Cross-fertilization with other fields

Bill Freeman,
Fredo Durand
Martial Herbert
Aaron Hertzmann
Dimitri Metaxas

August 22, 2011

Cross-fertilization with other fields 2:30 - 3:45

Fredo Durand

- Bill Freeman: Computer Science, Computational Photography
- Fredo Durand: Computer Graphics 1
- Aaron Hertzmann: Computer Graphics 2
- Martial Herbert: Robotics
- Dimitri Metaxas: Medicine

Computational Photography

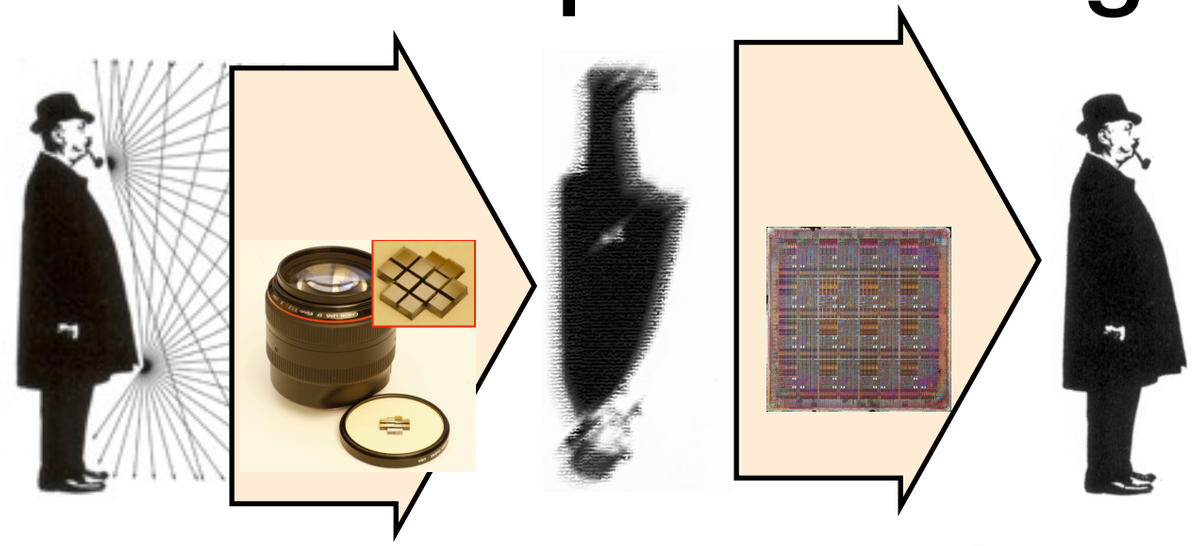
Frédo Durand
MIT CSAIL

sub for Bill Freeman

who will conveniently be back for beers

Computational Photography

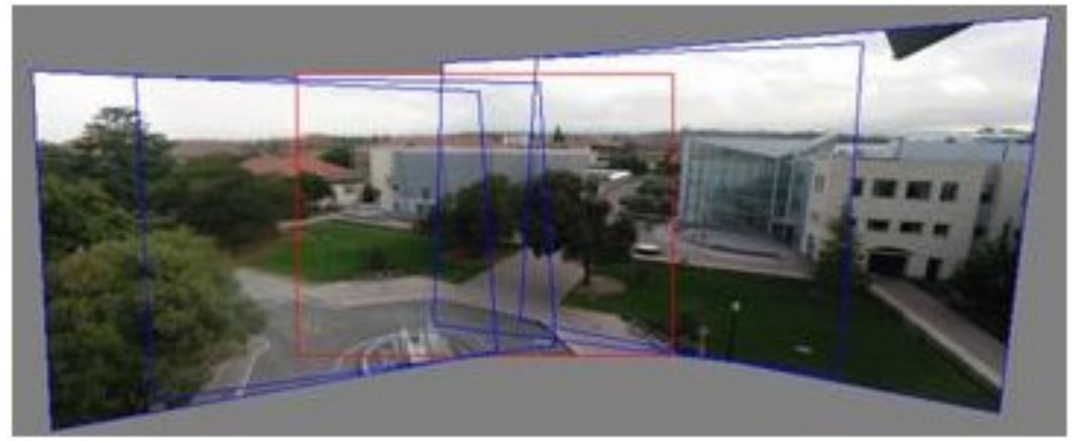
- Computation is an inherent part of image formation

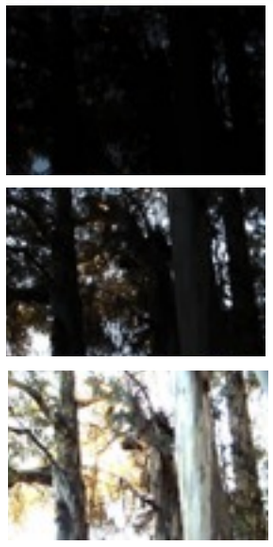


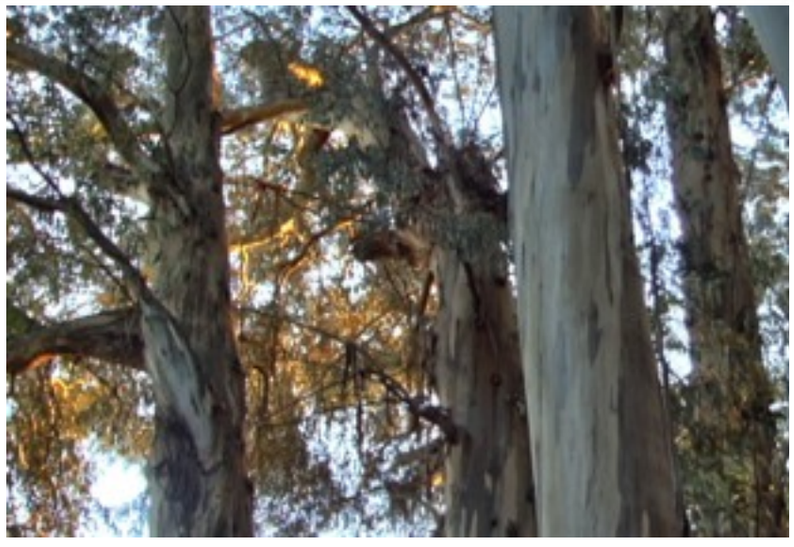
- ...or anything where computation helps with imaging
  - quantitively
  - qualitatively
  - automatically
  - user-assitedly

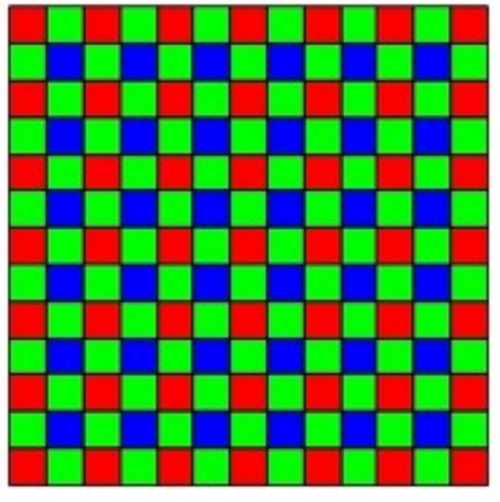
Multiple-exposure & multiplexing

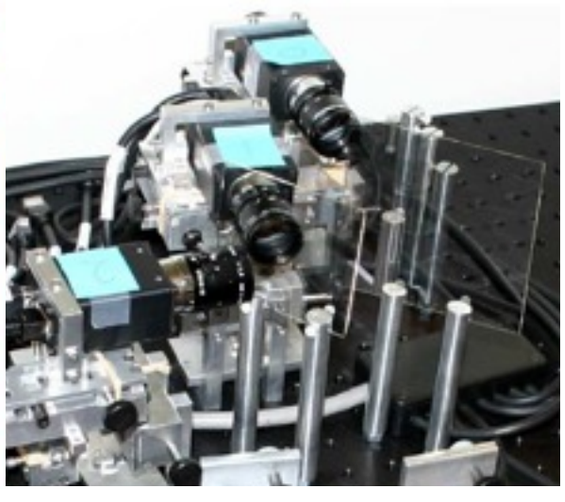
- Expand capabilities by combining multiple images
- Multiplex through time, assorted pixels, beam splitters, camera array
- e.g.
  - Panorama stitching
  - HDR imaging
  - Focus stacks
  - Photomontage
  - Super-resolution





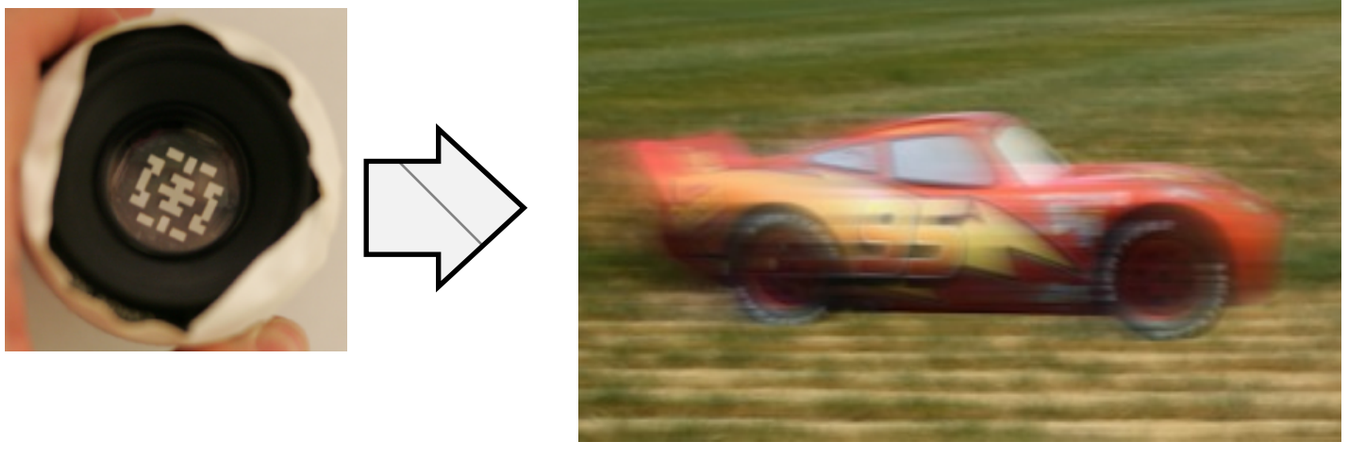




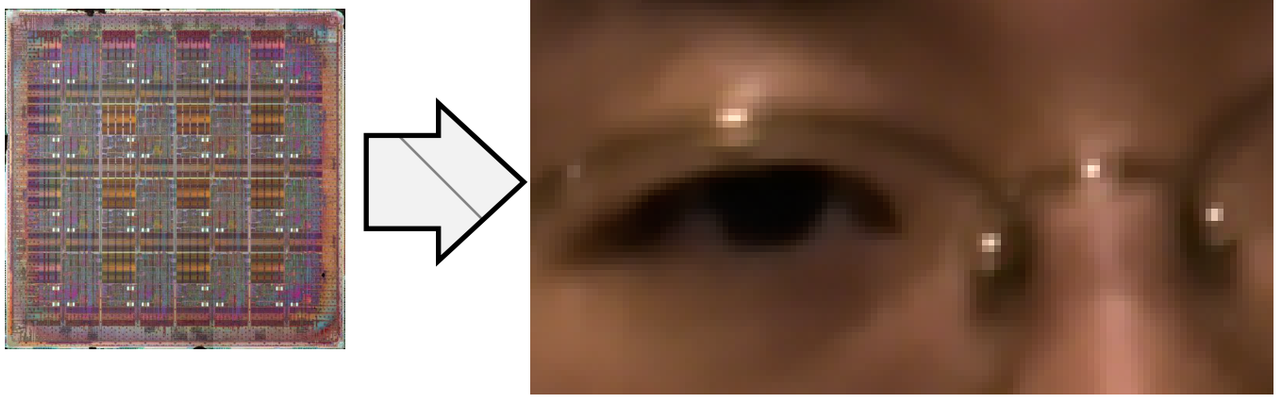


Coded Imaging

Optics encodes information



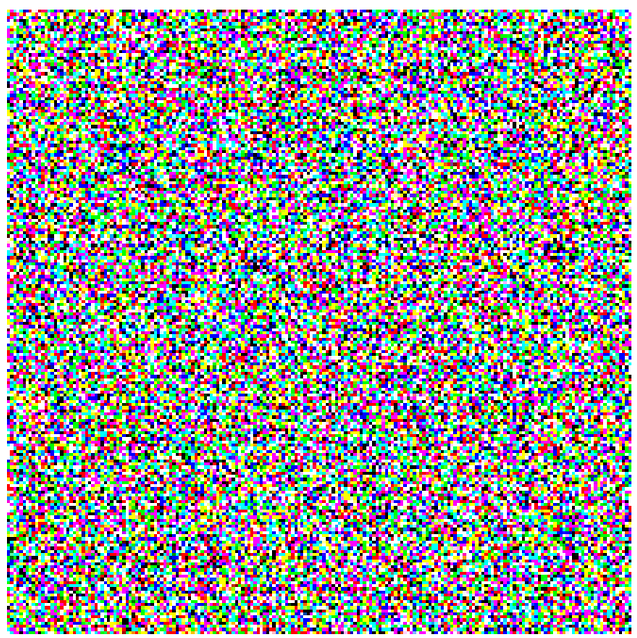
Computation decodes



- e.g.
  - motion-invariant
  - coded aperture
  - flutter shutter
  - wavefront coding
  - compressive sensing
  - heterodyning
  - warp-unwarp

Natural signal prior

- Statistics that distinguish images of the world from random signals
- Use to “bias” algorithms to output more likely results or to disambiguate ill-posed problems
- Extension of regularization
- **e.g.**
  - Denoising
  - Deconvolution
  - Compressive sensing
  - Light field prior

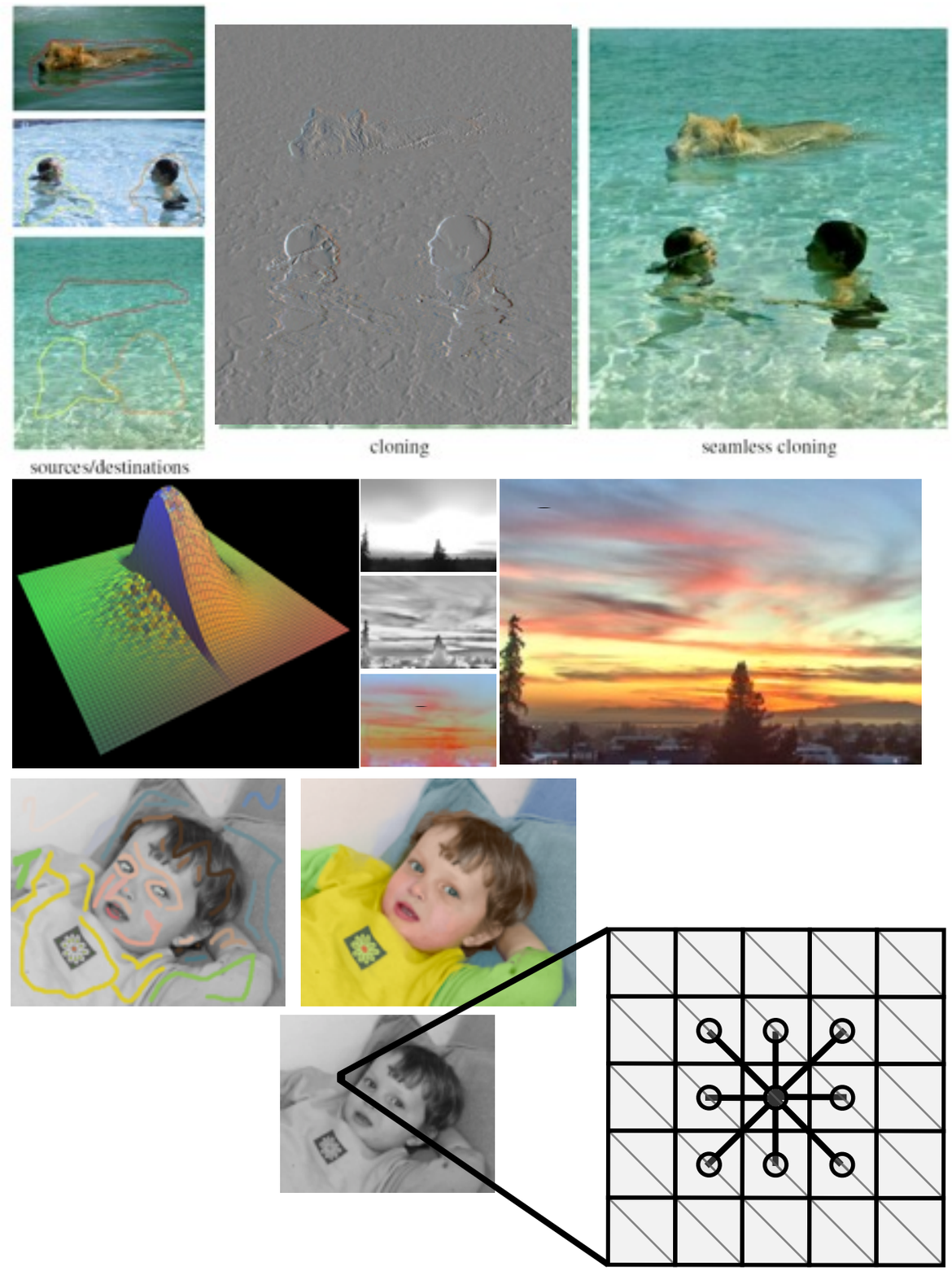


Random

“Natural” image

Edges matter but are not binary

- Sparse derivative image prior
- Gradient domain (seamless cloning, tone mapping, convert2gray)
- Bilateral filter for decomposition
- Non-homogenous regularization for scribble propagation



Leverage millions of images

- The ultimate prior?
- Reconstruct the world

Hays & Efros 07

The raw data is high dimensional

- Light field: 4D (space-angle)
- Time space: 3D
- +Fourier

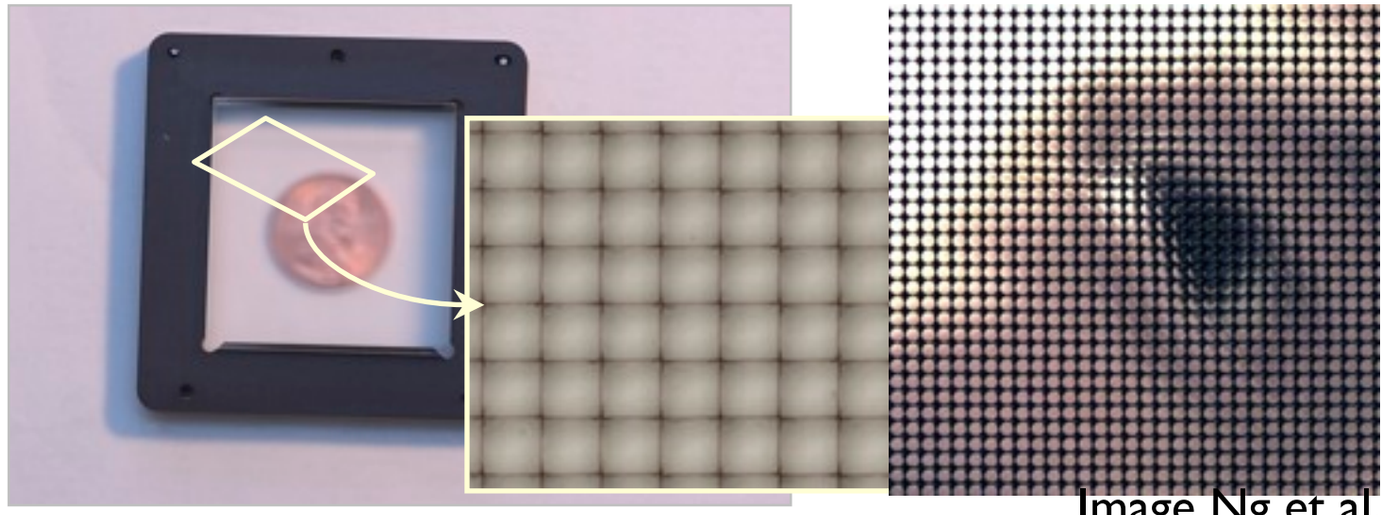
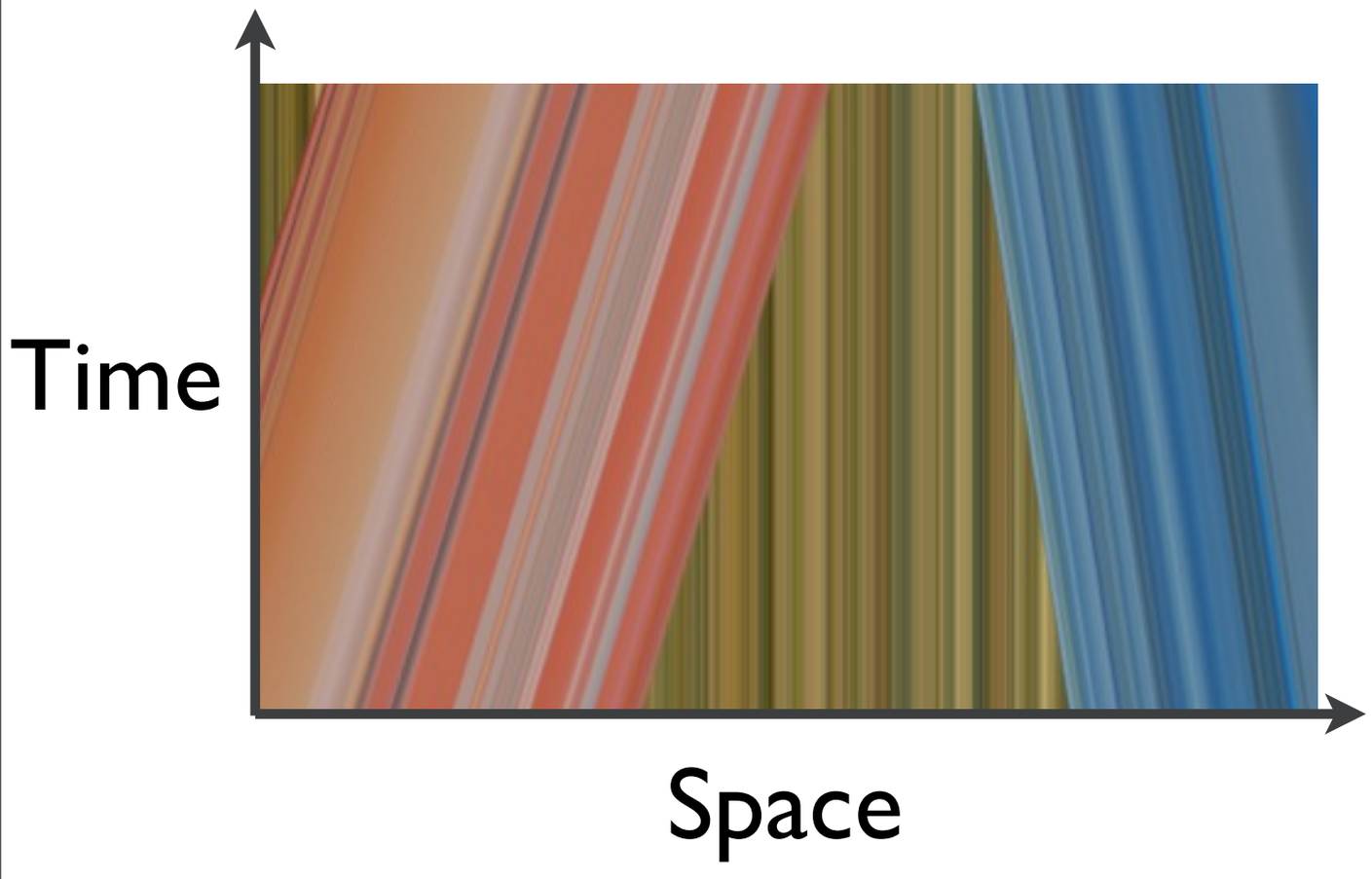
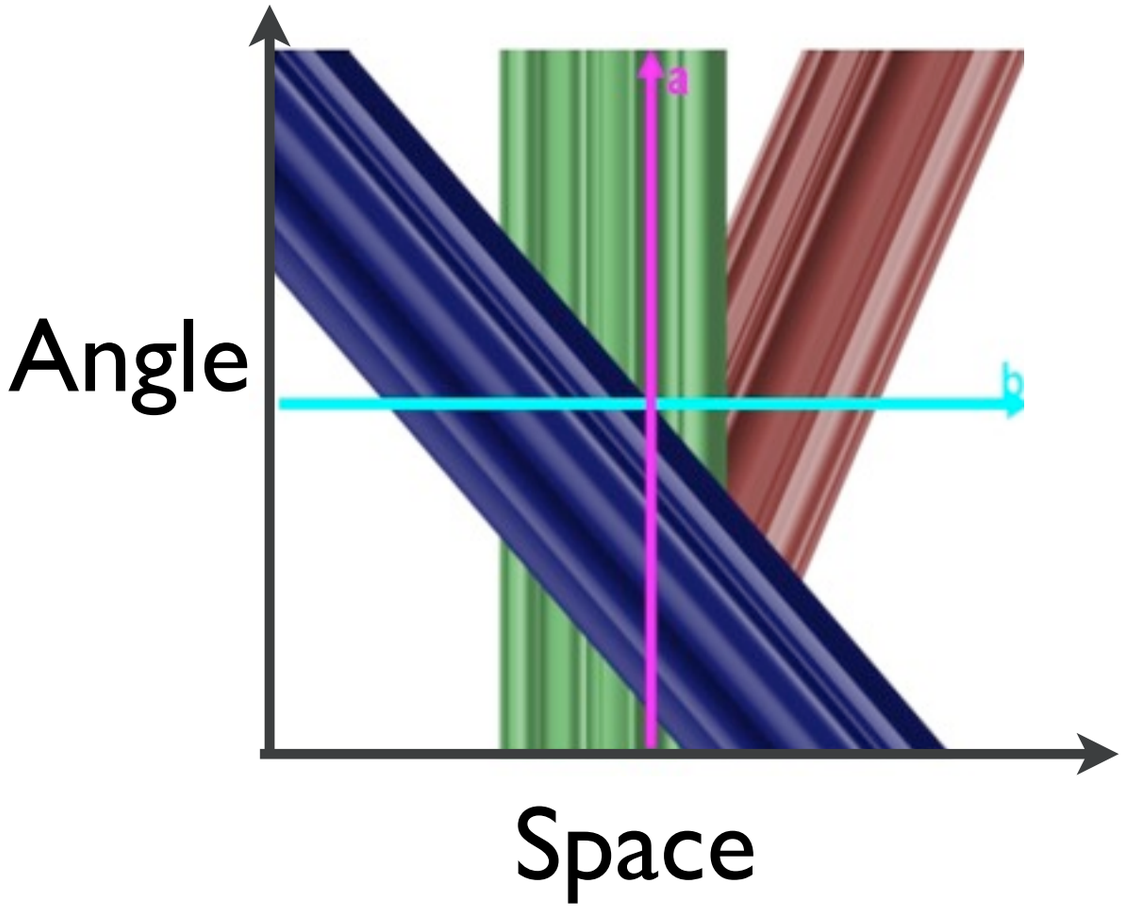


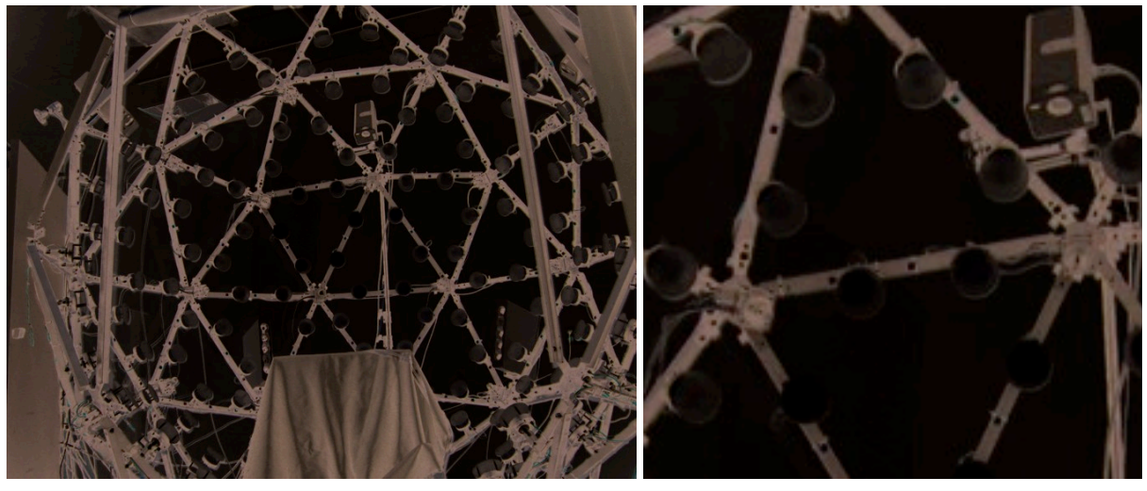
Image Ng et al.





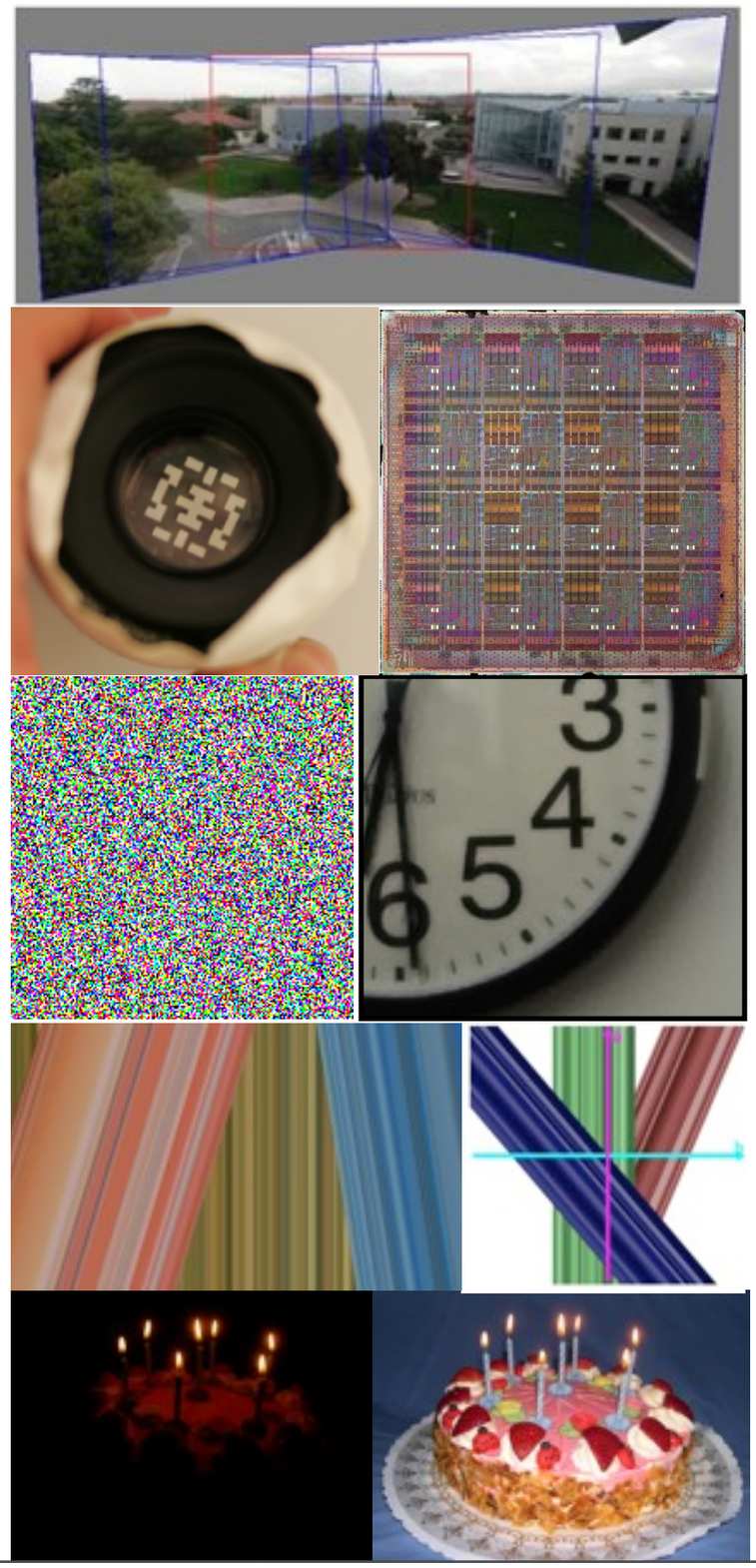
Active imaging

- Modulate light to facilitate information gathering
- e.g.
  - Flash/no flash
  - Light stages
  - Dual imaging
  - Structured-light scanning



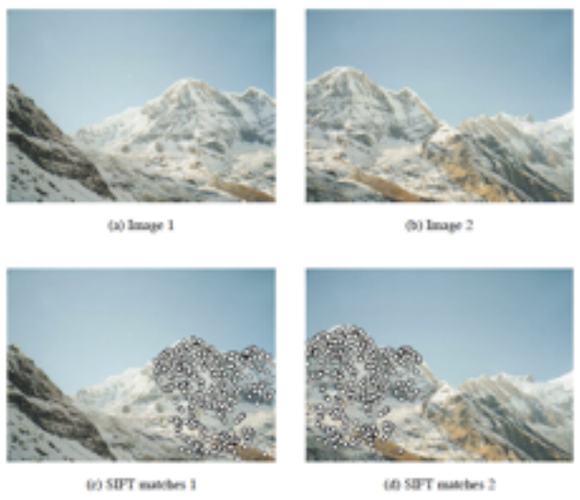
Recap: Big ideas in comp. photo.

- Multiplexing:
  quality through quantity
- Coded imaging
- Natural signal prior
- Edges matter but should not be detected
- Leverage millions of images
- Raw data is high-dimensional (ligh field, space-time)
- Active Imaging



Current successes

- Panorama stitching


(a) Image 1

(b) Image 2

(c) SIFT matches 1

(d) SIFT matches 2


(e) Images aligned according to a homography

- High-Dynamic-RangeImaging & tone mapping

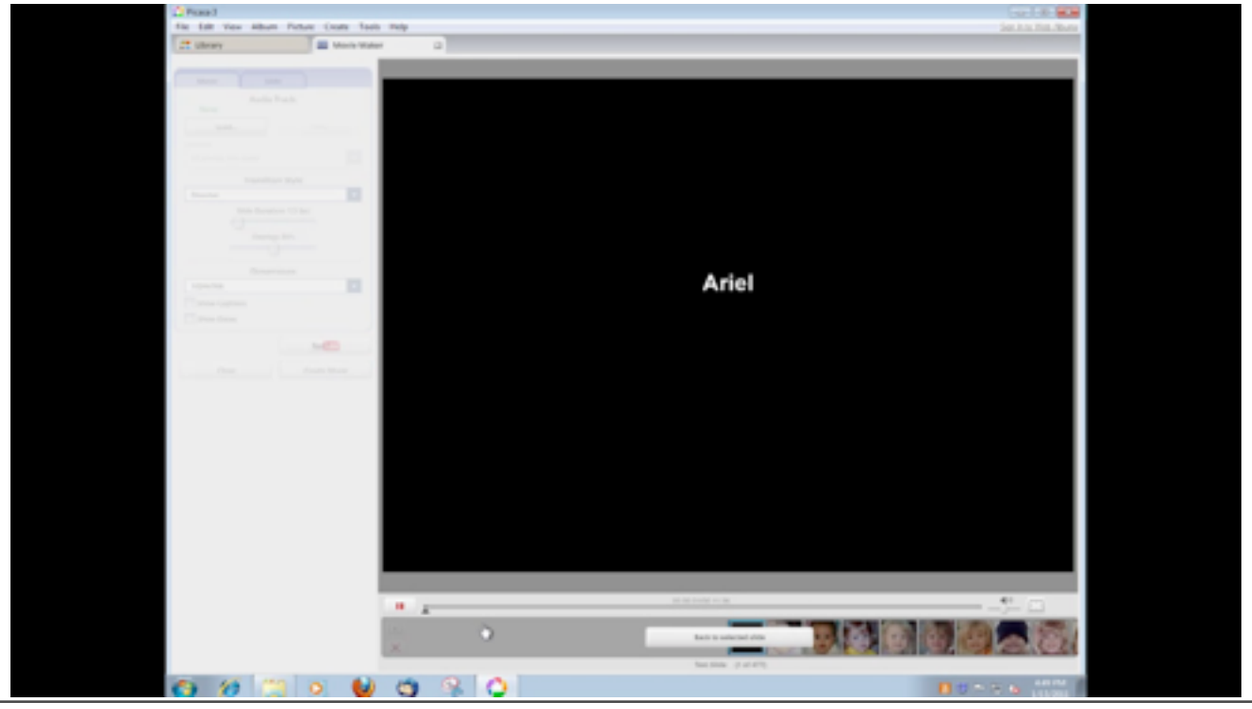

Top-left: iPhone photo #1. Bottom-left: iPhone photo #2. Right: **TrueHDR** result.

Current successes

- Face detection (+smile +blink)

Joz Wang

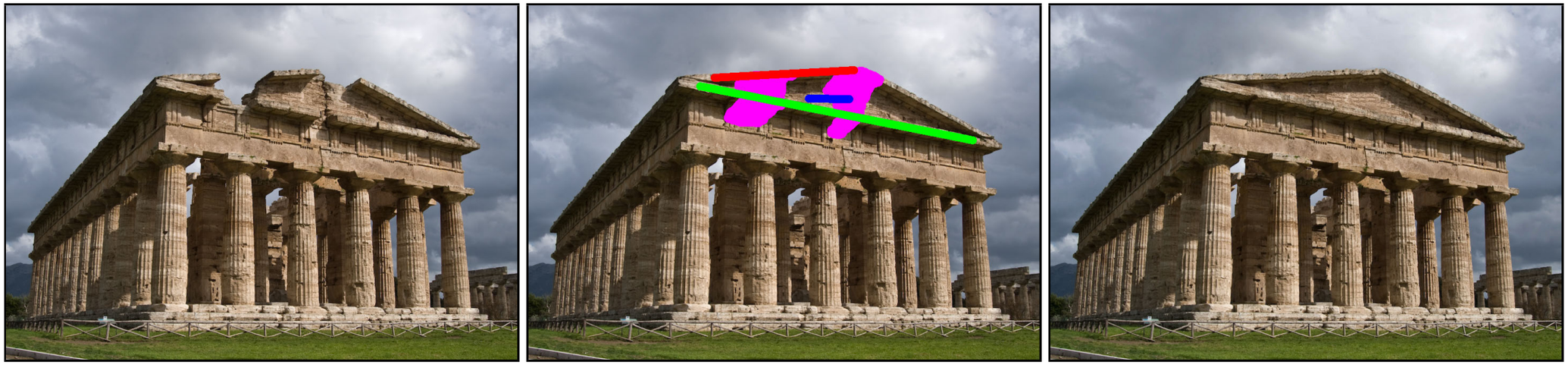
- Photo bios



Current successes

- Poisson image editing / healing brush

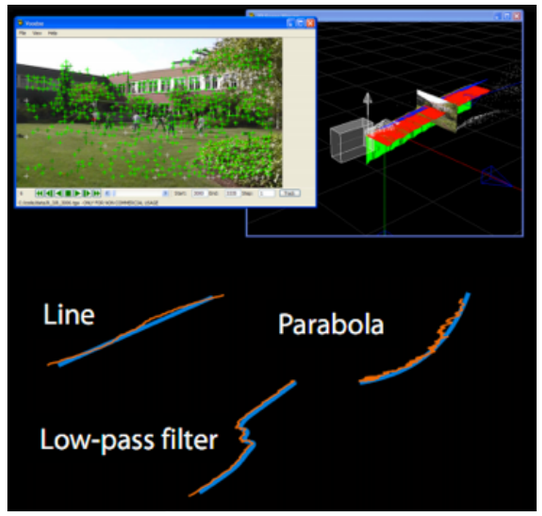
- Patch match (content-aware fill)

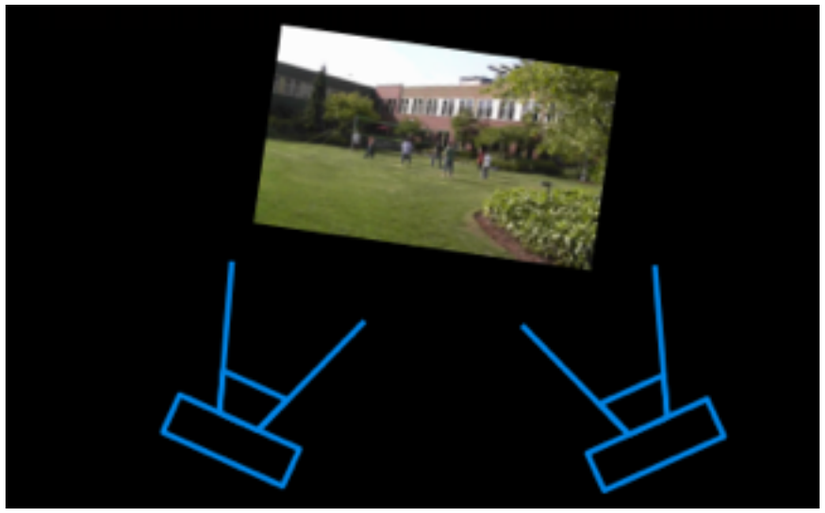


(a) original (b) hole+constraints (c) hole filled

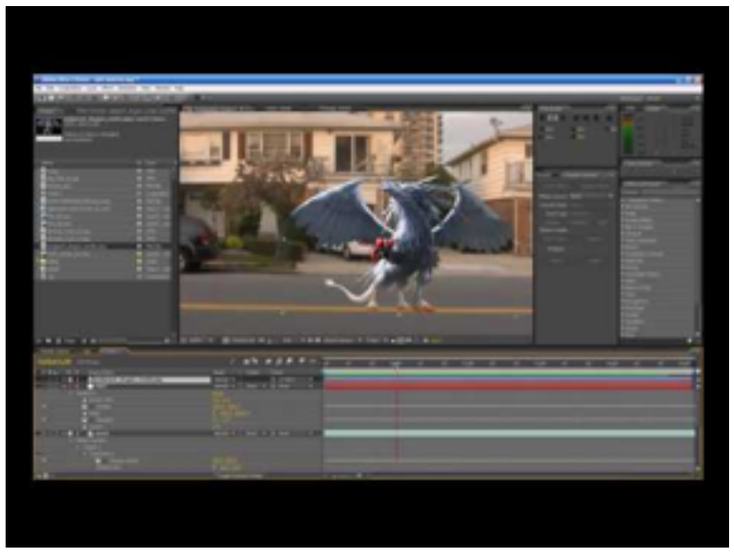
Current successes

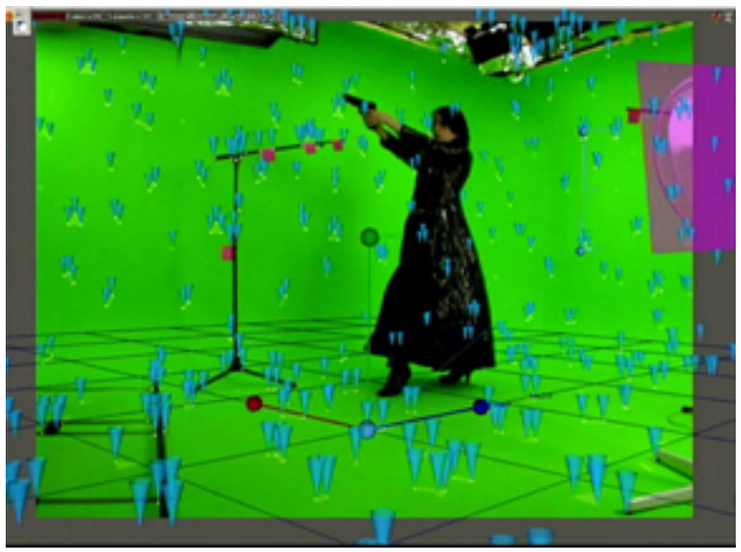
- Video stabilization





- match move

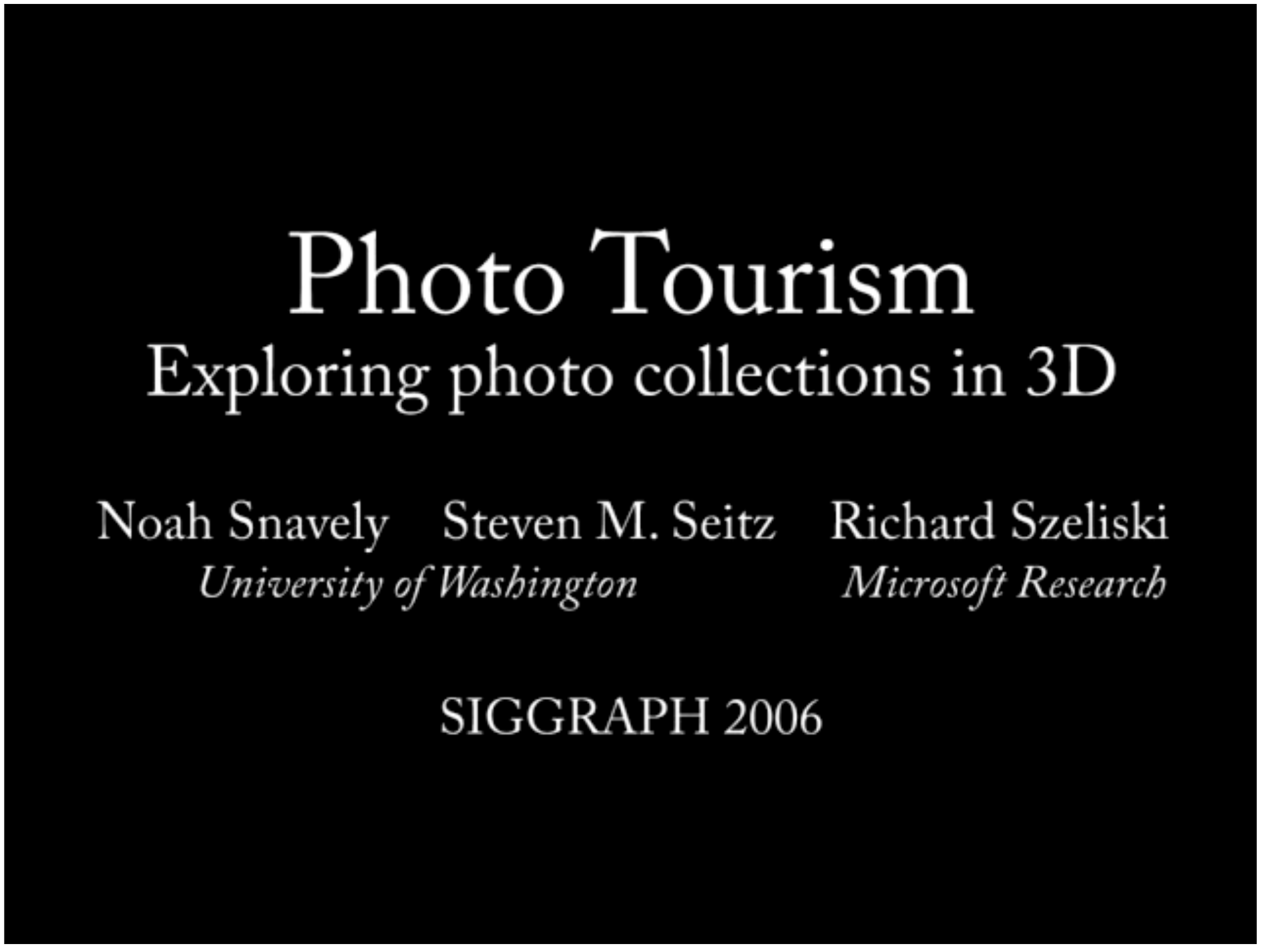




- Tracking

Current successes

- Photo tourism / Photosynth

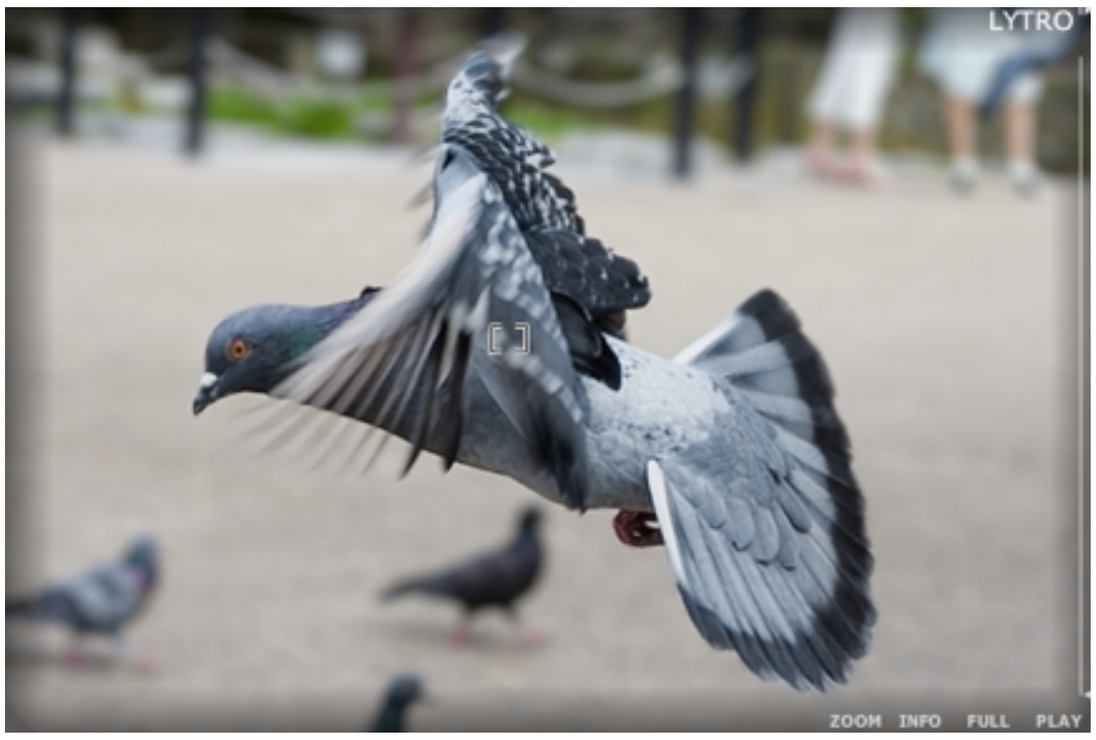


Current successes

- Calibrate & remove blur
- e.g. DXO, Adobe, Panasonic, Mamya

Current successes

- Light field cameras



Single Lens Stereo with a Plenoptic Camera

Edward H. Adelson and John Y.A. Wang

Abstract—Ordinary cameras gather light across the area of their lens aperture, and the light striking a given subregion of the aperture is structured somewhat differently than the light striking an adjacent subregion. By analyzing this optical structure, one can infer the depths of objects in the scene, i.e., one can achieve "single lens stereo." We describe a novel camera for performing this analysis. It incorporates a single main lens along with a lenticular array placed at the sensor plane. The resulting "plenoptic camera" provides information about how the scene would look when viewed from a continuum of possible viewpoints bounded by the main lens aperture. Deriving depth information is simpler than in a binocular stereo system because the correspondence problem is minimized. The camera extracts information about both horizontal and vertical parallax, which improves the reliability of the depth estimates.

I. INTRODUCTION

"EVERY BODY in the light and shade fills the surrounding air with infinite images of itself; and these, by infinite pyramids diffused in the air, represent this body throughout space and on every side." Leonardo da Vinci [1] uses

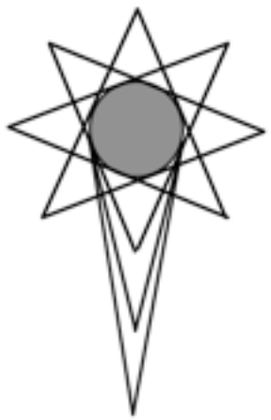
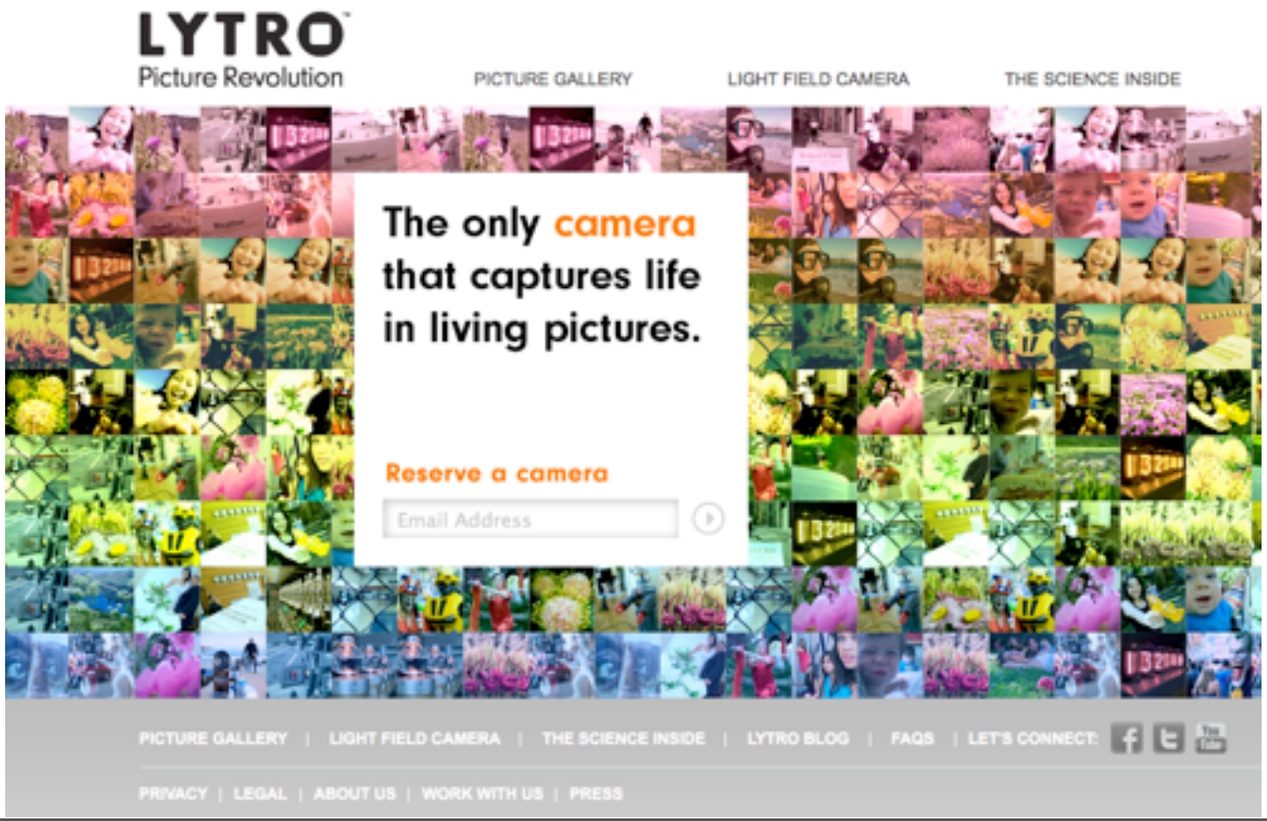


Fig. 1. Diagram from Leonardo's notebooks illustrating the fact that the light rays leaving an object's surface may be considered to form a collection of cones (which Leonardo calls "pyramids"), each cone constituting an image that would be seen by a pinhole camera at a given location.



Open Challenges

- Upper bounds on acquisition/reconstruction
- Natural image priors
- Light field, space time priors/reconstruction
- Computational illumination
- New modalities (coherent, femtosecond)
- Video mid-level representation

- Link to other fields
  - Astronomy, microscopy, medical, radar, science