Computational Photography

CSE P 576

Larry Zitnick (larryz@microsoft.com)

What is computational photography?

Def: The generation of an photograph requiring the use of a computer that enhances or extends the capabilities of photography.

Typically, multiple images are used to create a final "photograph."

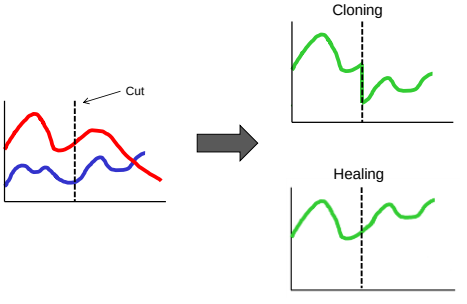
How can these images vary?

Panoramas

We already covered this…

Composite

How do Photoshop's cloning and healing brushes work? What is the difference?



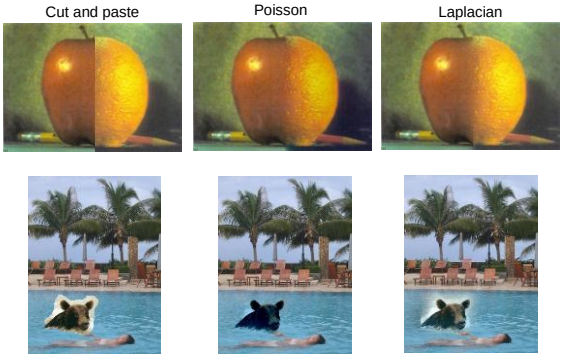
Poisson blending

Limitations:

- Can't do contrast reversal (gray on black -> gray on white)
- Colored backgrounds "bleed through"
- Images need to be very well aligned
- Textures may not perfectly agree

Perez et al, SIGGRAPH 2003

Comparison



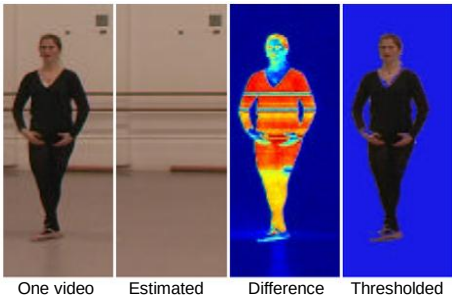
Steve Gomez http://www.cs.brown.edu/courses/csci1950-g/results/proj2/steveg/

Selecting regions

Use Photoshop's "magic wand."

Background Subtraction

A largely unsolved problem…



One video frame | Estimated background | Difference Image | Thresholded Foreground on blue

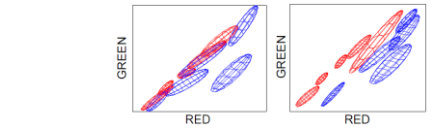
Background Subtraction

Harder than you think…

Jodoin et al., 2008

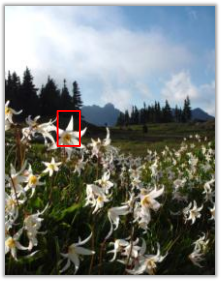
Grabcut

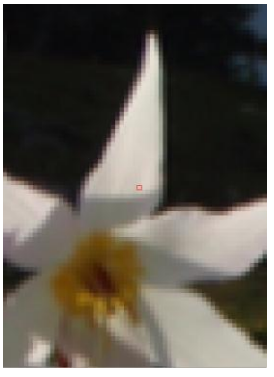
Compute color model for foreground and background using EM (k=5).



C. Rother, V. Kolmogorov, A. Blake. **GrabCut: Interactive Foreground Extraction using Iterated Graph Cuts.** *ACM Transactions on Graphics (SIGGRAPH'04), 2004*

Matting





Matting

Matting equation

observed color

blending

$$C = \alpha F + (1 - \alpha) B$$

foreground color

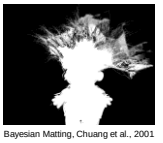
background color

$$\alpha \in [0, 1]$$

http://www.alphamatting.com/

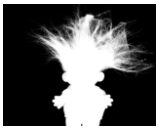
Solving

$$C = \alpha F + (1 - \alpha) B$$



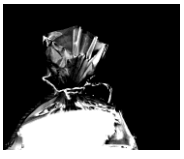
Bayesian Matting, Chuang et al., 2001

Robust Matting, Wang et al., 2007



Shared matting, Eduardo et al., 2010

Harder problems



Bayesian Matting, Chuang et al., 2001

Robust Matting, Wang et al., 2007

Shared matting, Eduardo et al., 2010

Movies

How do they perform matting in movies?

Blue Screen

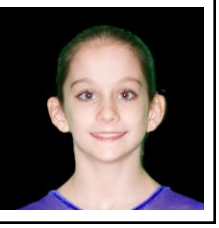
Blue Screen matting

Most common form of matting in TV studios & movies

Petros Vlahos invented blue screen matting in the 50s. His Ultimatte® is still the most popular equipment. He won an Oscar for lifetime achievement.

A form of background subtraction:

- Need a known background
- Foreground not equal to background
  - no blue ties!
- Why blue?
- Why uniform?

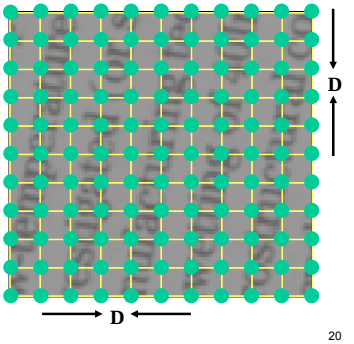


Improving resolution: super resolution

What if you don't have a zoom lens or a mega-pixel sensor?

Intuition (slides from Yossi Rubner & Miki Elad)

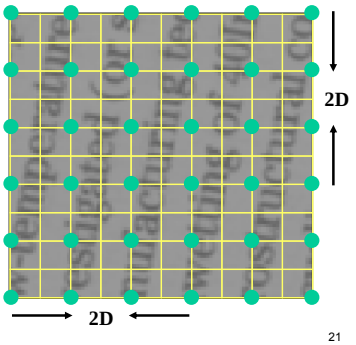
For a given band-limited image, the Nyquist sampling theorem states that if a uniform sampling is fine enough (≥**D**), perfect reconstruction is possible.



20

Intuition (slides from Yossi Rubner & Miki Elad)

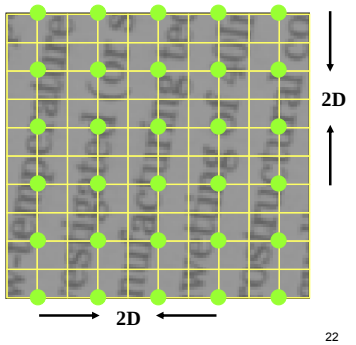
Due to our limited camera resolution, we sample using an insufficient 2D grid



21

Intuition (slides from Yossi Rubner & Miki Elad)

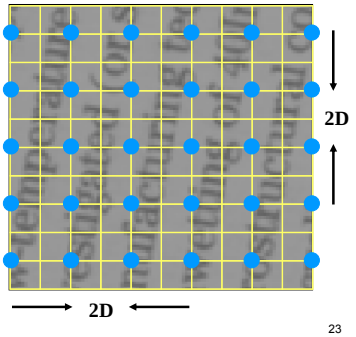
However, if we take a second picture, shifting the camera 'slightly to the right' we obtain:



22

Intuition (slides from Yossi Rubner & Miki Elad)

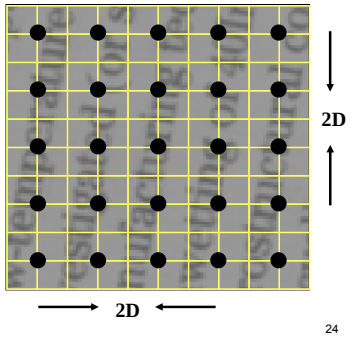
Similarly, by shifting down we get a third image:



23

Intuition (slides from Yossi Rubner & Miki Elad)

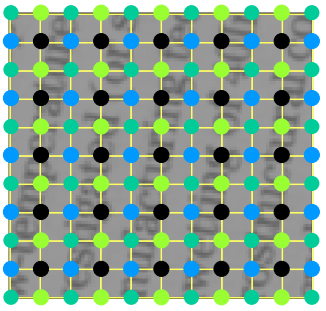
And finally, by shifting down and to the right we get the fourth image:



24

Intuition

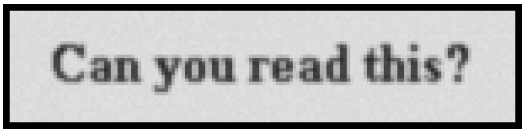
By combining all four images the desired resolution is obtained, and thus perfect reconstruction is guaranteed.



25

Example

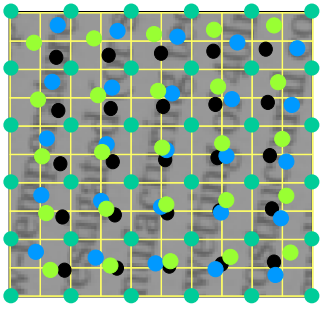
3:1 scale-up in each axis using 9 images, with pure global translation between them



26

Handling more general 2D motions

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



27

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]

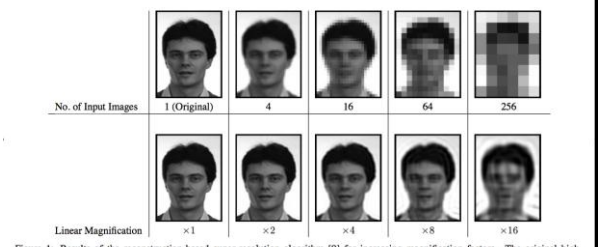


Figure 1: Results of the reconstruction-based super-resolution algorithm [9] for increasing magnification factors. The original high-resolution image (top left) is translated multiple times by random sub-pixel amounts, blurred with a Gaussian, and then down-sampled. Comparing the images in the right-most 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"

Noise

Many possible techniques:

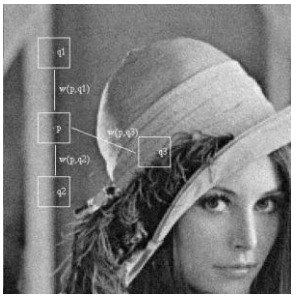
Bilateral filter and median filter are very common.

Blur color more than intensity. Why?

There are more advanced techniques…

Non-local means

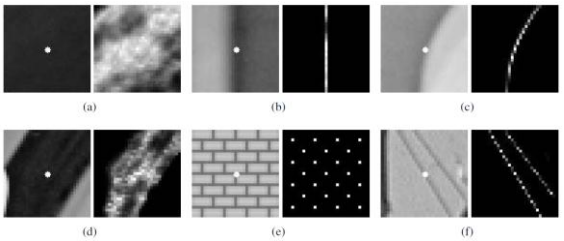
Look for similar patches in the image…



A. Buades, B. Coll, J.M. Morel"A non local algorithm for image denoising" IEEE Computer Vision and Pattern Recognition 2005, Vol 2, pp:60-65, 2005.

Non-local means

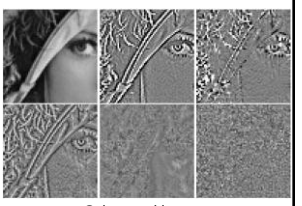
Look for similar patches in the image…



A. Buades, B. Coll, J.M. Morel"A non local algorithm for image denoising"
IEEE Computer Vision and Pattern Recognition 2005, Vol 2, pp:60-65, 2005.

Non-local means results

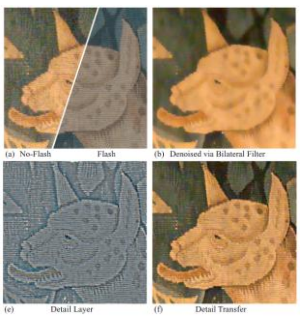
De-noising results



Subtracted images

Multiple images

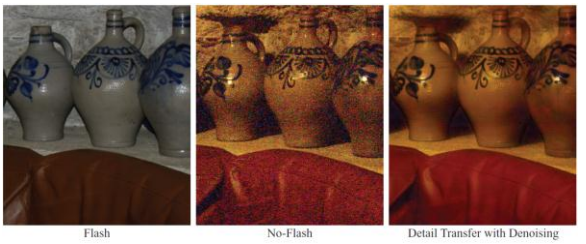
Use two images: Flash and no flash



G. Petschnigg, M. Agrawala, H. Hoppe, R. Szeliski, M. Cohen, K. Toyama. Digital Photography with Flash and No-Flash Image Pairs. *SIGGRAPH* 2004

Multiple images

Flash/no flash

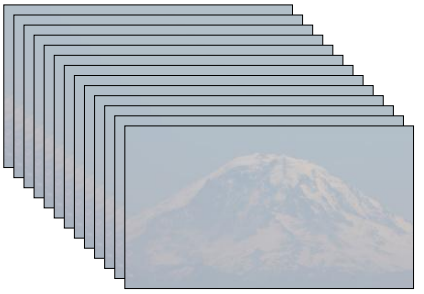


Flash | No-Flash | Detail Transfer with Denoising

Seeing Mt. Rainier

What if we want to take a picture of Mt. Rainier from Seattle?

Seeing Mt. Rainier



Joshi, N. and Cohen, M. *Seeing Mt. Rainier: Lucky Imaging for Multi-Image Denoising, Sharpening, and Haze Removal*. IEEE ICCP 2010.

Seeing Mt. Rainier

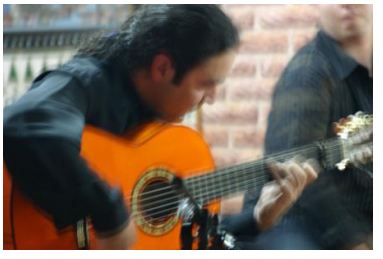
Deblurring

What causes blur?

Depth of field (focus)

Motion

Camera shake



Deblurring

Two main problems:

1. What is the blur kernel?
2. How do we de-convolve the image?

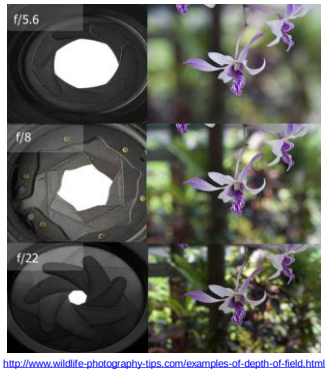
Chicken and egg problem.

Deblurring

What do blur kernels look like?

Gaussian?

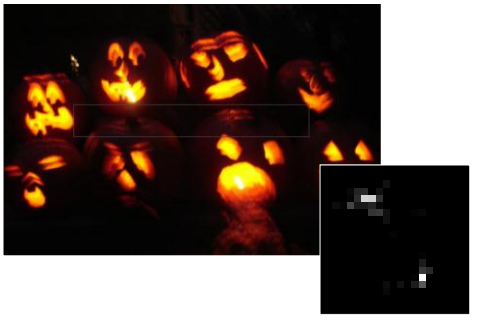
Blur kernel and the aperture



http://www.wildlife-photography-tips.com/examples-of-depth-of-field.html

Deblurring

What do blur kernels look like?

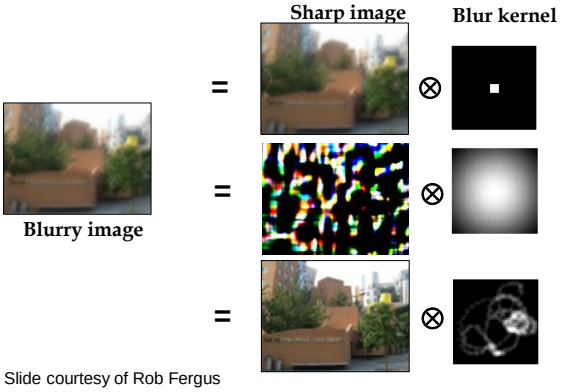


Deblurring

What do blur kernels look like?

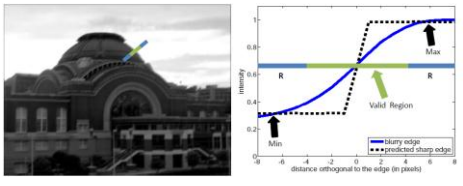
R. Fergus, B. Singh, A. Hertzmann, S. Roweis and W. Freeman, Removing Camera Shake From A Single Photograph, SIGGRAPH 2006

Multiple possible solutions



Slide courtesy of Rob Fergus

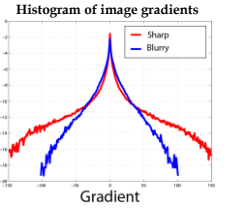
Finding the kernel



Joshi, N., Szeliski, R., and Kriegman, D. *PSF Estimation using Sharp Edge Prediction*. Proceedings of IEEE CVPR 2008.

Finding the kernel

Find the kernel that produces images with "natural" image statistics.



R. Fergus, B. Singh, A. Hertzmann, S. Roweis and W. Freeman, Removing Camera Shake From A Single Photograph, SIGGRAPH 2006

Slide courtesy of Rob Fergus

Deconvolution

Solve the following:

$$\underset{I}{\operatorname{argmax}}\, P(I|B) \quad = \quad \underset{I}{\operatorname{argmin}}[L(B|I) + L(I)]$$

Blur kernel

Data term: $L(B|I) = ||\vec{B} - A(d)\vec{I}||^2/\sigma^2$

Sparse gradient prior: $L(I) = \lambda||\nabla I||^{0.8}$

Deblurring: Special hardware

Can we make the deblurring process easier using special hardware?

Image stabilization

Image stabilization can be done using a floating lens.

Vibration is detected using gyroscopic sensors and compensated for.

Mainly on high-end lenses. $$$

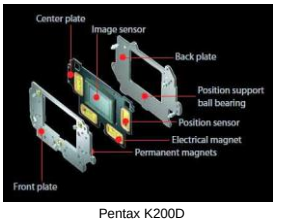
Canon EF 300mm

Image stabilization

Image stabilization can be done by moving the sensor.

This "Mechanical image stabilization" also uses gyros to sense motion. Actuators move the sensor.

Cheaper than lens stabilization and any lens may be used.



Pentax K200D

http://en.wikipedia.org/wiki/Image_stabilization

Multiple images

Long exposure/short exposure

Figure 11: *Large kernel. Left: blurred image, noisy image, denoised image, and our result. Top right: two image patches in the light-orange boxes in blurred/noisy images reveal the kernel shape. Note that the highlight point in the noisy patch is an ellipse-like shape. Bottom right: estimated 87 × 87 kernel.*

YUAN, L., SUN, J., QUAN, L., AND SHUM, H.-Y. Image deblurring with blurred/noisy image pairs. Siggraph, 2007.

Coded aperture

How can we get rid of depth of field de-focusing?

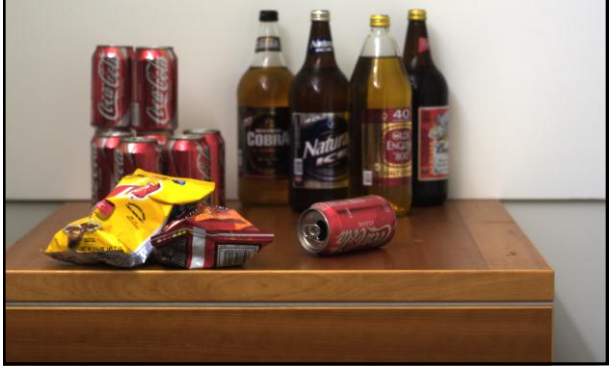
Build your own coded aperture

Levin et al., SIGGRAPH 2007

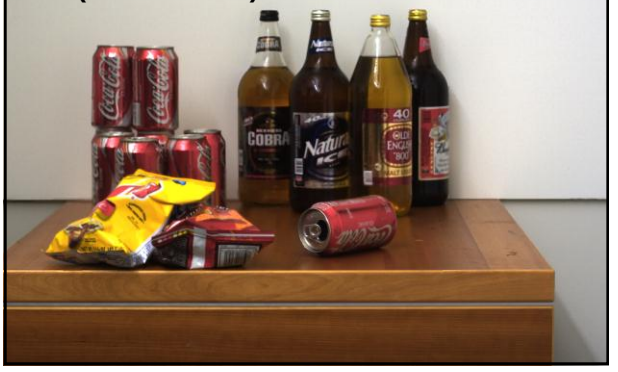
Voila!

Levin et al., SIGGRAPH 2007

Input



All-focused
(deconvolved)



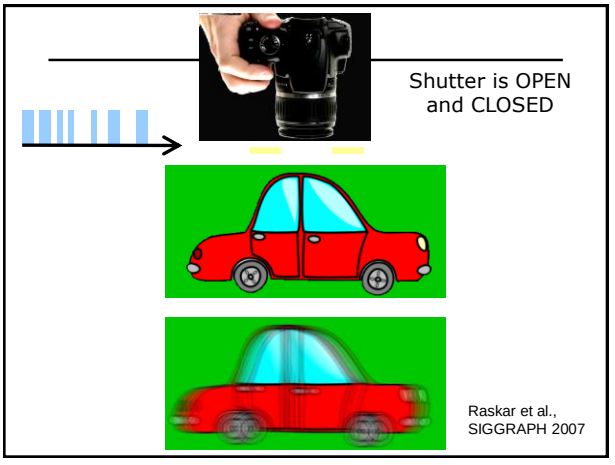
Close-up

Original image

All-focus image

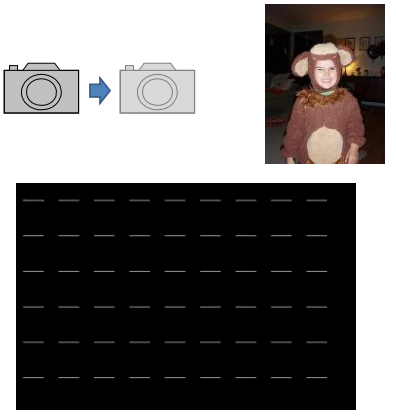
Motion blur removal

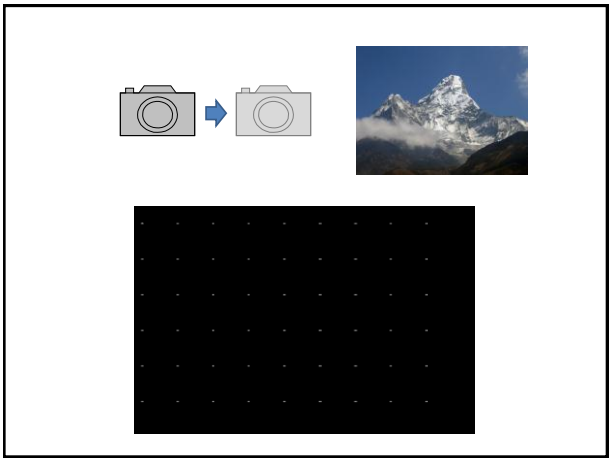
Instead of coding the aperture, code the...

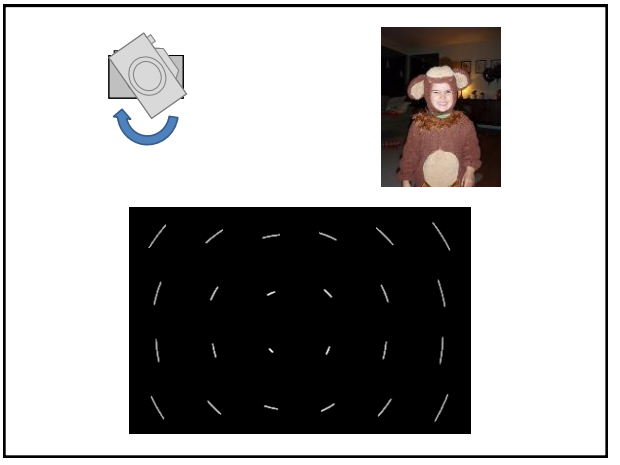


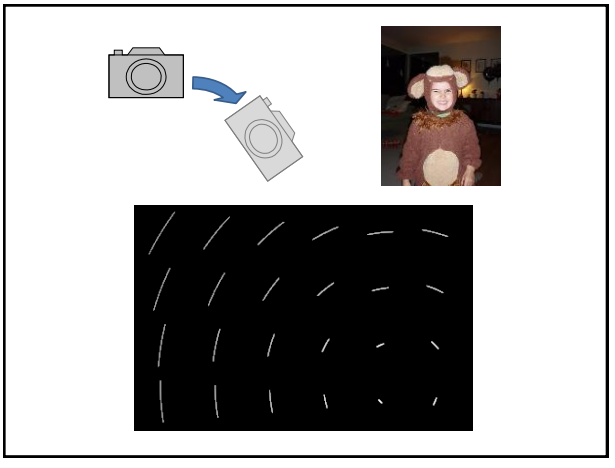
Spatially variant blurs

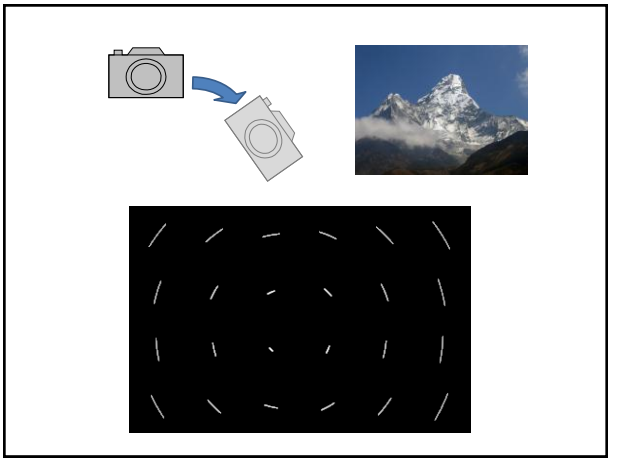
Why do the blur kernels vary across the image?



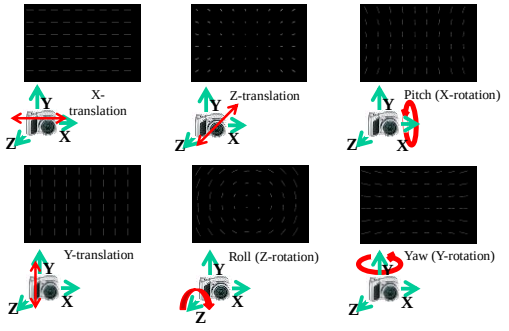








6 degrees = many motions



* Also spatially variant due to depth (DoF).

Inertial Measurement Sensors

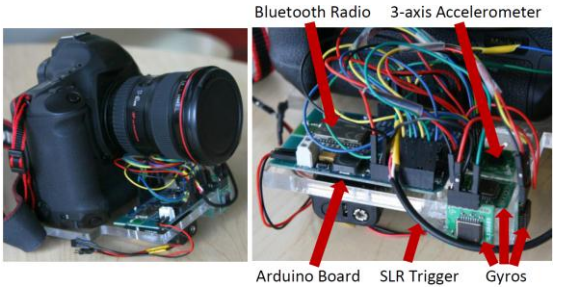
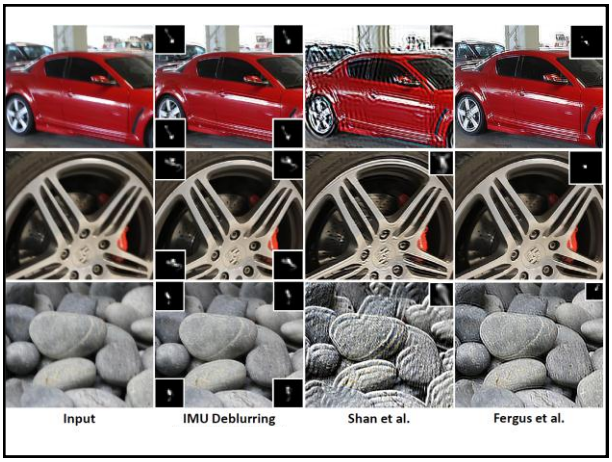


Image deblurring with inertial measurement sensors, N. Joshi, S.B. Kang, C.L. Zitnick, and R. Szeliski, *SIGGRAPH*, 2010



Focus

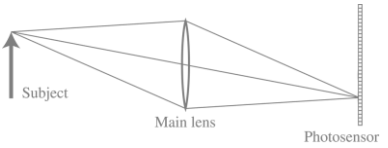
Suppose we want to produce images where the desired object is ***guaranteed*** to be in focus?

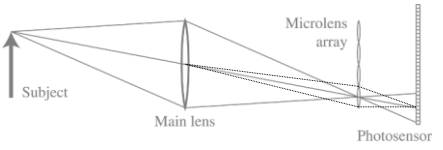
Or suppose we want ***everything*** to be in focus?

Light field camera [Ng et al., 2005]

http://lytro.com/gallery/

Conventional vs. light field camera

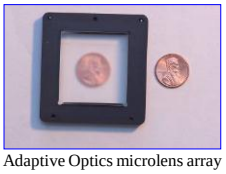




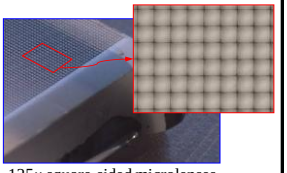
Prototype camera

Contax medium format camera

Kodak 16-megapixel sensor



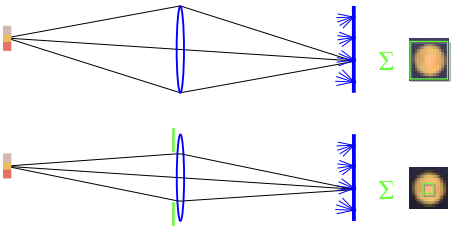
Adaptive Optics microlens array



125μ square-sided microlenses

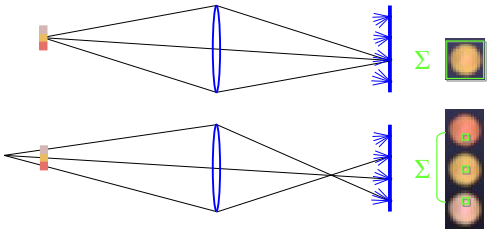
4000 × 4000 pixels ÷ 292 × 292 lenses = 14 × 14 pixels per lens

Simulating depth of field



stopping down aperture = summing only
the central portion of each microlens

Digital refocusing



refocusing = summing windows extracted from several microlenses

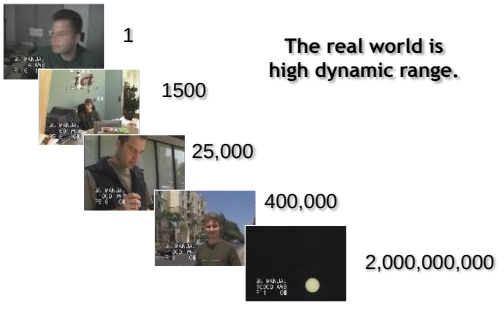
Example of digital refocusing

What is the tradeoff with using microlenses?

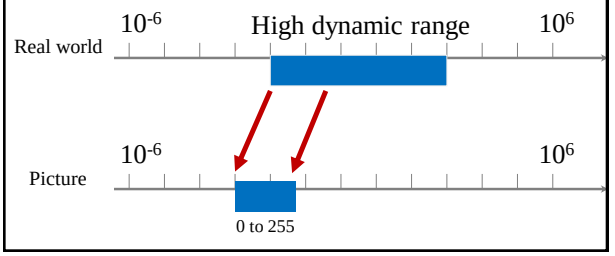
High Dynamic Range Images

Slides from Alexei A. Efros and Paul Debevec

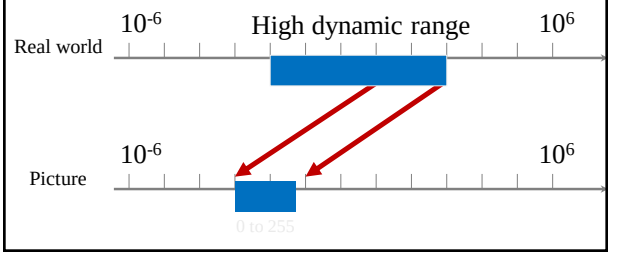
Problem: Dynamic Range



Long Exposure



Short Exposure



Camera Calibration

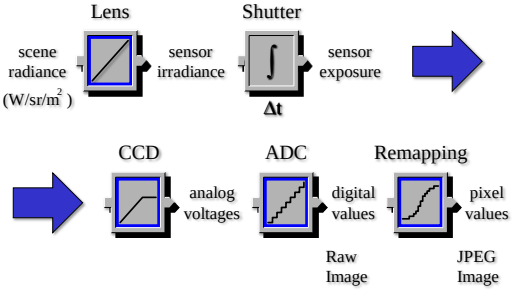
Geometric

- How pixel **coordinates** relate to **directions** in the world

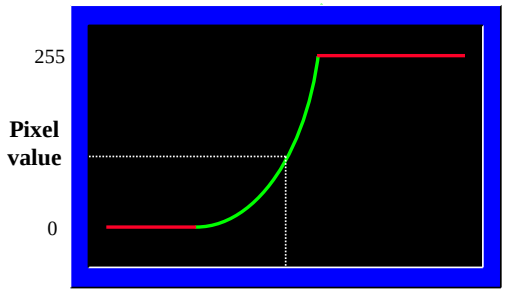
Photometric

- How pixel **values** relate to **radiance** amounts in the world

The Image Acquisition Pipeline



Imaging system response function



log Exposure = log (Radiance * Δt)

(CCD photon count)

Varying Exposure

Shutter Speed

Ranges: **Canon D30: 30 to 1/4,000 sec.**
Sony VX2000: ¼ to 1/10,000 sec.

Pros:

- **Directly varies the exposure**
- **Usually accurate and repeatable**

Issues:

- **Noise in long exposures**

Shutter Speed

Note: shutter times usually obey a power series – each "stop" is a factor of 2

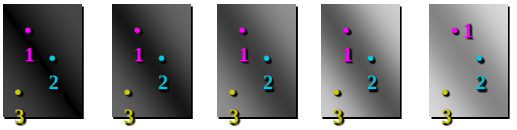
¼, 1/8, 1/15, 1/30, 1/60, 1/125, 1/250, 1/500, 1/1000 sec

Usually really is:

¼, 1/8, 1/16, 1/32, 1/64, 1/128, 1/256, 1/512, 1/1024 sec

The Algorithm

Image series



Δt = 1/64 sec | Δt = 1/16 sec | Δt = 1/4 sec | Δt = 1 sec | Δt = 4 sec

Pixel Value Z = f(Exposure)

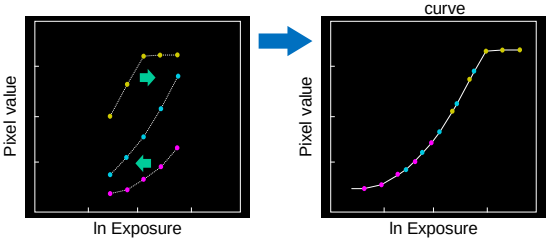
Exposure = Radiance ∗ Δt

log Exposure = log Radiance + log Δt

Response Curve

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:

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

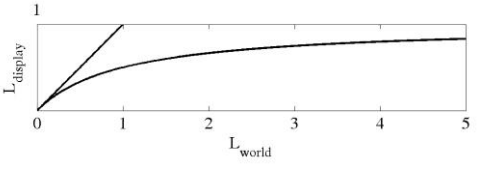
- Solve the overdetermined linear system:

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

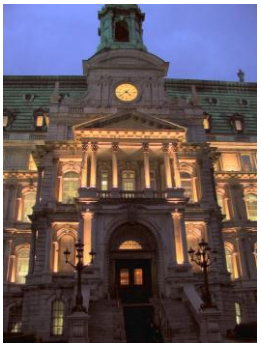
fitting term | smoothness term

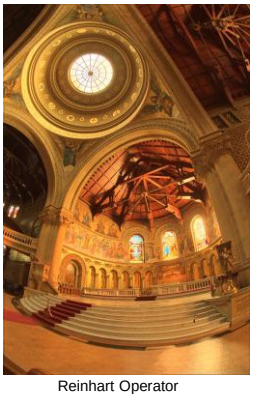
Global Operator (Reinhart et al)

$$L_{display} = \frac{L_{world}}{1 + L_{world}}$$

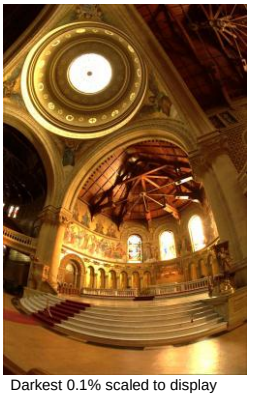


Global Operator Results





Reinhart Operator



Darkest 0.1% scaled to display device

Local tone mapping

3 exposures, Keith Moyer

Local tone mapping

"Do it right, it works. Do it wrong, it's brutal."

Keith Moyer

Tonal Aspects of Look

Ansel Adams Kenro Izu

Slides courtesy of Sylvain Paris

Tonal aspects of Look - Global Contrast

Ansel Adams Kenro Izu

High Global Contrast **Low Global Contrast**

SIGGRAPH2006

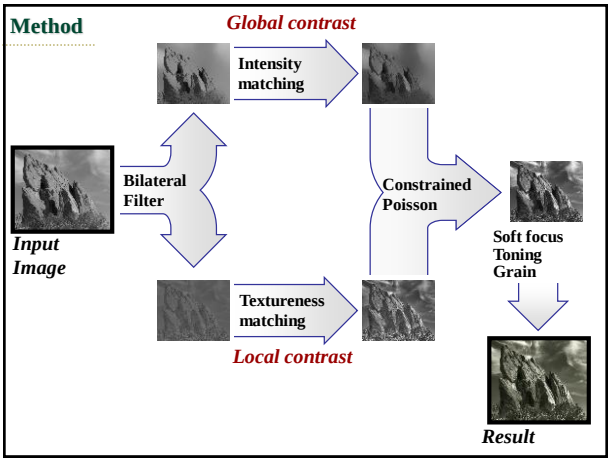
Tonal aspects of Look - Local Contrast

Ansel Adams Kenro Izu

Variable amount of texture **Texture everywhere**

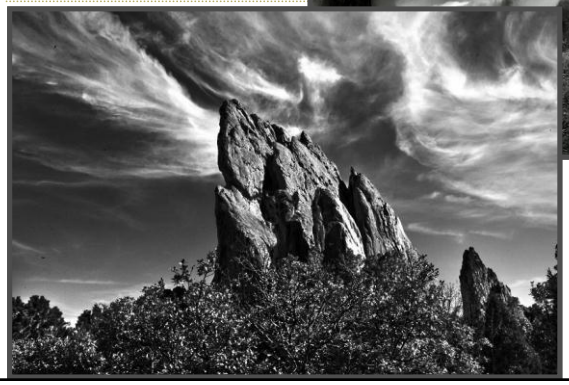
SIGGRAPH2006

Method



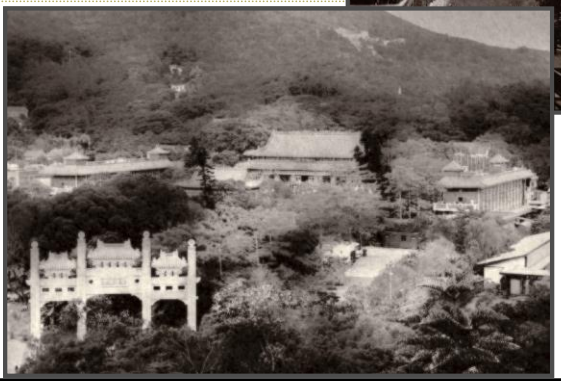
Result

Model

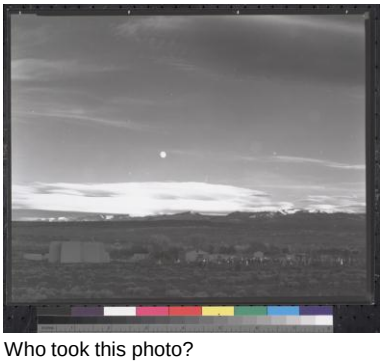


Result

Model



Professional examples



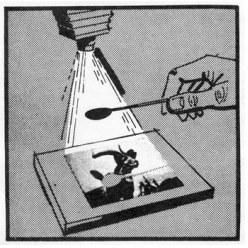
Who took this photo?

Ansel Adams

Contact print

Final

Dodge and burn



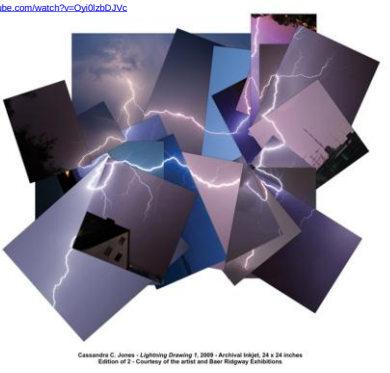
Is this dodging or burning?

Ernst Haas

Ernst Haas

Cassandra Jones

http://www.youtube.com/watch?v=Oyi0IzbDJVc



Cassandra C. Jones - *Lightning Drawing 7*, 2009 - Archival Inkjet, 24 x 24 inches
Edition of 2 - Courtesy of the artist and Baer Ridgway Exhibitions

Peter Funch

Peter Funch

Peter Funch

http://www.peterfunch.com/

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)