Computational Photography

Image from Durand & Freeman's MIT Course on Computational Photography

Today's reading
- Szeliski Chapter 9

The ultimate camera

What does it do?

The ultimate camera

Infinite resolution

Infinite zoom control

Desired object(s) are in focus

No noise

No motion blur

Infinite dynamic range (can see dark and bright things)

...

Creating the ultimate camera

The "analog" camera has changed very little in >100 yrs
- we're unlikely to get there following this path

More promising is to combine "analog" optics with computational techniques
- "Computational cameras" or "Computational photography"

This lecture will survey techniques for producing higher quality images by combining optics and computation

Common themes:
- take multiple photos
- modify the camera

Noise reduction

Take several images and
average them

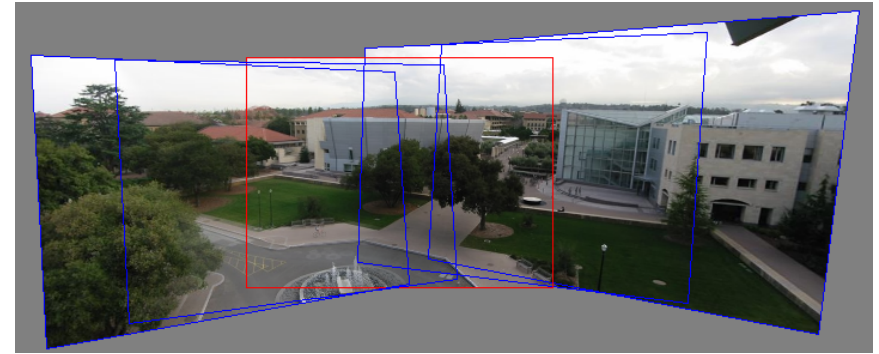Why does this work?

Basic statistics:
- variance of the mean
decreases with n:

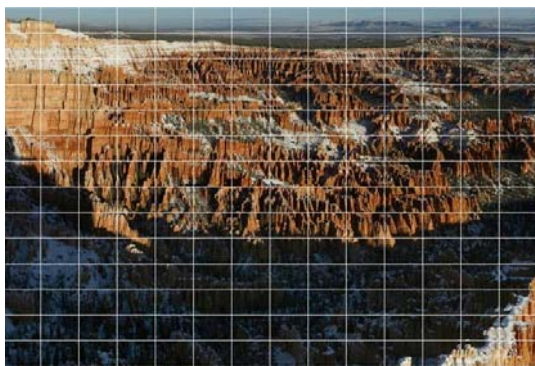$$Var(\overline{X}) = \frac{\sigma^2}{n}$$

Field of view

We can artificially increase the field of view by
compositing several photos together (project 2).

Improving resolution: Gigapixel images

Max Lyons, 2003
fused 196 telephoto shots

A few other notable examples:
- Obama inauguration (gigapan.org)
- HDView (Microsoft Research)

Improving resolution: super resolution

What if you don't have a zoom lens?

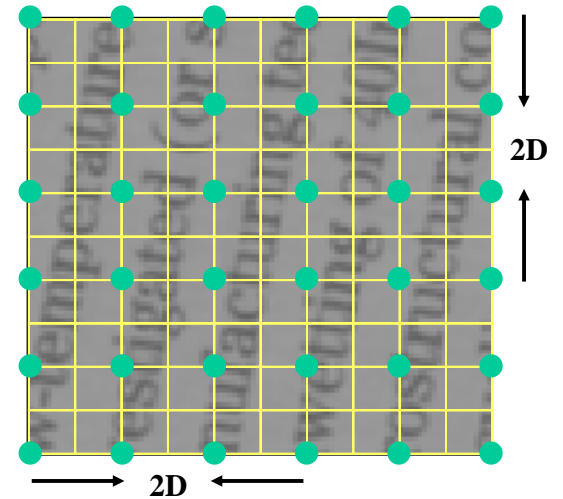## Intuition (slides from Yossi Rubner & Miki Elad)

For a given band-limited image, the Nyquist sampling theorem states that if a uniform sampling is fine enough ($\geq$**D**), perfect reconstruction is possible.

D

D
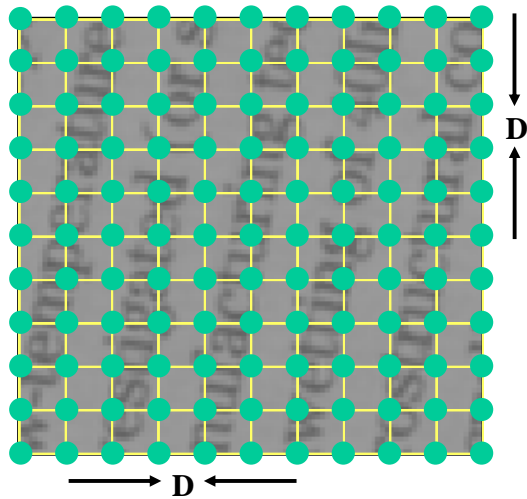
9

Intuition (slides from Yossi Rubner & Miki Elad)
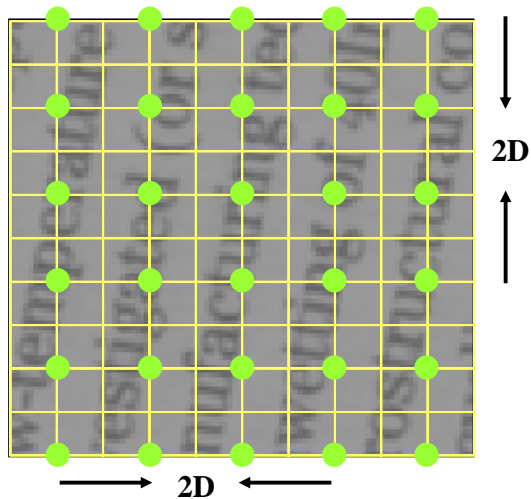

Due to our limited camera resolution, we sample using an insufficient 2D grid

2D

2D

10

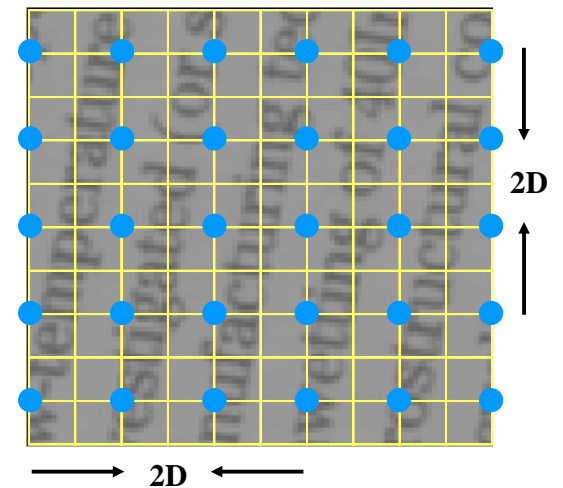## Intuition (slides from Yossi Rubner & Miki Elad)

However, if we take a second picture, shifting the camera 'slightly to the right' we obtain:

2D

2D

11

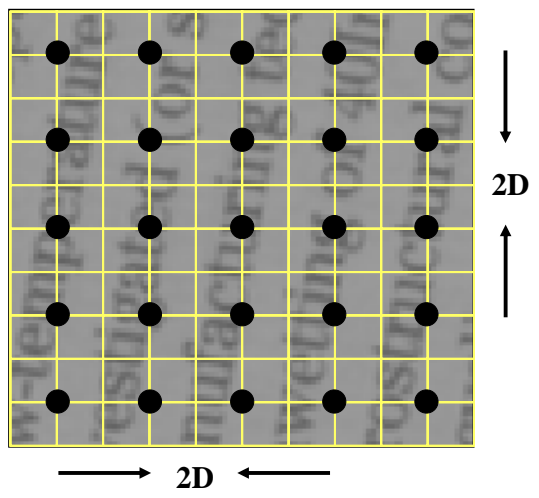## Intuition (slides from Yossi Rubner & Miki Elad)

Similarly, by shifting down we get a third image:
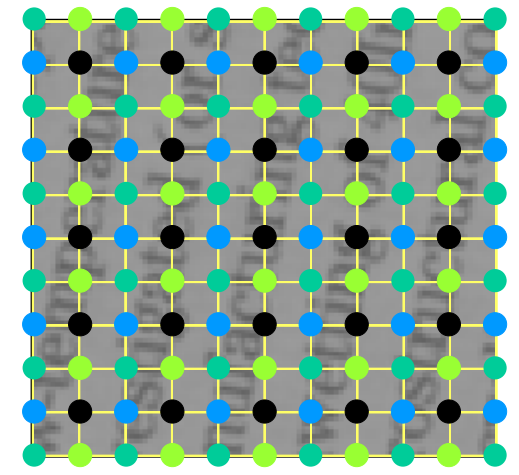
2D

2D

12

Intuition (slides from Yossi Rubner & Miki Elad)

And finally, by shifting down and to the right we get the fourth image:

2D

2D

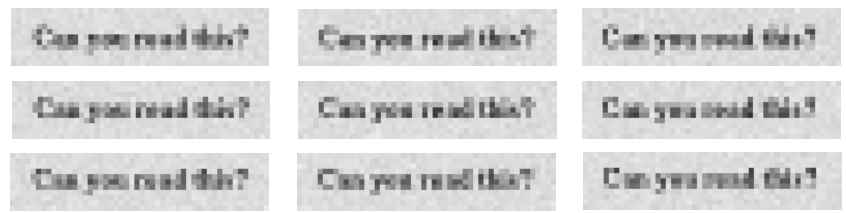## Intuition

By combining all four images the desired resolution is obtained, and thus perfect reconstruction is guaranteed.

Example

Can you read this? | Can you read this? | Can you read this?
Can you read this? | Can you read this? | Can you read this?
Can you read this? | Can you read this? | Can you read this?

3:1 scale-up in each axis using 9 images, with pure global translation between them

Can you read this?

Handling more general 2D motions

What if the camera displacement is Arbitrary ?
What if the camera rotates? Gets closer to the object (zoom)?
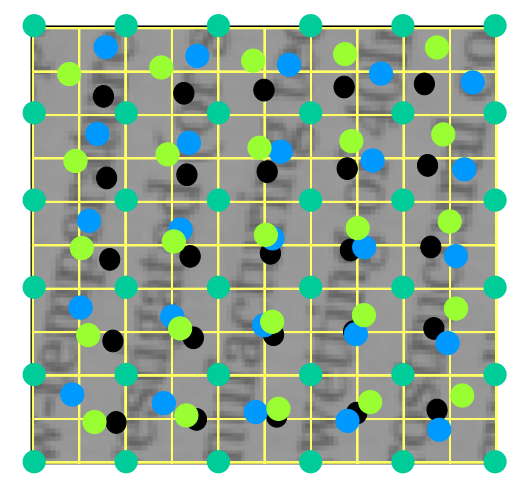
Super-resolution

Basic idea:

- define a destination (dst) image of desired resolution
- assume mapping from dst to each input image is known
  - usually a combination of a 2D motion/warp and an average (point-spread function)
  - can be expressed as a set of linear constraints
  - sometimes the mapping is solved for as well
- add some form of regularization (e.g., "smoothness assumption")
  - can also be expressed using linear constraints
  - but L1, other nonlinear methods work better

How does this work? [Baker & Kanade, 2002]

| No. of Input Images | 1 (Original) | 4 | 16 | 64 | 256 |

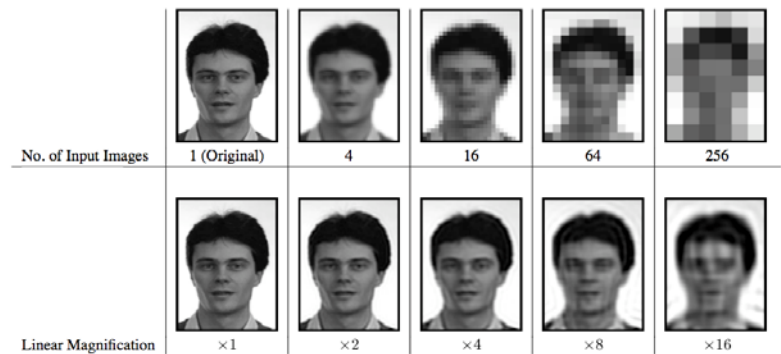| Linear Magnification | ×1 | ×2 | ×4 | ×8 | ×16 |

Figure 1: Results of the reconstruction-based super-resolution algorithm [9] for increasing magnification factors. The original high-resolution image (top left) is translated multiple times by random sub-pixel amounts, blurred with a Gaussian, and then down-sampled. Comparing the images in the right-most column, we see that the reconstruction algorithm does quite well given the very low resolution of the input. The degradation in performance as the magnification increases from left to right is very dramatic, however.

Limits of super-resolution [Baker & Kanade, 2002]

Performance degrades significantly beyond 4x or so

Doesn't matter how many new images you add

- space of possible (ambiguous) solutions explodes quickly

Major cause

- quantizing pixels to 8-bit gray values
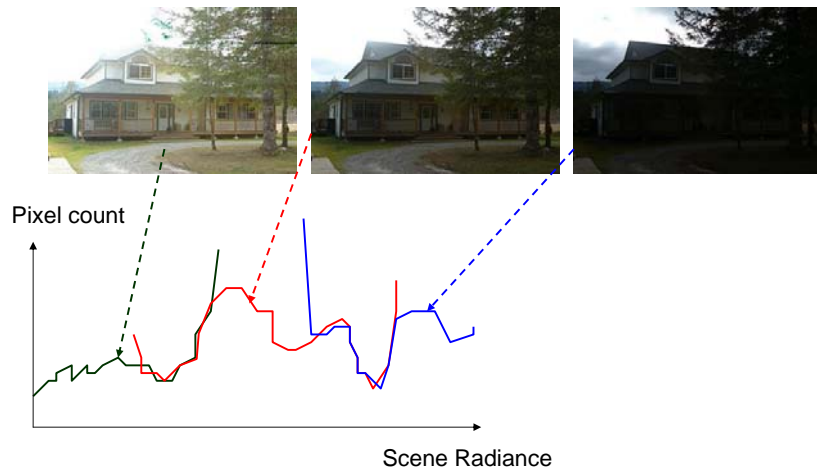
Possible solutions:

- nonlinear techniques (e.g., L1)
- better priors (e.g., using domain knowledge)
  - Baker & Kanade "Hallucination", 2002
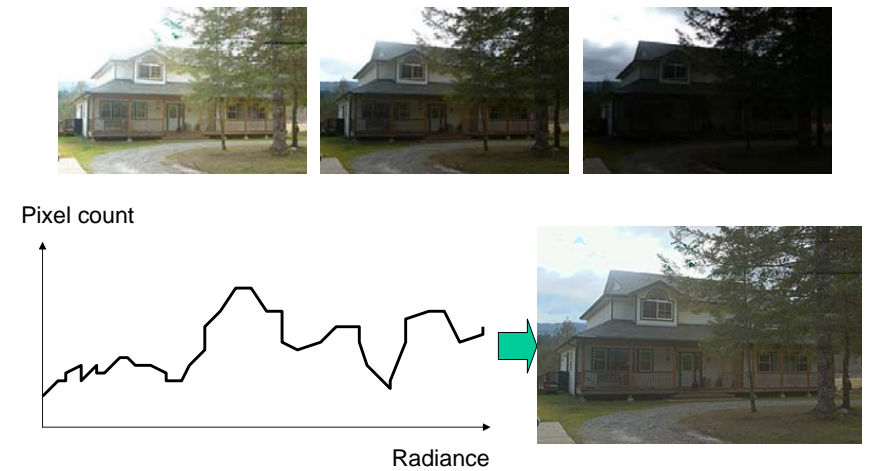 - Freeman et al. "Example-based super-resolution"

Dynamic Range

Typical cameras have limited dynamic range

HDR images — merge multiple inputs

Pixel count

Scene Radiance

HDR images — merged

Pixel count

Radiance

Camera is not a photometer!

Limited dynamic range
- 8 bits captures only 2 orders of magnitude of light intensity
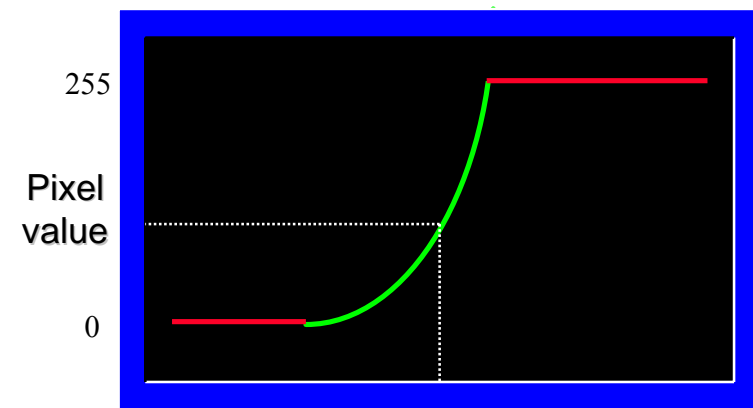- We can see ~10 orders of magnitude of light intensity

Unknown, nonlinear response
- pixel intensity ≠ amount of light (# photons, or "radiance")

Solution:
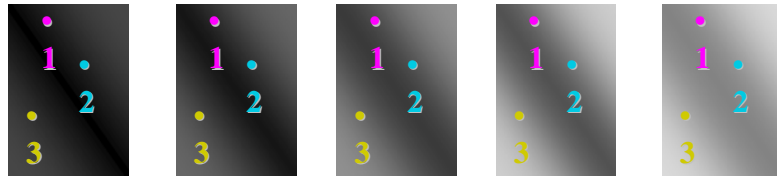- Recover response curve from multiple exposures, then reconstruct the *radiance map*

Camera response function

255

Pixel value

0

$\log$ Exposure = $\log$ (Radiance * Δt)

(CCD photon count)

Calculating response function

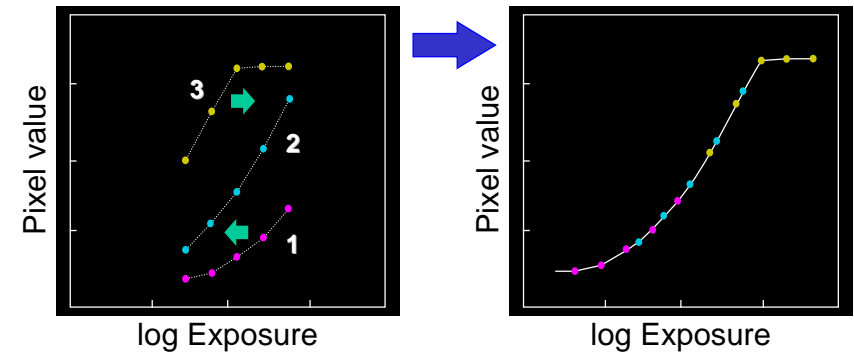| $\Delta t =$ 1/64 sec | $\Delta t =$ 1/16 sec | $\Delta t =$ 1/4 sec | $\Delta t =$ 1 sec | $\Delta t =$ 4 sec |

Pixel Value $Z = f(Exposure)$

Exposure = Radiance $\times \Delta t$

log Exposure = log Radiance + log Δt

Debevec & Malik [SIGGRAPH 1997]

Assuming unit radiance for each pixel

After adjusting radiances to obtain a smooth response curve

The Math

Let *g(z)* be the *discrete* inverse response function

For each pixel site *i* in each image *j*, want:

$$\ln Radiance_i + \ln \Delta t_j = g(Z_{ij})$$

Solve the over-determined linear system:

$$\sum_{i=1}^{N} \sum_{j=1}^{P} \left[\ln Radiance_i + \ln \Delta t_j - g(Z_{ij}) \right]^2 + \lambda \sum_{z=Z_{min}}^{Z_{max}} g''(z)^2$$

fitting term   smoothness term

Capture and composite several photos

Same trick works for

- field of view
- resolution
- signal to noise
- dynamic range
- Focus

But sometimes you can do better by modifying the camera…

Focus

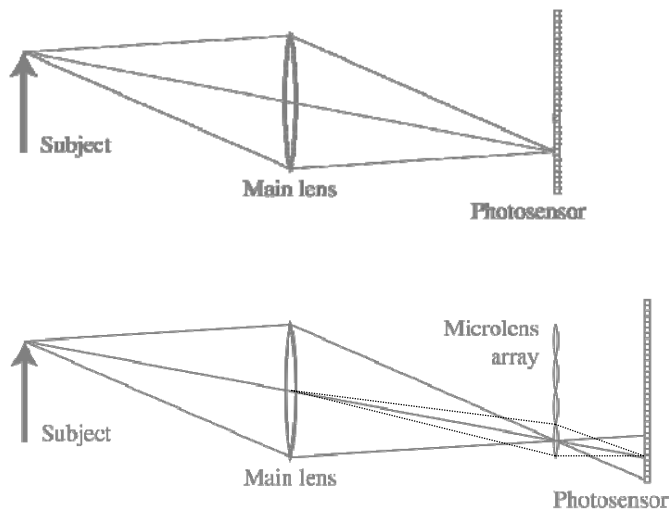Suppose we want to produce images where the desired object is *guaranteed* to be in focus?

Or suppose we want *everything* to be in focus?

Light field camera [Ng et al., 2005]

http://www.refocusimaging.com/gallery/
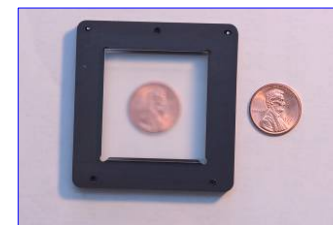
Conventional vs. light field camera

Prototype camera
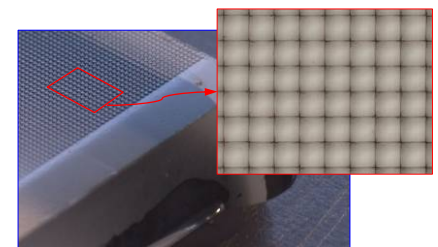

Contax medium format camera   Kodak 16-megapixel sensor

Adaptive Optics microlens array   125μ square-sided microlenses

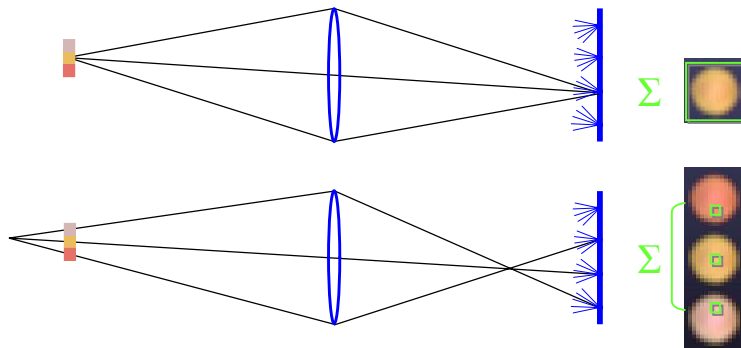*4000 × 4000 pixels ÷ 292 × 292 lenses = 14 × 14 pixels per lens*
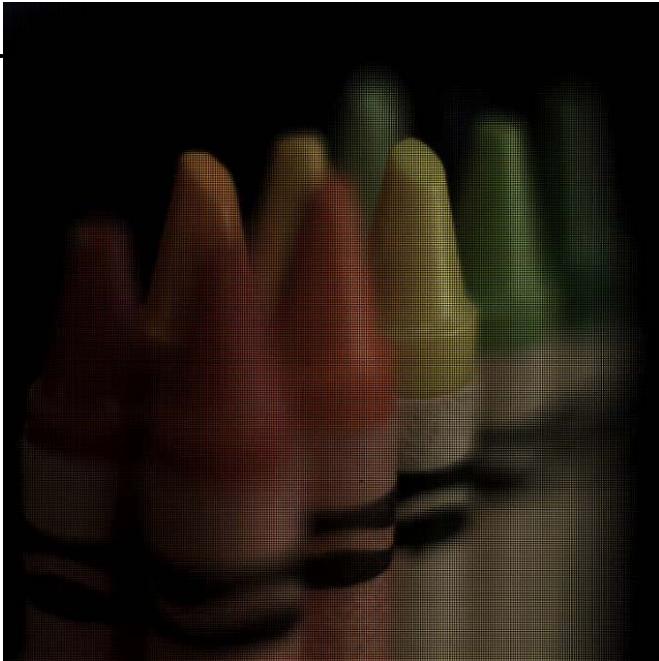
Simulating depth of field
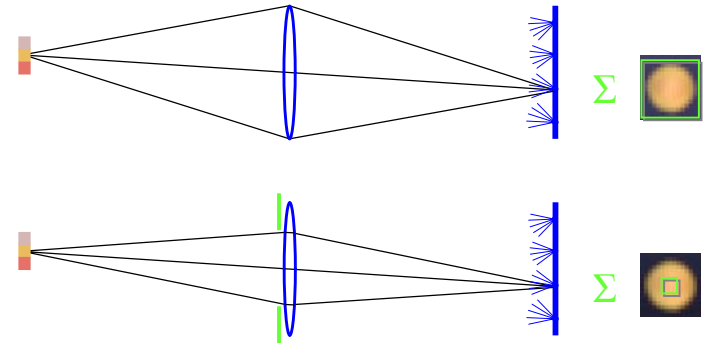
stopping down aperture = summing only
the central portion of each microlens

Digital refocusing

refocusing = summing windows extracted
from several microlenses

Example of digital refocusing

All-in-focus

If you only want to produce an all-focus image, there are simpler alternatives

E.g.,
- Wavefront coding [Dowsky 1995]
- Coded aperture [Levin SIGGRAPH 2007], [Raskar SIGGRAPH 2007]
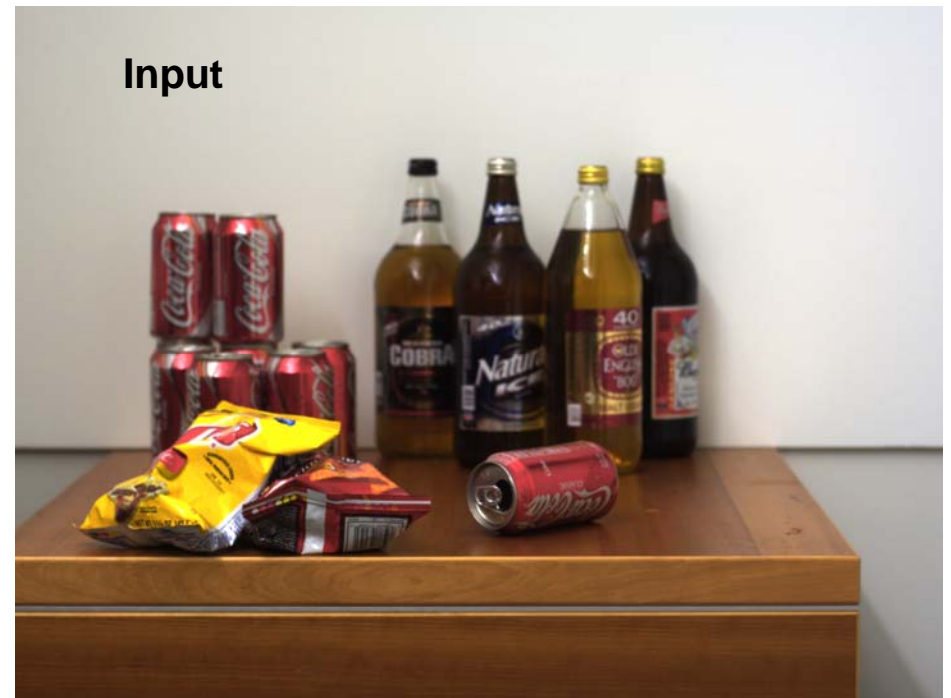 - can also produce change in focus (ala Ng's light field camera)

Build your own coded aperture

Levin et al., SIGGRAPH 2007

Voila!

Levin et al., SIGGRAPH 2007

Input

**All-focused
(deconvolved)**

Original image

All-focus image

Motion blur removal

Instead of coding the aperture, code the...

Raskar et al.,
SIGGRAPH 2007
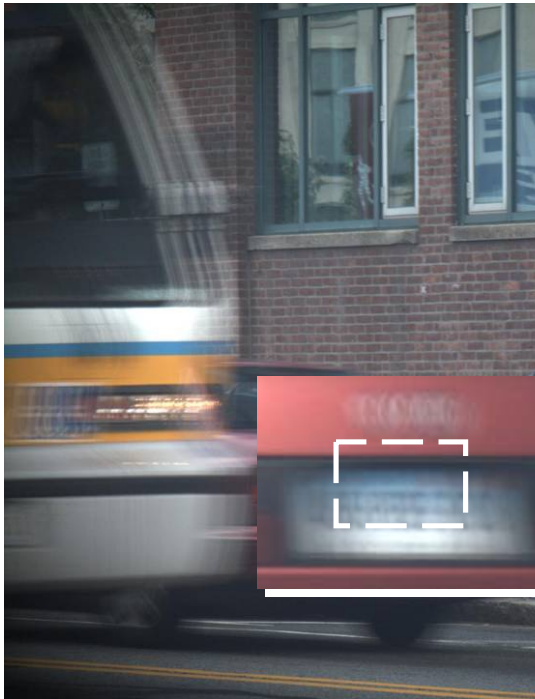
Raskar et al.,
SIGGRAPH 2007

License Plate Retrieval

Many more possibilities

Seeing through/behind objects
- Using a camera array ("synthetic aperture")
- Levoy et al., SIGGRAPH 2004

Removing interreflections
- Nayar et al., SIGGRAPH 2006

Family portraits where everyone's smiling
- Photomontage (Agarwala at al., SIGGRAPH 2004)

…

More on computational photography

SIGGRAPH course notes and video
Other courses
- MIT course
- CMU course
- Stanford course
- Columbia course

Wikipedia page

Symposium on Computational Photography

ICCP 2009 (conference)