Depth and defocus

Lecture 11

Admin

- Next week:
  - Anat Levin on Matting
  - Tuesday 11am on her SIGGRAPH 2008 paper
- Projects
  - Can everyone send me ½ page telling me where they are/what they've done
  - Projects due Friday 9th 4pm (last class)

Obtaining depth

- Multiple viewpoints (Stereo)
- Special camera (last lecture on May 9th)
- Active methods
  - Change scene illumination
- Passive methods (require scene texture)
  - Defocus analysis
  - Multiple images
  - Single image

Depth from Stereo

- Huge amount of literature in Computer Vision
- Good survey paper: A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. Daniel Scharstein & Richard Szeliski

1 of 2 input images

Ground truth Depth map

Estimated Depth Map

Active methods

- Illuminate scene with light
  - Laser range scan
  - Binary pattern

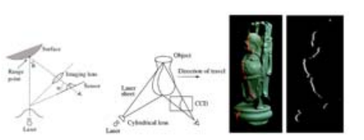
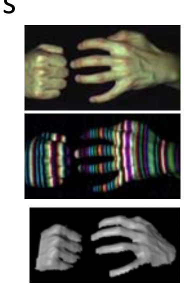


Figure 1. Optical triangulation. (a) 2D triangulation using a laser beam for illumination. (b) Extension to 3D. (c) Red laser line projected onto small (20cm) statuette. (d) Reflected light seen by CCD camera.

http://www.cs.washington.edu/homes/curless/publications/cg99.pdf

Active methods

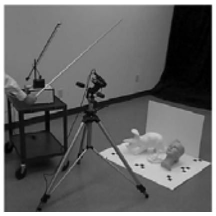
- Illuminate scene with light
  - Binary or colored patterns
  - Can use LCD projectors



Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programming by Li Zhang, Brian Curless, Steven M. Seitz

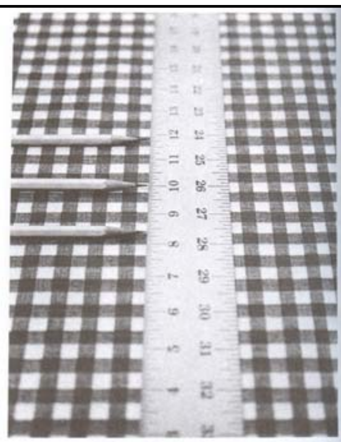
Active methods

- Use shadows:
  - 3D Photography using shadows in dual-space geometry (Bouguet, J.Y & Perona, P.)



Defocus & Depth of field

Slides from Fredo Durand (MIT)



Circle of confusion

From Basic Photographic Materials and Processes, Stroebel et al.

Depth of focus

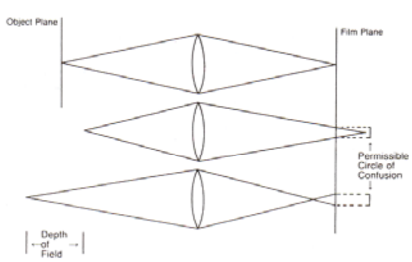


Figure 5–33A Depth of field is the range of distances within which objects are imaged with acceptable sharpness. At the limits, object points are imaged as permissible circles of confusion.

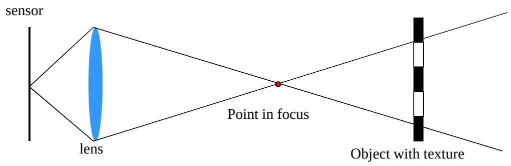
From Basic Photographic Materials and Processes, Stroebel et al.

Size of permissible circle?

- **Assumption on print size, viewing distance, human vision**
  - Typically for 35mm film: diameter = 0.02mm
- **Film/sensor resolution (8μ photosites for high-end SLR)**
- **Best lenses are around 60 lp/mm**
- **Diffraction limit**

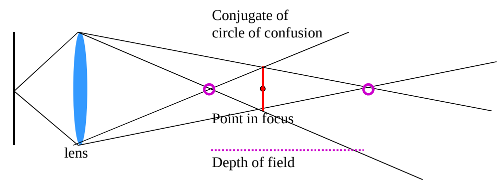
Depth of field: Object space

- **Simplistic view: double cone**
  - Only tells you about the value of one pixel
  - Things are in fact a little more complicated to asses circles of confusion across the image
  - We're missing the magnification factor (proportional to 1/distance and focal length)



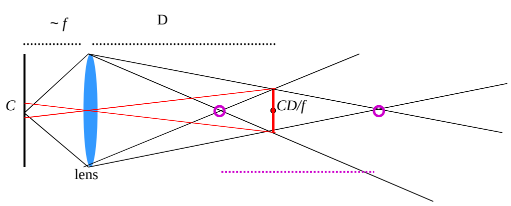
Depth of field: more accurate view

- **Backproject the image onto the plane in focus**
  - Backproject circle of confusion
  - Depends on *magnification factor*
- **Depth of field is slightly asymmetrical**



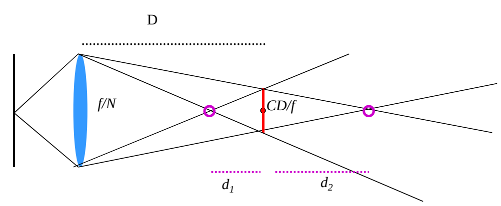
Depth of field: more accurate view

- **Backproject the image onto the plane in focus**
  - Backproject circle of confusion
  - Depends on *magnification factor* ~ *f/D*



Deriving depth of field

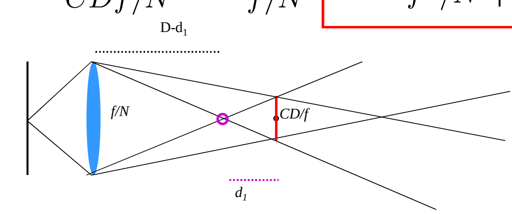
- **Circle of confusion *C*, magnification *m***
- **Simplification: *m=f/D***
- **Focusing distance D, focal length *f*, aperture *N***
- **As usual, similar triangles**



Deriving depth of field

$$\frac{fd_1}{CD} = \frac{D-d_1}{f/N} \qquad \frac{fd_1}{CD} + \frac{d_1}{f/N} = \frac{D}{f/N}$$

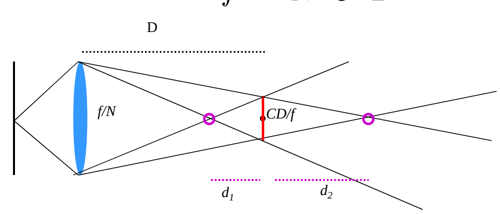
$$d1\frac{f^2/N+CD}{CDf/N} = \frac{D}{f/N} \qquad \boxed{d1 = \frac{CD^2}{f^2/N+CD}}$$



Deriving depth of field

$$d1 = \frac{NCD^2}{f^2+NCD} \qquad d2 = \frac{NCD^2}{f^2-NCD}$$

$$d = d_1 + d_2 = \frac{2NCD^2\, f^2}{f^4 - N^2C^2D^2}$$

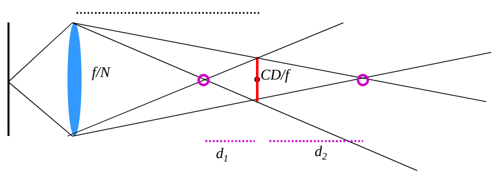


Deriving depth of field

$$d = \frac{2NCD^2 f^2}{f^4 - N^2C^2D^2}$$

$N^2C^2D^2$ term can often be neglected when DoF is small
(conjugate of circle of confusion is smaller than lens aperture)

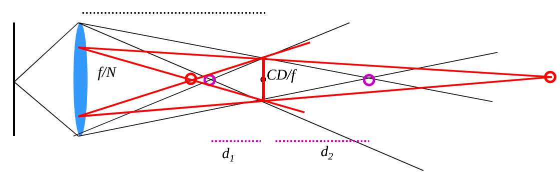
$$d = \frac{2NCD^2}{f^2}$$



Depth of field and aperture

- Linear: proportional to f number
- Recall: big f number N means small physical aperture

$$d = \frac{2NCD^2}{f^2}$$



DoF & aperture

- http://www.juzaphoto.com/eng/articles/depth_of_field.htm

f/2.8

f/32

Depth of field & focal length

- Recall that to get the same image size, we can double the focal length and the distance
- Recall what happens to physical aperture size when we double the focal length for the same f number?
  - It is doubled

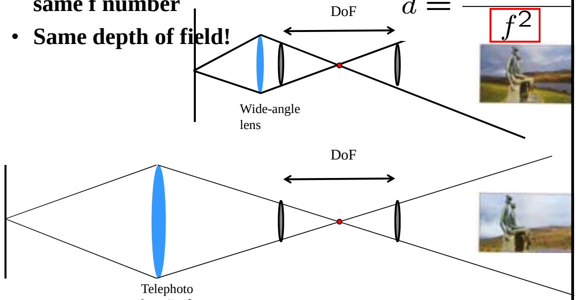
24mm

50mm

Depth of field & focal length

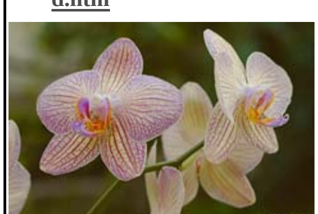
- Same image size (same magnification), same f number
- Same depth of field!

$$d = \frac{2NCD^2}{f^2}$$

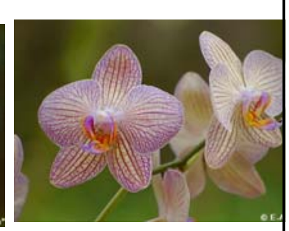


DoF & Focal length

- http://www.juzaphoto.com/eng/articles/depth_of_field.htm



50mm f/4.8



200mm f/4.8
(from 4 times farther)

See also http://luminous-landscape.com/tutorials/dof2.shtml

Important conclusion

- **For a given image size and a given f number, the depth of field (in object space) is the same.**
- **Might be counter intuitive.**
- **Very useful for macro where DoF is critical. You can change your working distance without affecting depth of field**
- **Now what happens to the background blur far far away?**

Important conclusion

- **For a given image size and a given f number, the depth of field (in object space) is the same.**
  - The depth of acceptable sharpness is the same
- **But background far far away looks more blurry Because it gets magnified more**
- **Plus, usually, you don't keep magnification constant**

Effect of parameters

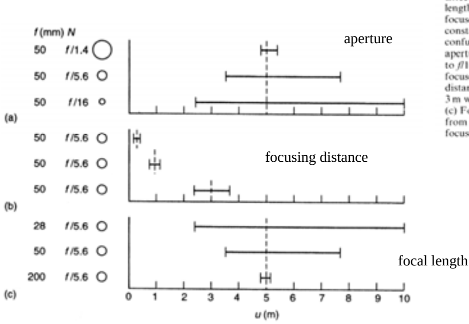


Figure 22.1 Depth of field Effect of the variables focal length (*f*), *f*-number (*N*) and focused distance (*u*) at constant value for circle of confusion (*C*). (a) Lens aperture varying from *f*/1.4 to *f*/16 with 50 mm lens focused on 5 m. (b) Focused distance varying from 0.5 to 3 m with 50 mm lens at *f*/5.6. (c) Focal length varying from 28 to 200 mm at *f*/5.6 focused on 5 m.

From applied photographic optics

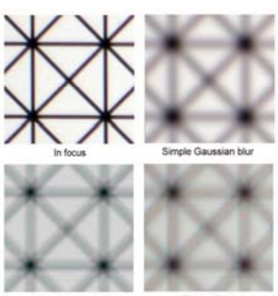
Is depth of field a blur?

- **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 4D space… more about this in ten days)**

From Macro Photography

Bokeh

- Pattern of out-of-focus blur



http://www.bobatkins.com/photography/technical/bokeh.html

Shape depends on aperture

- Also on location within image

Defocused Star Image

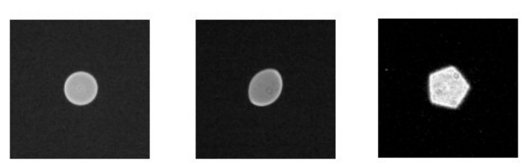


Image center ar f1.8 — Image corner at f1.8 — Image center at f8

http://www.bobatkins.com/photography/technical/bokeh.html

Mirror Lens

- Blur pattern

http://photo.net/learn/optics/mirrors/tamron500-8a.jpg

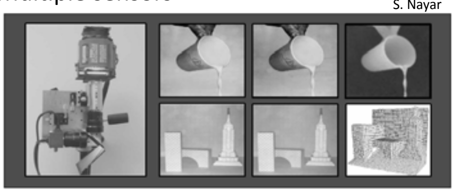
Comparison of lenses

http://www.bobatkins.com/photography/technical/bokeh.html

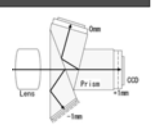
Depth from defocus – Multiple images

- Multiple sensors

S. Nayar



- Time multiplex
  - Need static scene



Confocal Stereo

- Hasinoff & Kutulakos ECCV'06
- Vary aperture and focus of lens
  - Multiple images

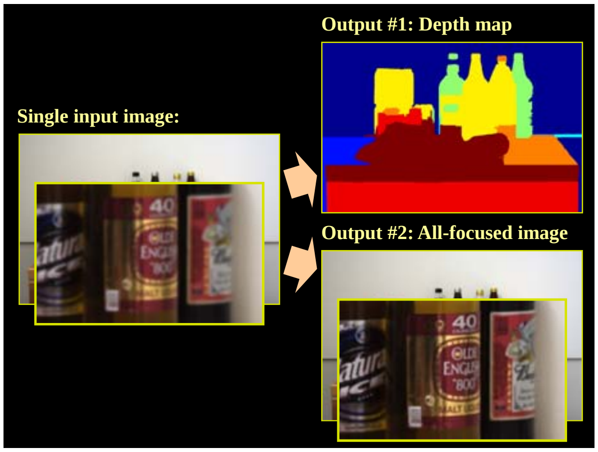
Image and Depth from a Conventional Camera with a Coded Aperture

Anat Levin, Rob Fergus, Frédo Durand, William Freeman

MIT CSAIL

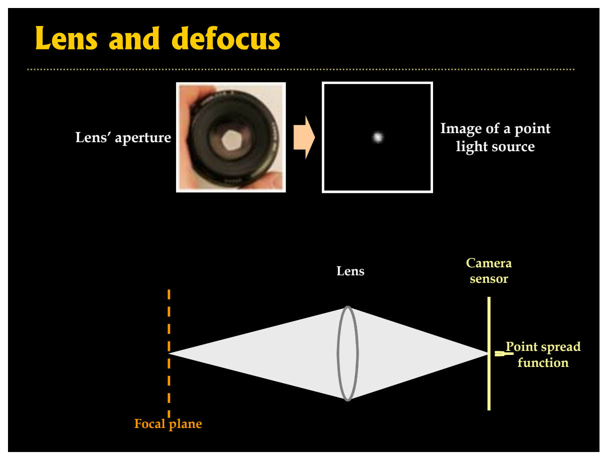
Output #1: Depth map

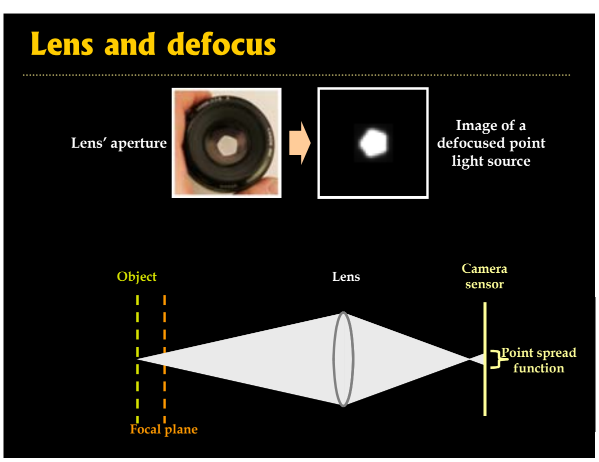
Single input image:

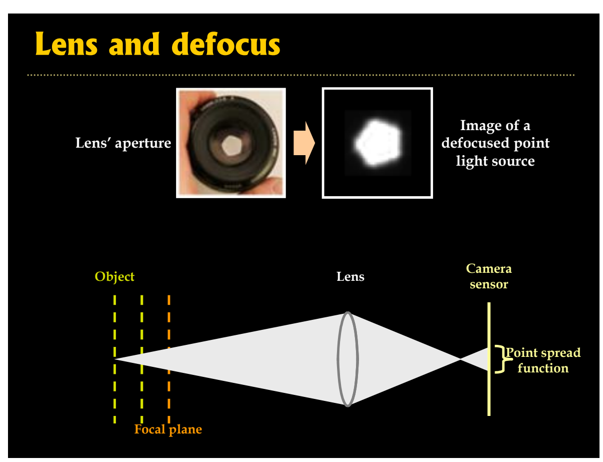

Output #1: Depth map
Single input image:
Output #2: All-focused image

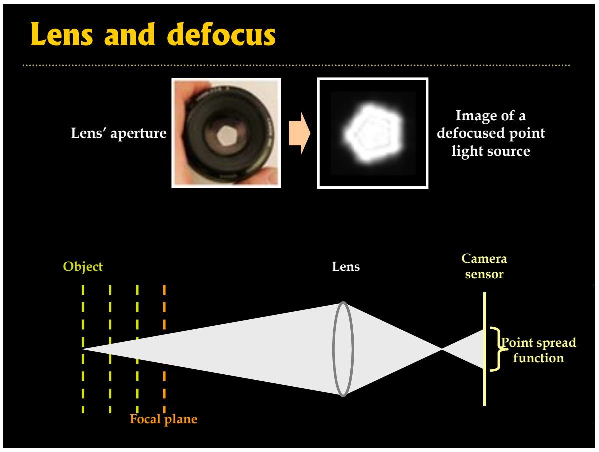

Build your own coded aperture

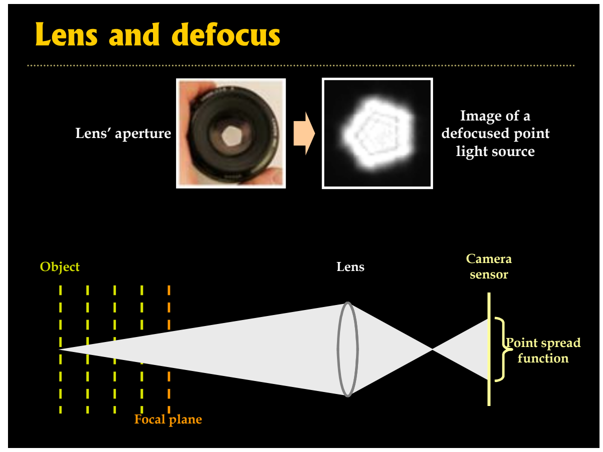

Voila!

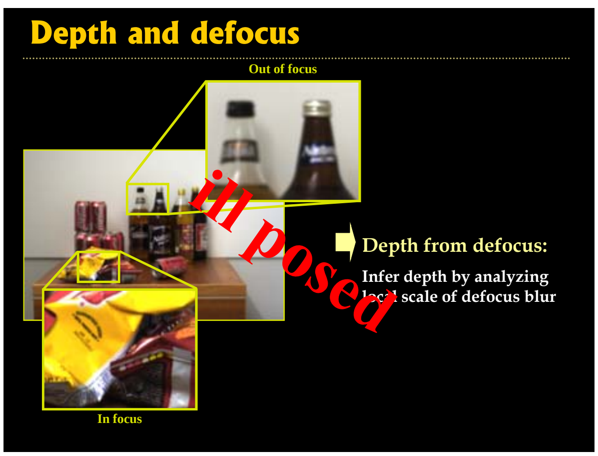

Lens and defocus
Lens' aperture
Image of a point light source
Lens
Camera sensor
Point spread function
Focal plane

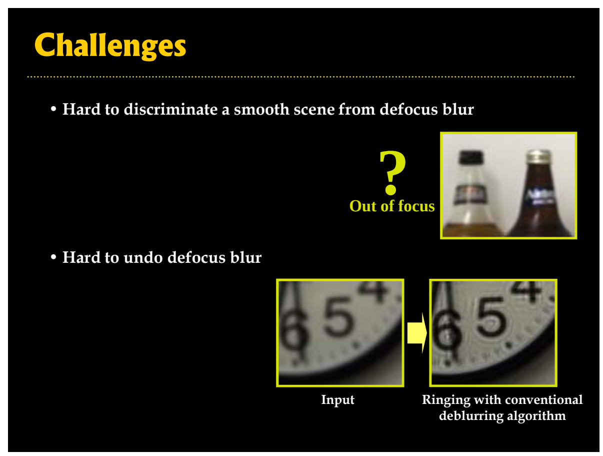

Lens and defocus
Lens' aperture
Image of a defocused point light source
Object
Lens
Camera sensor
Point spread function
Focal plane

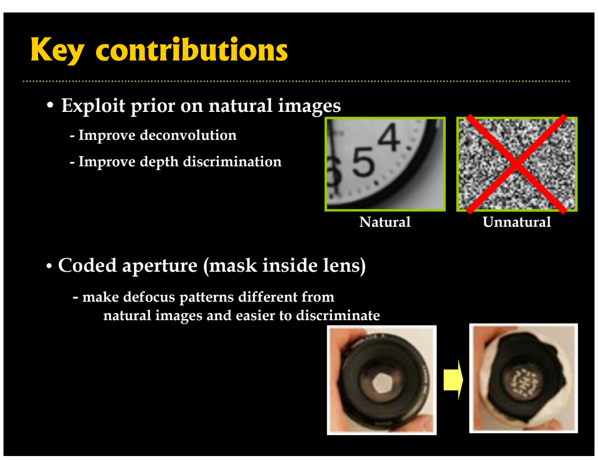

Lens and defocus
Lens' aperture
Image of a defocused point light source
Object
Lens
Camera sensor
Point spread function
Focal plane

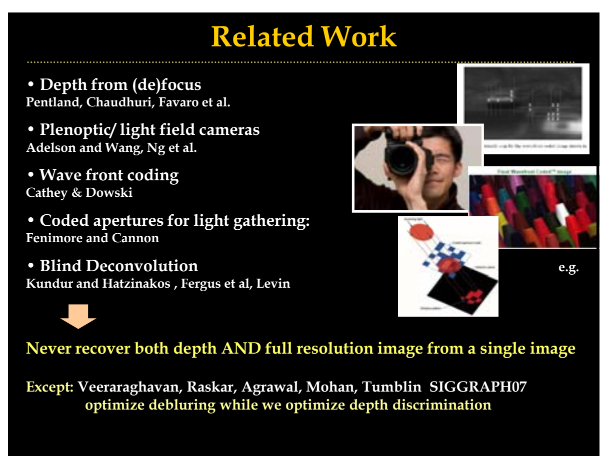

Lens and defocus
Lens' aperture
Image of a defocused point light source
Object
Lens
Camera sensor
Point spread function
Focal plane

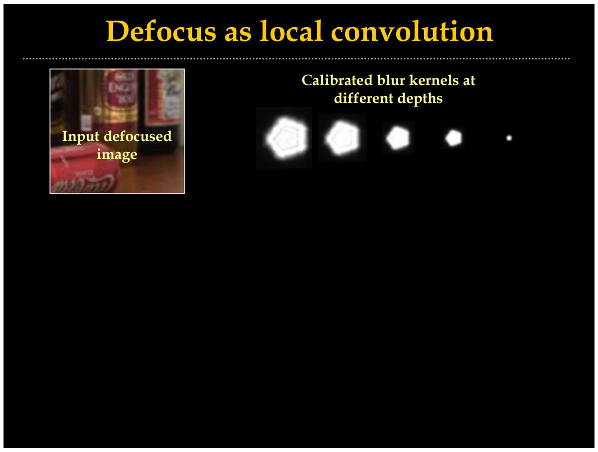

Lens and defocus
Lens' aperture
Image of a defocused point light source
Object
Lens
Camera sensor
Point spread function
Focal plane

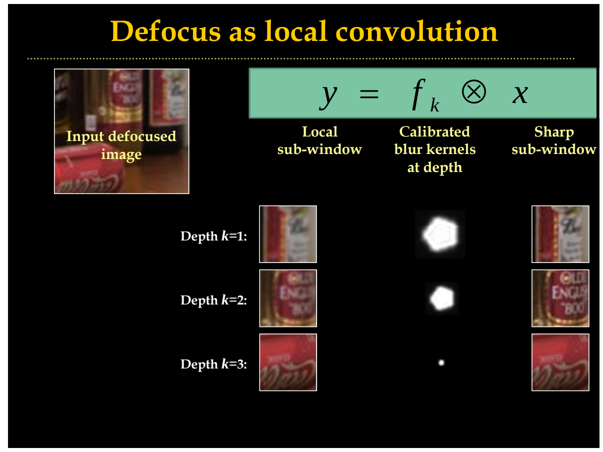

Depth and defocus
Out of focus
ill posed
Depth from defocus:
Infer depth by analyzing local scale of defocus blur
In focus

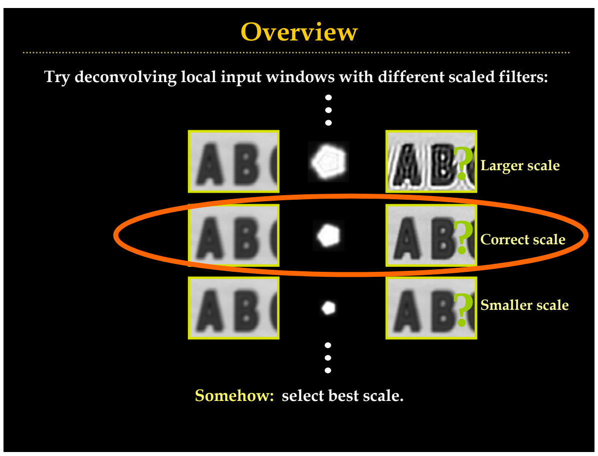

Challenges
• Hard to discriminate a smooth scene from defocus blur
?
Out of focus
• Hard to undo defocus blur
Input
Ringing with conventional deblurring algorithm

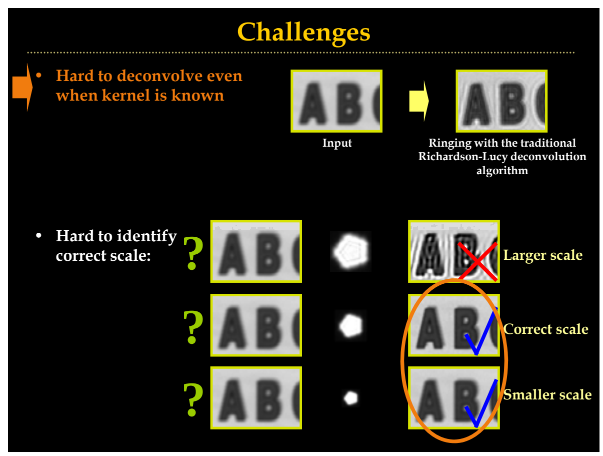

Key contributions
• Exploit prior on natural images
- Improve deconvolution
- Improve depth discrimination
Natural
Unnatural
• Coded aperture (mask inside lens)
- make defocus patterns different from natural images and easier to discriminate

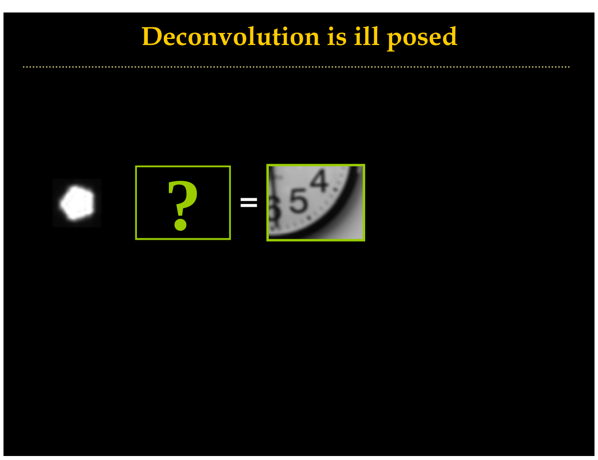

Related Work
• Depth from (de)focus
Pentland, Chaudhuri, Favaro et al.
• Plenoptic/ light field cameras
Adelson and Wang, Ng et al.
• Wave front coding
Cathey & Dowski
• Coded apertures for light gathering:
Fenimore and Cannon
• Blind Deconvolution
Kundur and Hatzinakos , Fergus et al, Levin
e.g.
Never recover both depth AND full resolution image from a single image
Except: Veeraraghavan, Raskar, Agrawal, Mohan, Tumblin SIGGRAPH07
optimize debluring while we optimize depth discrimination

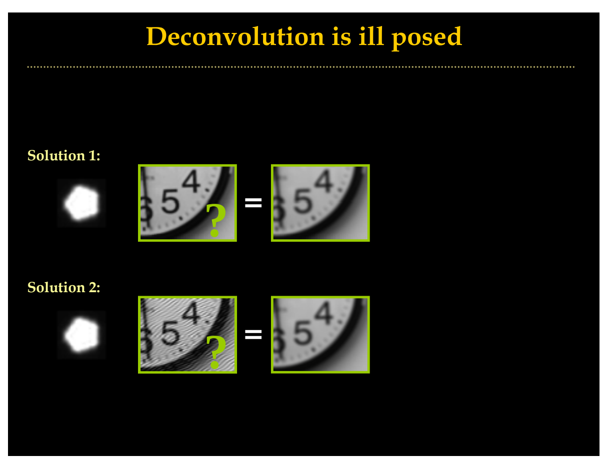

Defocus as local convolution
Input defocused image
Calibrated blur kernels at different depths

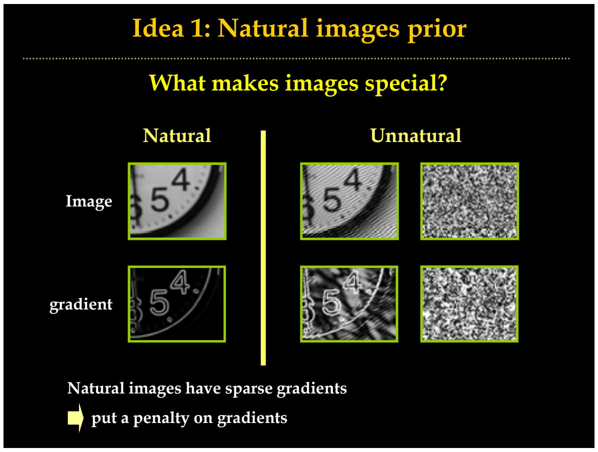

Defocus as local convolution
Input defocused image
$y = f_k \otimes x$
Local sub-window
Calibrated blur kernels at depth
Sharp sub-window
Depth k=1:
Depth k=2:
Depth k=3:

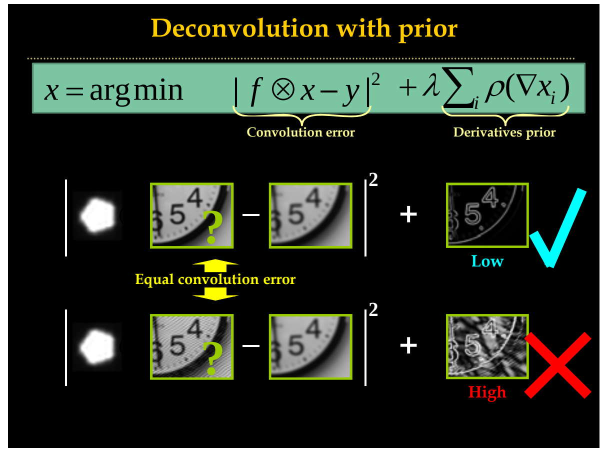

Overview
Try deconvolving local input windows with different scaled filters:
Larger scale
Correct scale
Smaller scale
Somehow: select best scale.

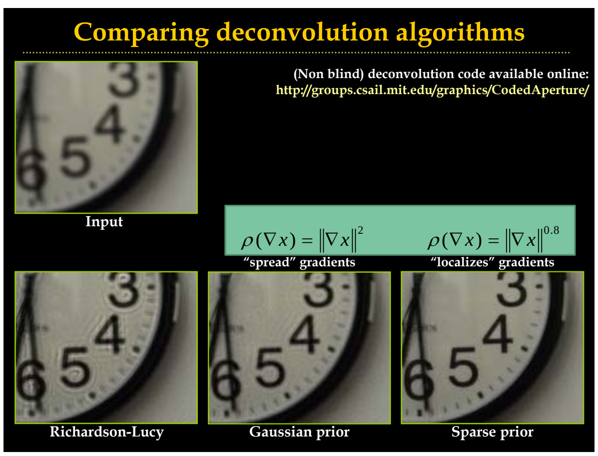

Challenges
Hard to deconvolve even when kernel is known
Input
Ringing with the traditional Richardson-Lucy deconvolution algorithm
Hard to identify correct scale:
Larger scale
Correct scale
Smaller scale

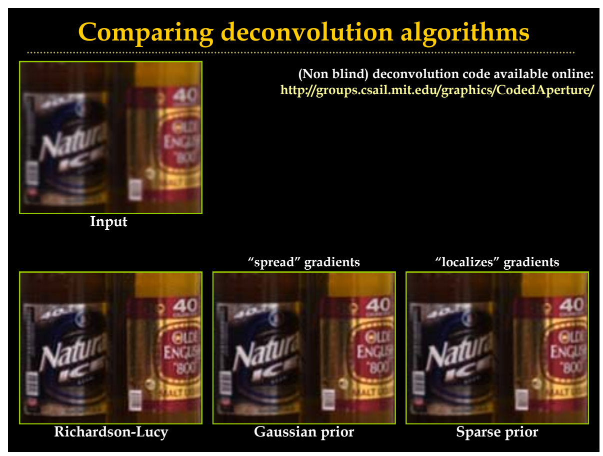

Deconvolution is ill posed
?
=

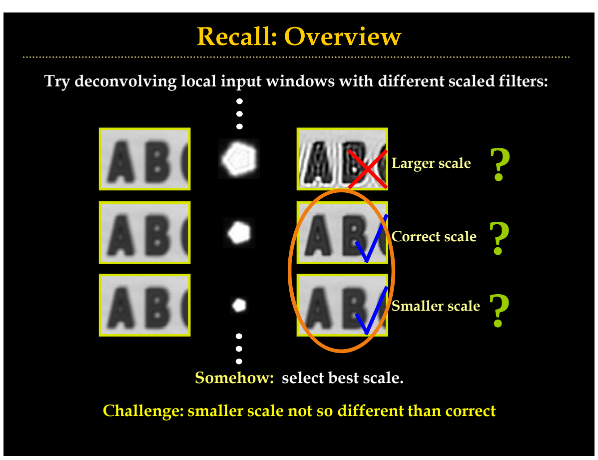

Deconvolution is ill posed
Solution 1:
=
Solution 2:
=

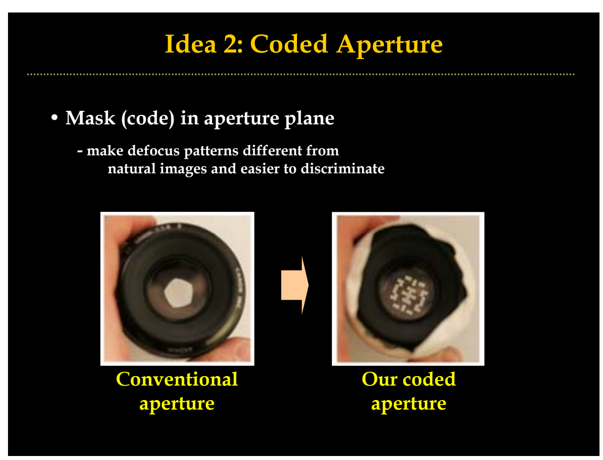

Idea 1: Natural images prior
What makes images special?
Natural
Unnatural
Image
gradient
Natural images have sparse gradients
put a penalty on gradients

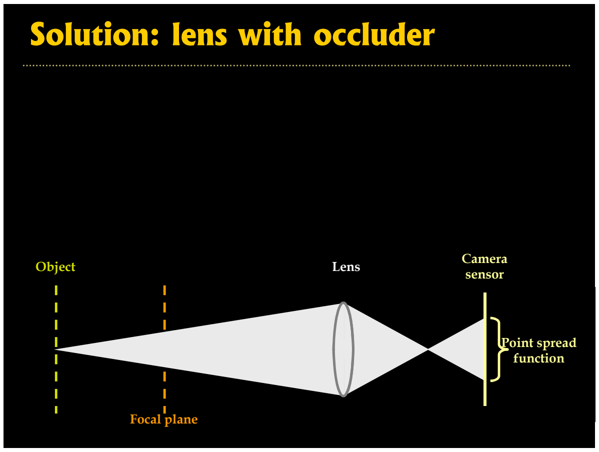

Deconvolution with prior
$x = \arg\min |f \otimes x - y|^2 + \lambda \sum_i \rho(\nabla x_i)$
Convolution error
Derivatives prior
Equal convolution error
Low
High

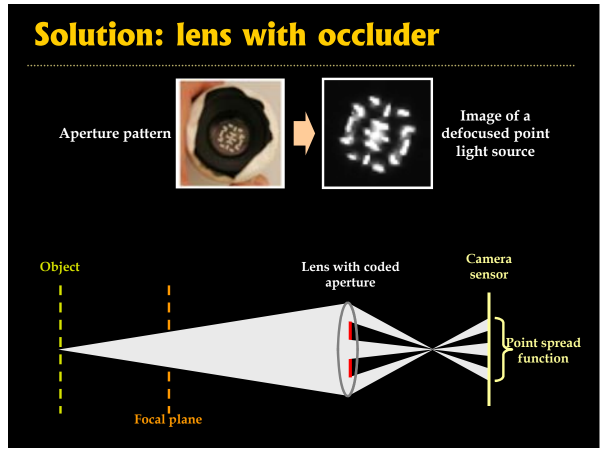

Comparing deconvolution algorithms
(Non blind) deconvolution code available online:
http://groups.csail.mit.edu/graphics/CodedAperture/
Input
$\rho(\nabla x) = \|\nabla x\|^2$
$\rho(\nabla x) = \|\nabla x\|^{0.8}$
"spread" gradients
"localizes" gradients
Richardson-Lucy
Gaussian prior
Sparse prior

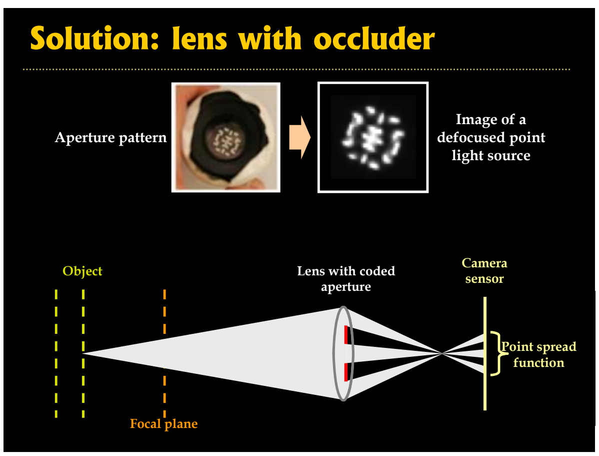

Comparing deconvolution algorithms
(Non blind) deconvolution code available online:
http://groups.csail.mit.edu/graphics/CodedAperture/
Input
"spread" gradients
"localizes" gradients
Richardson-Lucy
Gaussian prior
Sparse prior

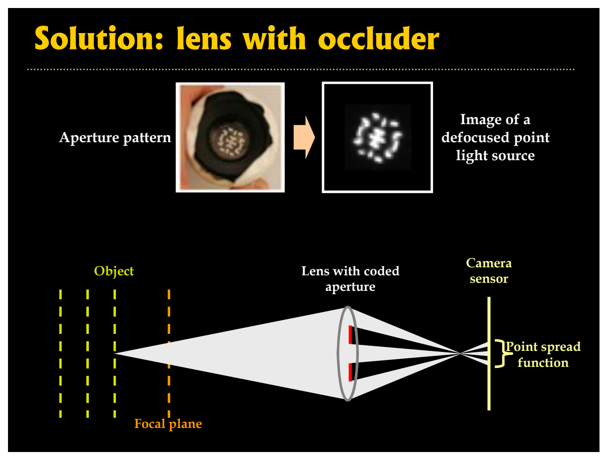

Recall: Overview
Try deconvolving local input windows with different scaled filters:
Larger scale
Correct scale
Smaller scale
Somehow: select best scale.
Challenge: smaller scale not so different than correct

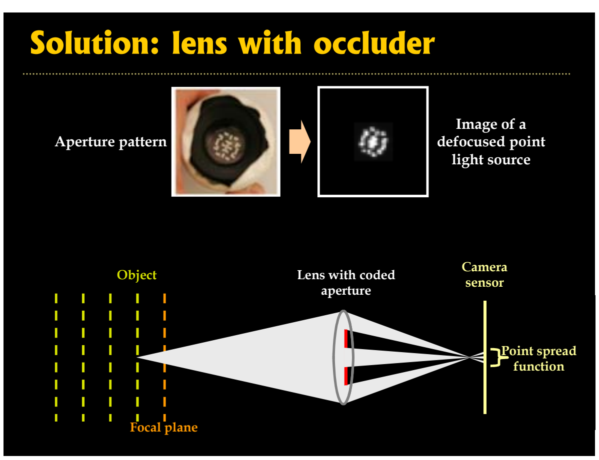

Idea 2: Coded Aperture
• Mask (code) in aperture plane
- make defocus patterns different from natural images and easier to discriminate
Conventional aperture
Our coded aperture

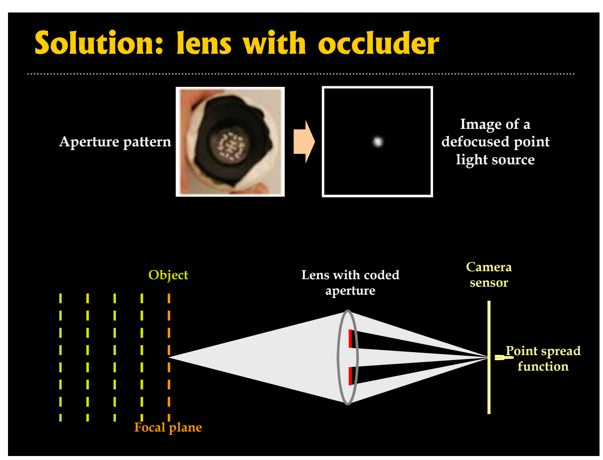

Solution: lens with occluder
Object
Lens
Camera sensor
Point spread function
Focal plane

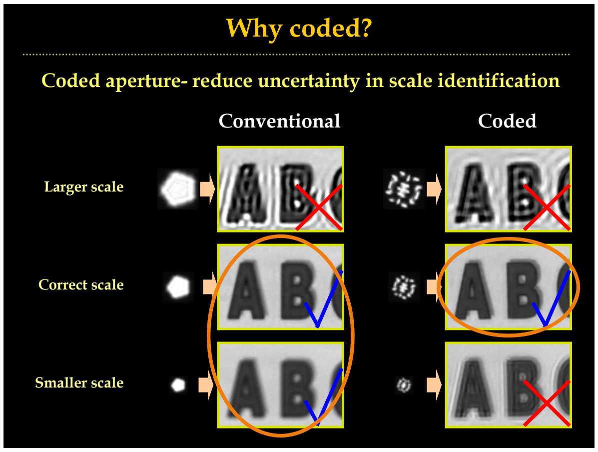

Solution: lens with occluder
Aperture pattern
Image of a defocused point light source
Object
Lens with coded aperture
Camera sensor
Point spread function
Focal plane

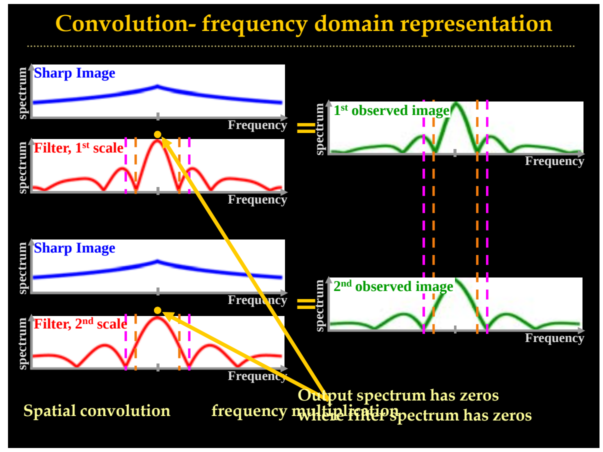

Solution: lens with occluder
Aperture pattern
Image of a defocused point light source
Object
Lens with coded aperture
Camera sensor
Point spread function
Focal plane

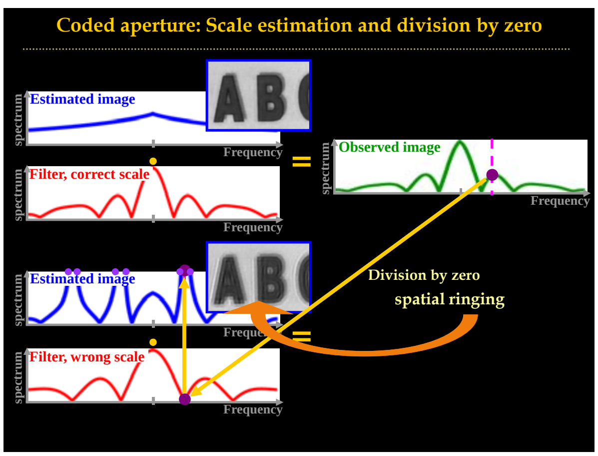

Solution: lens with occluder
Aperture pattern
Image of a defocused point light source
Object
Lens with coded aperture
Camera sensor
Point spread function
Focal plane

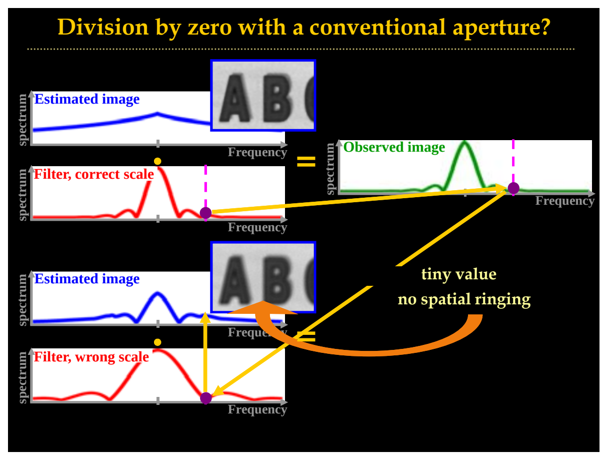

Solution: lens with occluder
Aperture pattern
Image of a defocused point light source
Object
Lens with coded aperture
Camera sensor
Point spread function
Focal plane

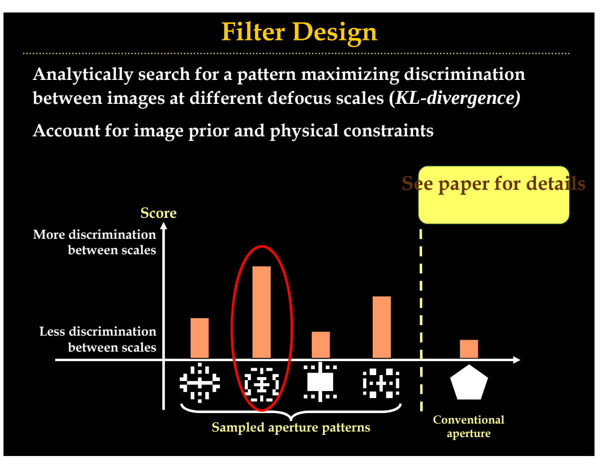

Solution: lens with occluder
Aperture pattern
Image of a defocused point light source
Object
Lens with coded aperture
Camera sensor
Point spread function
Focal plane

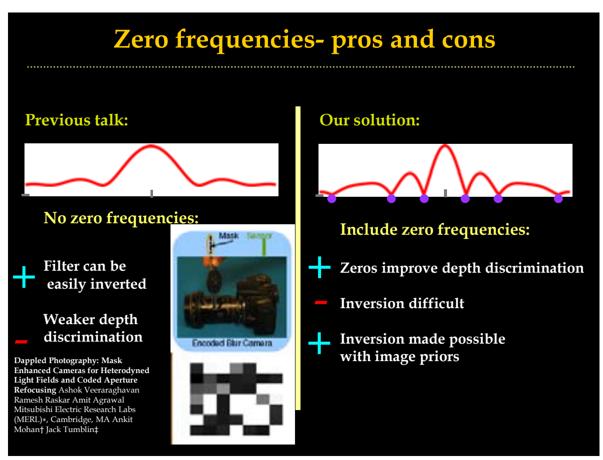

Why coded?
Coded aperture- reduce uncertainty in scale identification
Conventional
Coded
Larger scale
Correct scale
Smaller scale


Convolution- frequency domain representation
Sharp Image
Frequency
Filter, 1st scale
1st observed image
Sharp Image
Filter, 2nd scale
2nd observed image
spectrum
Spatial convolution
frequency multiplication
Output spectrum has zeros
where filter spectrum has zeros


Coded aperture: Scale estimation and division by zero
Estimated image
Filter, correct scale
Observed image
Estimated image
Filter, wrong scale
Frequency
spectrum
Division by zero
spatial ringing


Division by zero with a conventional aperture?
Estimated image
Filter, correct scale
Observed image
Estimated image
Filter, wrong scale
Frequency
spectrum
tiny value
no spatial ringing


Filter Design
Analytically search for a pattern maximizing discrimination between images at different defocus scales (KL-divergence)
Account for image prior and physical constraints
See paper for details
Score
More discrimination between scales
Less discrimination between scales
Sampled aperture patterns
Conventional aperture


Zero frequencies- pros and cons
Previous talk:
No zero frequencies:
Filter can be easily inverted
Weaker depth discrimination
Our solution:
Include zero frequencies:
Zeros improve depth discrimination
Inversion difficult
Inversion made possible with image priors
Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing
Ashok Veeraraghavan Ramesh Raskar Amit Agrawal
Mitsubishi Electric Research Labs (MERL)*, Cambridge, MA
Ankit Mohan† Jack Tumblin‡
Encoded Blur Camera

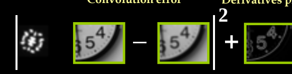
Depth results

Regularizing depth estimation

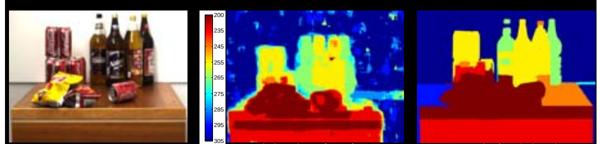
Try deblurring with 10 different aperture scales

$$x = \arg\min |f \otimes x - y|^2 + \lambda \sum_i \rho(\nabla x_i)$$

Convolution error — Derivatives prior

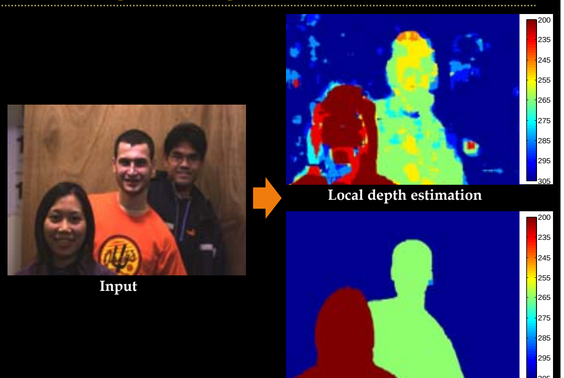


Keep minimal error scale in each local window + regularization



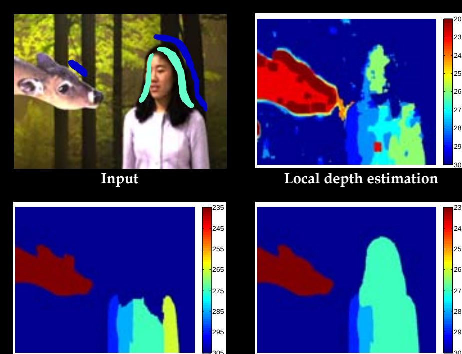
Input — Local depth estimation — Regularized depth

Regularizing depth estimation



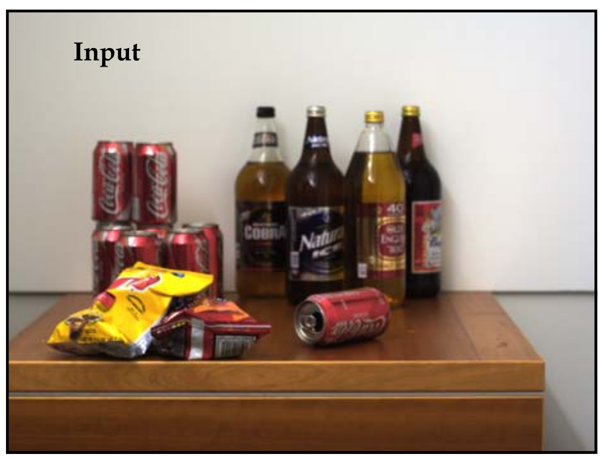
Input — Local depth estimation — Regularized depth

Sometimes, manual intervention



Input — Local depth estimation — Regularized depth — After user corrections

All focused results

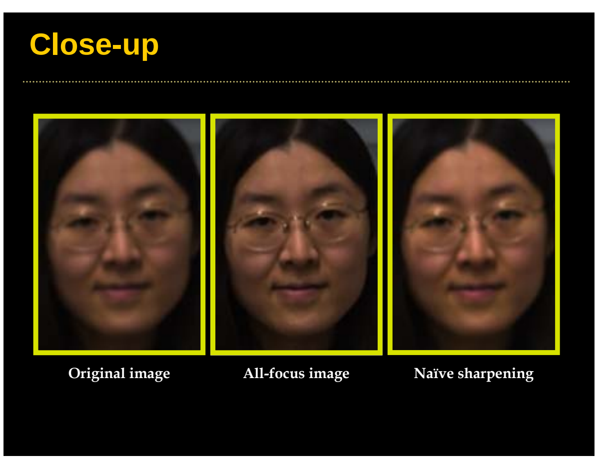


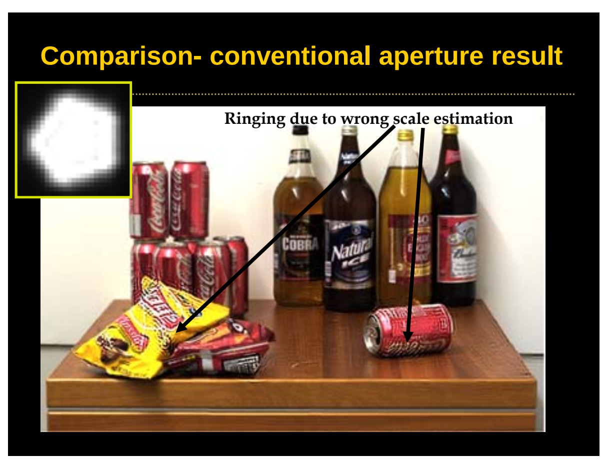

All-focused
(deconvolved)


Close-up
Original image
All-focus image

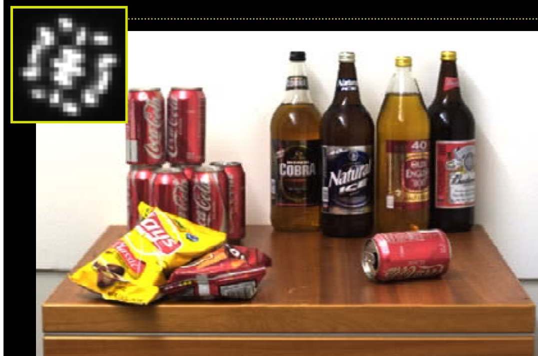

Input


All-focused
(deconvolved)


Close-up
Original image
All-focus image
Naïve sharpening


Comparison- conventional aperture result
Ringing due to wrong scale estimation

Comparison- coded aperture result



Application: Digital refocusing from a single image

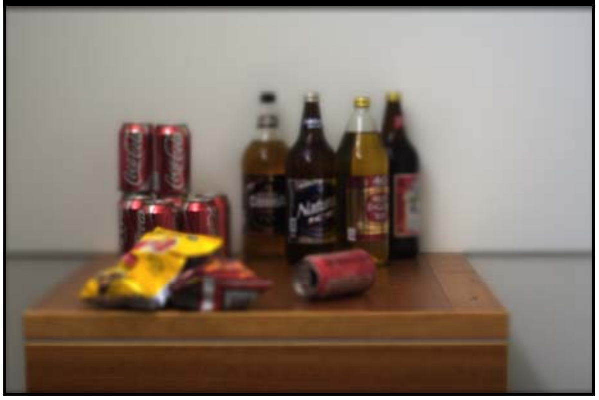
Application: Digital refocusing from a single image

Application: Digital refocusing from a single image

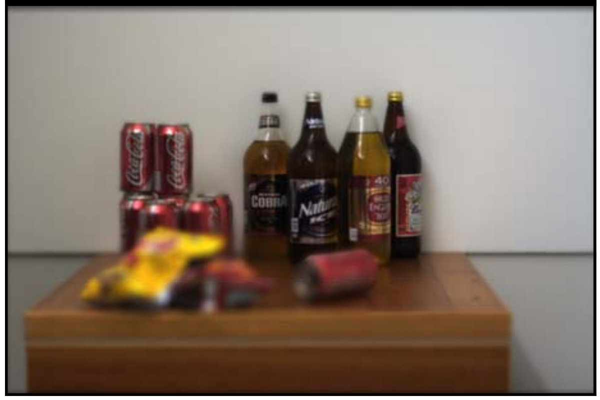
Application: Digital refocusing from a single image

Application: Digital refocusing from a single image

Application: Digital refocusing from a single image



Application: Digital refocusing from a single image



Coded aperture: pros and cons

- \+ Image AND depth at a single shot
- \+ No loss of image resolution
- \+ Simple modification to lens
- \- Depth is coarse
  - unable to get depth at untextured areas, might need manual corrections.
- \+ But depth is a pure bonus
- \- Loss some light
- \+ But deconvolution increases depth of field

Deconvolution code available

http://groups.csail.mit.edu/graphics/CodedAperture/

50mm f/1.8:	$79.95
Cardboard:	$1
Tape:	$1
Depth acquisition:	priceless

CodedAperture

Admin

- Fill in feedback forms
  - Can someone collect and return to 305WWH