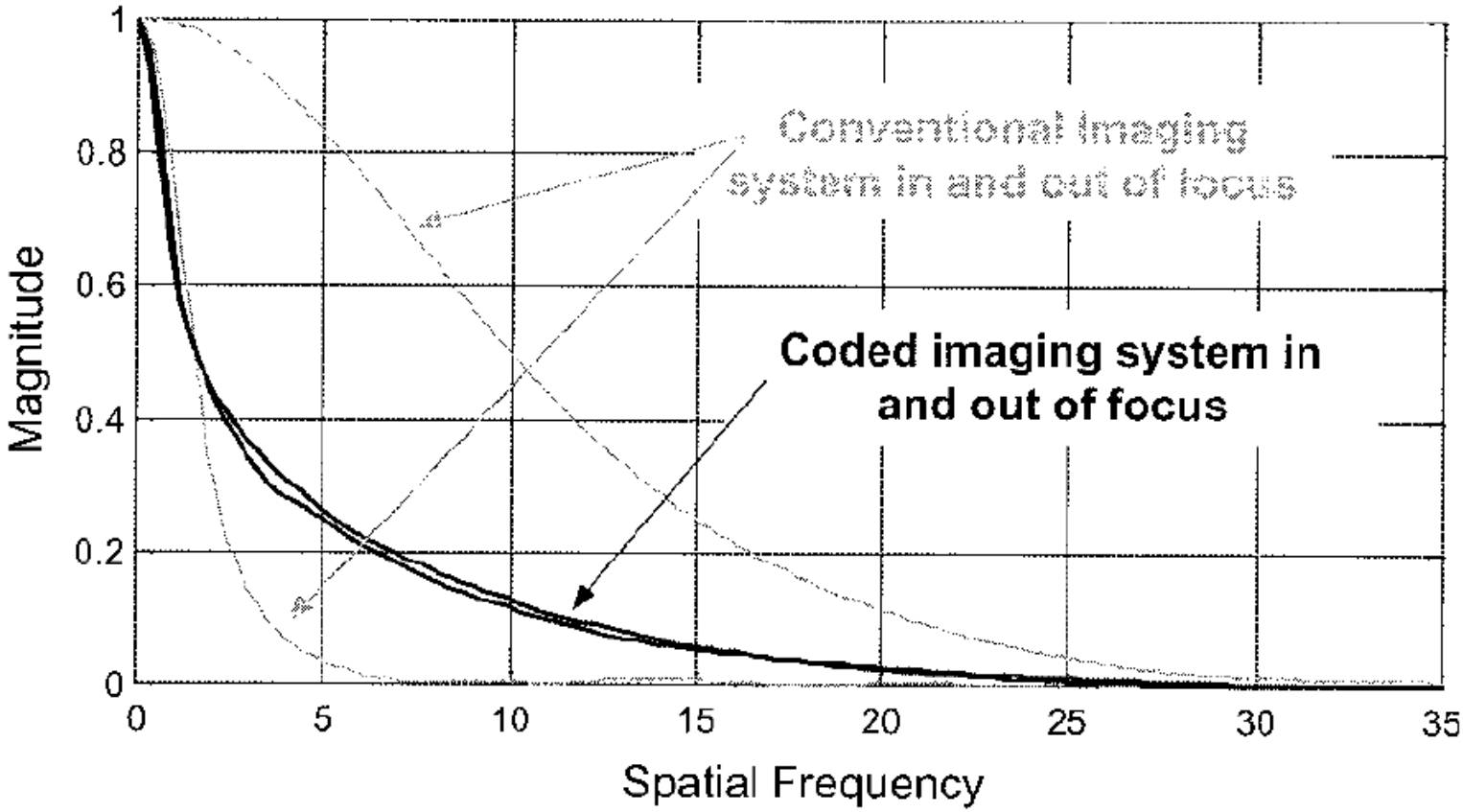


Fig. 5. MTFs corresponding with the PSFs of Fig. 3 for a conventional image in and out of focus and a coded image for the same misfocus values.

Results

Traditional Optical System Image



Intermediate Extended Depth of Field Image



Stopped Down Traditional System Image



Final Wavefront Coded™ Image



Questions?

CSAIL

In General: Coded Imaging

Philosophy: Image capture

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

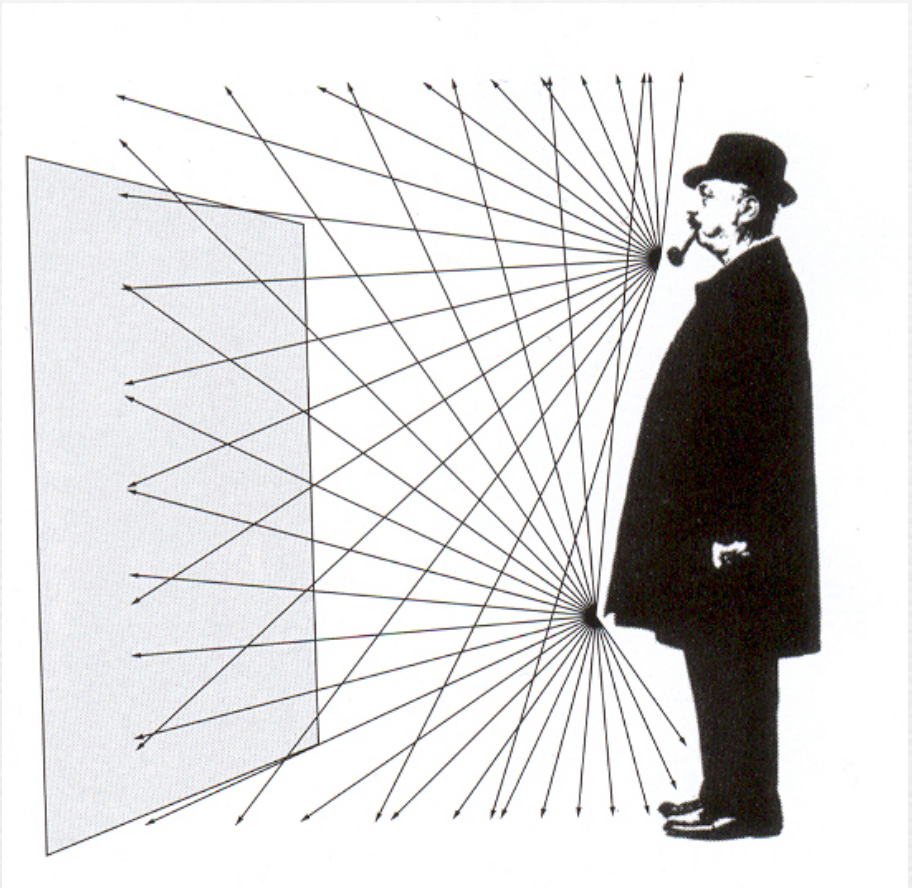


Image capture

- Pinhole allows you to select light rays

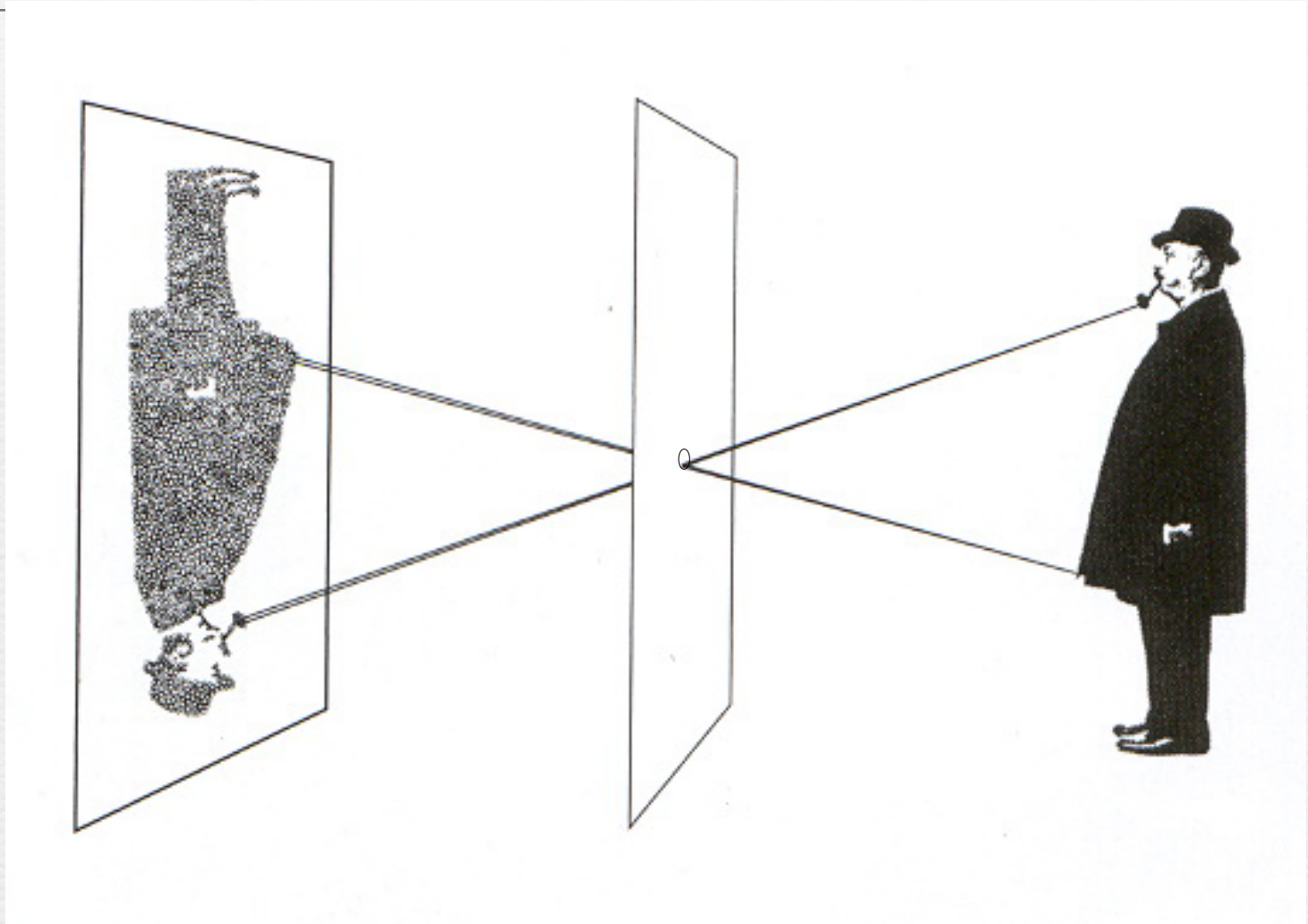


Image formation: optics

- Optics forms an image: selects and integrates light rays

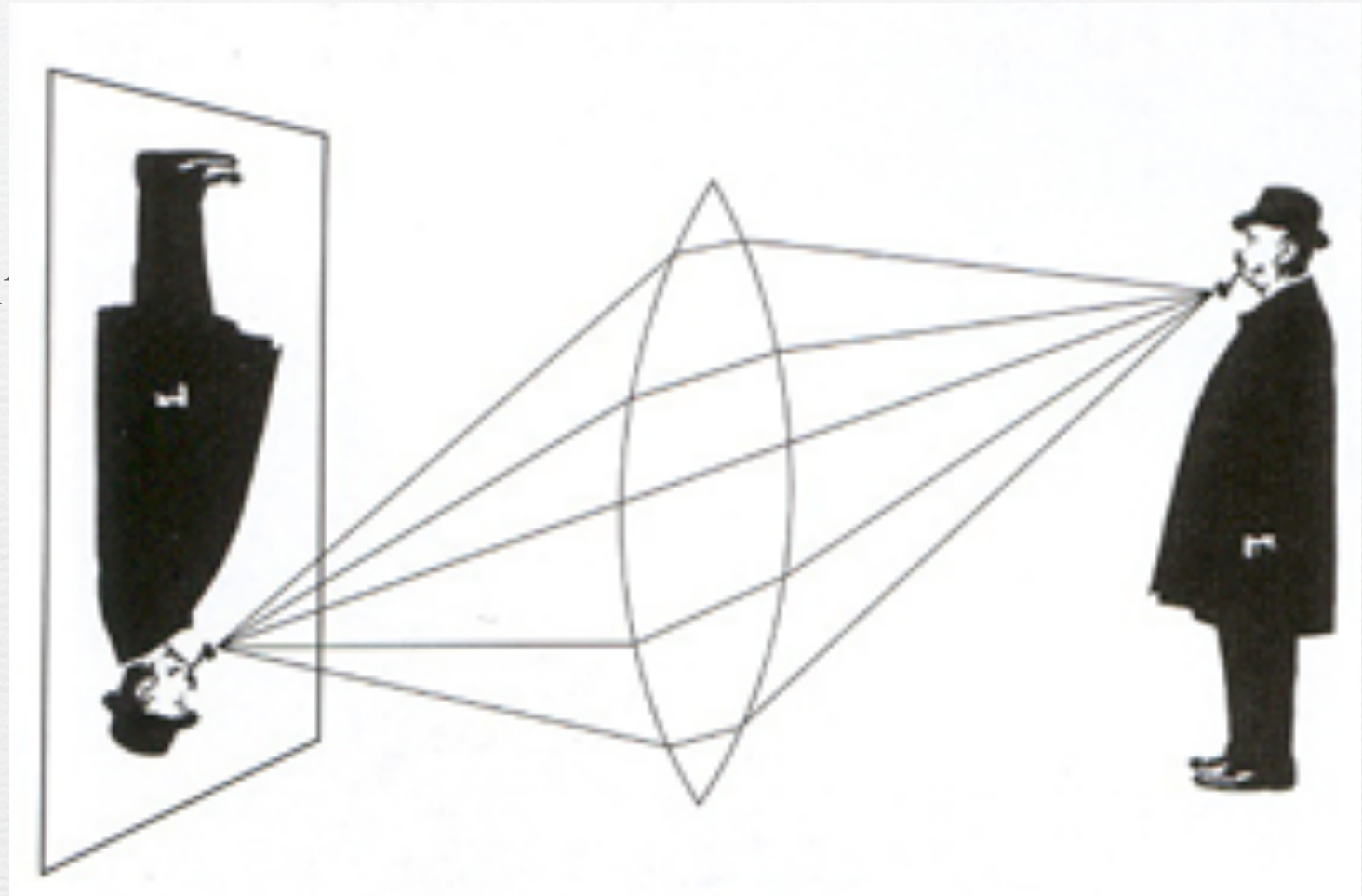


Image formation: computation

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



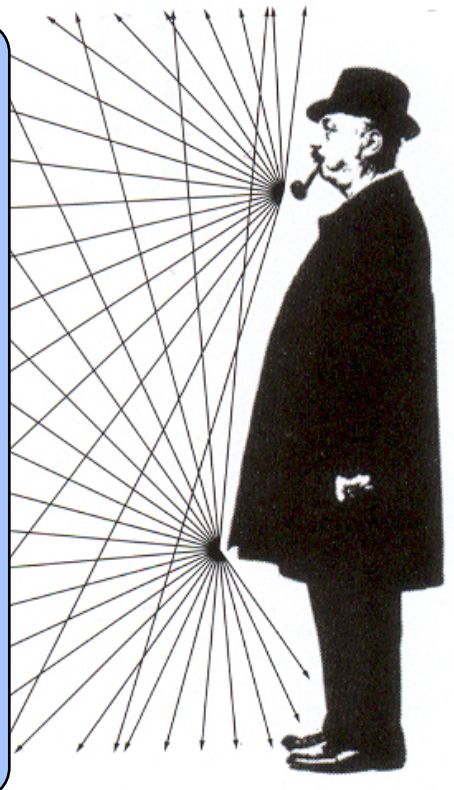
Final image

Computation



Intermediate optical image

Generalized optics



Computational imaging goals

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



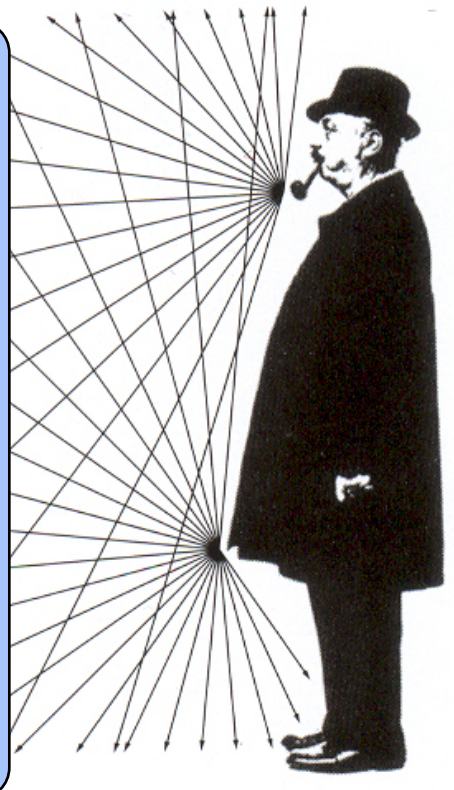
Final image

Computation



Intermediate optical image

Generalized optics



Better capture information

- Same as communication theory:
  optics encodes , computation decodes
- Code seeks to minimize distortion



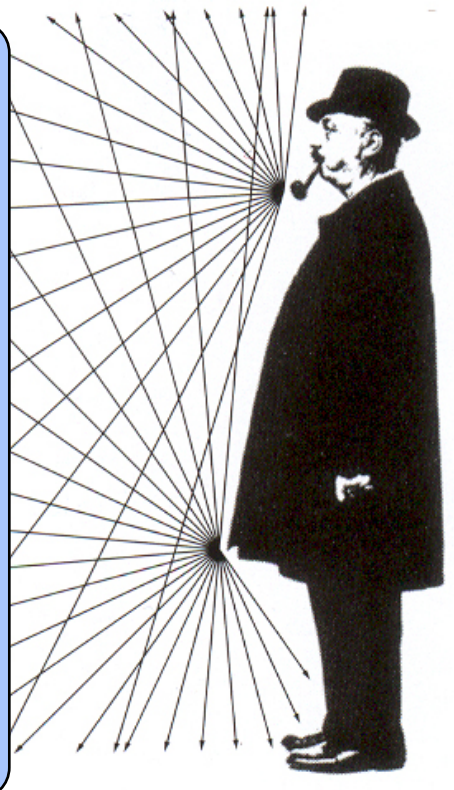
Final image

Computation



Intermediate optical image

Generalized optics



Form images as a post-process

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



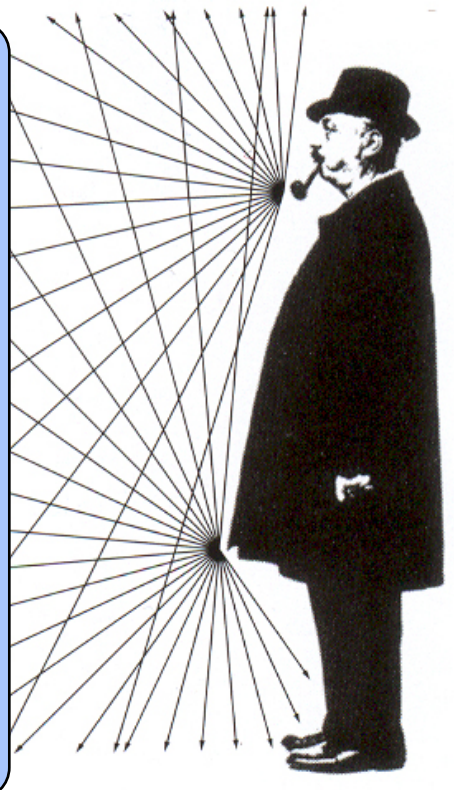
Final image

Computation



Intermediate optical image

Generalized optics

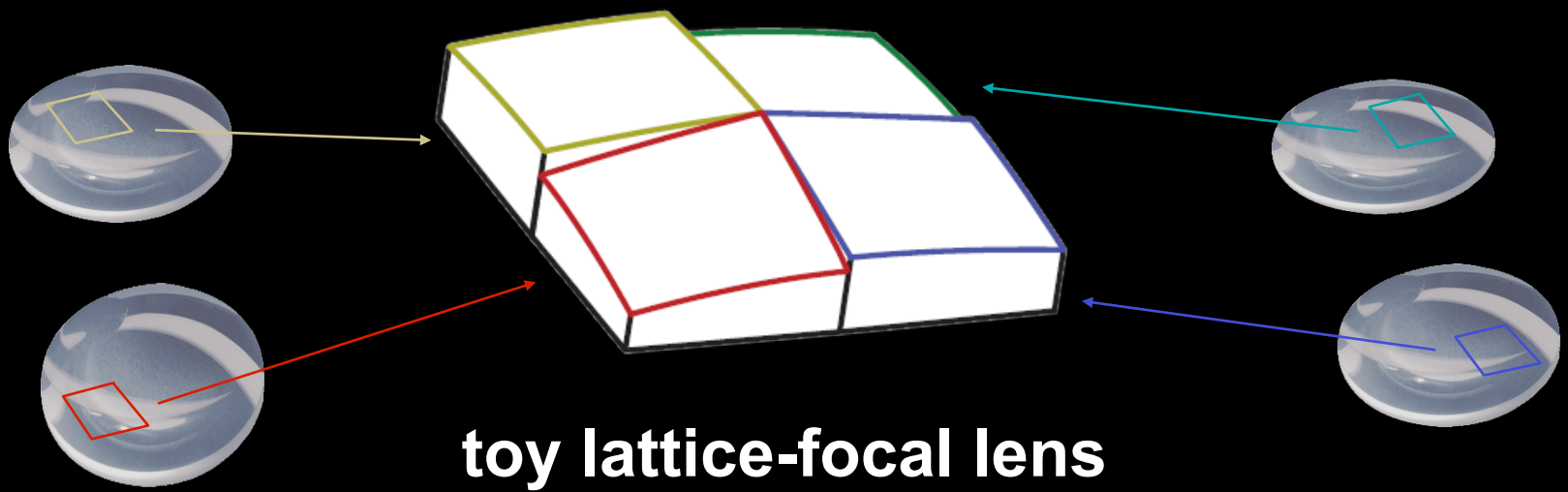


Lattice focal lens

New solution: The lattice-focal lens

Levin et al. 09: assembly of subsquares with different focal powers

➡ **each element focuses on a different depth**



toy lattice-focal lens with 4 elements

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

Hardware construction



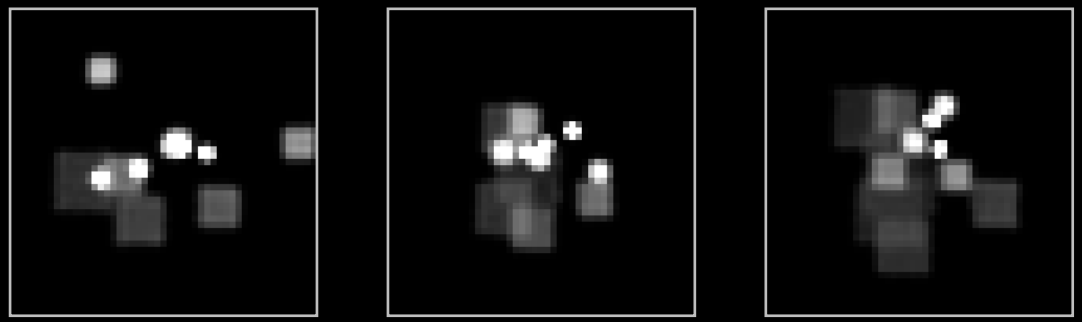
Proof of concept

- 12 subsquares cut from plano-convex spherical lenses
- Attached to main lens

  extra focal power
  eded very low
- Modest DOF extension with only 12 subsquares

Depth estimation

- Defocus kernels vary with depth

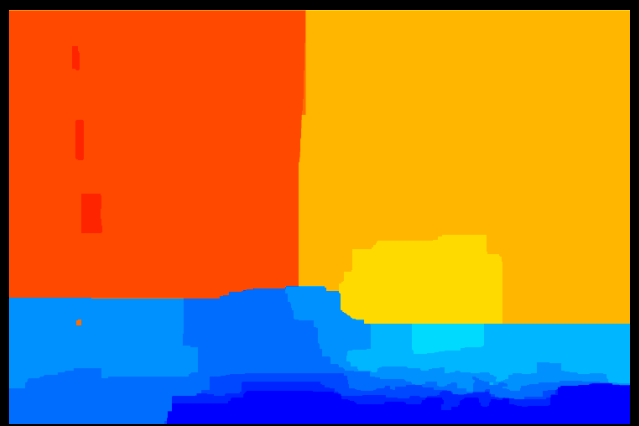


defocus kernels at different depths

- Depth estimation as for the coded aperture camera [Levin et al. 07]



input



depth map

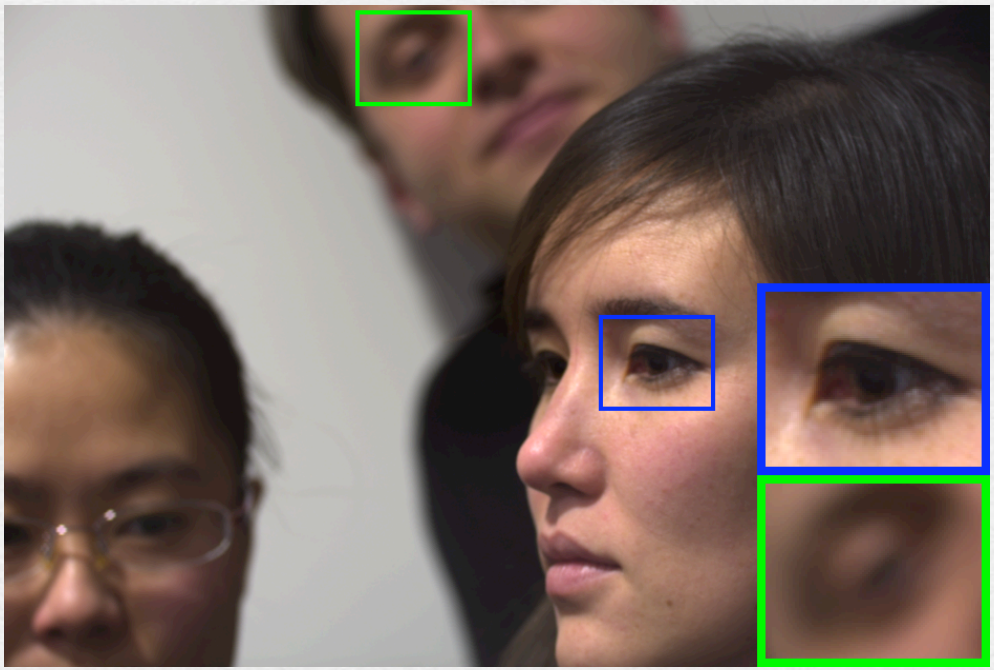
Standard lens reference



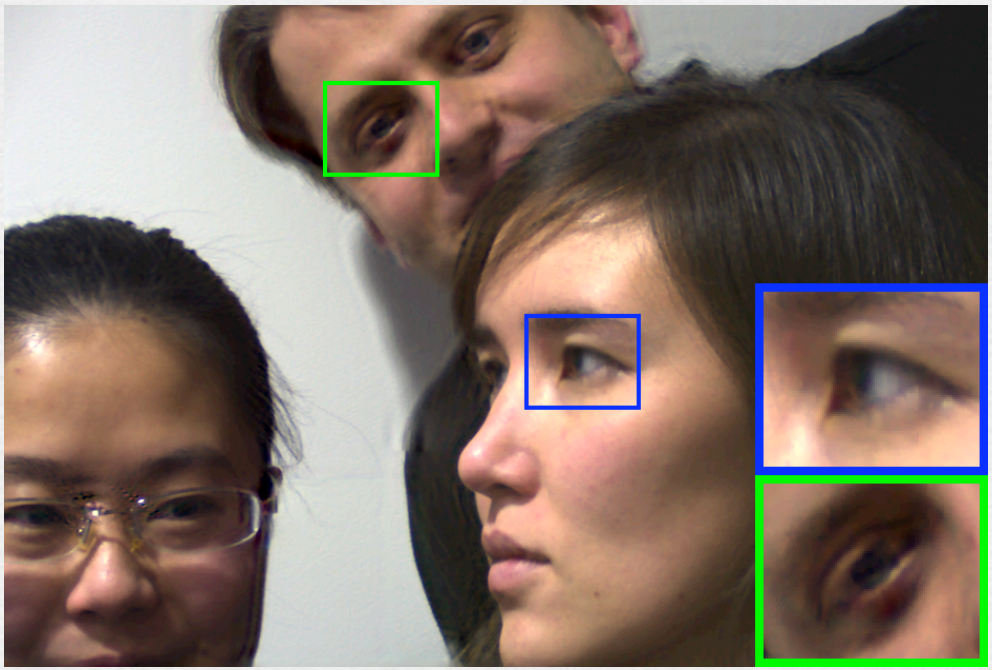
Lattice-focal lens



Lattice focal lens



Standard lens image



Lattice-focal lens: all-focused output

Application: Refocusing from single captured image



Application: Refocusing from single captured image



Application: Refocusing from single captured image

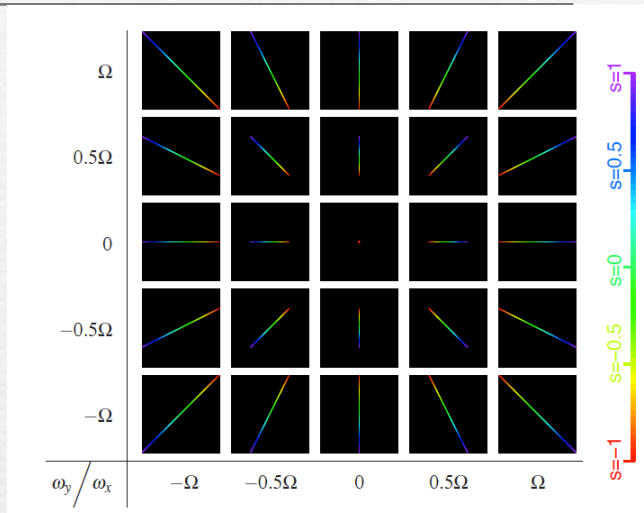


Depth of field analysis

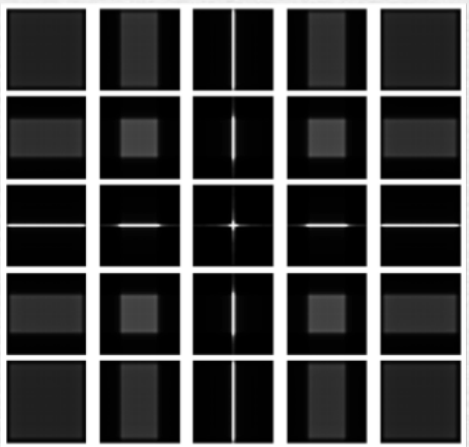
- How do different cameras compare?
- What is the best that can be done?

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

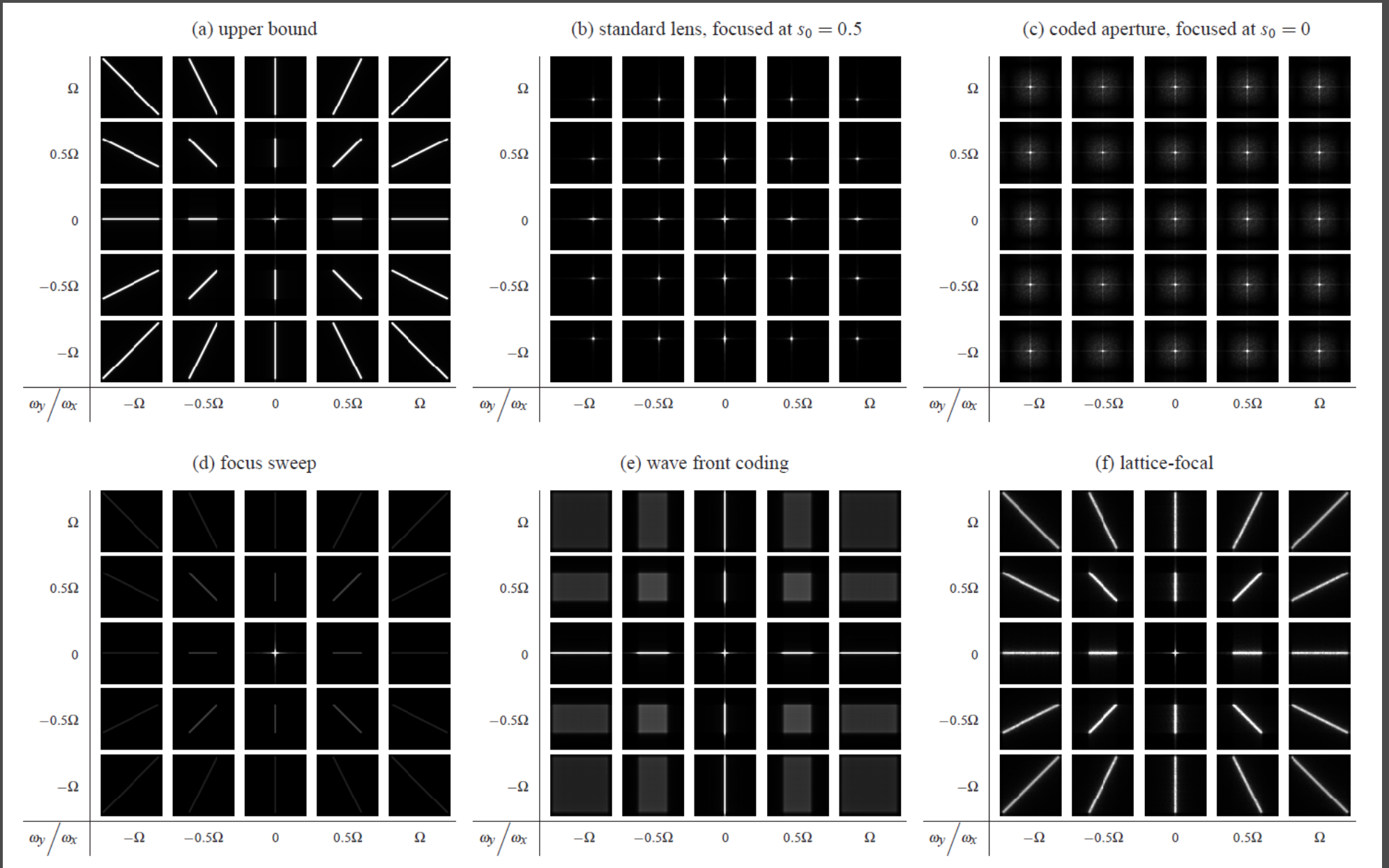


Only a 3D subset of the 4D spectrum is useful



Previous designs spend energy outside the useful subset

Comparison of different cameras

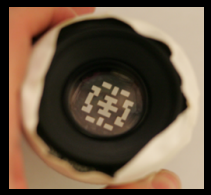


Comparing image reconstruction (simulation)



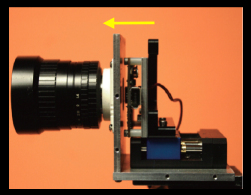
standard lens

<



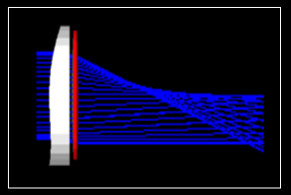
coded aperture

<<



focus sweep

<



wavefront coding

<



lattice-focal lens

Object at in-focus depth



Object at extreme depth



References

- http://www.cdm-optics.com/site/publications.php
- http://www.wisdom.weizmann.ac.il/~levina/papers/lattice/
- http://www1.cs.columbia.edu/CAVE/publications/pdfs/Nagahara_ECCV08.pdf
- http://groups.csail.mit.edu/graphics/CodedAperture/
- http://web.media.mit.edu/~raskar/Mask/

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.tessera.com/Pages/tessera.aspx
- http://www.tessera.com/technologies/imagingandoptics/Documents/OptiML_Focus.pdf