SIGGRAPH 2008 Course:
Computational Photography: Advanced Topics

http://computationalphotography.org

Speakers

- Paul Debevec (USC, USA) (debevec (at) ict.usc.edu)
- Ramesh RASKAR (MIT Media Lab, USA) (raskar (at) media.mit.edu)
- Jack TUMBLIN (Northwestern University, USA) (jet (at) cs.northwestern.edu)

Course Abstract

Computational photography combines plentiful computing, digital sensors, modern optics, many varieties of actuators, probes and smart lights to escape the limitations of traditional film cameras and enables novel imaging applications. Unbounded dynamic range, variable focus, resolution, and depth of field, hints about shape, reflectance, and lighting, and new interactive forms of photos that are partly snapshots and partly videos, performance capture and interchangeably relighting real and virtual characters are just some of the new applications emerging in Computational Photography. The computational techniques encompass methods from modification of imaging parameters during capture to sophisticated reconstructions from indirect measurements.

We will bypass basic and introductory material presented in earlier versions of this course (Computational Photography 2005,6,7) and expand coverage of more recent topics. Emphasizing more recent work in computational photography and related fields (2006 or later) this course will give more attention to advanced topics only briefly touched before, including tomography, heterodyning and Fourier Slice applications, inverse problems, gradient illumination, novel optics, emerging sensors and social impact of computational photography. With this deeper coverage, the course offers a diverse but practical guide to topics in image capture and manipulation methods for generating compelling pictures for computer graphics and for extracting scene properties for computer vision, with several examples.

Speaker Info

Paul Debevec
Research Associate Professor, USC

Paul Debevec is a research associate professor at the University of Southern California and the associate director of graphics research at USC's Institute for Creative Technologies. Debevec's Ph.D. thesis (UC Berkeley, 1996) presented Façade, an image-based modeling and rendering system for creating photoreal architectural models from photographs. Using Facade he led the creation of virtual cinematography of the Berkeley campus for his 1997 film The Campanile Movie whose techniques were used to create virtual backgrounds in the 1999 film The Matrix. Subsequently, Debevec developed techniques for illuminating computer-generated scenes with real-world lighting captured through high dynamic range photography, demonstrating new image-based lighting techniques in his films Rendering with Natural Light (1998), Fiat Lux (1999), and The Parthenon (2004); he also led the design of HDR Shop, the first high dynamic range image editing program. At USC ICT, Debevec has led the development of a series of Light Stage devices for capturing and simulating how objects and people reflect light, recently used to create realistic digital actors in films such as Spider Man 2 and Superman Returns. He is the recipient of ACM SIGGRAPH's first Significant New Researcher Award and a co-author of the 2005 book High Dynamic Range Imaging from Morgan Kaufmann.

Ramesh Raskar
Associated Professor, Media Lab, MIT

Ramesh Raskar joined the Media Lab in spring 2008 as head of the Camera Culture research group. He was a a Senior Research Scientist at MERL. The group focuses on developing tools to help us capture and share the visual experience. This research involves developing novel cameras with unusual optical elements, programmable illumination, digital wavelength control, and femtosecond analysis of light transport, as well as tools to decompose pixels into perceptually meaningful components. He is a member of the ACM and IEEE.

Jack Tumblin
Associate Professor, EECS Dept. Northwestern University

Jack Tumblin is an Associate Professor of Computer Science at Northwestern University. His interests include novel photographic sensors and lighting devices to assist museum curators in historical preservation, computer graphics and visual appearance, and image-based modeling and rendering. During his doctoral studies at Georgia Tech and post-doc at Cornell, he investigated tone-mapping methods to depict high-contrast scenes. His MS in Electrical Engineering (December 1990) and BSEE (1978), also from Georgia Tech, bracketed his work as co-founder of IVEX Corp., (>45 people as of 1990) where he designed flight simulators. He was co-organizer of Computational Photography courses at Siggraph 2005 and 2006. He was an Associate Editor of ACM Transactions on Graphics 2001-2006, and holds 9 patents.

Schedule

Module 1: 90 minutes

9:00:	A.1 Introduction and Overview	(Raskar, 15 minutes)
9:15:	A.2 Concepts in Computational Photography	(Tumblin, 15 minutes)
9:30:	A.3 Optics: Computable Extensions	(Raskar, 30 minutes)
10:00:	A.4 Sensor Innovations	(Tumblin, 30 minutes)
10:30:	Q & A	(5 minutes)

10:35: Break: 25 minutes

Module 2: 90 minutes

11:00:	B.1 Illumination As Computing	(Debevec, 25 minutes)
11:25:	B.2 Scene and Performance Capture	(Debevec, 20 minutes)
11:45:	B.3 Image Aggregation & Sensible Extensions	(Tumblin, 20 minutes)
12:05:	B.4 Community and Social Impact	(Raskar, 20 minutes)
12:25:	B.4 Summary and Discussion, Q&A	(All, 10 minutes)

Course: Computational Photography Advanced Topics

Module 1: 90 minutes

9:00:	A.1	Introduction and Overview	(Raskar, 15 minutes)
9:15:	A.2	Concepts in Computational Photography	(Tumblin, 15 minutes)
9:30:	A.3	Optics: Computable Extensions	(Raskar, 30 minutes)
10:00:	A.4	Sensor Innovations	(Tumblin, 30 minutes)
10:30:	Q & A		(5 minutes)

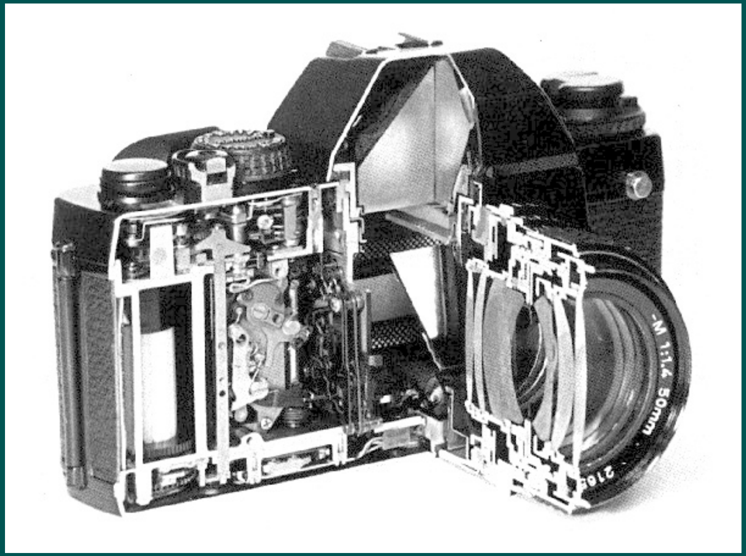
10:35: Break: 25 minutes

Module 2: 90 minutes

11:00:	B.1	Illumination As Computing	(Debevec, 25 minutes)
11:25:	B.2	Scene and Performance Capture	(Debevec, 20 minutes)
11:45:	B.3	Image Aggregation & Sensible Extensions	(Tumblin, 20 minutes)
12:05:	B.4	Community and Social Impact	(Raskar, 20 minutes)
12:25:	B.4 Summary and Discussion, Q&A		(All, 10 minutes)

Course Page : http://computationalphotography.org/

Class: Computational Photography



Organisers

Ramesh Raskar

MIT – Media Lab

Jack Tumblin

Northwestern University

What *is* Photography?

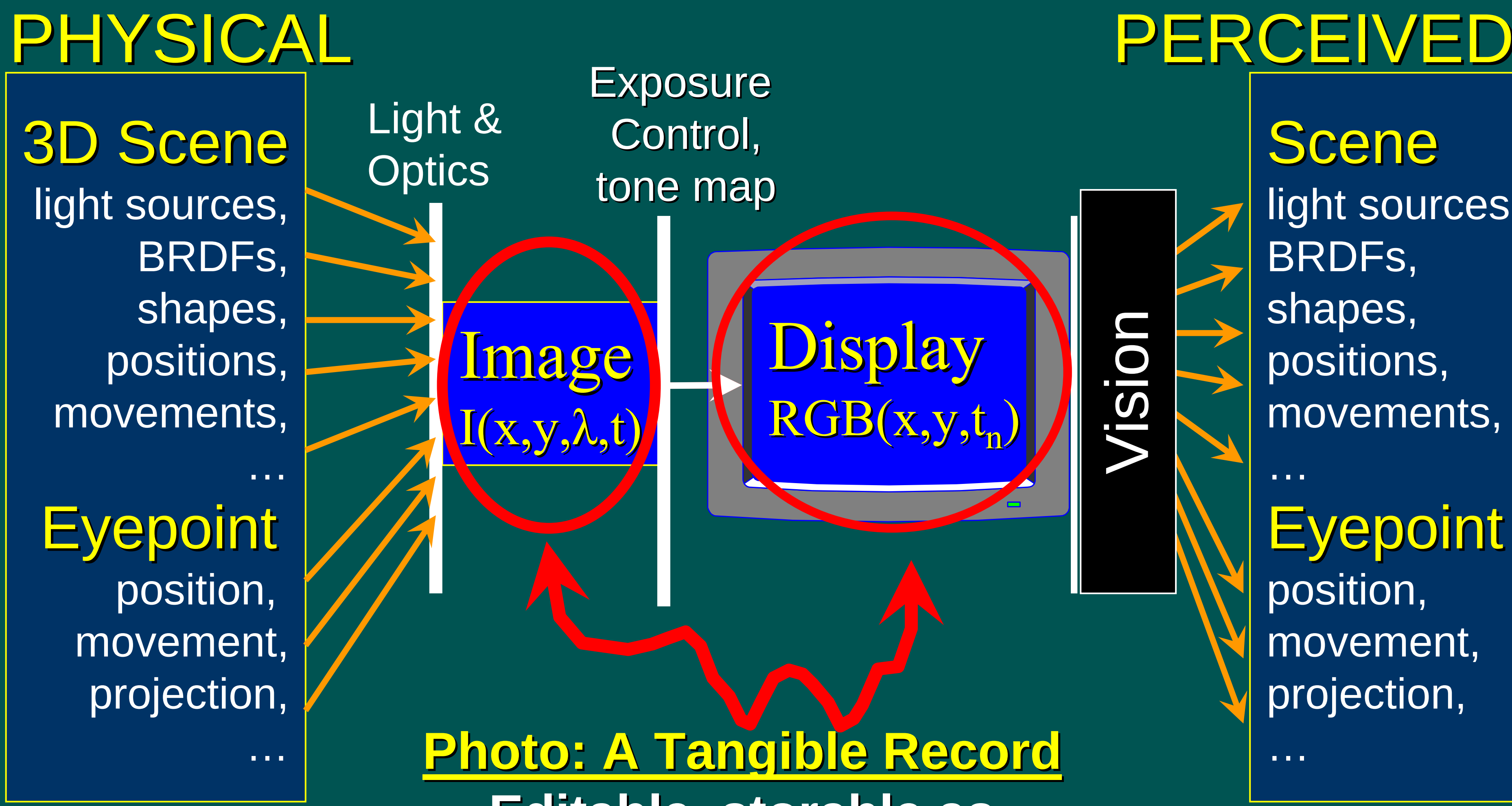
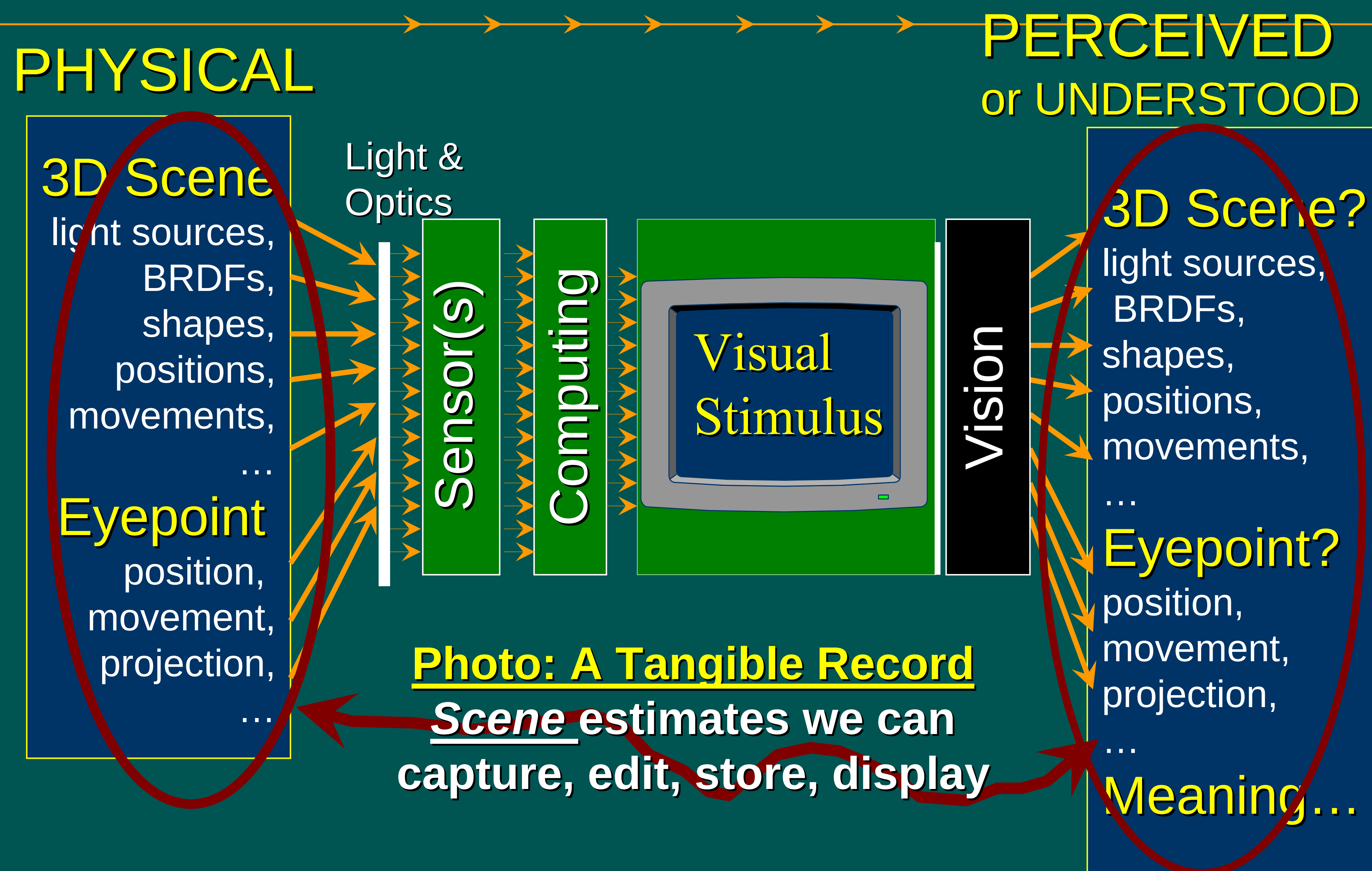


Photo: A Tangible Record

Editable, storable as
Film or Pixels

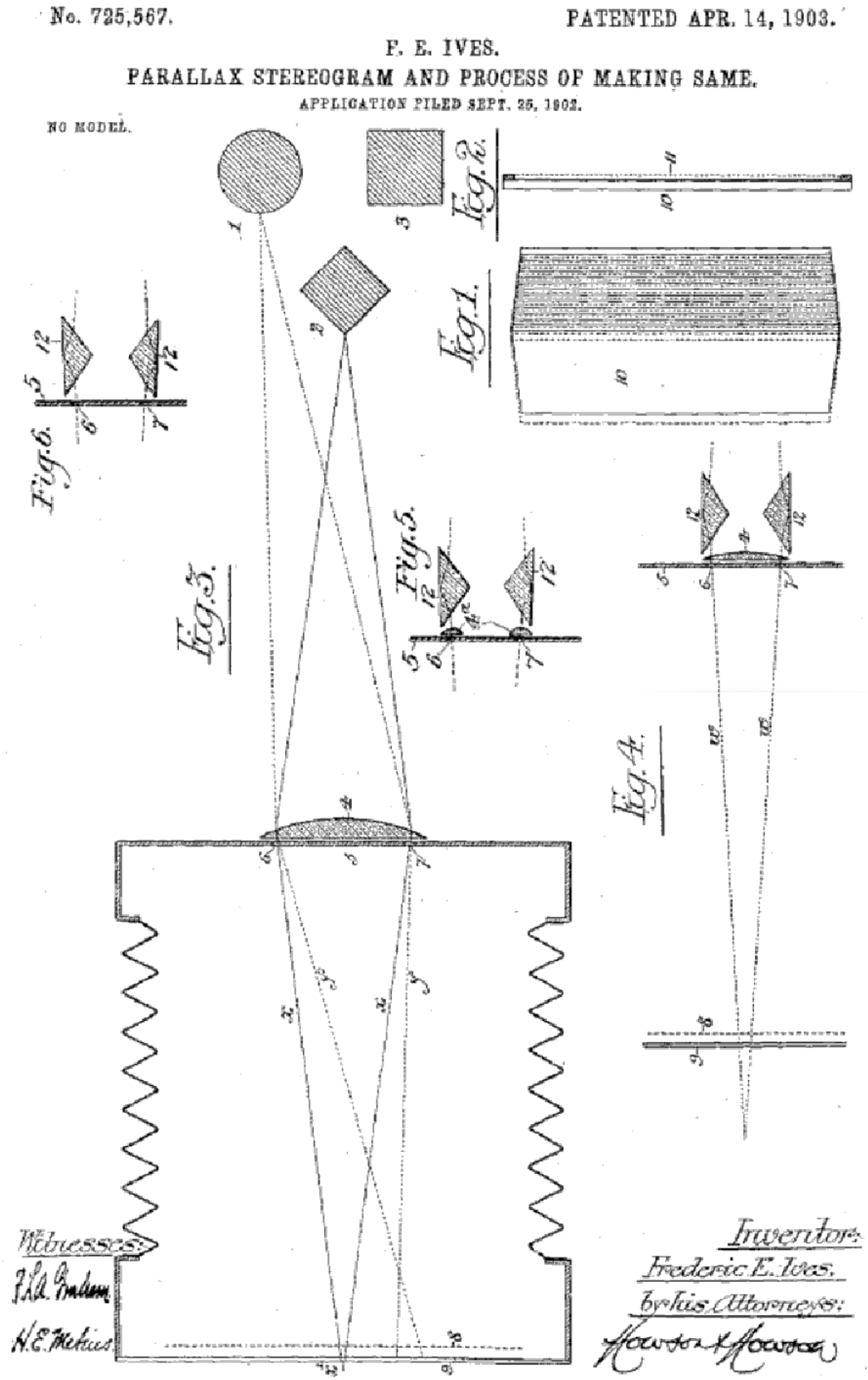
Ultimate Photographic Goals



Ives' Camera

Patented 1903

Array of pinholes near image plane

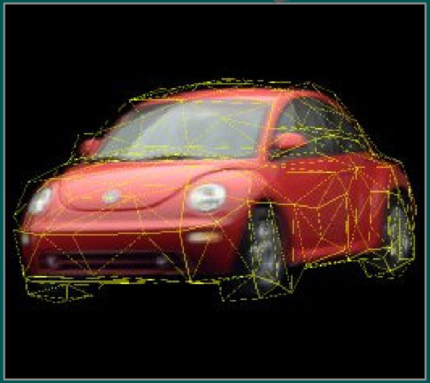


Devices for recording light fields (using geometrical optics)

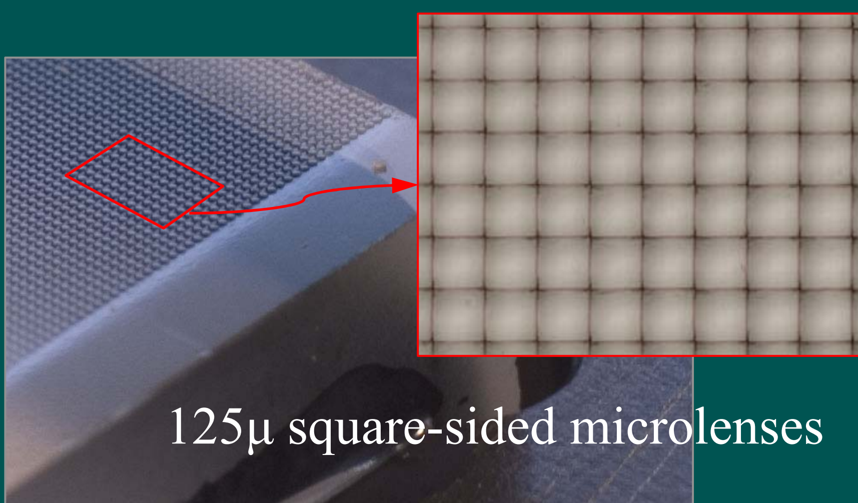
big baseline

small baseline

- handheld camera [Buehler 2001]
- camera gantry [Stanford 2002]
- array of cameras [Wilburn 2005]
- plenoptic camera [Ng 2005]
- light field microscope [Levoy 2006]



Digital Refocusing using Light Field Camera

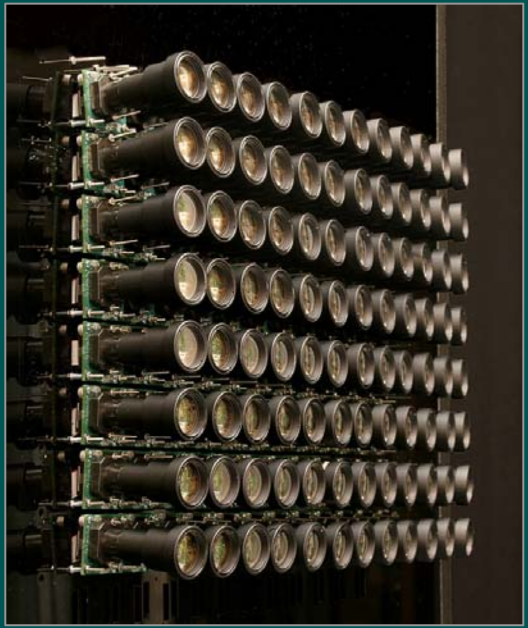


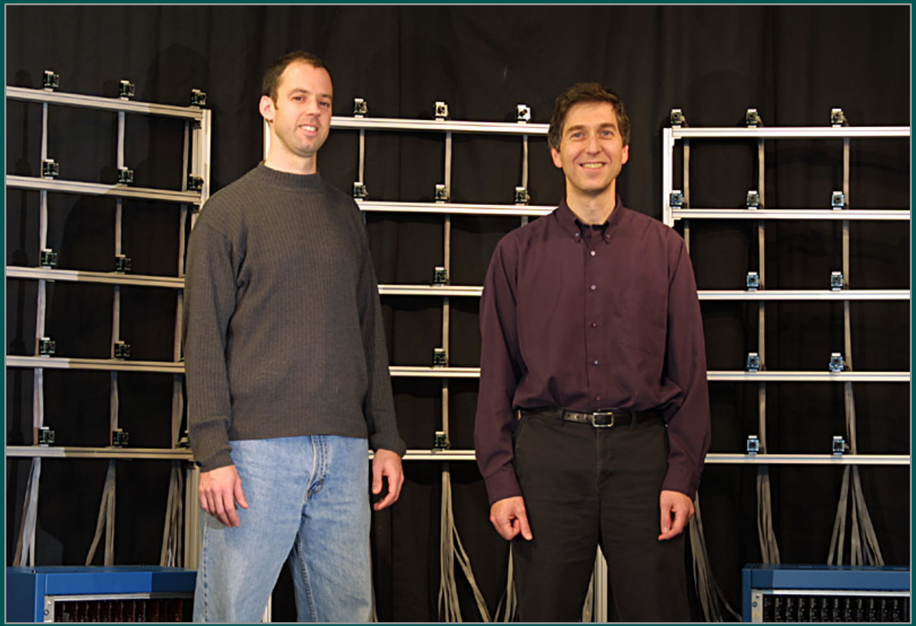
125μ square-sided microlenses

High performance imaging using large camera arrays

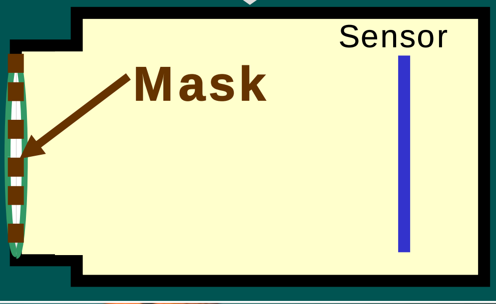
Bennett Wilburn, Neel Joshi, Vaibhav Vaish, Eino-Ville Talvala, Emilio Antunez, Adam Barth, Andrew Adams, Mark Horowitz, Marc Levoy

(Proc. SIGGRAPH 2005)

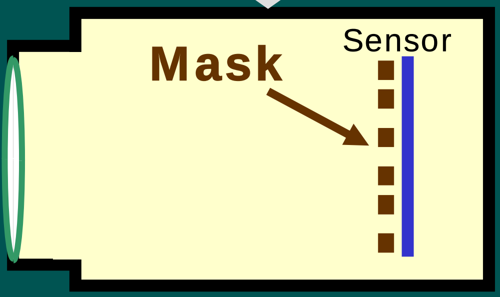


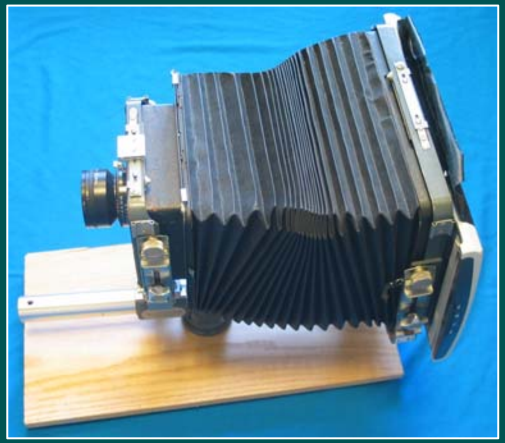


Coding and Modulation in Camera Using Masks



Coded Aperture for
Full Resolution
Digital Refocusing

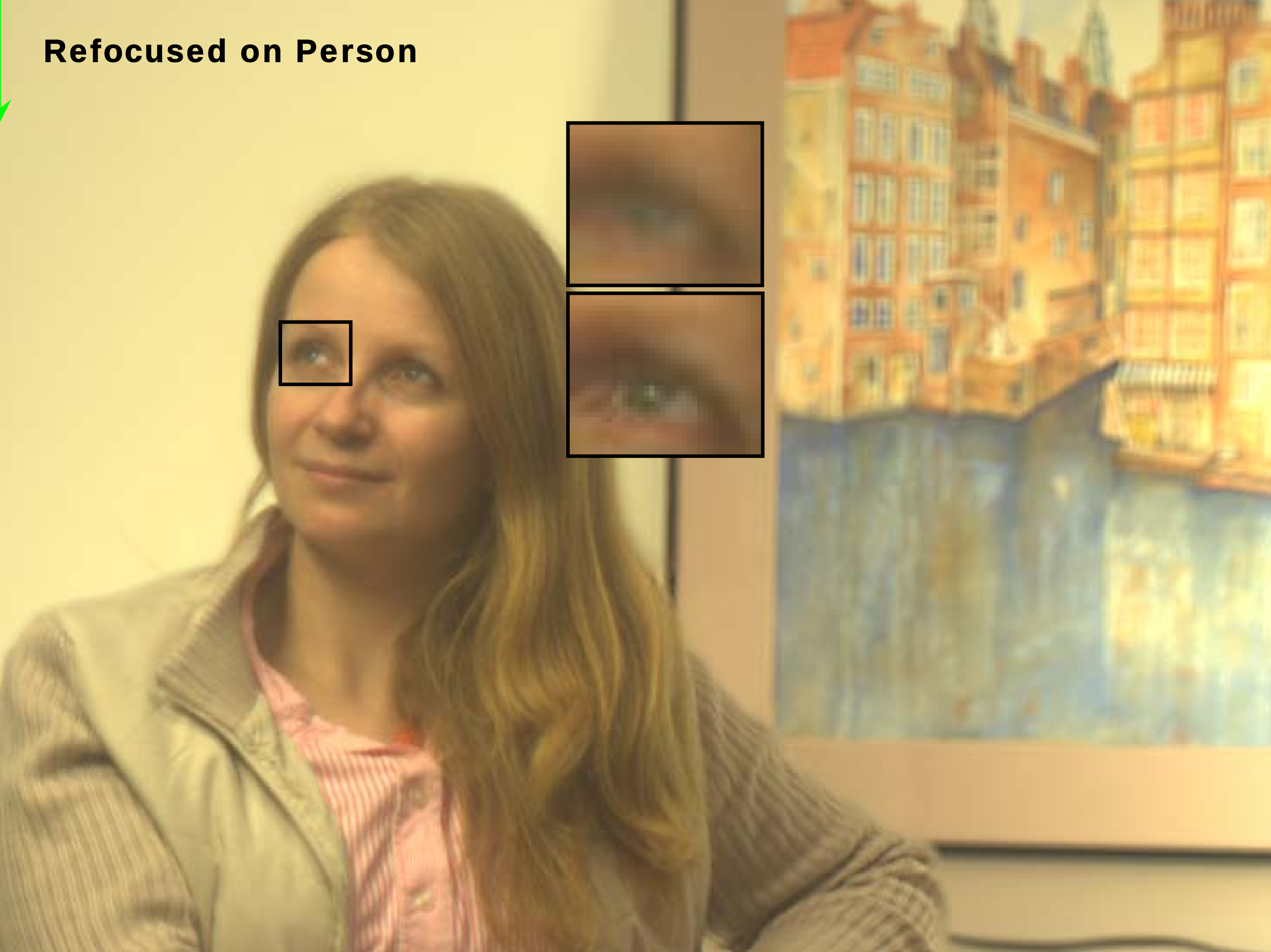




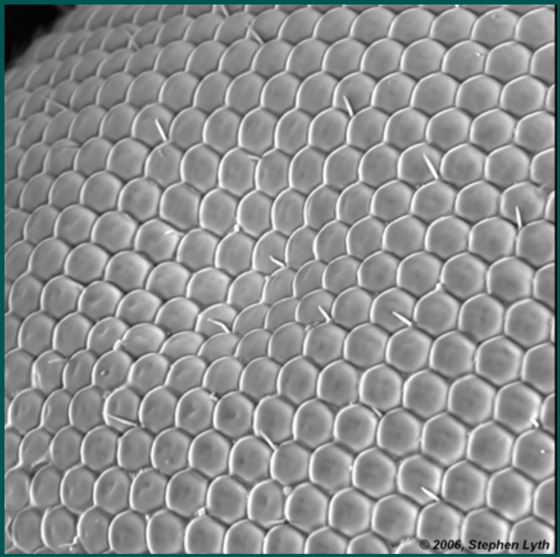
Heterodyne Light
Field Camera

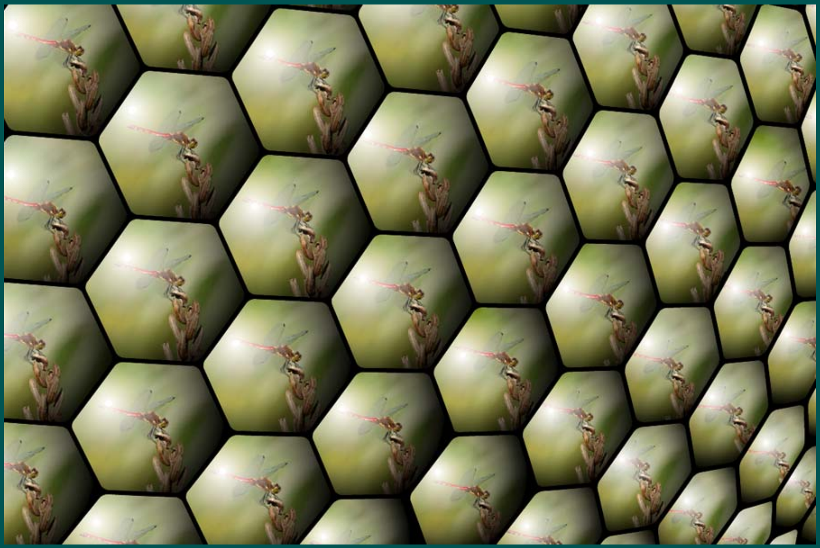

Captured Blurred
Photo

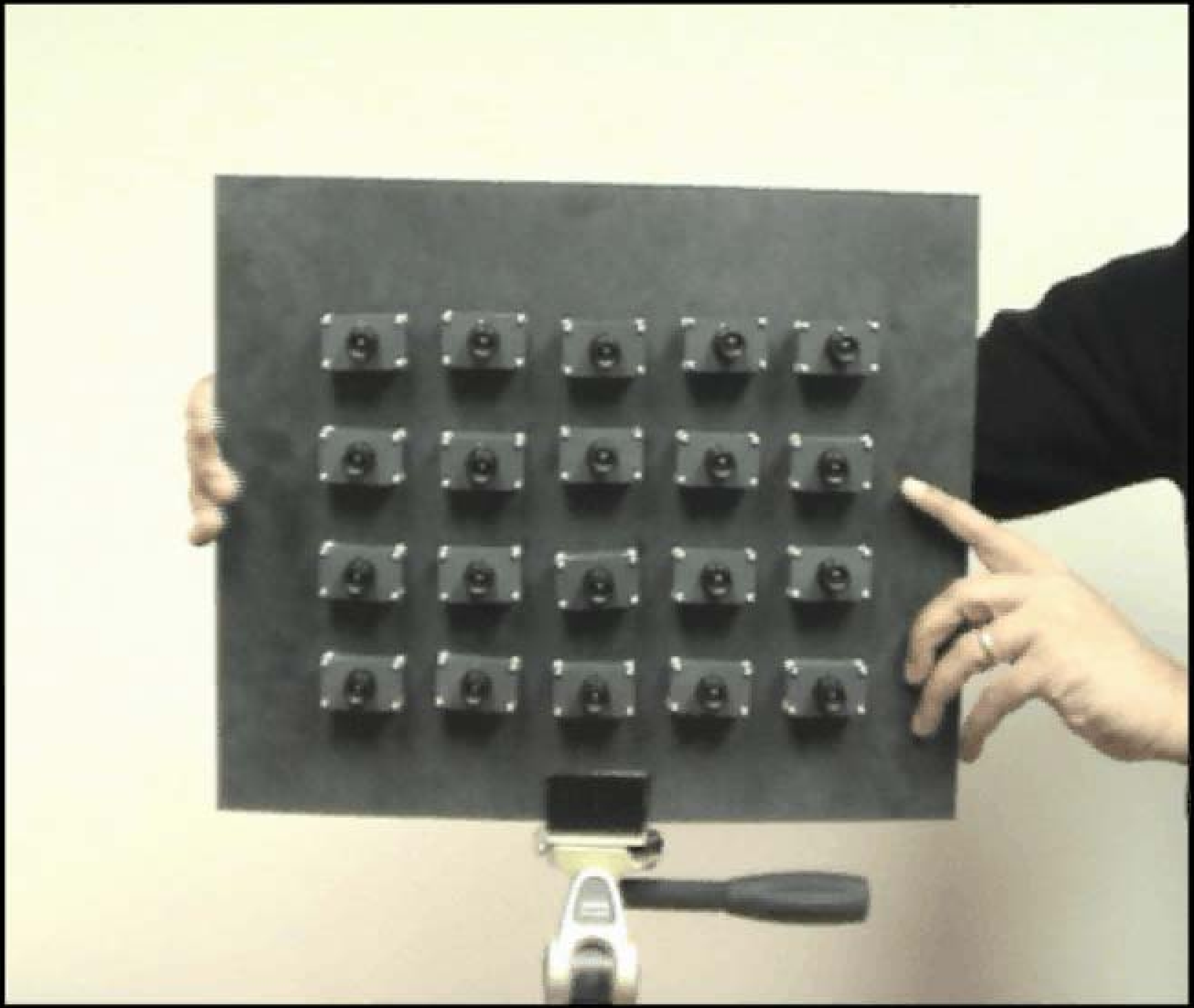
Refocused on Person



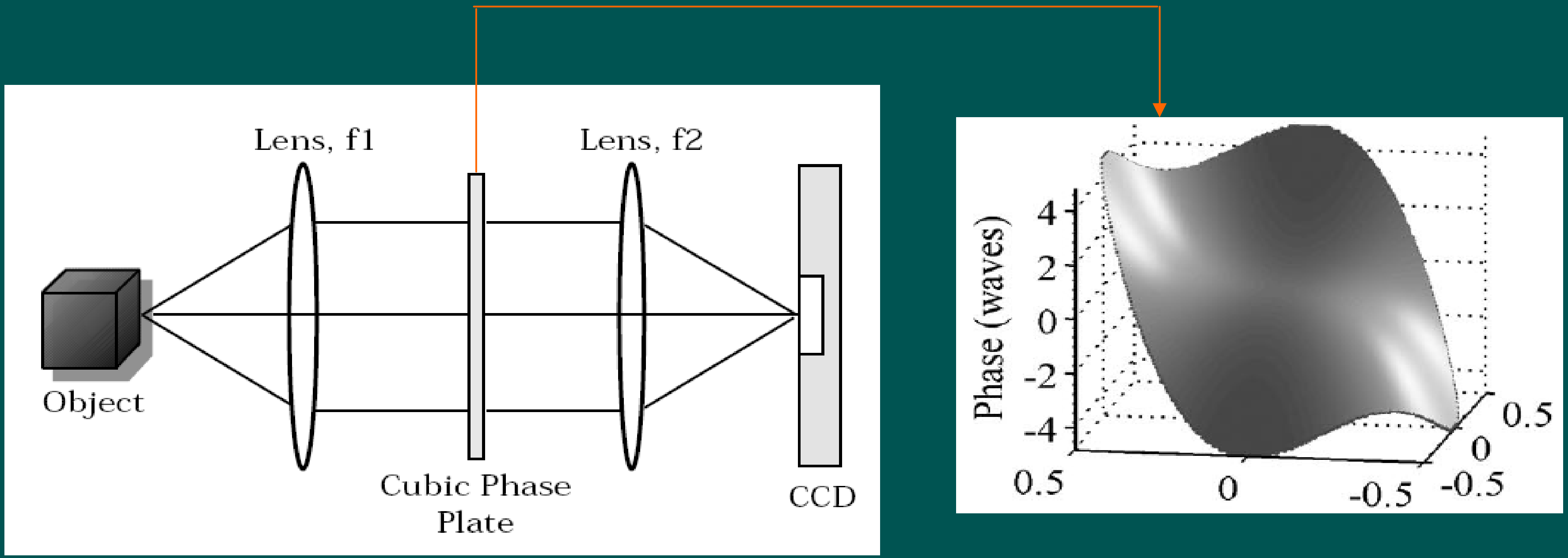
Compound Lens of Dragonfly







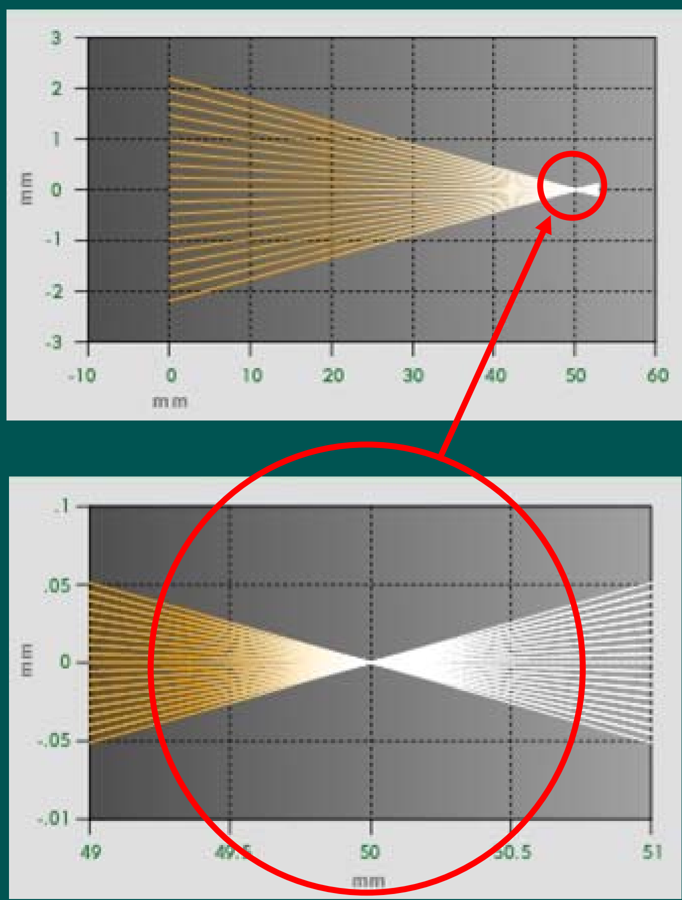
Wavefront Coding using Cubic Phase Plate



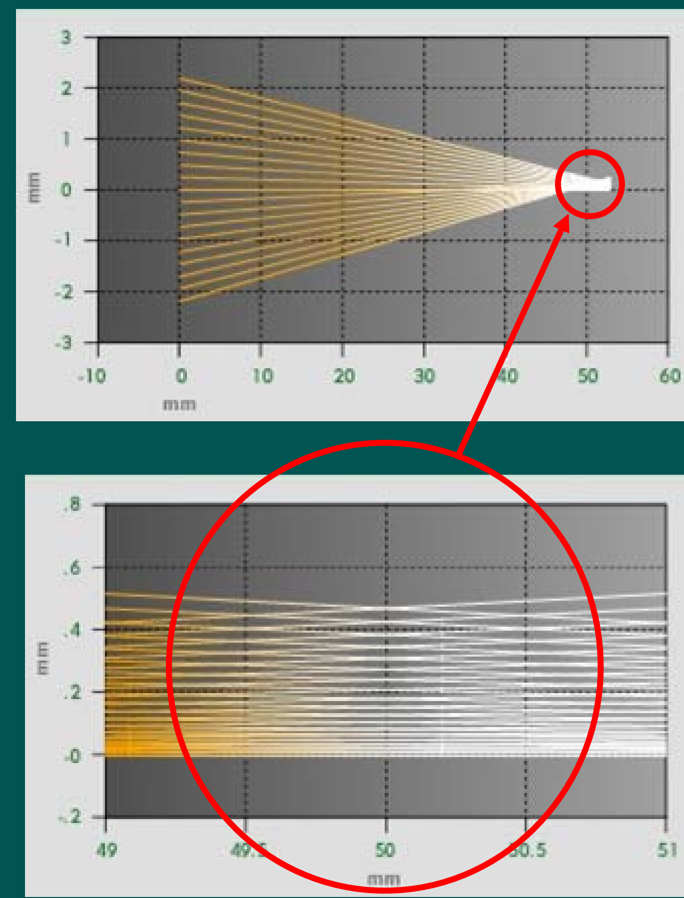
"*Wavefront Coding: jointly optimized optical and digital imaging systems*",
E. Dowski, R. H. Cormack and S. D. Sarama ,
Aerosense Conference, April 25, 2000

Depth Invariant Blur

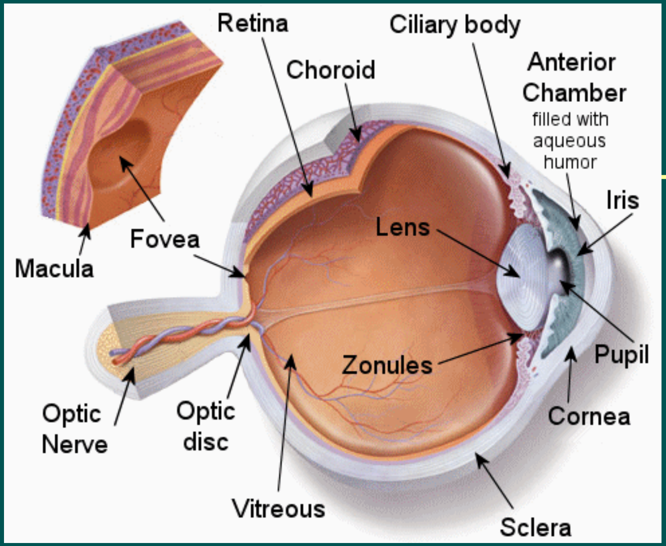
Conventional System



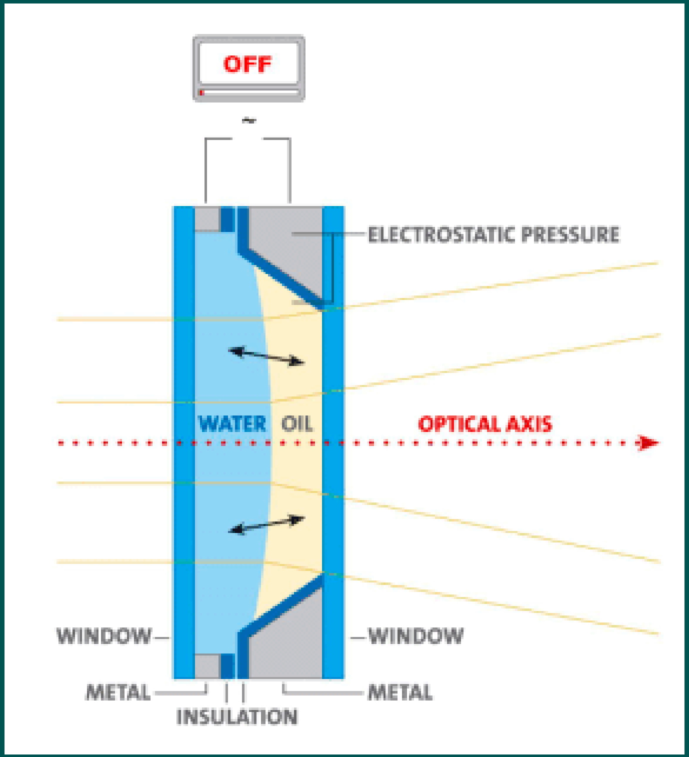
Wavefront Coded System

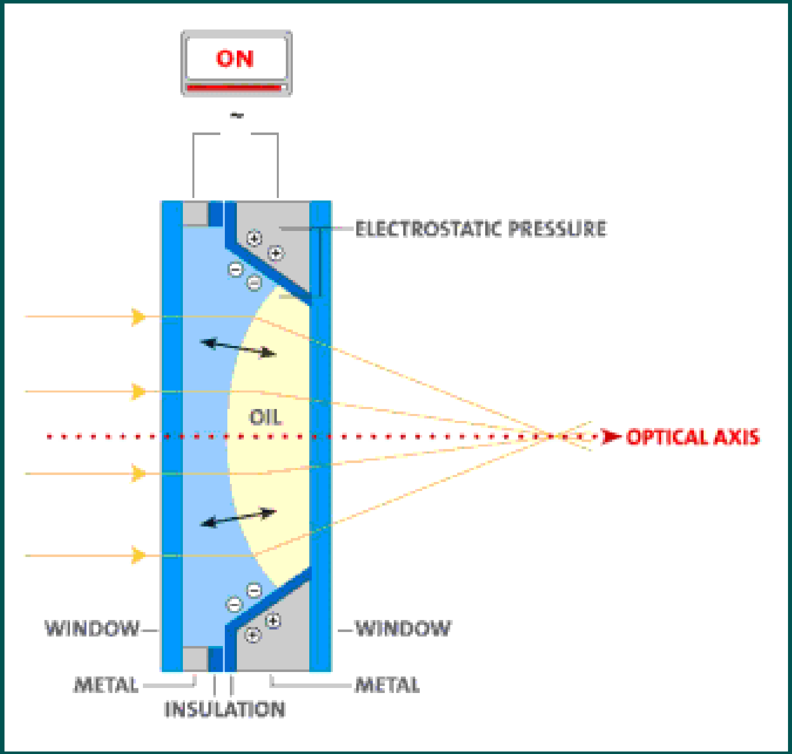


The Eye's Lens



Varioptic Liquid Lens: Electrowetting

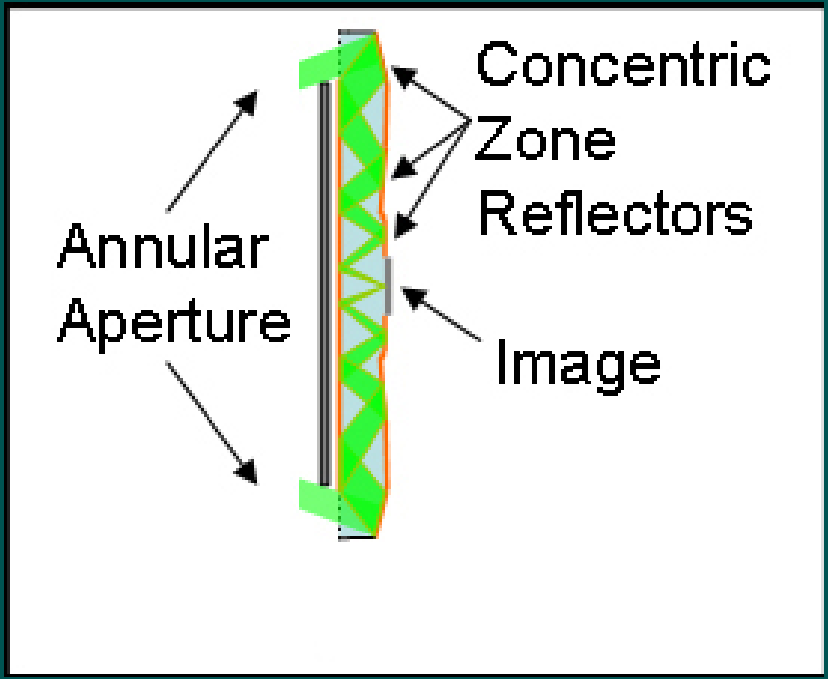


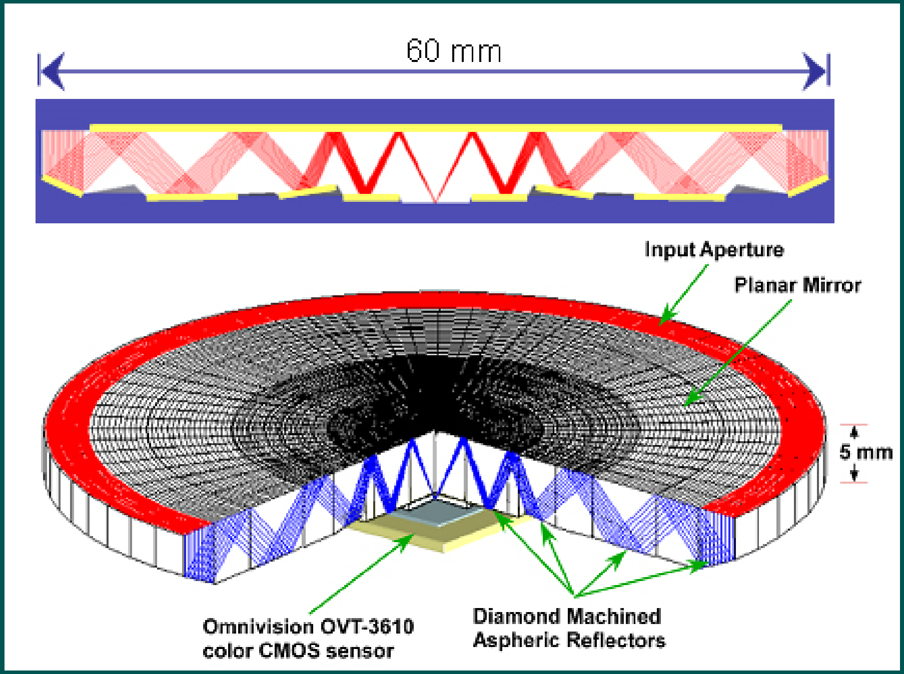


Varioptic Liquid Lens

(Courtesy Varioptic Inc.)

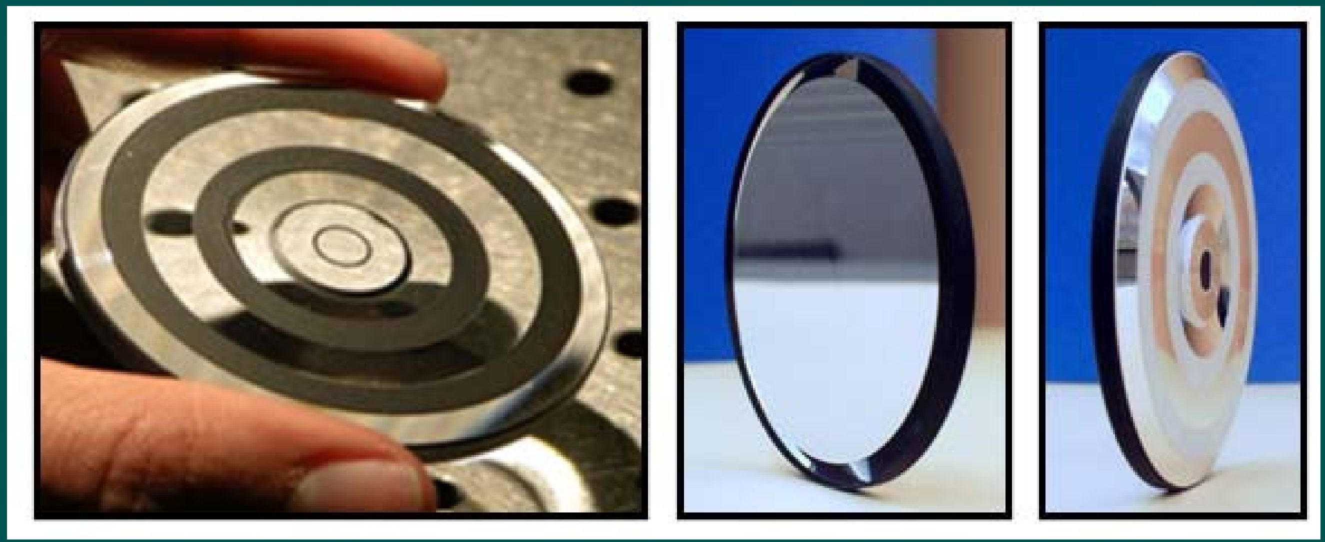
"Origami Lens": Thin Folded Optics (2007)





"Ultrathin Cameras Using Annular Folded Optics, "
E. J. **Tremblay**, R. A. Stack, R. L. Morrison, J. E. **Ford**
Applied Optics, 2007 - OSA

Origami Lens



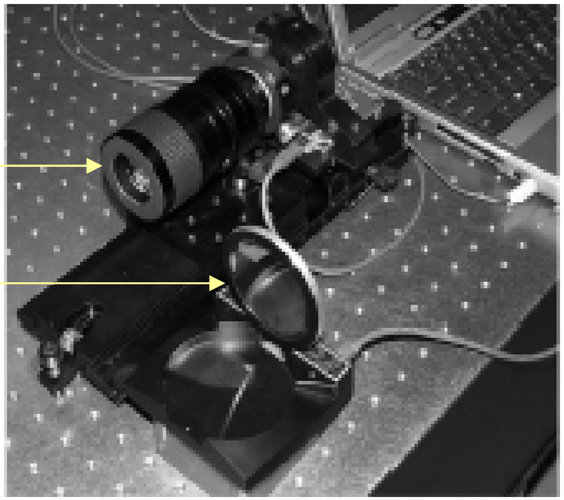
Conventional Lens

Origami Lens

Optical Performance

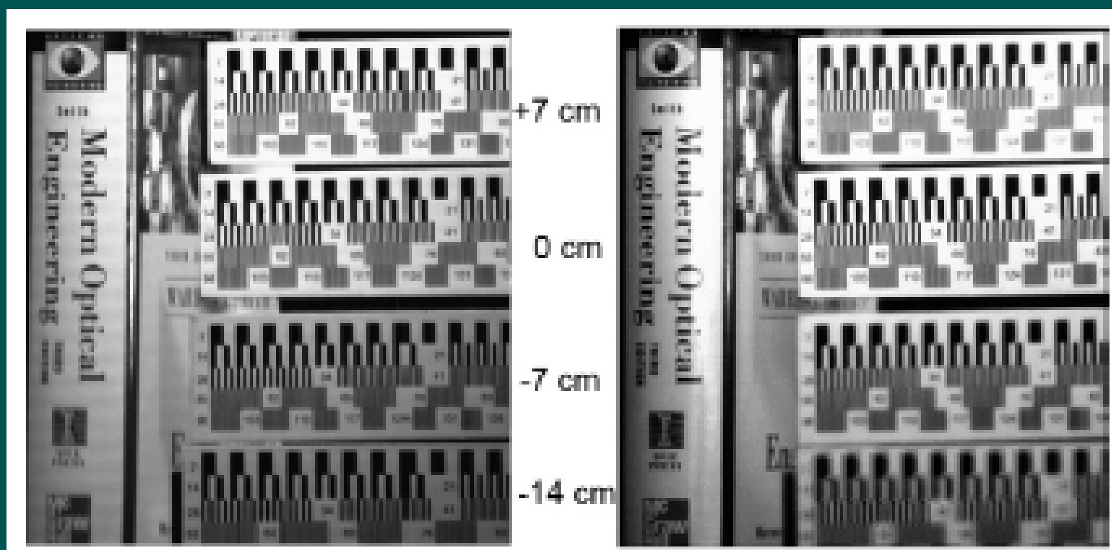
Conventional

Origami



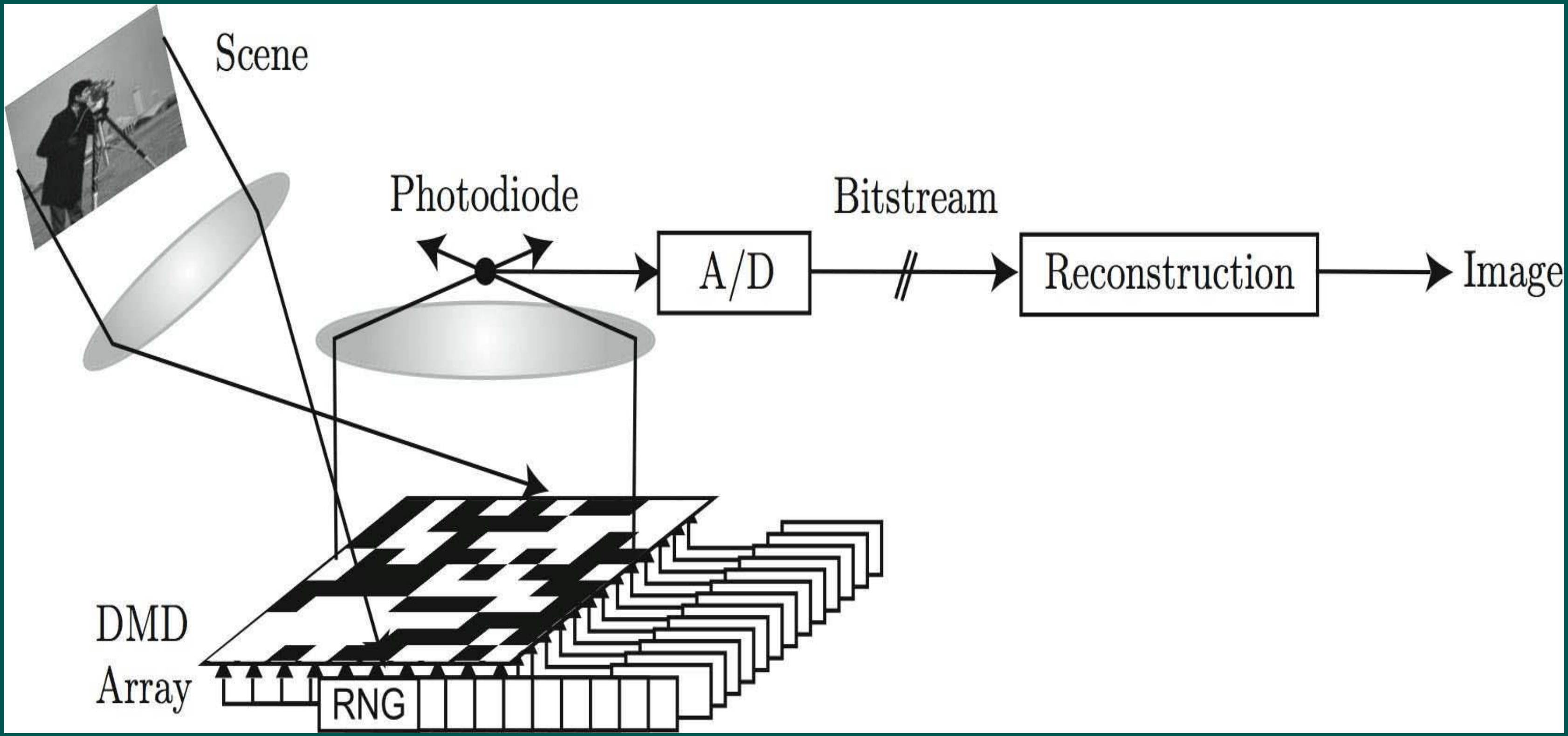
Scene

Conventional
Lens Image

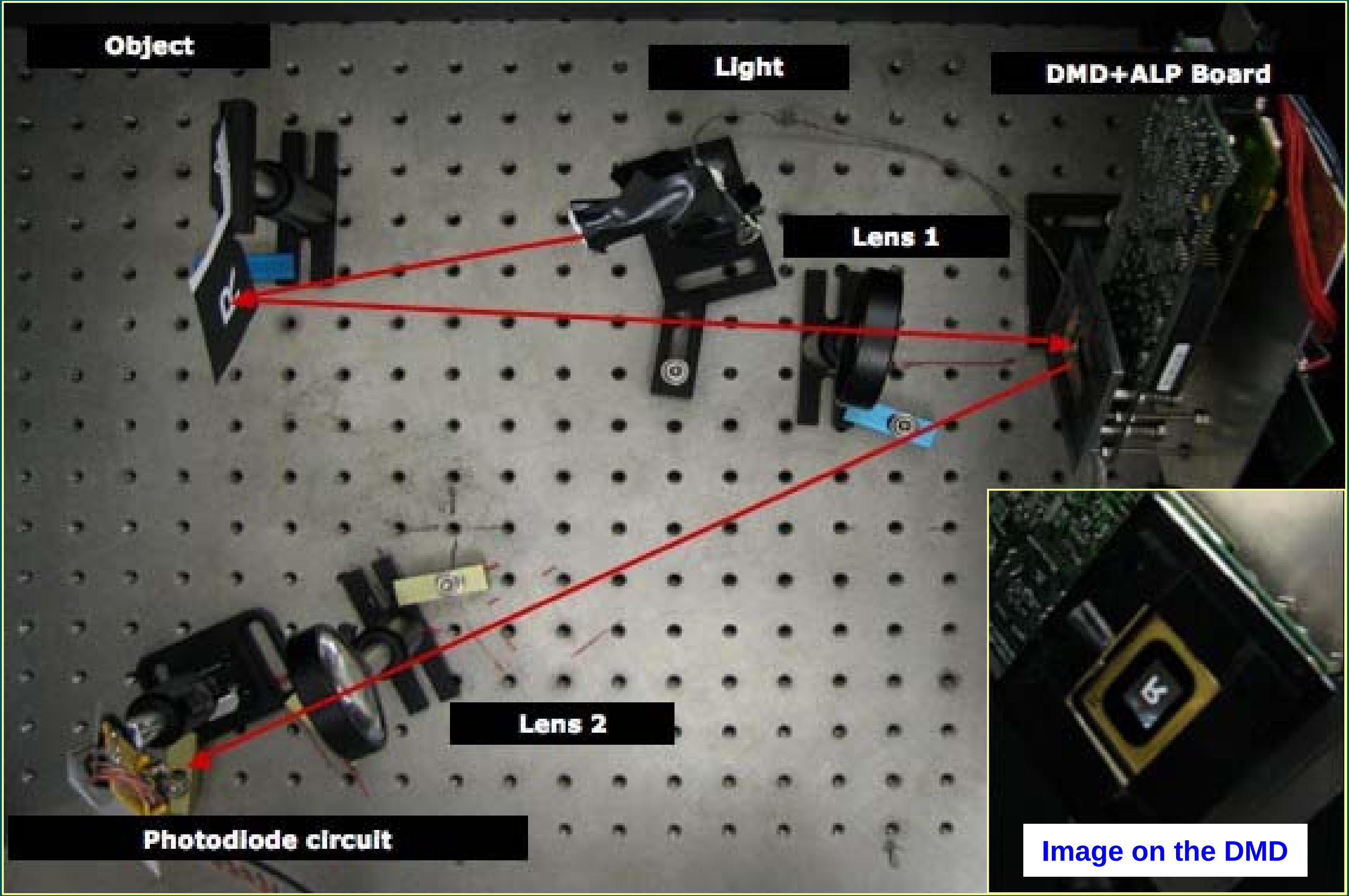


Origami
Lens Image

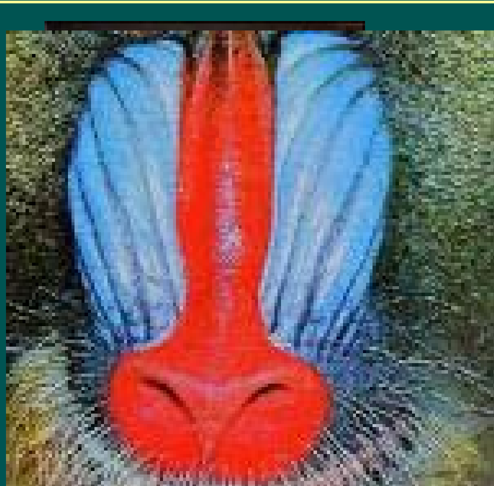
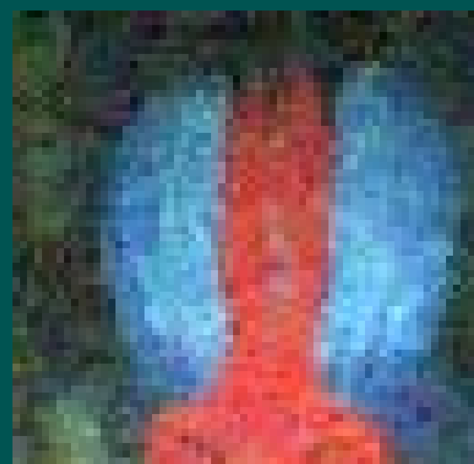
Single Pixel Camera



Single Pixel Camera



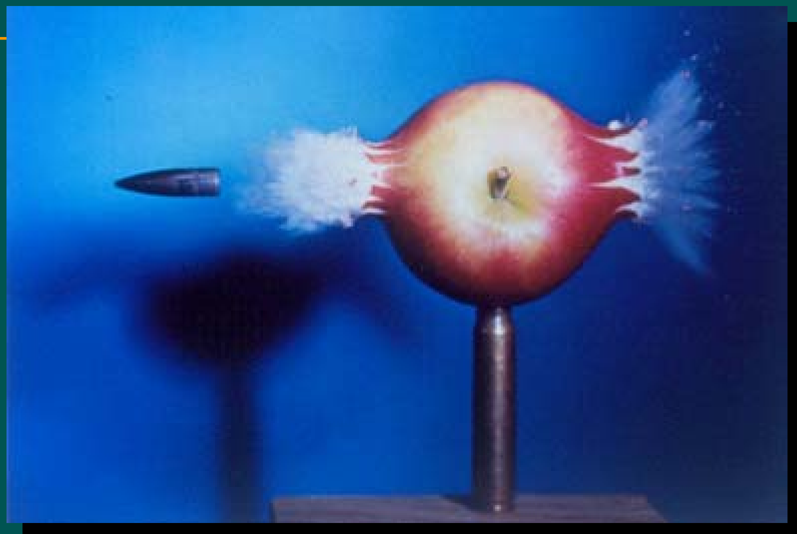
Example

Original	Compressed Imaging	
		

4096 Pixels
1600 Measurements
(40%)

65536 Pixels
6600 Measurements
(10%)

Edgerton 1930's

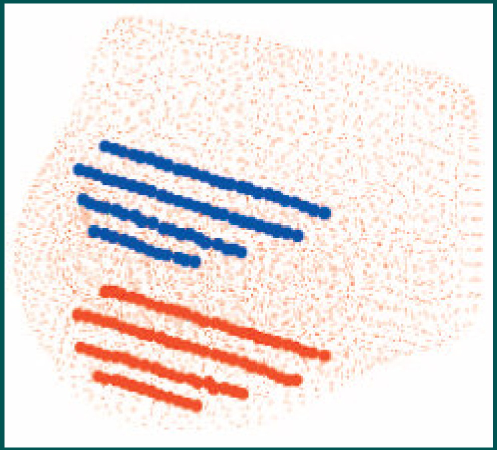


Stroboscope
(Electronic Flash)

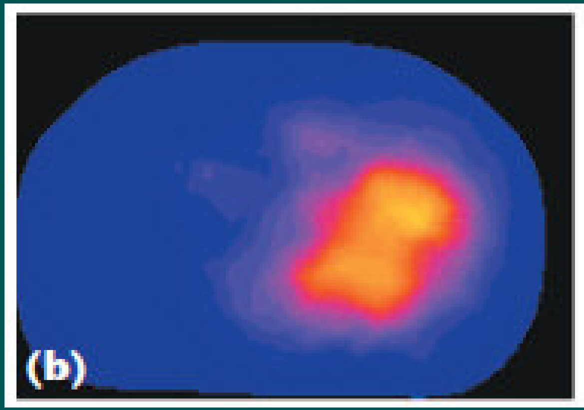
Multi-flash
Sequential Photography

Diffuse optical tomography

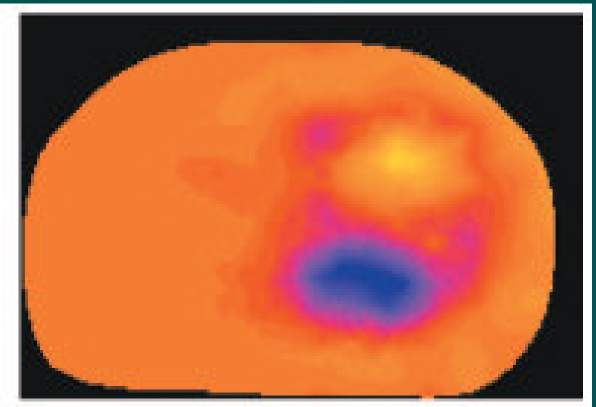
[Arridge 2003]



female breast with
sources (red) and
detectors (blue)



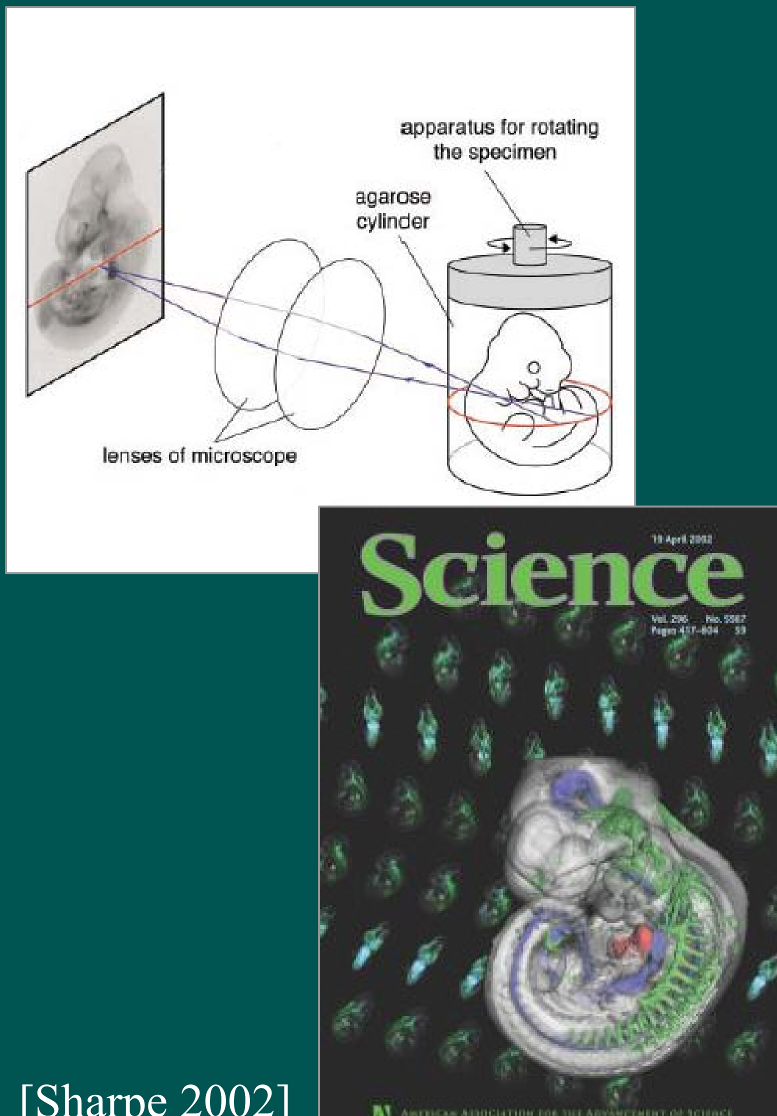
absorption
(yellow is high)



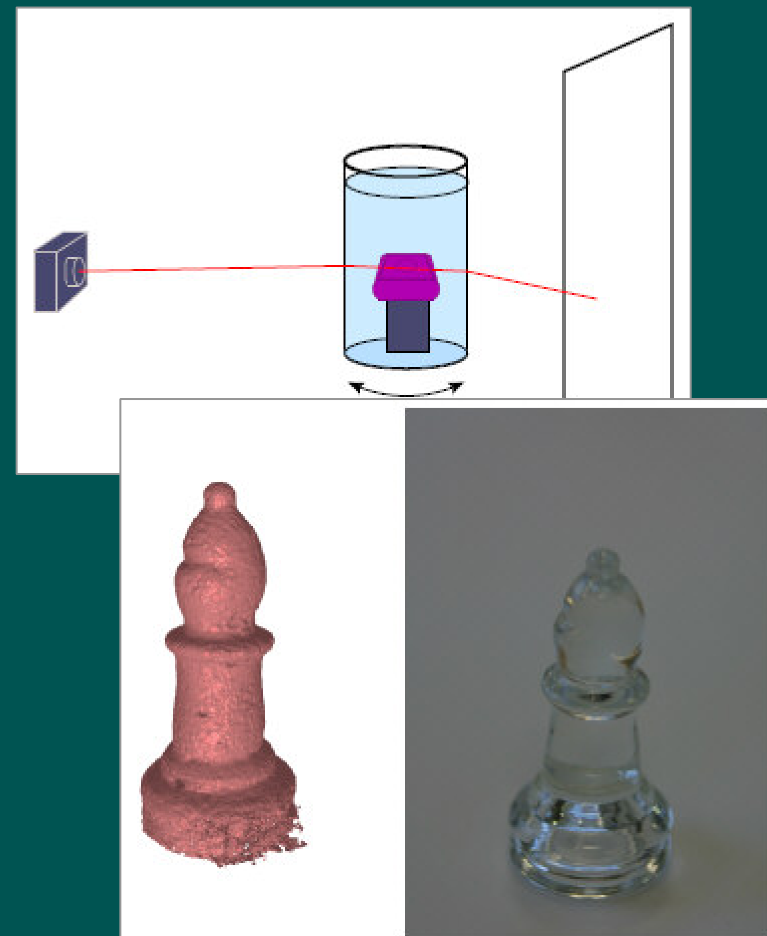
scattering
(yellow is high)

- assumes light propagation by multiple scattering
- model as diffusion process
- inversion is non-linear and ill-posed
- solve using optimization with regularization (smoothing)

Optical Projection Tomography (OPT)

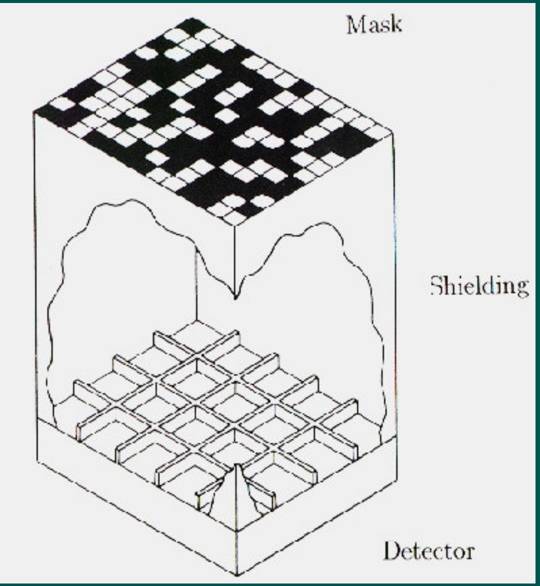


[Sharpe 2002]



[Trifonov 2006]

Coded aperture imaging

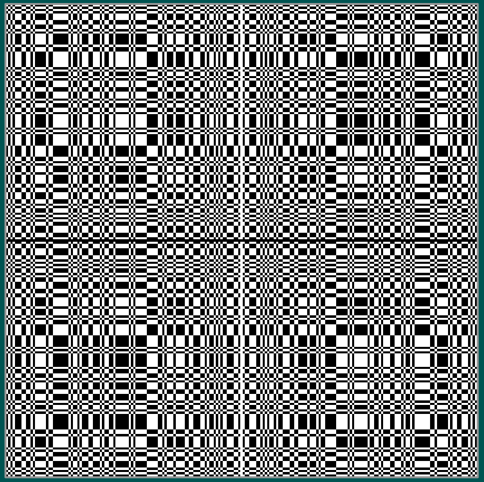


(from Zand)

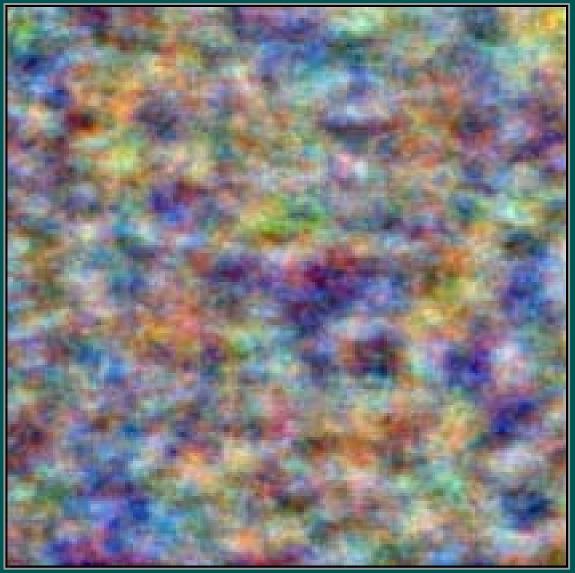
- optics cannot bend X-rays, so they cannot be focused
- pinhole imaging needs no optics, but collects too little light
- use multiple pinholes and a single sensor

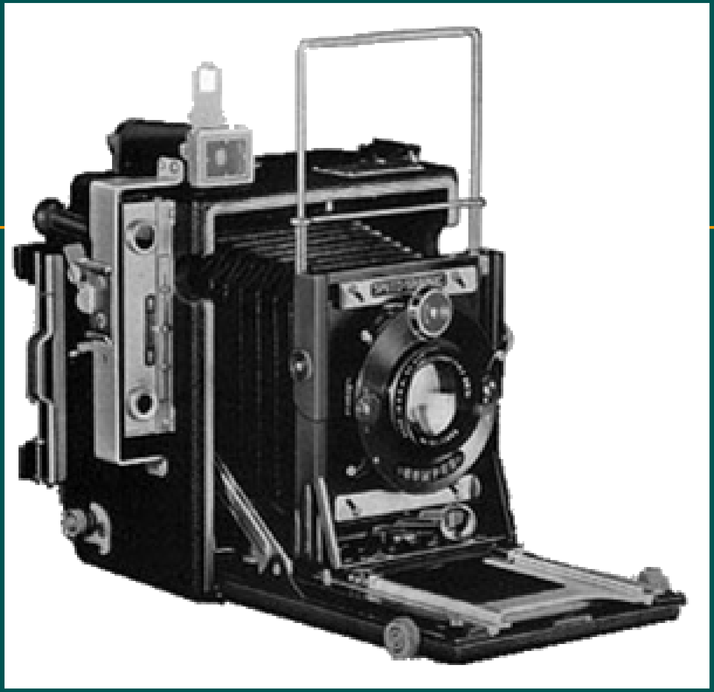
Example using 2D images
(Paul Carlisle)

*



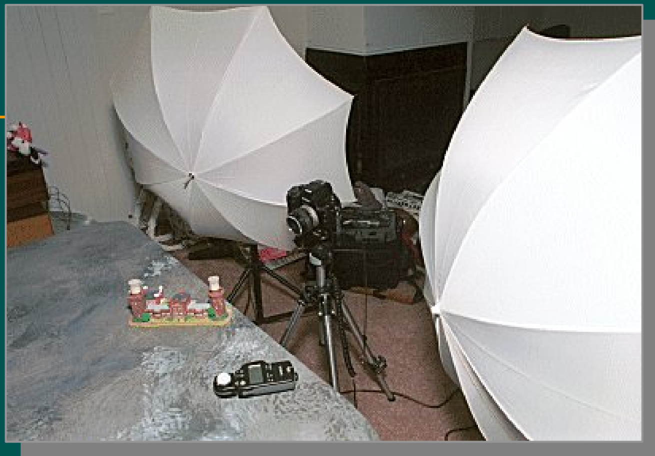
=





Computational Illumination

'Smarter' Lighting Equipment



What Parameters Can We Change ?

Image-Based Actual Re-lighting

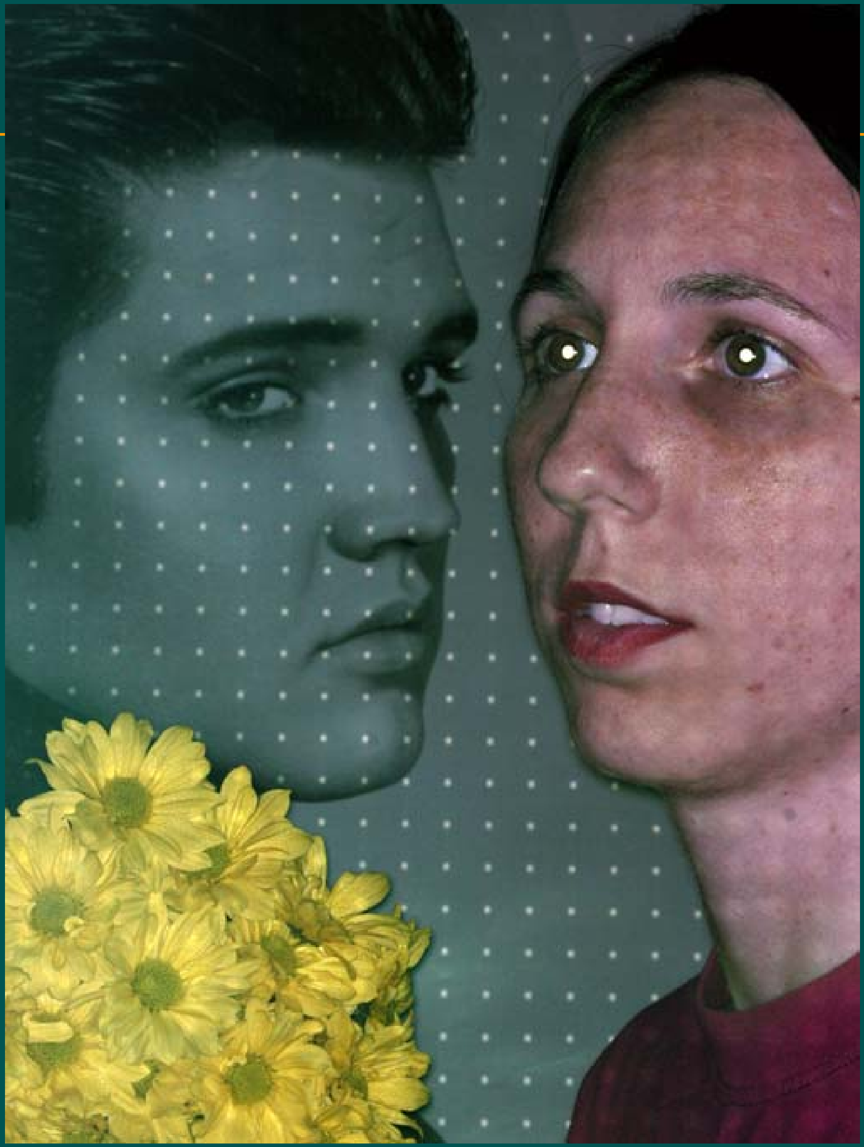
Debevec et al., SIGG2001

Light the actress in Los Angeles

Film the background in Milan,
Measure incoming light,

Matched LA and Milan lighting.

Matte the background



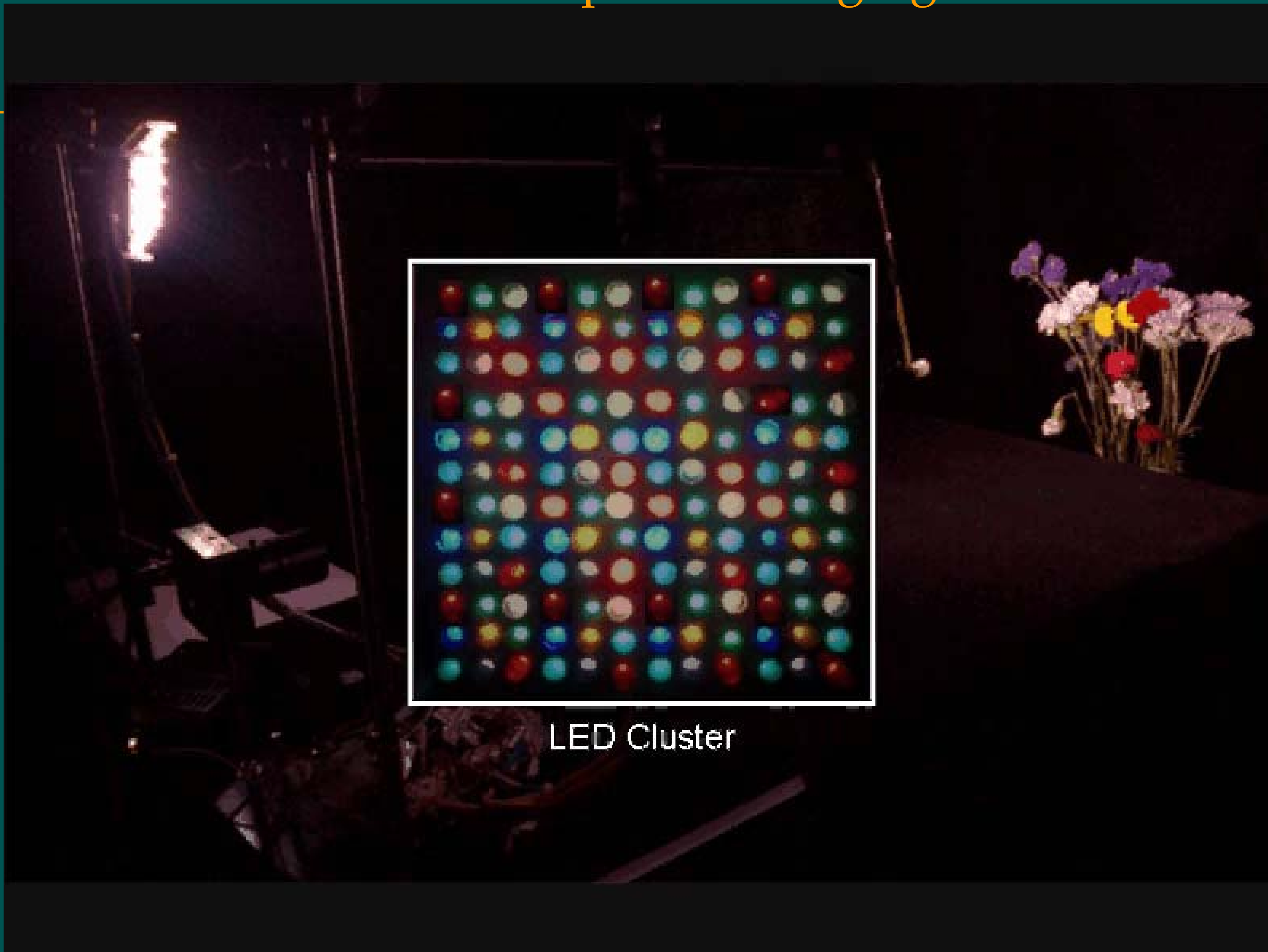
Acquired Image

Dots Removed

Depth Map Completion

(with Francesc Moreno and Peter Belhumeur 07)

Fast Multispectral Imaging



(with J. Park, M. Lee, M. Grossberg)

A.2 Concepts in Computational Photography (Tumblin, 15 minutes)

- The 'Photographic Signal'
- What is the ideal photograph?
- Ray-based versus pixel-based concepts
- Understanding dimensionality of rays outside and inside the camera

Course: Computational Photography Advanced Topics

Module 1: 90 minutes

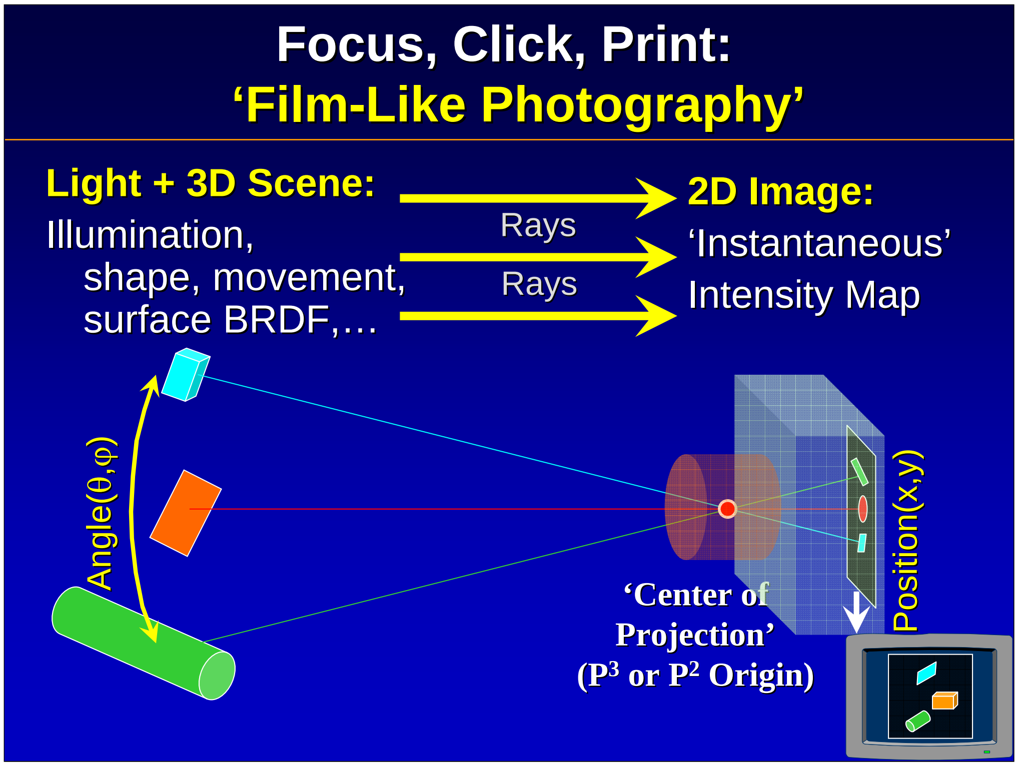
9:00:	A.1	Introduction and Overview	(Raskar, 15 minutes)
9:15:	A.2	Concepts in Computational Photography	(Tumblin, 15 minutes)
9:30:	A.3	Optics: Computable Extensions	(Raskar, 30 minutes)
10:00:	A.4	Sensor Innovations	(Tumblin, 30 minutes)
10:30:	Q & A		(5 minutes)

10:35: Break: 25 minutes

Module 2: 90 minutes

11:00:	B.1	Illumination As Computing	(Debevec, 25 minutes)
11:25:	B.2	Scene and Performance Capture	(Debevec, 20 minutes)
11:45:	B.3	Image Aggregation & Sensible Extensions	(Tumblin, 20 minutes)
12:05:	B.4	Community and Social Impact	(Raskar, 20 minutes)
12:25:	B.4	Summary and Discussion, Q&A	(All, 10 minutes)

Course Page : http://computationalphotography.org/



We still hang on to the mistaken notion that we're 'copying' the image formed by the lens to the image formed by the display, an intrinsically 2D process to approximate the appearance of a 3D scene.

We've confused 'the PROCESS of photography with its PURPOSE and GOALS.

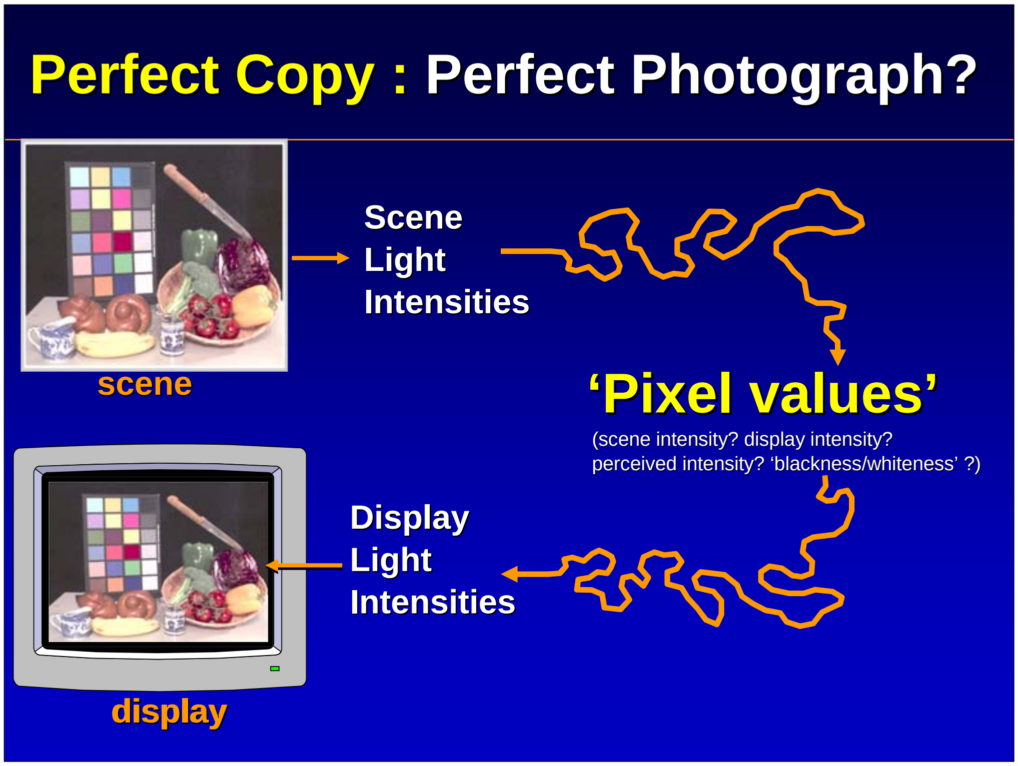
At first, it was a wonder we could do it at all:

Now it's a wonder how easily we take (bad) photos,

how many choices and adjustments we can make to our cameras to make them better, but even more importantly, how many OTHER CHOICES we have besides a lens and a box holding a sensitized plate. We have many other choices for image formation (tomography, coded image methods, structured lighting, coded aperture, etc. etc.) for lighting (projectors, movable sources, multispectral sources, tuneable lasers, flash, strobe, reflectors, Schlieren retro-reflectors), and for display (interactive devices; light-sensitive displays, HDR, etc.)

.

Yet look at how much of high-quality photography is dominated by overcoming device limitations, artful choices of lighting, and adjusting the myriad settings our cameras and digital darkrooms offer to us.



Digital Photography is almost entirely a matter of copying---just like film!

The underlying assumption is that we copy a 2D scene to a 2D display, and if we do it accurately, we're done.

‘Film-Like’ Photography

Ideals, Design Goals:

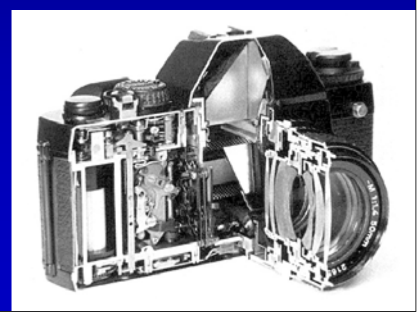
– ‘Instantaneous’ light measurement…
– Of focal plane image behind a lens.
– Reproduce those amounts of light.

Implied:

“What we see is ≅
focal-plane intensities.”

well, no…we see *much* more!

(seeing is *deeply* cognitive)



A common misconception:

Our Definitions

- **'Film-like' Photography:**

  Displayed image ≅ sensor image

- **'Computational' Photography:**

  Displayed image ≠ sensor image

  ≅ visually meaningful scene contents

A more expressive & controllable displayed result, transformed, merged, decoded data from compute-assisted sensors, lights, optics, displays

What *is* Photography?

Safe answer:

A wholly new, expressive medium (ca. 1830s)

- Manipulated display of what we think, feel, want, …
  - Capture a memory, a visual experience in tangible form
  - 'painting with light'; express the subject's visual essence
  - *"Exactitude is not the truth."* –Henri Matisse

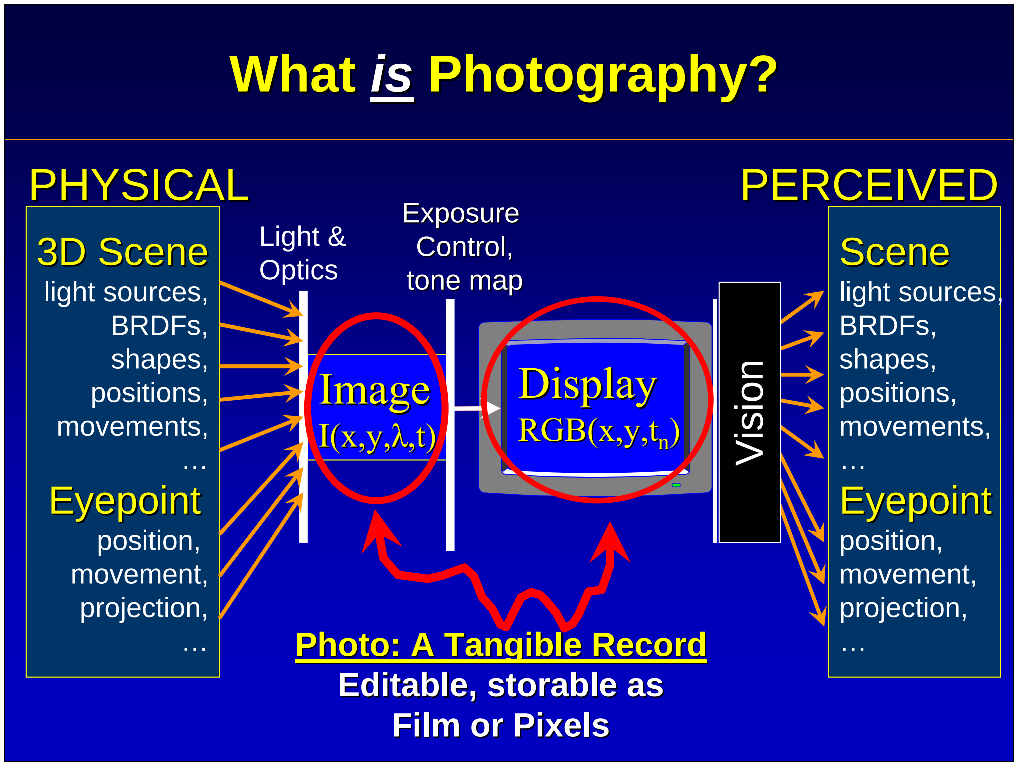
It's

What *is* Photography?

- A 'bucket' word: a neat container for messy notions (e.g. aviation, music, comprehension)
- A record of what we see,
  or would like to see,
  in tangible form.
- **Does 'film' photography always capture it?** Um, no...
- **What do we see?**

Harold 'Doc' Edgerton 1936

Um, er. This isn't

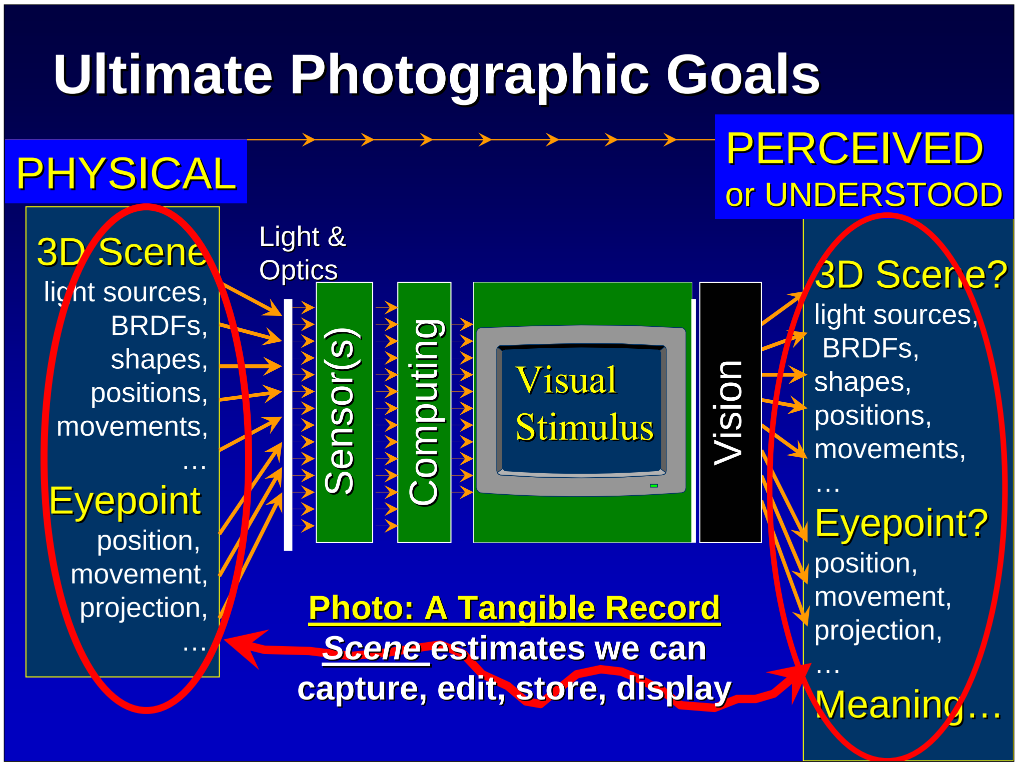


Humans see basic, partial information about boundaries, shape, occlusion, lighting, shadows and texture, with few discernible difficulties with high dynamic range, resolution, or noise, lighting, or exposure.

This basic data is usually difficult or impossible to reliably extract from pixels.

But why require extraction? Instead, we should encode this information as part of the image itself. Towards this goal, Bixels offer a straightforward way to represent intensity and gradient discontinuities within images with subpixel precision, at a fixed cost an additional 8 bits per pixel.

‘BLACKEST OF BLACK BOXES’



What we would like is something that more directly describes the visual experience,

--something that, **with some computing**, would allow a computer-equipped display to **construct** a display image,

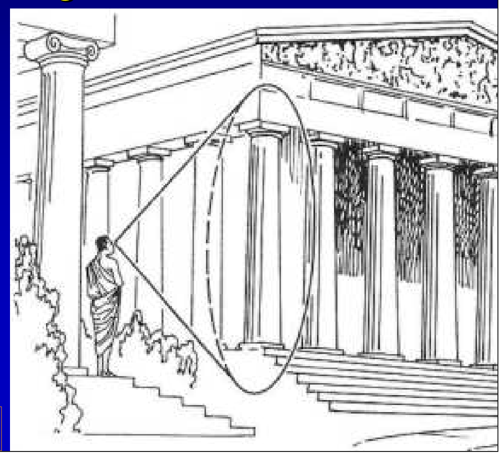
one that, based on the viewing conditions, has the best chance of evoking the desired perceptions of the original scene.

Photographic Signal: ~~Pixels~~ Rays

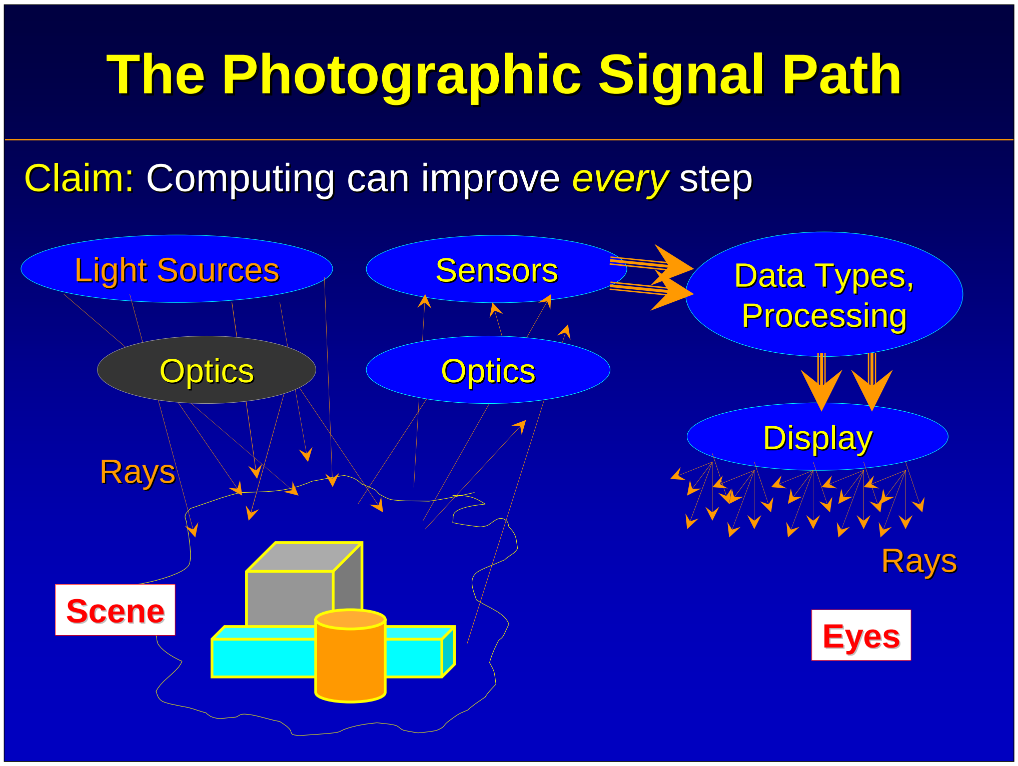
- Core ideas are ancient, simple, seem obvious:
  - **Lighting:** ray sources
  - **Optics:** ray bending/folding devices
  - **Sensor:** measure light
  - **Processing:** assess it
  - **Display:** reproduce it

- Ancient Greeks:
  'eye rays' wipe the world
  to feel its contents…

http://www.mlahanas.de/Greeks/Optics.htm



GREEKS: Photog. SEEMS obvious because what we gather can be described by ray geometry—if we think of our retina as a sensory organ, we 'WIPE' it across the scene, as if light let our retina 'reach out' and touch' what is around us. So let's look further into that:; lets consider light as a way of exploring our surroundings without contact, a magical way of transporting the the perceivable properties of our surroundings into our brain. EVEN THE GREEKS knew this idea well—they used RAYS in exploration of vision, and described how rays going through a small aperture mapped angle to position…



We tend to think of photography as capturing light, not visual impressions.

BUT VISUAL IMPRESSIONS DEPEND ON EVERY STAGE OF ‘The Photographic Signal Path’

If we **REPLACE 2D PIXELS WITH NOTIONS OF <u>MEANINGFUL CHANGES IN SETS OF RAYS</u>, then ..**

remember LIGHT IS LINEAR…

Review: How many Rays in a 3-D Scene?

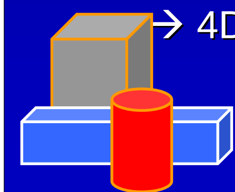
A 4-D set of infinitesimal members.

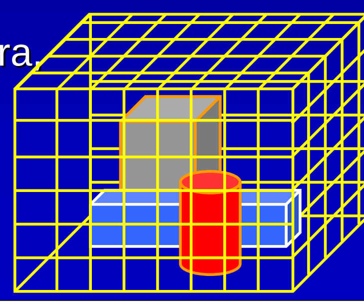
(Levoy et al. SIGG'96) (Gortler et al. '96)

Imagine:

- Convex Enclosure of a 3D scene
- Inward-facing ray camera at every surface point
- Pick the rays you need for ANY camera outside.

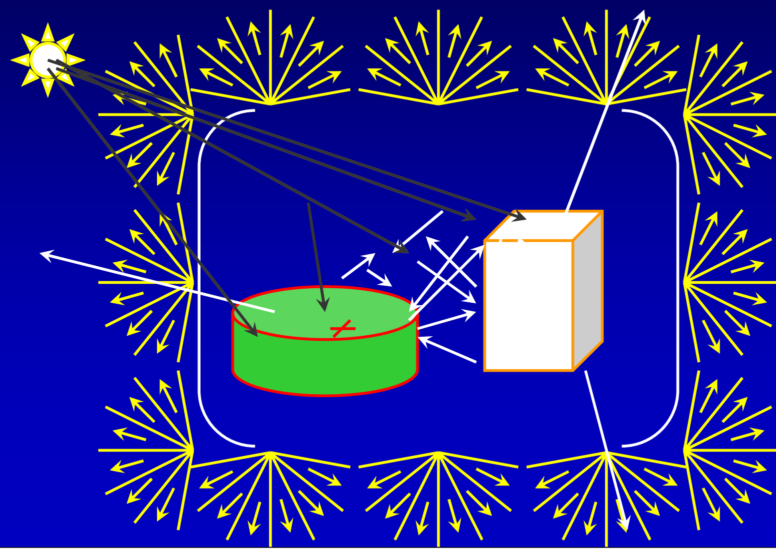
2D surface of cameras,
+ 2D ray set for each camera,
→ 4D set of rays.





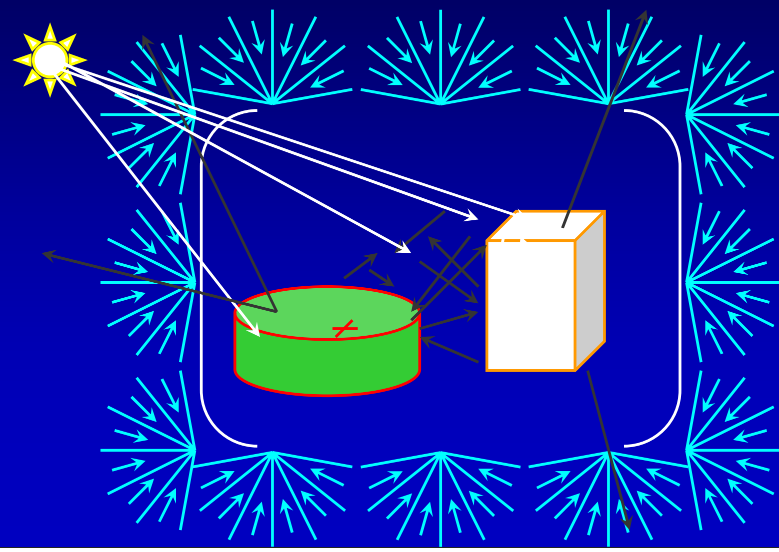
4-D Light Field / Lumigraph

Measure all the ***outgoing*** light rays.



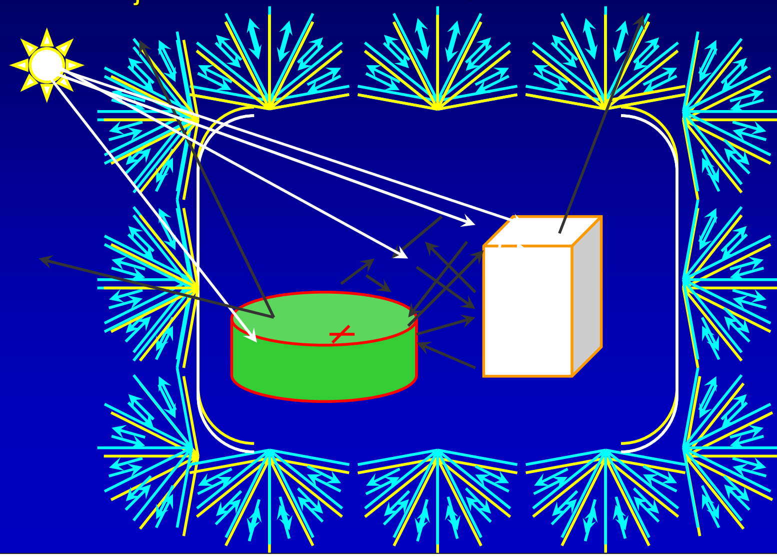
4-D Illumination Field

Same Idea: Measure all the ***incoming*** light rays



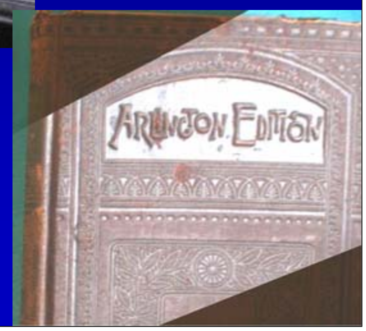
4D x 4D = 8-D Reflectance Field

Ratio: R_{ij} = (outgoing ray$_i$) / (incoming ray$_j$)



Because Ray *Changes* Convey Appearance

- These rays + all these rays give me…



- MANY more useful details I can examine…

Missing:
Expressive Time Manipulations

What other ways better reveal appearance to human viewers?

(Without direct shape measurement?)

Can you understand this shape better?

Time for space wiggle. Gasparini, 1998.

Occlusion often hides visually important features that help us understand what we see.

Missing: Interaction…

Adjust everything: lighting, pose, viewpoint, focus, FOV,…

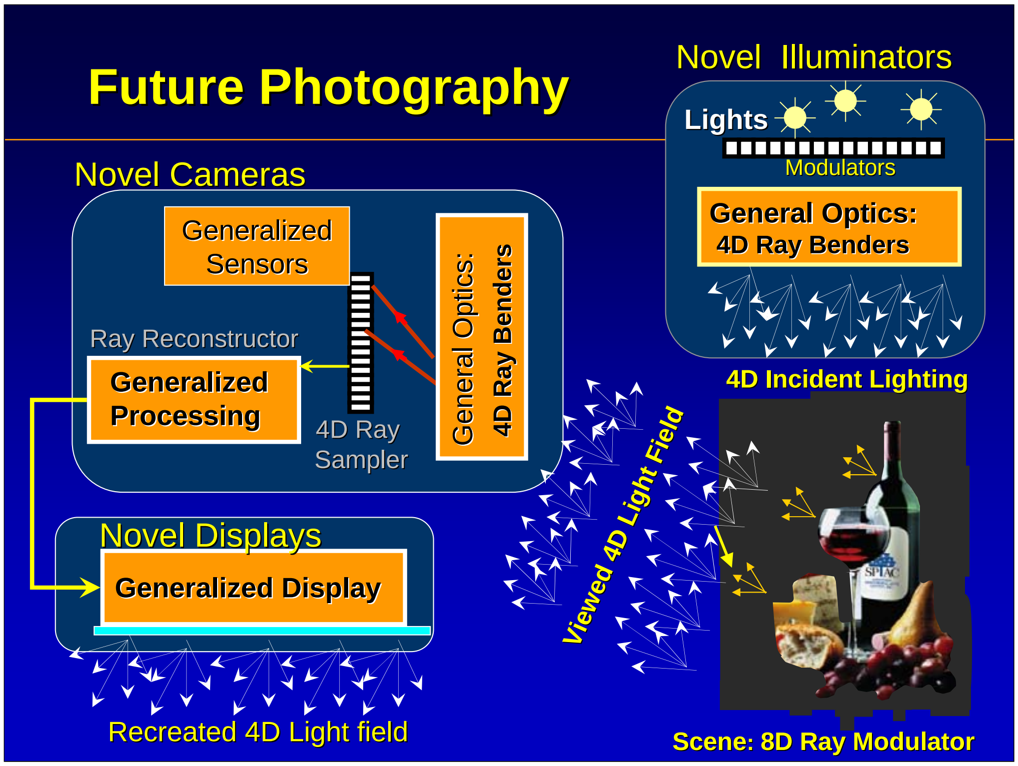
Winnemoller EG 2005: after Malzbender, SIGG2001

Mild Viewing & Lighting Changes;
(is true 3D shape necessary?)

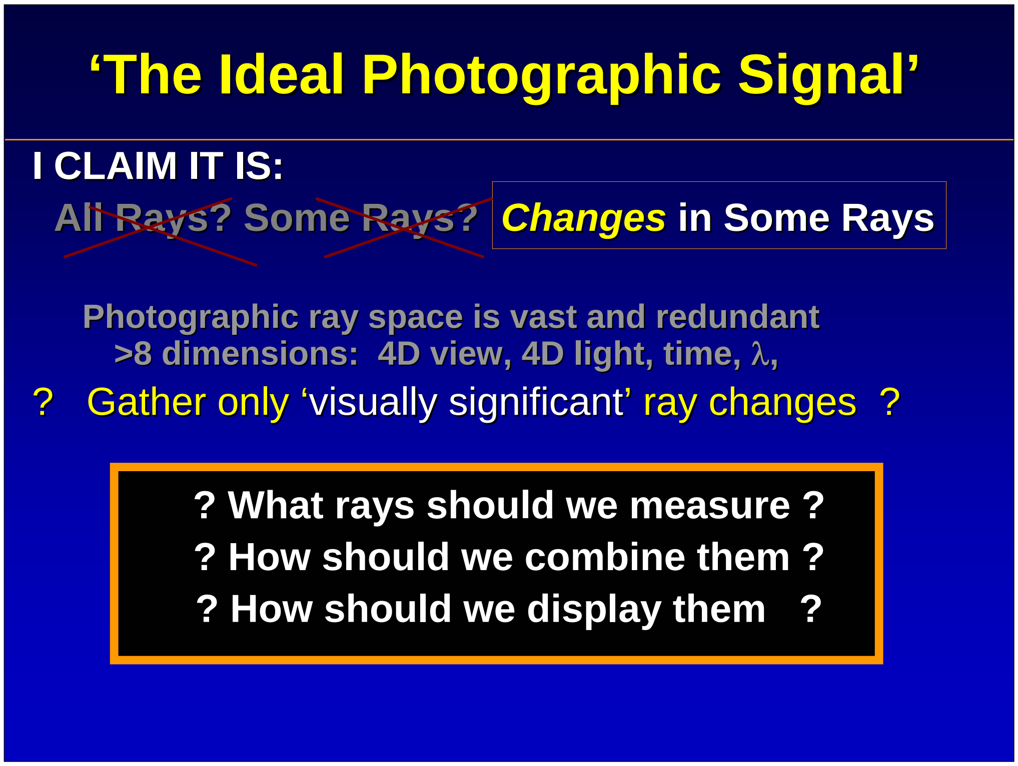
Convicing visual appearance:
Is Accurate Depth really necessary?

a few good 2-D images may be enough…

"Image jets, Level Sets, and Silhouettes"
Lance Williams,
talk at Stanford, 1998.



THERE ARE AT **LEAST 4 blocks** that we can **generalize and improve:**
lighting, optics, sensors, processing, (display: light sensitive display)



Rays are an infinitesimal discrete, computed abstraction—they match what we perceive (an infinitely sharp world of disjoint objects), and they also escape a great deal inconvenient physics that entangles photography in practical difficulties– They ignore rarely-perceived effects (diffraction, noise, fluorescence) that are computationally MUCH more difficult.

> ASIDE: Rays largely abandoned in modern optics & lens design—replaced by `Fourier Optics' methods that properly account for diffraction effects, coherent (laser) light and nearly all wave propagation effects (see the classic textbook by Goodman, 1968). WHY USE Rays? They are ENOUGH…

Up until the time of machine-assisted image making, **none** of these efx of physics were a problem—human perception guided image making instead.

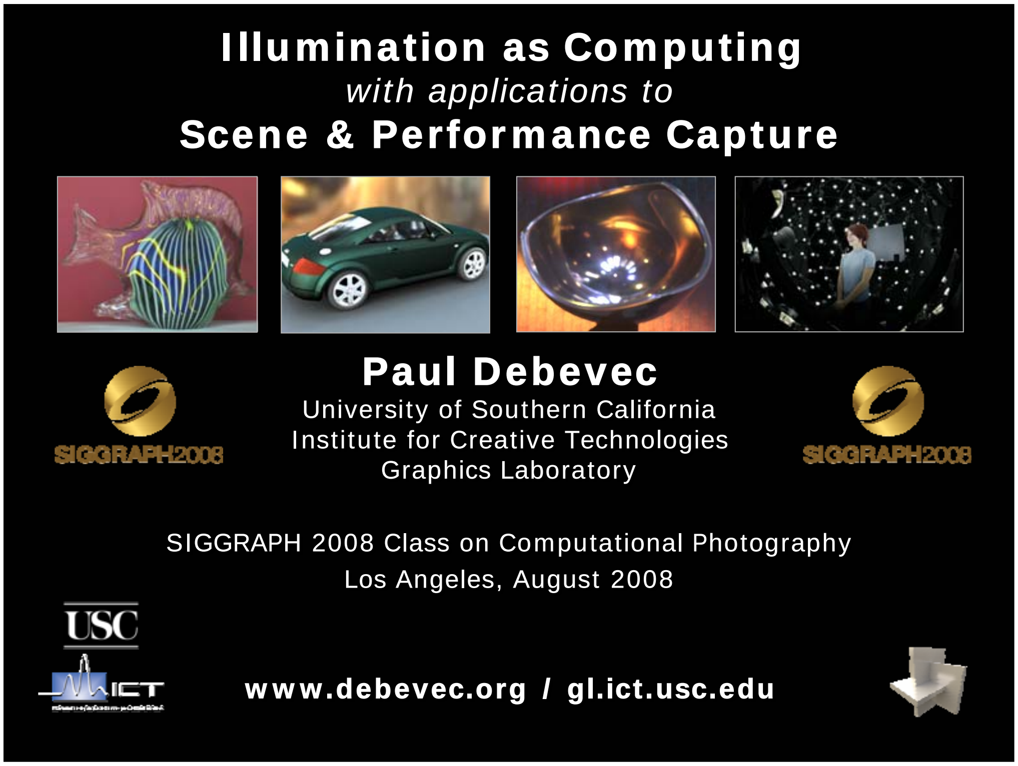
Beyond 'Film-Like' Photography

Call it **'Computational Photography':**

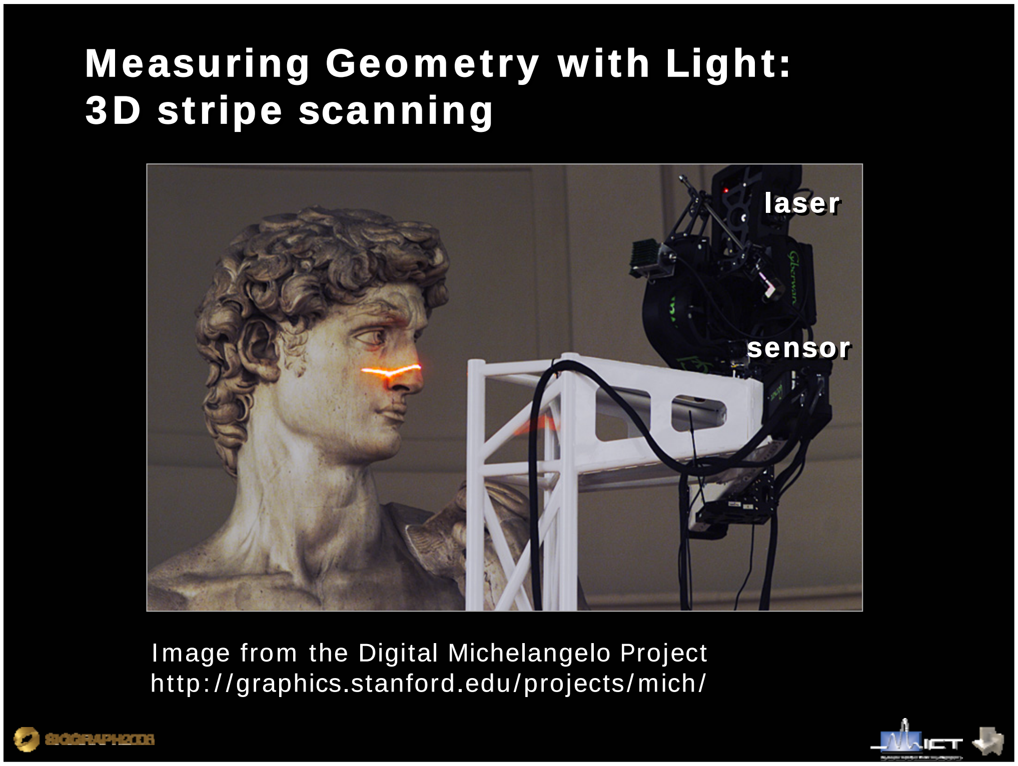
To make 'meaningful ray changes' tangible,

- **Optics** can do more…
- **Sensors** can do more…
- **Light Sources** can do more…
- **Processing** can do more…

by applying low-cost storage,
computation, and control.

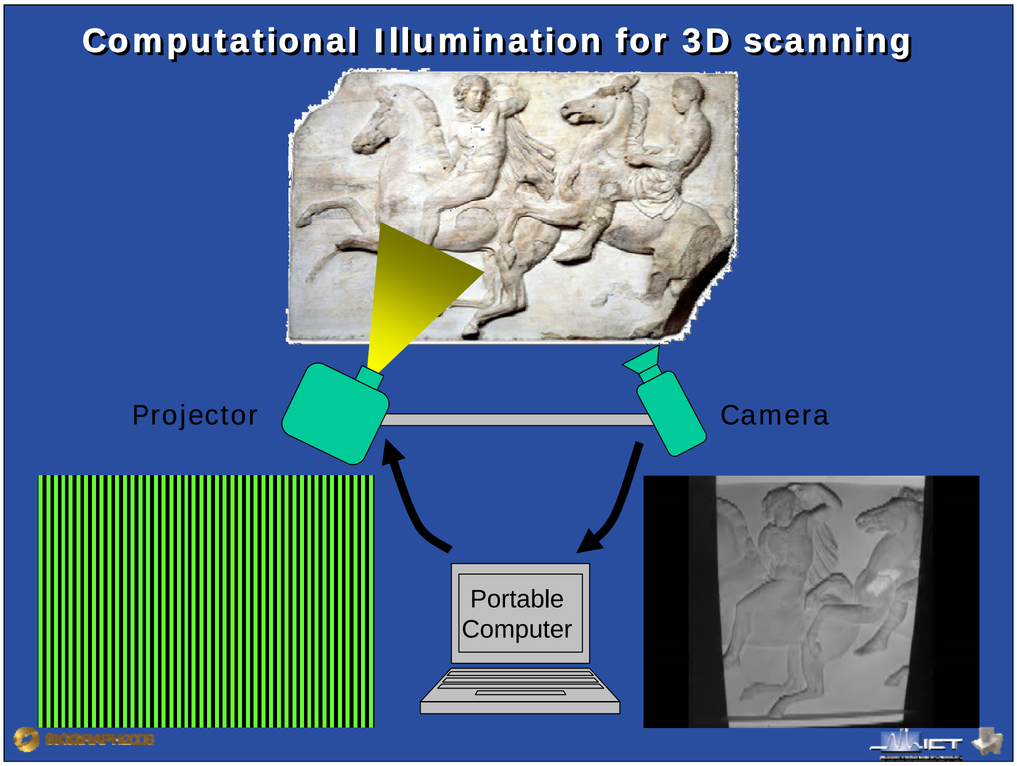


In this presentation I'll be speaking about some techniques that use Computational Photography to measure aspects of the lighting and reflectance of real scenes. There's been a lot of recent work in this area, and I'll only have a chance to give an overview of some of the projects, but hopefully what I have to say will give a reasonably clear path through a significant variety of material which will serve as a good primer to explore this area further.



The most traditional 3D scanners use a laser stripe which scans over the object. That's why we traditionally 3D scene capture as scanning, even if nothing actually scans across the scene. The laser hits the object and is imaged back onto a sensor, forming a triangle. The optics of the sensor are calibrated so that triangulation allows an entire line of scene points to be constructed in 3D. The sensors (such as this one custom-made by Cyberware) are usually designed so that the laser peak is detected for each pixel column in hardware, so that the images do not need to be processed for each laser stripe position.

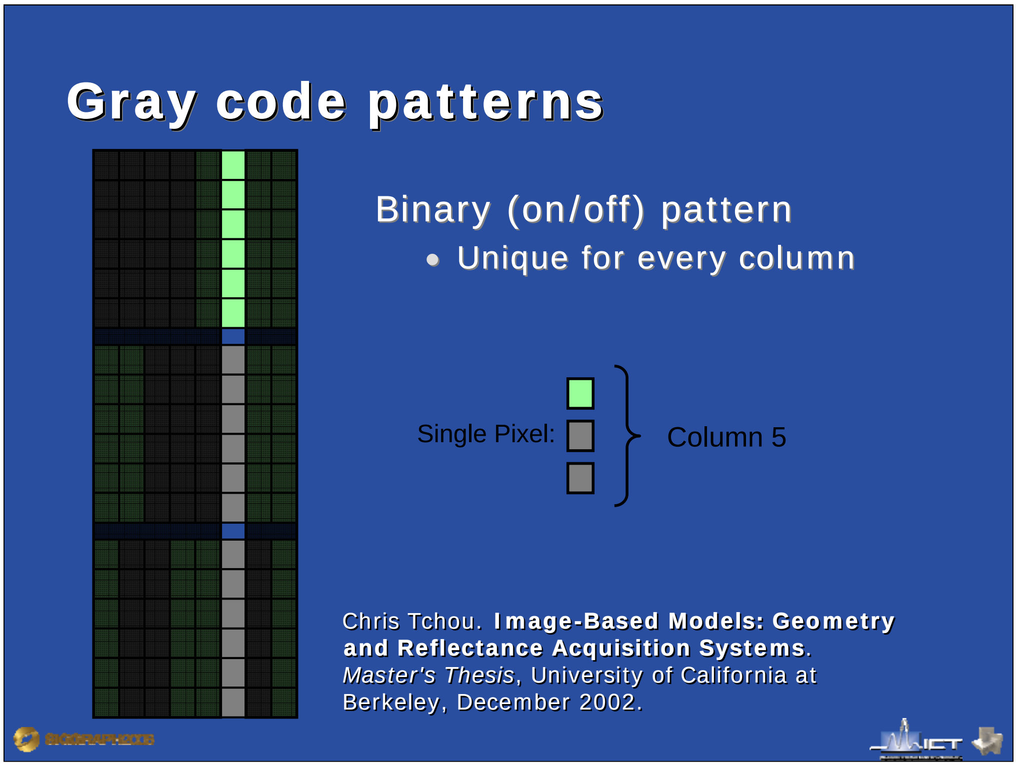
Without such peak detection in hardware, this isn't very practical since you have to take a whole image every time the laser moves. What if you would prefer to build your own scanner with just a video projector and a video camera?

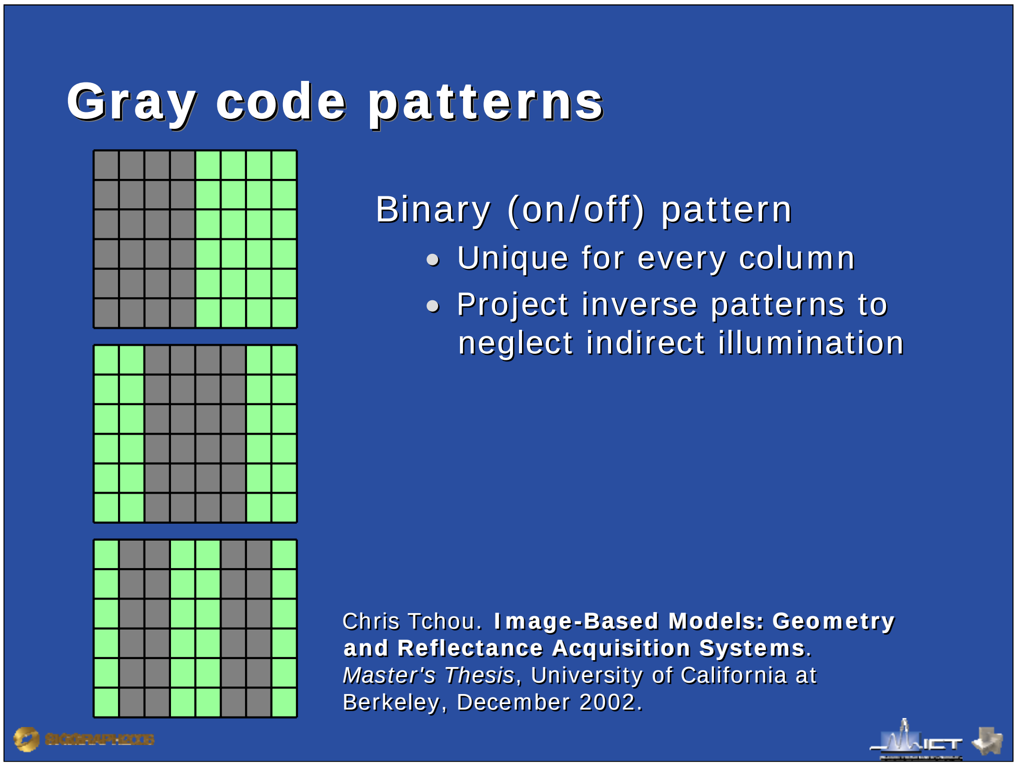


It turns out this isn't that difficult, and you don't even need to take all that many pictures!

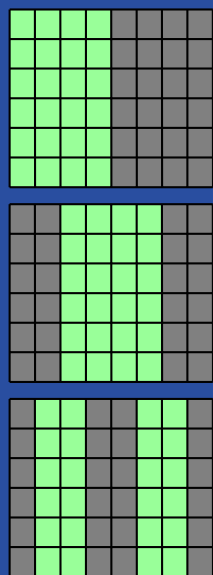
Many "computational illumination" techniques make use of video projectors to emit various types of coded illumination. A classic application of coded illumination is for 3D scanning using structured light patterns.

Now, as we all know scenes don't just consist of geometry, they also consist of reflectance properties and illumination.


Gray code patterns
Binary (on/off) pattern
• Unique for every column
Single Pixel:
Column 5
Chris Tchou. Image-Based Models: Geometry and Reflectance Acquisition Systems. Master's Thesis, University of California at Berkeley, December 2002.


Gray code patterns
Binary (on/off) pattern
Unique for every column
Project inverse patterns to neglect indirect illumination
Chris Tchou. Image-Based Models: Geometry and Reflectance Acquisition Systems. Master's Thesis, University of California at Berkeley, December 2002.

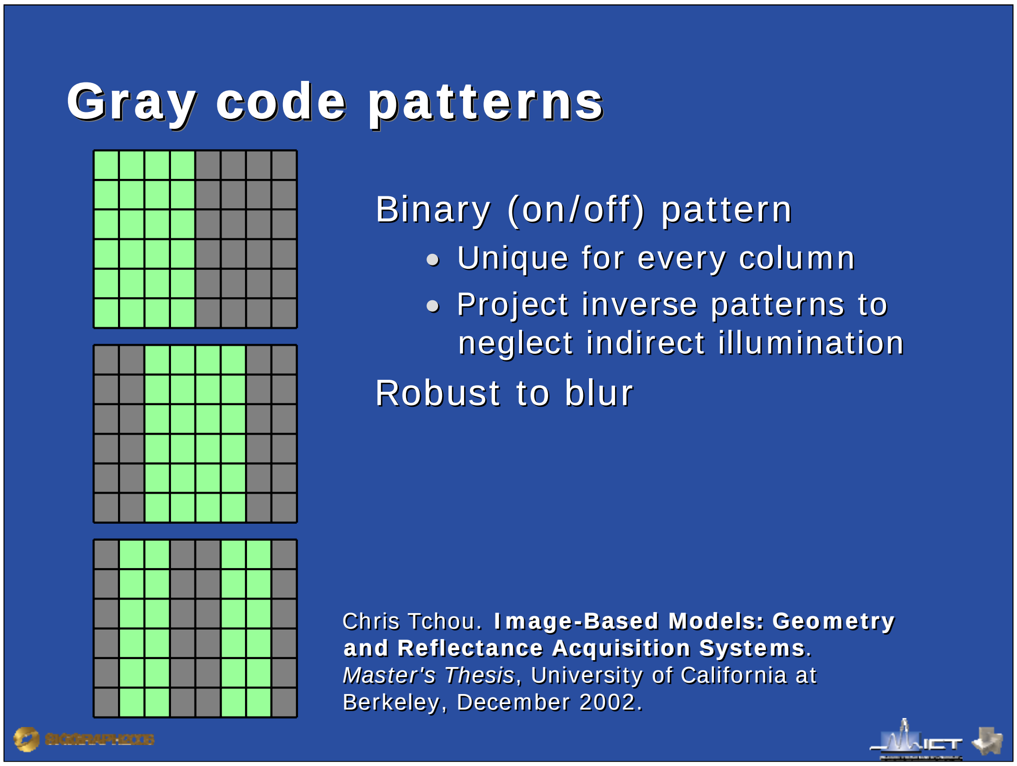
Gray code patterns

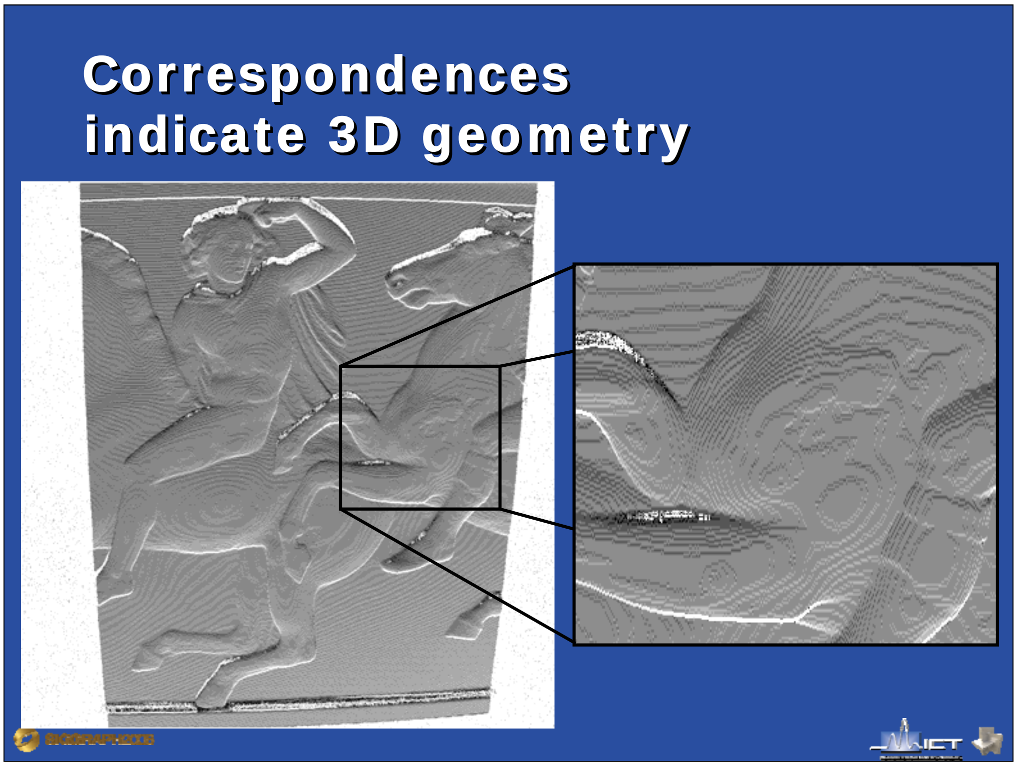


Binary (on/off) pattern

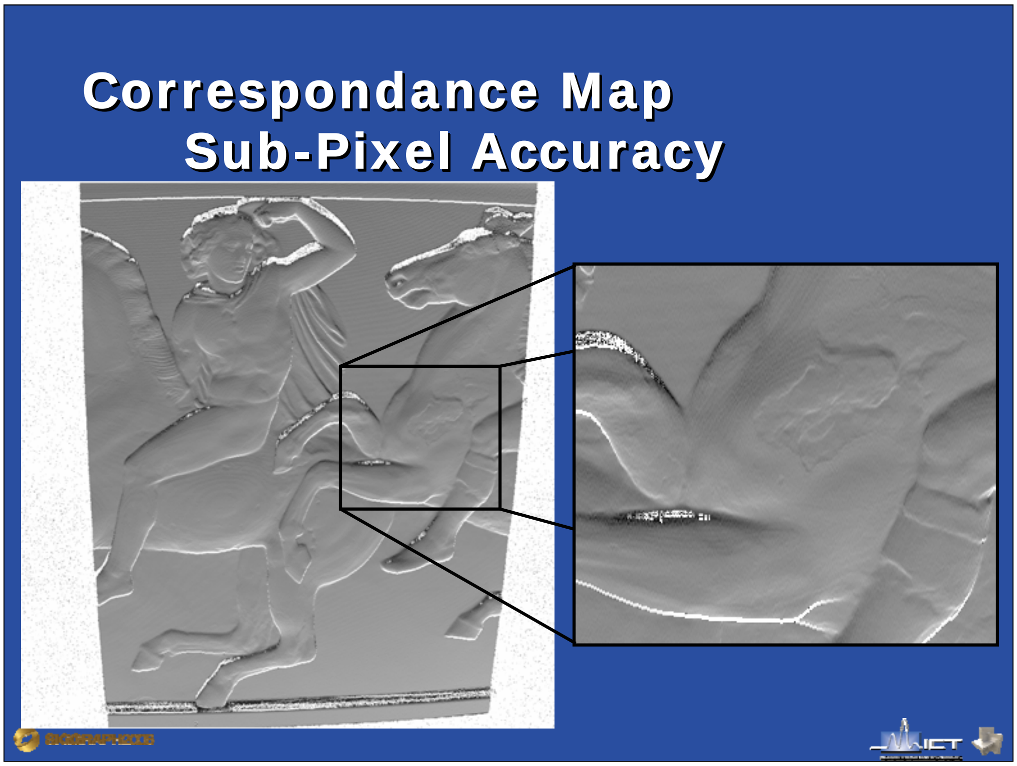
- Unique for every column
- Project inverse patterns to neglect indirect illumination

Chris Tchou. **Image-Based Models: Geometry and Reflectance Acquisition Systems**. *Master's Thesis*, University of California at Berkeley, December 2002.

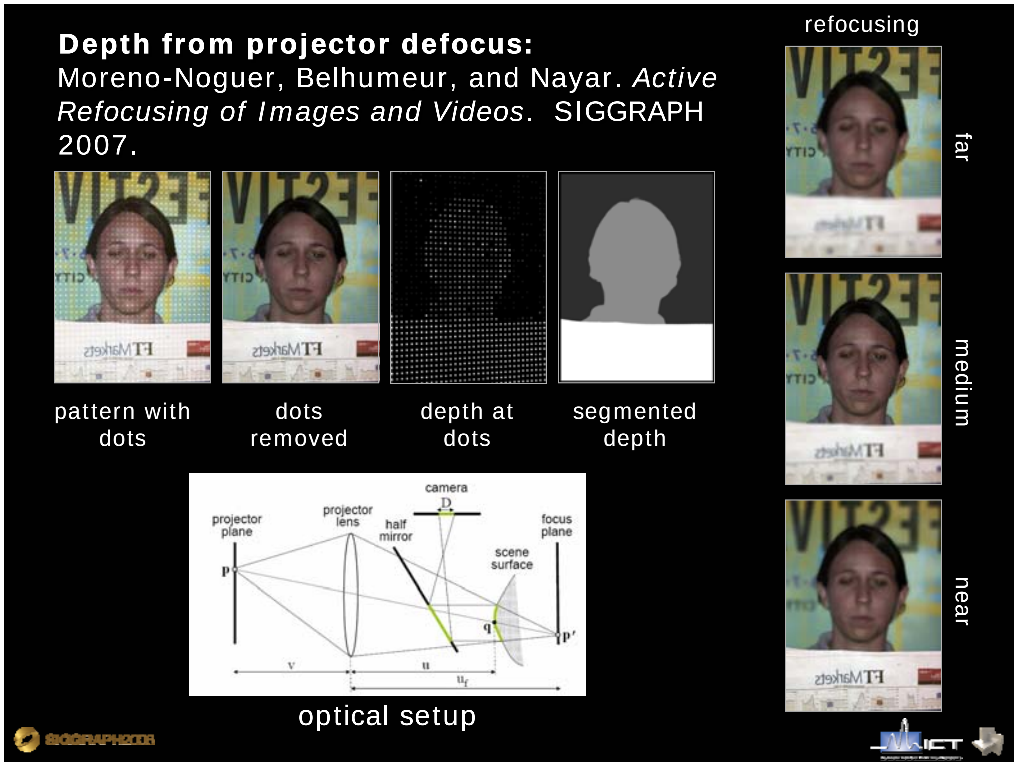

Gray code patterns
Binary (on/off) pattern
Unique for every column
Project inverse patterns to neglect indirect illumination
Robust to blur
Chris Tchou. Image-Based Models: Geometry and Reflectance Acquisition Systems. Master's Thesis, University of California at Berkeley, December 2002.



Finding out which camera pixels correspond to which projector pixels produces a correspondence map, which can be turned into a 3D point cloud or geometric mesh using triangulation. Unfortunately, the geometry can appear aliased to the discretization of pixel coordinates.

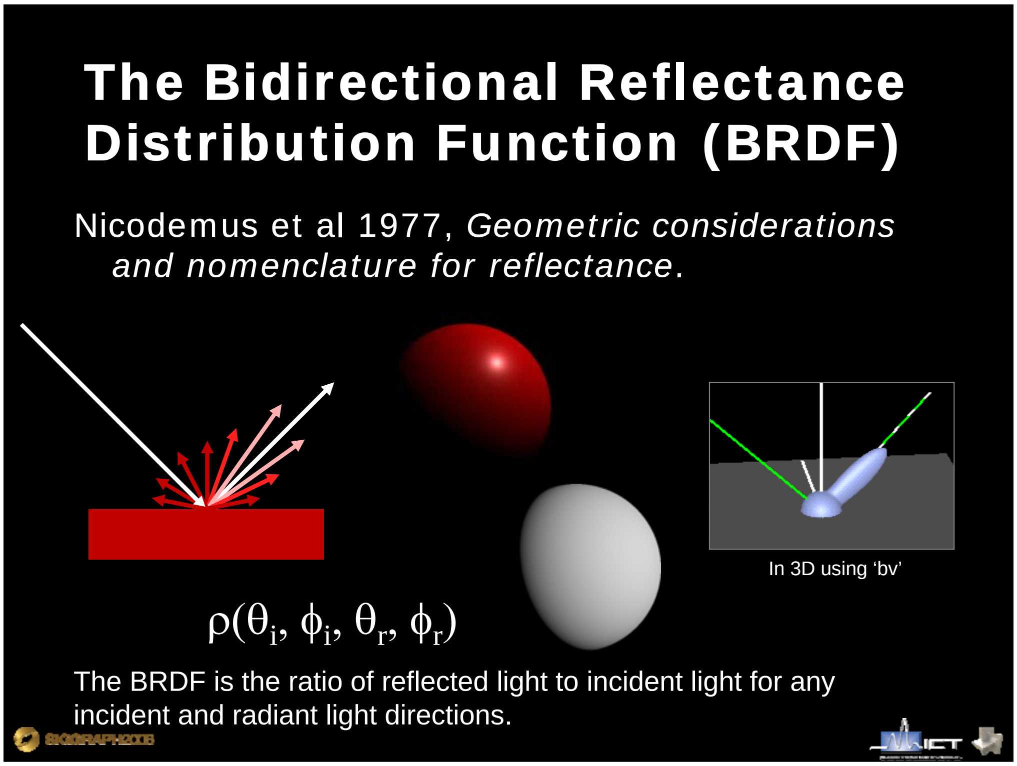


Much smoother geometry can be obtained by slightly blurring the projector and analyzing the grey levels at pixel boundaries, as described in Chris Tchou's Master's thesis.



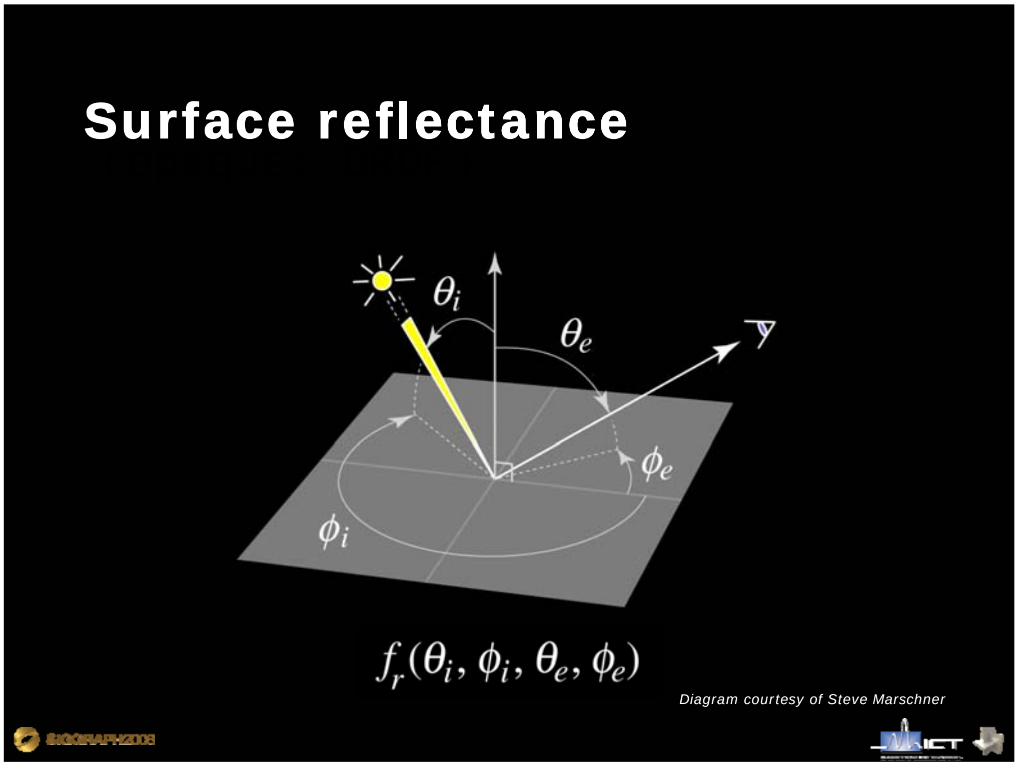
Here's a computational illumination technique for obtaining depth using just one video projector pattern. This SIGGRAPH 2007 paper from EPFL and Columbia aligns a video projector and a video camera using a beam splitter. They then project a grey pattern into the scene with a grid of white dots. The projector is focused behind everything, so the dots are the sharpest (and smallest) when they hit further away objects and larger and appear as larger out-of-focus circles when they hit nearer objects. This gives a depth estimate at each dot position, which can be turned into a depth estimate at each camera pixel based on region segmentation. The dots can also be removed from the digitally projected image since their locations are known.

This computed depth map does not have a great deal of depth fidelity (the person's face reads as a flat card), but it's enough to actively refocus the otherwise in-focus camera image.

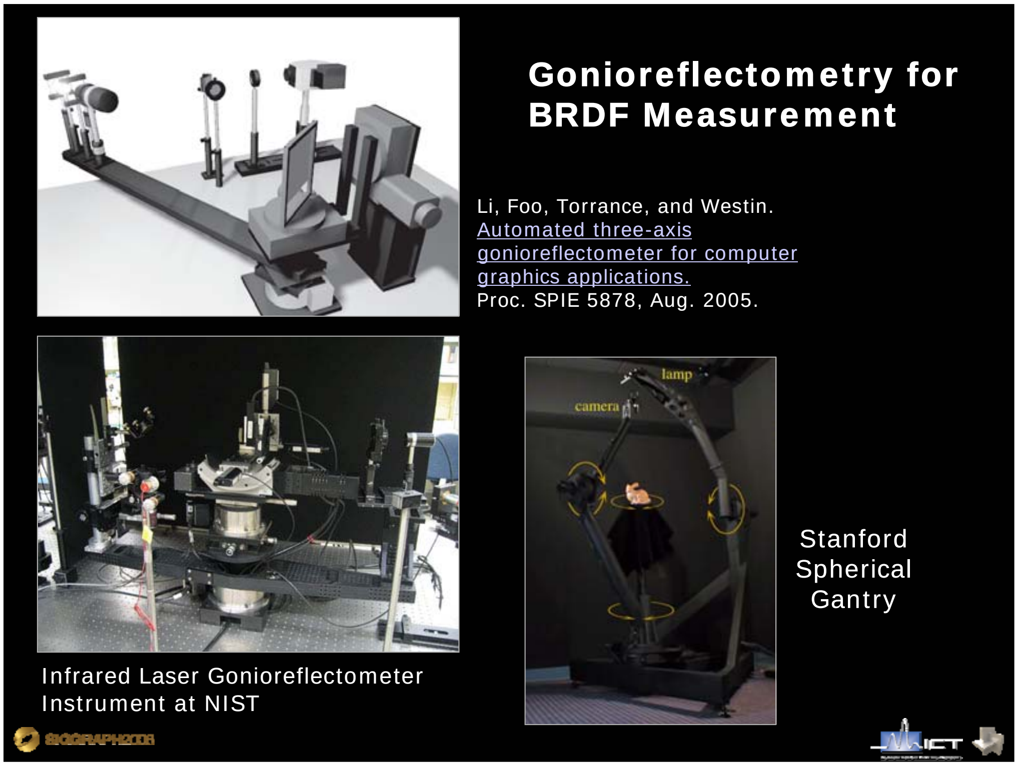


But scanning 3D geometry with computational illumination techniques is not the main topic today. Instead, we’re more interested capturing the *reflectance properties* of objects.

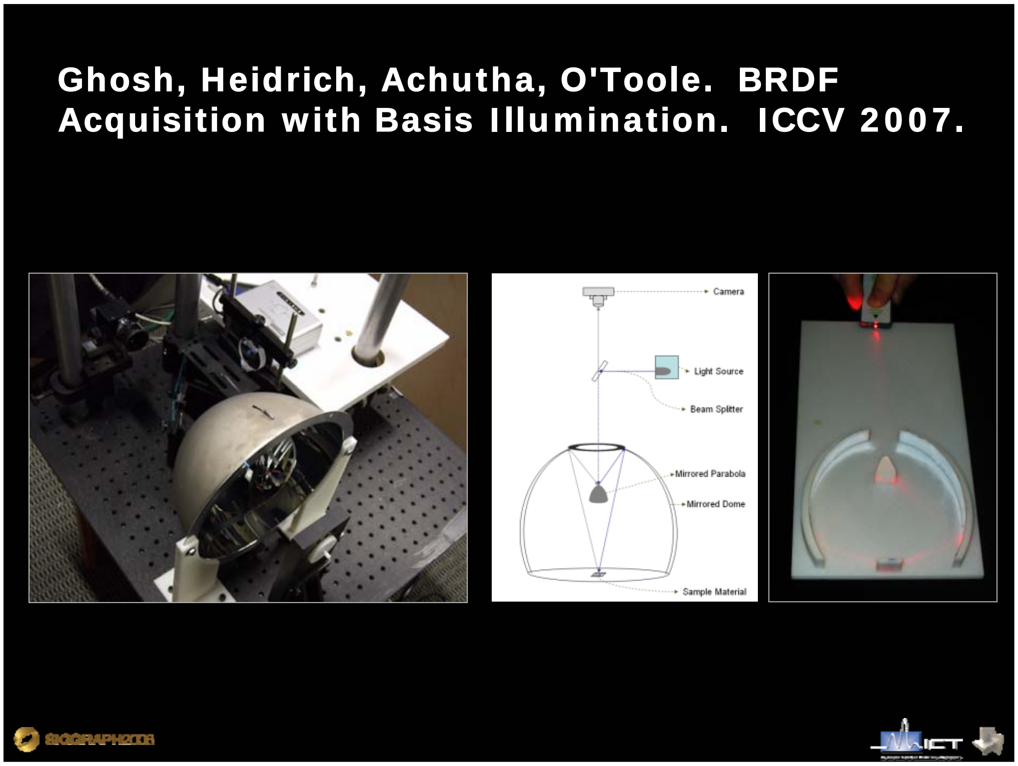
When we traditionally think of reflectance, we think of diffuse and specular components and the various reflectance models which have been proposed for them, all of which generalize to what are known as Bidirectional Reflectance Distribution Functions, or BRDFs. These say for any incident direction of illumination on the hemisphere, what the outgoing distribution of reflected light over the hemisphere is. Mirrors, which simply reflect rays, and diffuse Lambertian surfaces, have particularly simple forms of the BRDF.



Here's a nice graph of how a BRDF is typically parameterized courtesy of Steve Marschner.



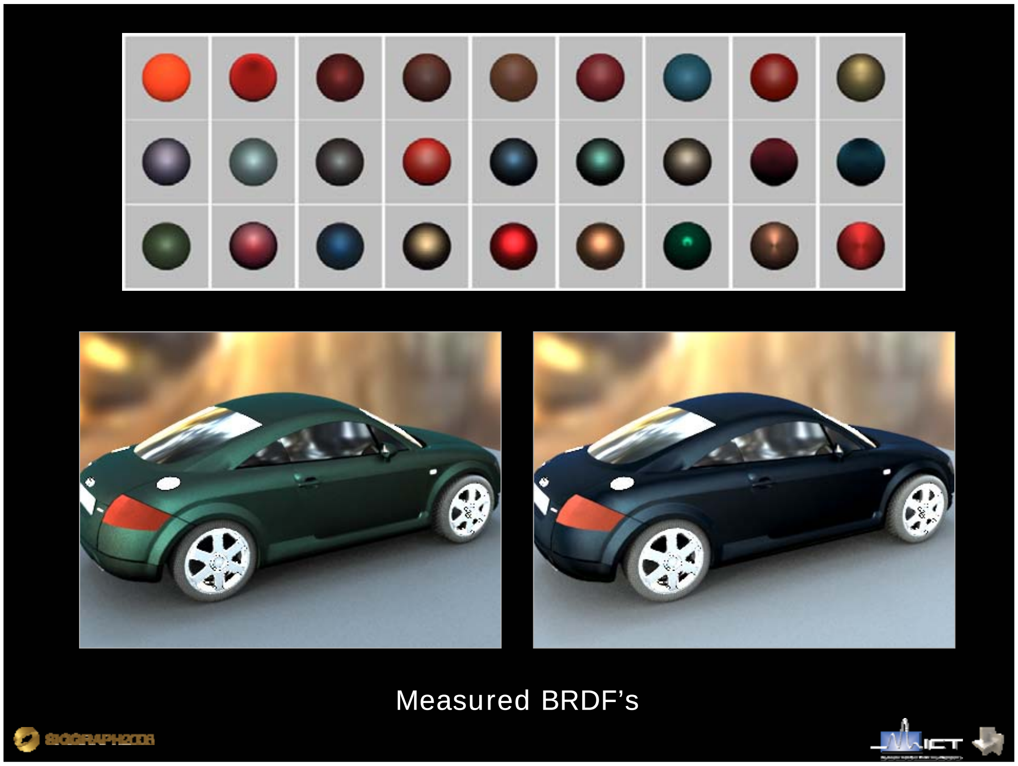
Measuring BRDF's of real materials traditionally requires complex equipment with well-calibrated moving parts and lots and lots of measurements to capture the 4D BRDF of a reflectance sample. Here are a few successful examples of from academia and government.



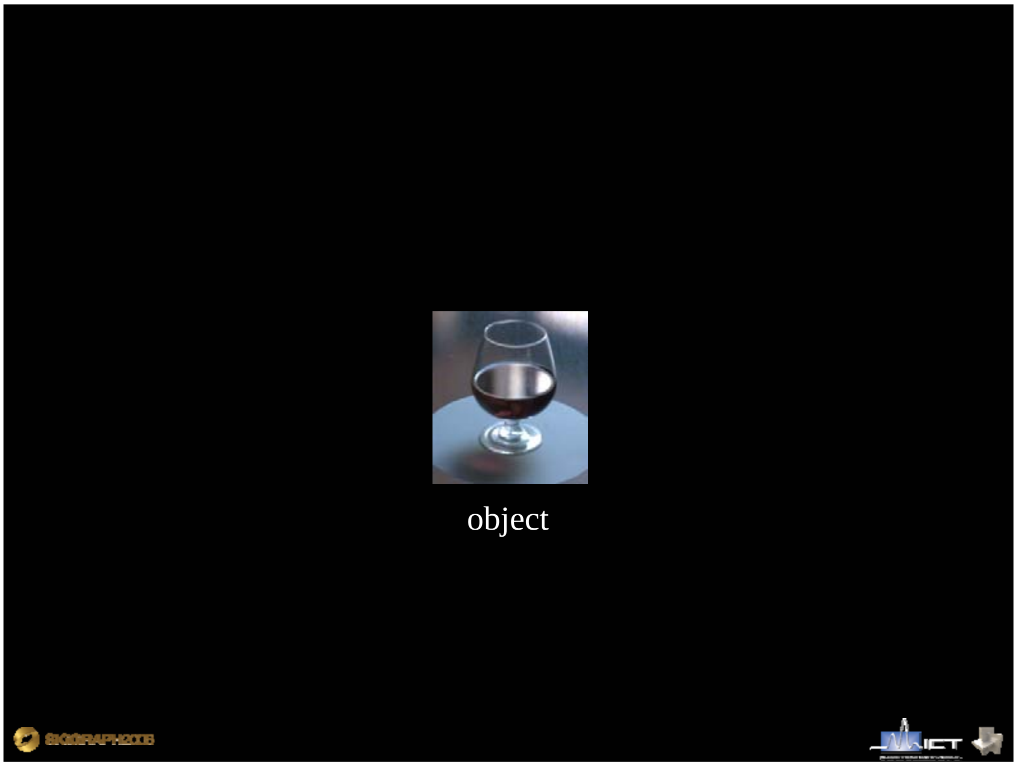
This project from UBC captures BRDFs using a small video projector, a video camera, and custom reflective optics to illuminate and image a material sample over (most of) the hemisphere with no moving parts. That makes measurement potentially much faster. More importantly, the authors do not just project point samples of incident illumination onto the scene. Instead, they project basis illumination functions, which directly measure the surface's response to basis illumination conditions. This allows the full BRDF, as projected onto a set of Zonal basis functions, to be captured in far fewer images than exhaustive BRDF measurement.

From: http://www.cs.ubc.ca/labs/imager/tr/2007/BRDFAcquisition/ :

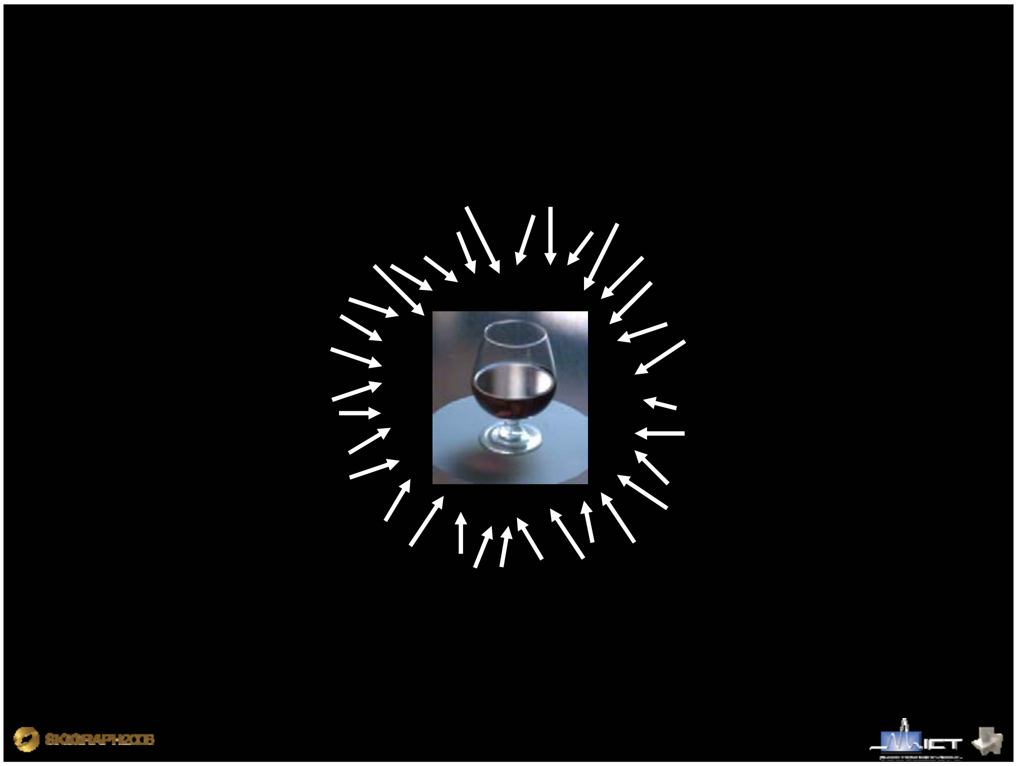
The distinguishing characteristic of our BRDF measurement approach is that it captures the response of the surface to illumination in the form of smooth basis functions, while existing methods measure impulse response using thin pencils of light that approximate Dirac peaks. For this concept to be practical, we require an optical setup that allows us to simultaneously project light onto the sample from a large range of directions, and likewise to measure the reflected light distribution over a similarly large range of directions. Developing such optics also has the advantage that no moving parts are required, which is one reason for the speed of our acquisition. In this work, we choose a spherical zone of directions as the acquisition region for both incident and exitant light directions. Spherical zones have several advantages over regions of other shape. First, they allow us to develop basis functions that align nicely with the symmetries present in many BRDFs, thus minimizing the number of basis functions required to represent a given BRDF. Alignment also simplifies extrapolation of data into missing regions. Second, a zonal setup allows us to design optics that could, in principle, cover over 98% of the hemisphere, with only a small hole near the zenith, where BRDF values are usually smoother compared to more tangential directions.



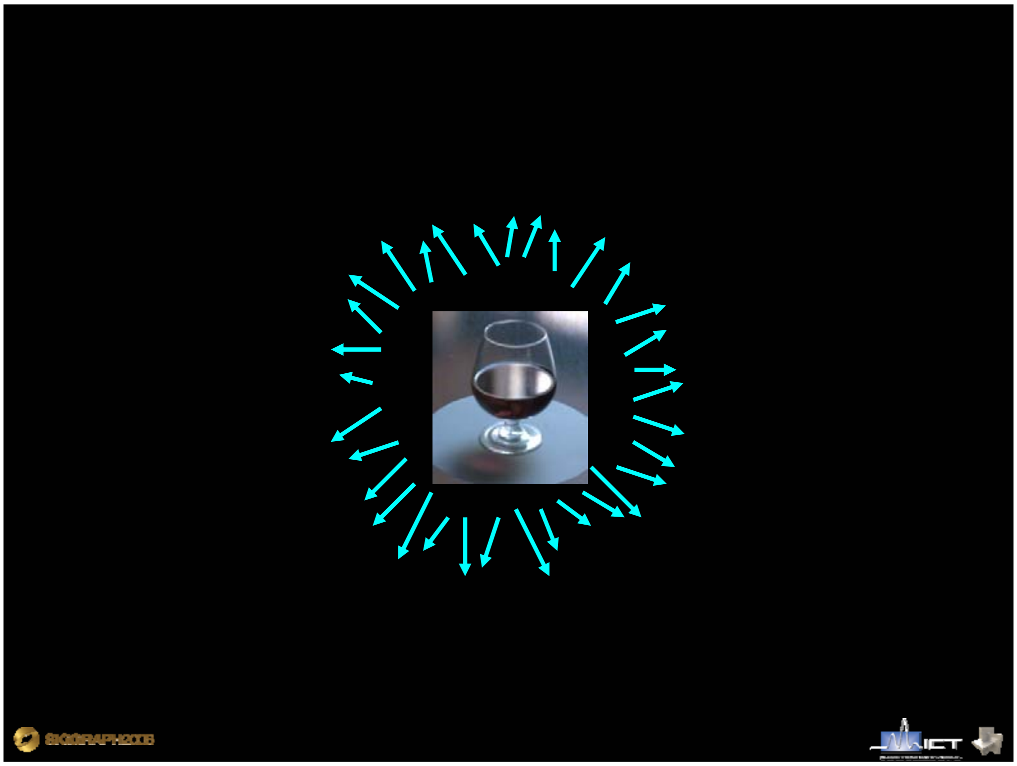
This technique leverages the fact that BRDF's can be represented as a sum of relatively simple basis functions. The projector emits a set of Zonal Basis Function Illumination conditions, and the camera picks up the result of this light when it is reflected. As a result, BRDF models can be fit to the data. Here are some of the BRDFs which were captured with relatively few measurements.



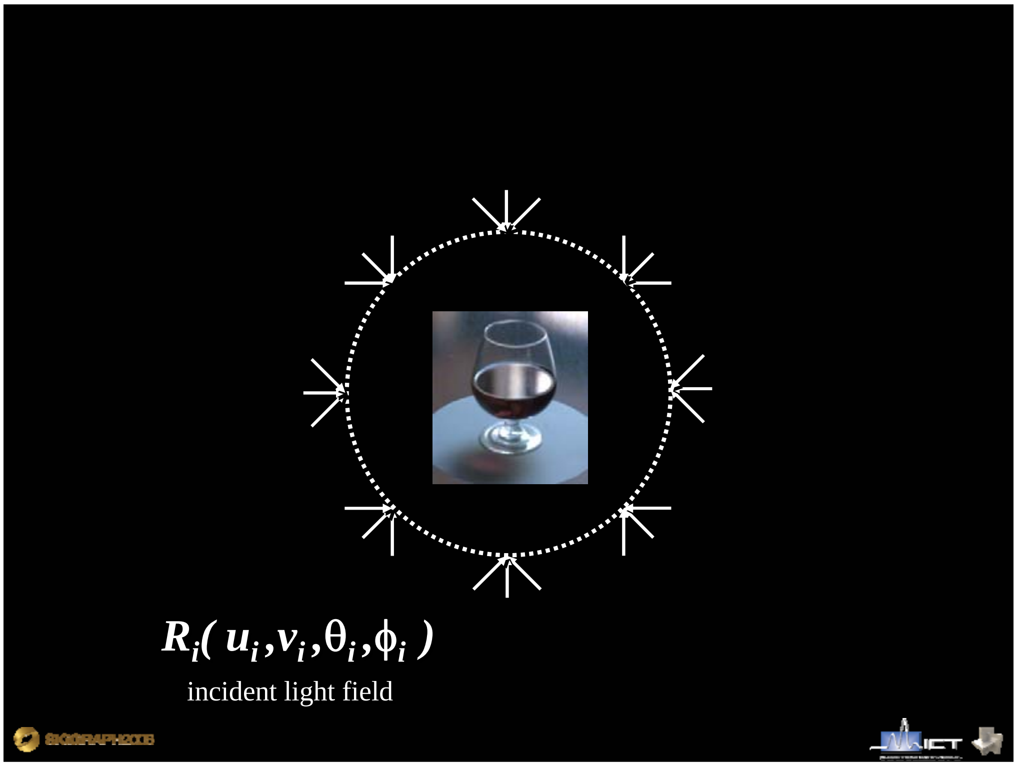
Now, suppose that we want to capture how a whole object reflects light, instead of just a material sample.



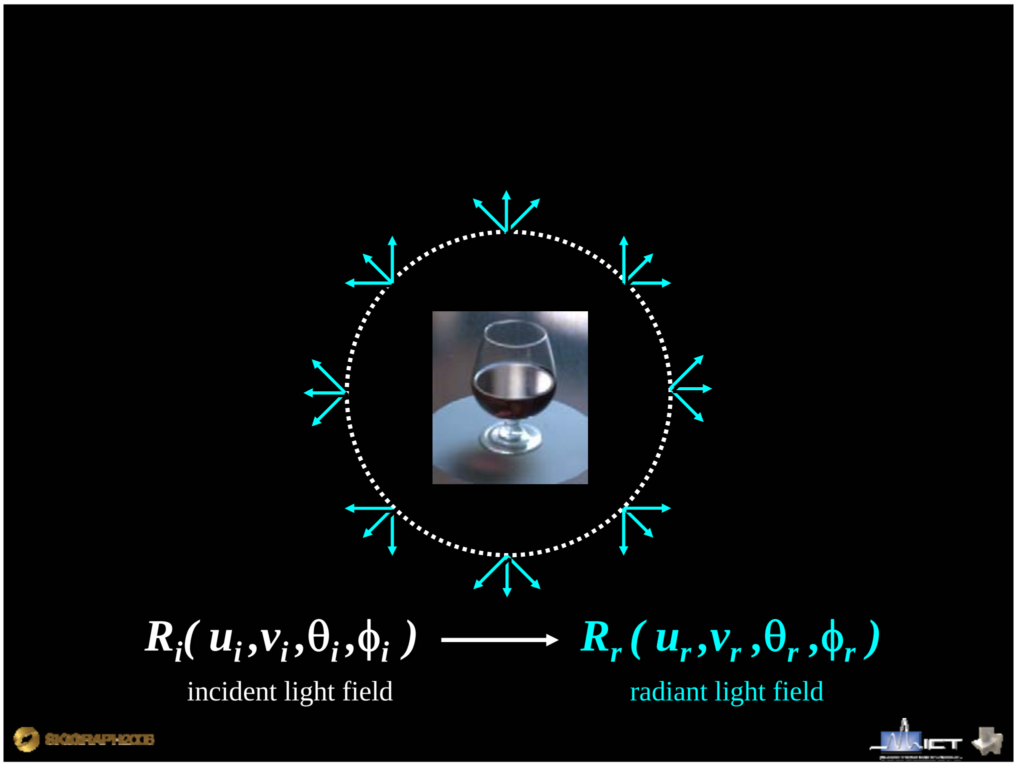
Objects, photometrically, are simply volumes of space which transform a field of incident illumination …



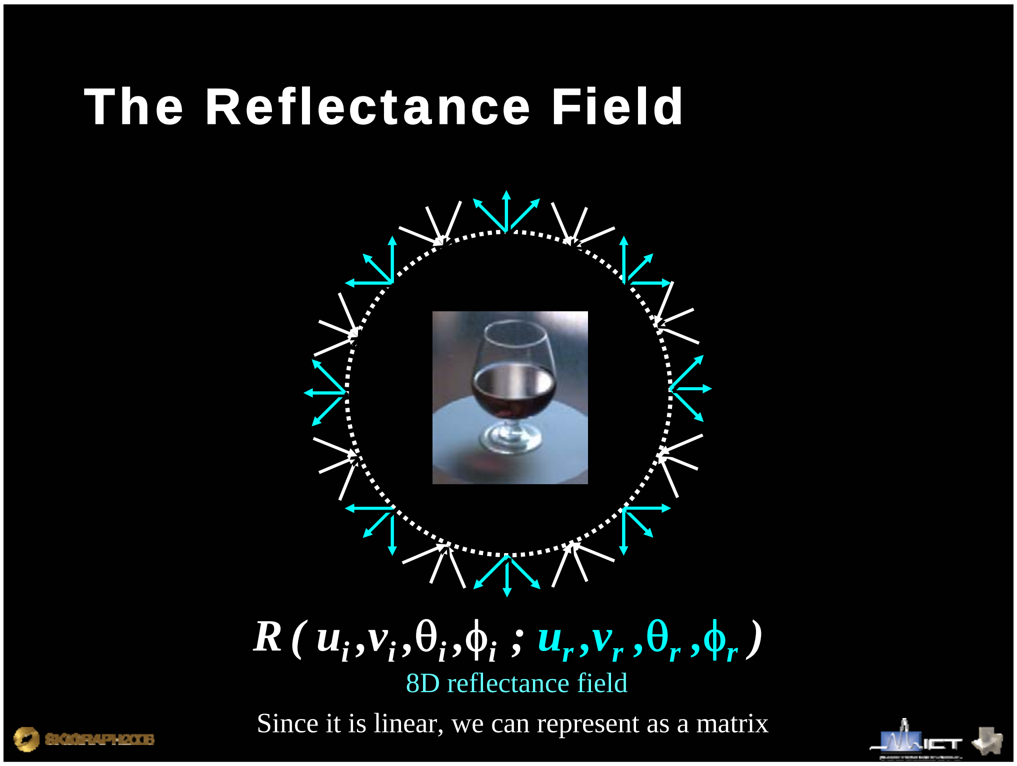
… into a field radiant illumination, reflected back from the object.



We know that incident illumination can be parameterized as a 4D incident light field. To do this we conceptually enclose the object within a convex surface such as a sphere, and we use (u,v) to indicate the position on the surface where the light enters, and (theta,phi) to indicate the direction in which it enters.



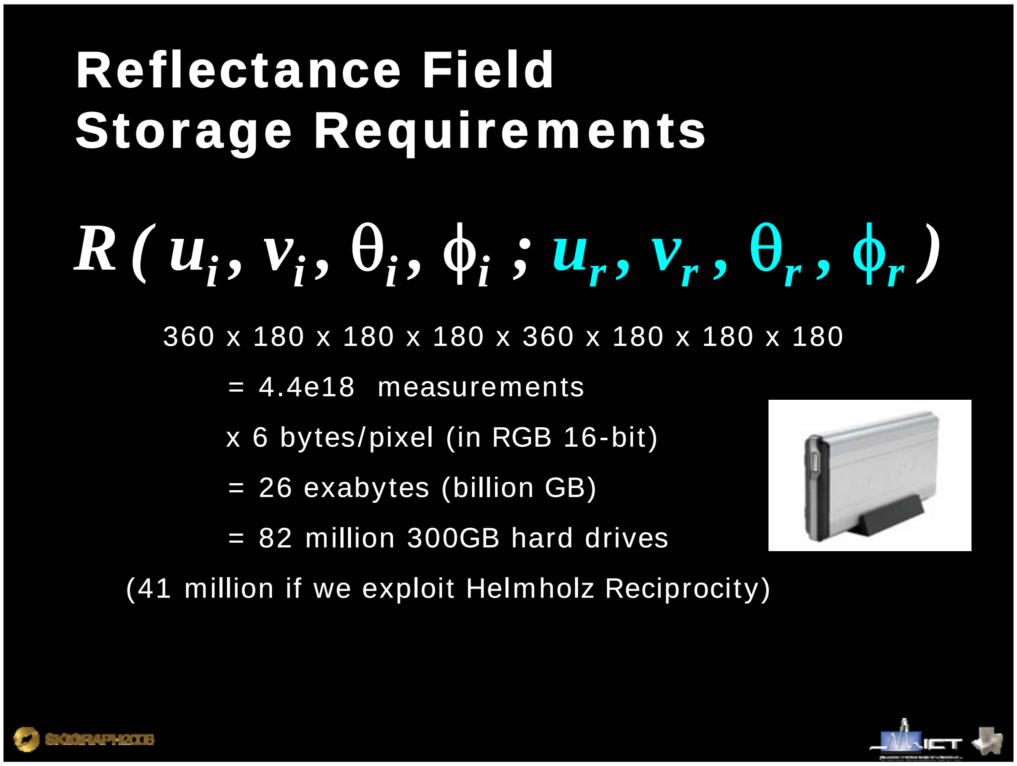
The radiant light can be described similarly as a *radiant light field*. It can be parameterized the same way, except we look at how light is leaving the surface that surrounds the object.



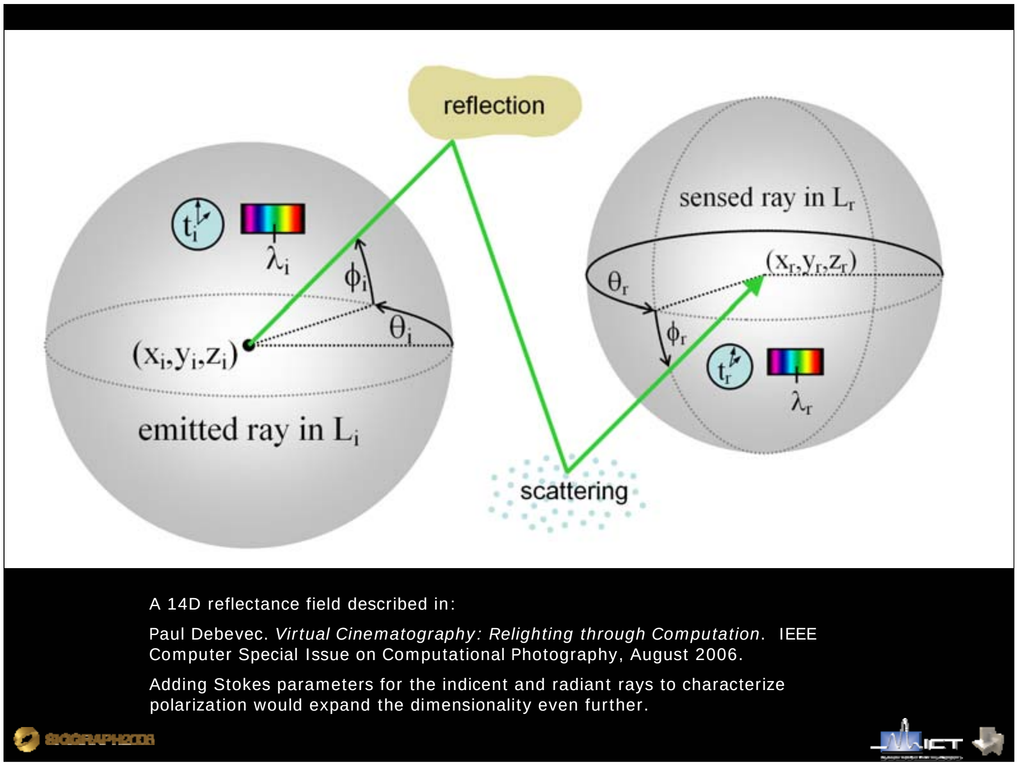
We can thus characterize how an object reflects light as an eight-dimensional function called the *reflectance field*. For any incident ray of light, it encodes the 4D radiant light distribution resulting from the object being illuminated by that ray. The reflectance field thus contains the information necessary to rendering the object under any illumination condition, from environmental to spatially-varying lighting, and seen from any viewpoint.

The reflectance field's form is similar to that of the BRDF, and it's almost as if we have promoted the BRDF from characterizing light reflection at a point to characterizing light transport into and out of a region of space. The reflectance field in fact has the same basic form as the BSSRDF, which represents how light diffuses through an inhomogeneous translucent surface such as skin. However, the surface upon which light impinges is not assumed to be coincident with the actual surface of the material.

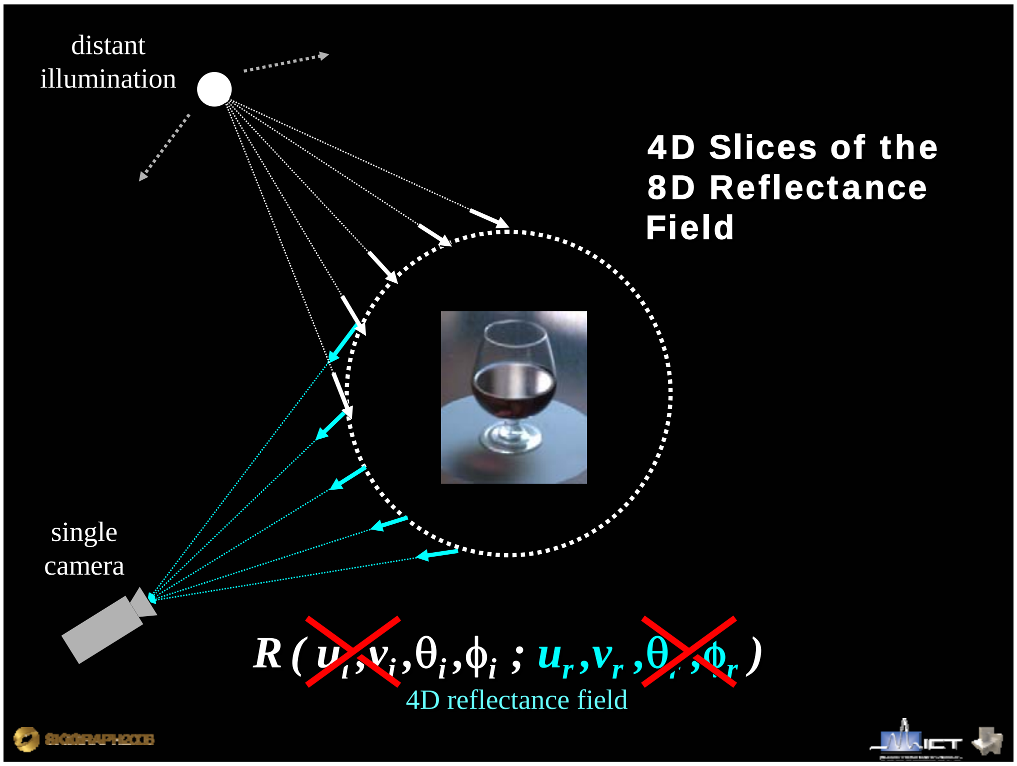
Since light is linear, the radiant distribution of any two simultaneous incident rays is the sum of the distributions of the individual rays. This means that the reflectance field is linear, and thus its transport of light can be represented as a matrix operation from a vector representing the incident light field to a vector representing the radiant light field. This is sometimes called the *transport matrix*.



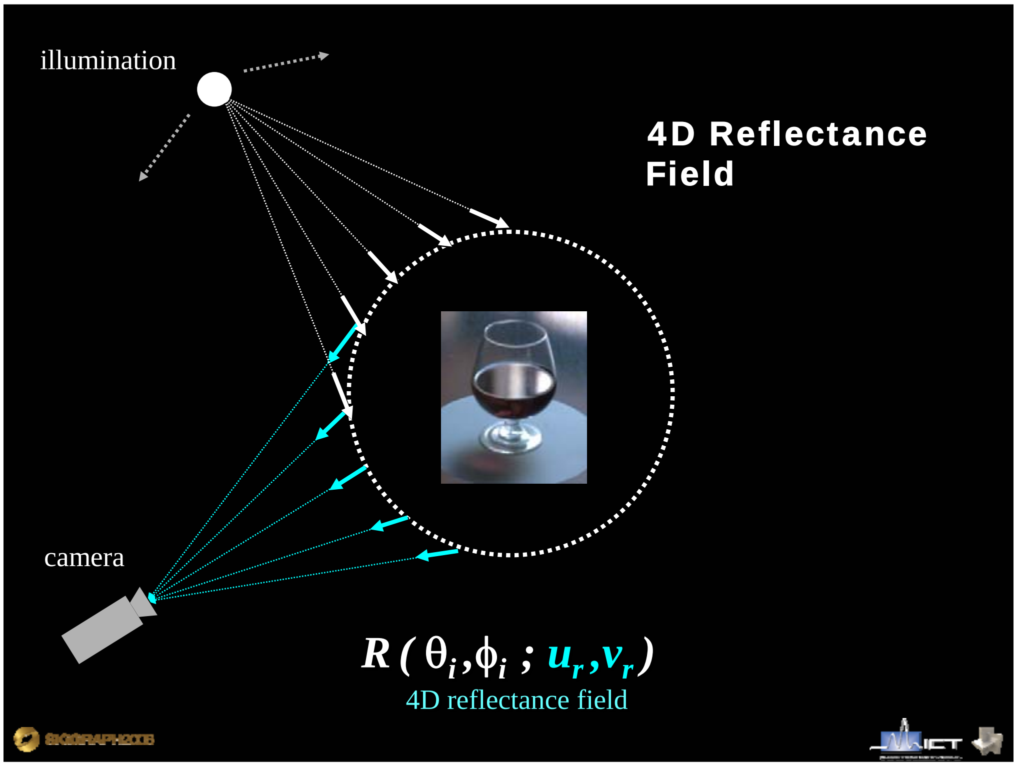
Compared to BRDF's, the reflectance field is even more daunting to capture and store exhaustively.



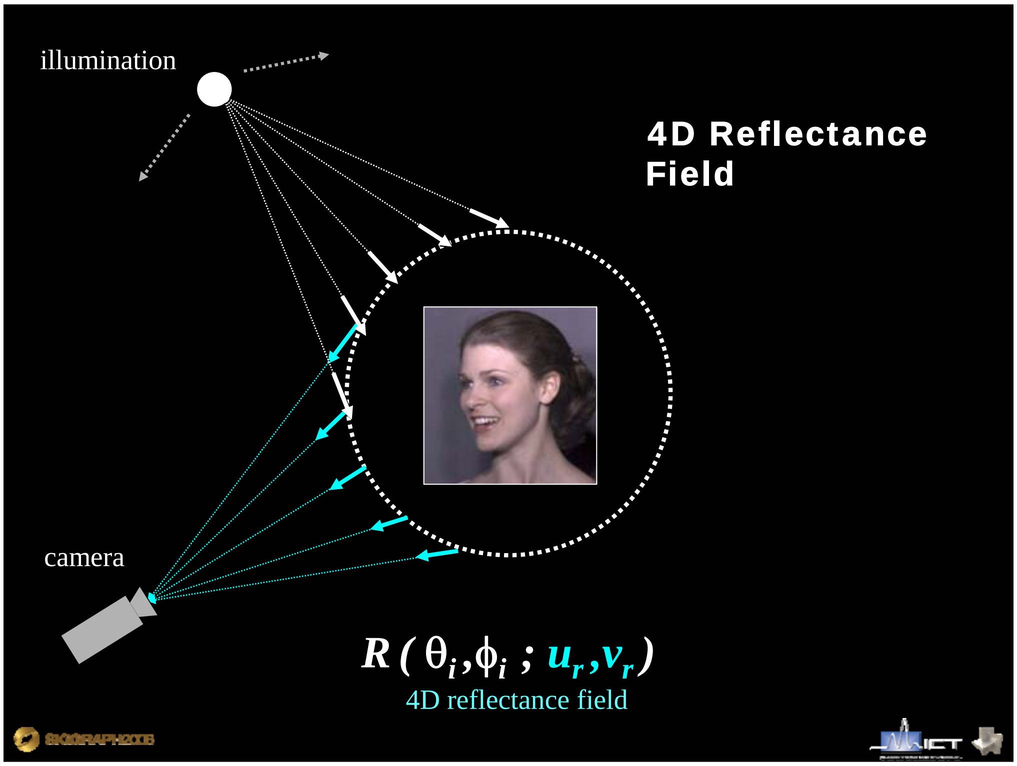
The reflectance field can even be considered more generally. Parameters (on both the incident and radiant rays) for time, wavelength, and 3D position yield a 14D function. Adding Stokes parameters for the incident and radiant rays to characterize polarization would expand the dimensionality even further.



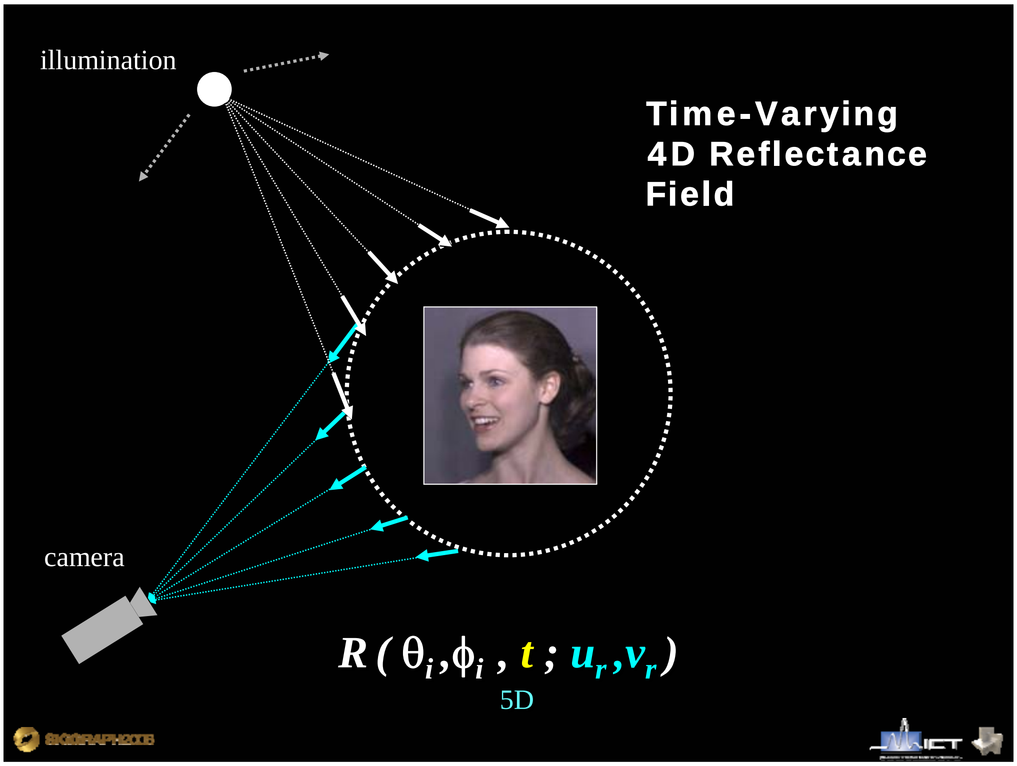
More often, we actually want to simplify reflectance the consideration of the reflectance field. It is often reduced to a 4D function wherein the viewpoint is fixed at a particular camera location, and rays of light are assumed to emanate from far away from the object. This precludes recording the effects of spatially-varying illumination, such as dappled light or partial shadow.



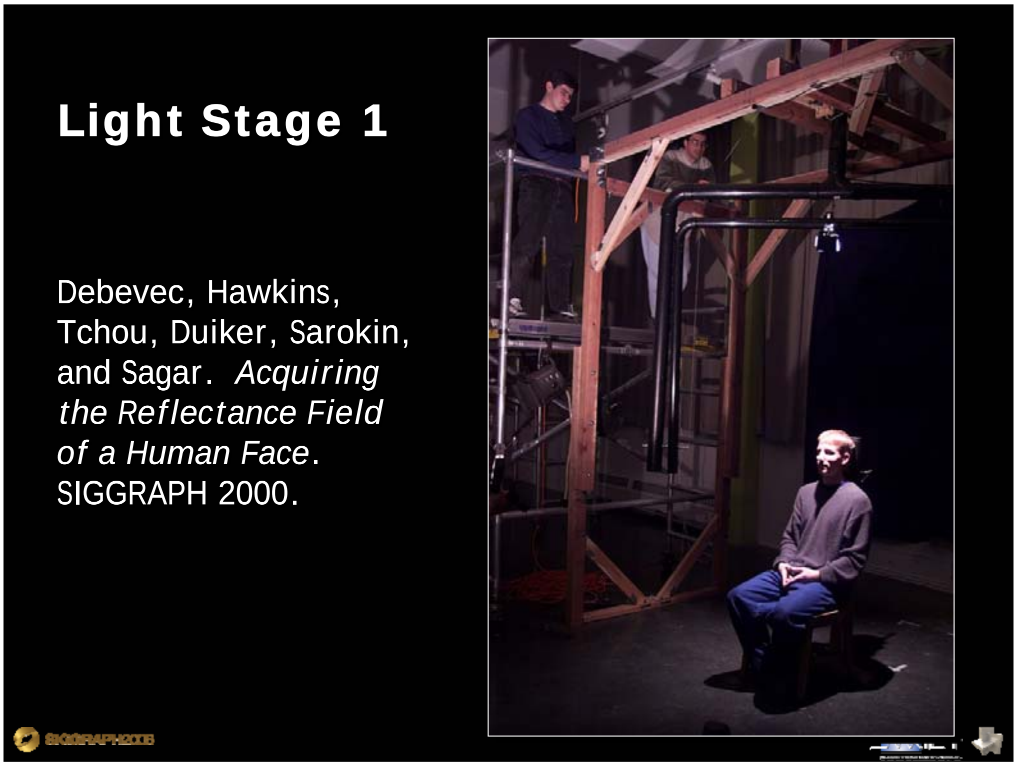
In this form, the coordinates on the surface of the reflectance field has a one-to-one relationship with camera pixels, so (u,v) is usually thought of as the particular camera pixel viewing the radiant illumination.



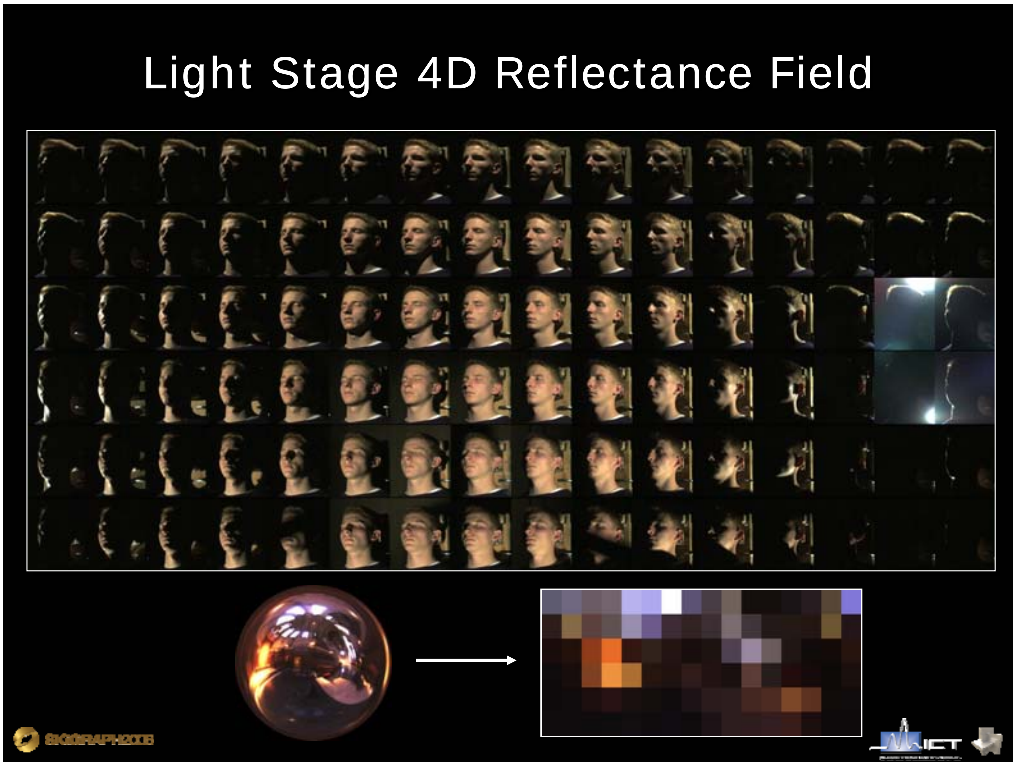
Since it is easier to capture, this 4D version is often the preferred form of reflectance field capture for human subjects.



Of course, people move, so recording a time-varying 4D reflectance field is of interest.



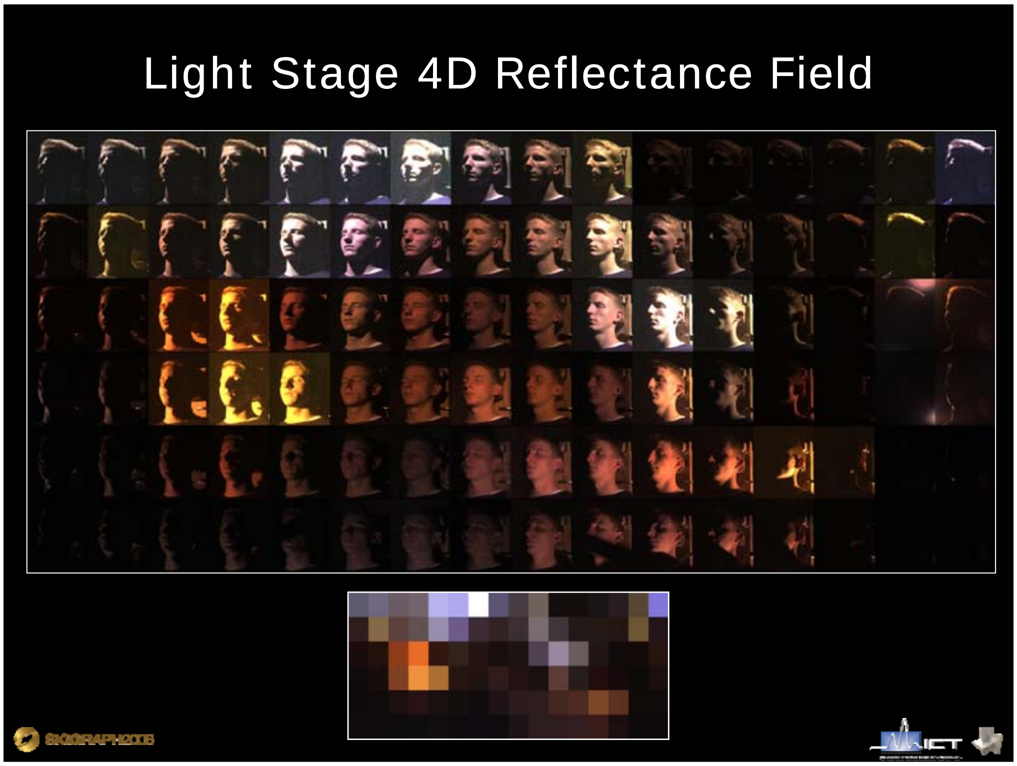
Light Stage 1 was designed to capture 4D reflectance fields of human faces in a tractable amount of time with low-cost equipment, with relatively high lighting resolution.



Spinning the light around in a spiral over the course of minute yields images of the face illuminated in nearly 2,000 lighting conditions. Here we see a sub-sampled version of such a dataset – about 1/16th of the total number of images acquired.

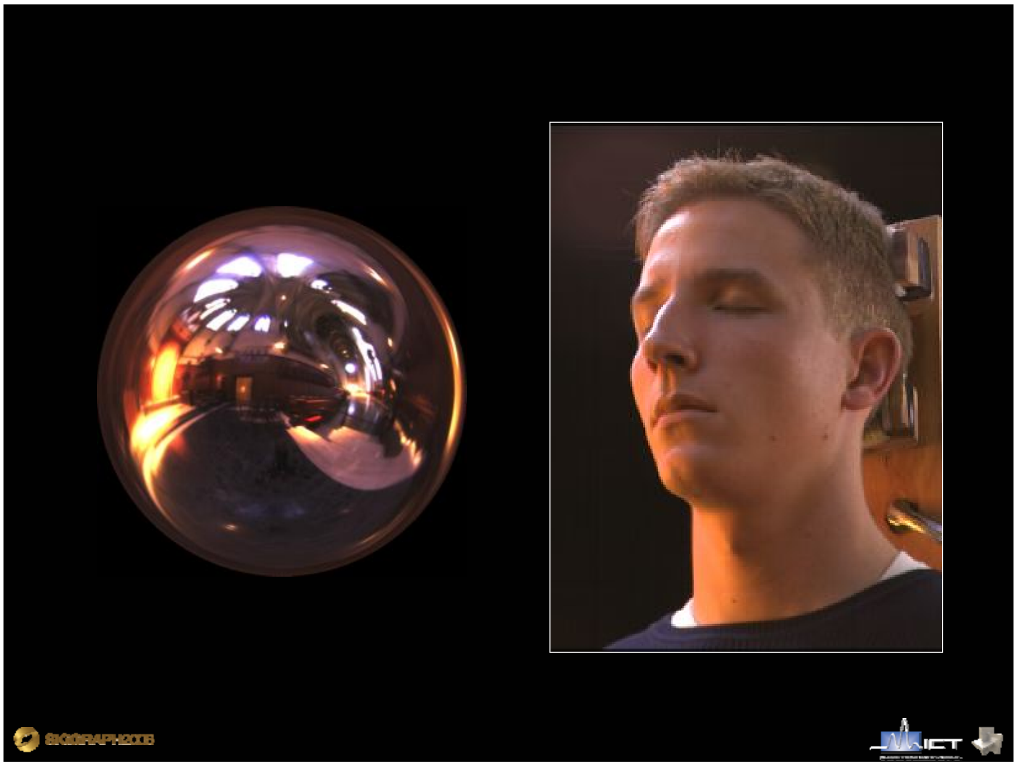
The data shows the face lit from every direction that light can come from. Technically, the light is always just 5 feet away, but since the head is small we assume that this represents the response to a distant lighting environment. It shows what the fact would look lik with a unit intensity white light source from every (theta, phi) direction.

If we want to show the face under a different lighting environment, we first need to resample the lighting environment to be in the same coordinate space and lighting resolution as the facial reflectance field dataset. You can see that at the bottom of this slide.



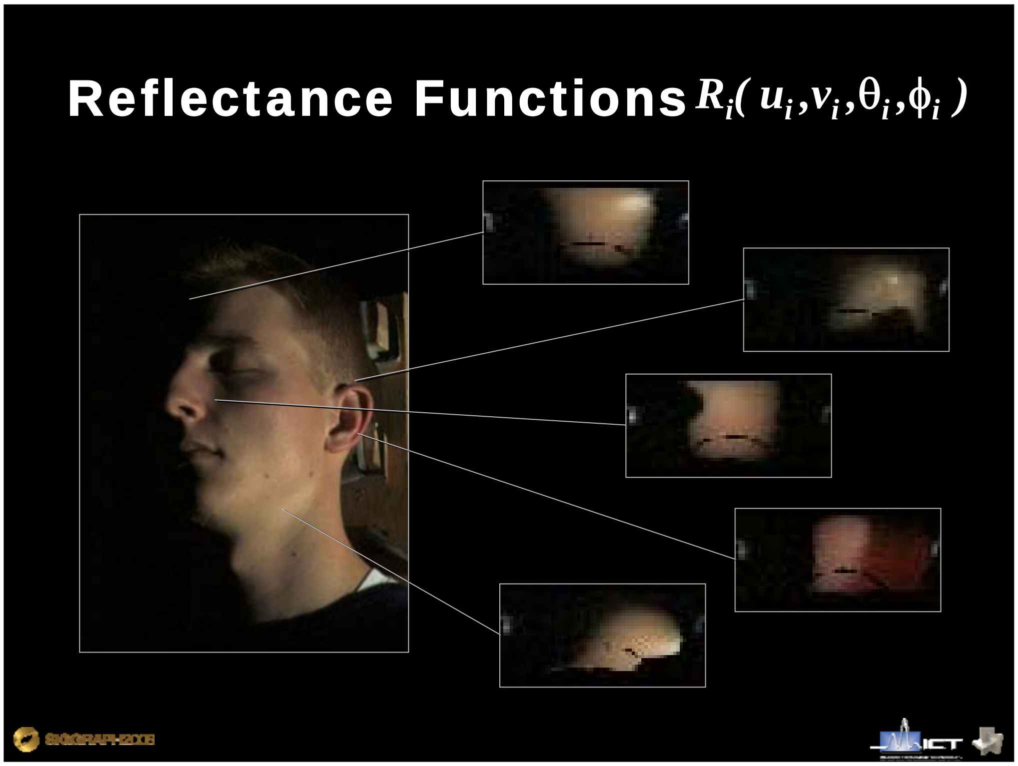
By multiplying the lighting and reflectance datasets together, we get a mosaic of images where each face has been tinted to be the color and intensity of the illumination coming from that direction in the environment. For example, the faces in the center left of the mosaic are bright and yellow since there is bright and yellow light coming from that direction in the environment.

Essentially, we have lit the face by the HDR lighting environment one piece of the environment at a time. To show the face in the entire environment at once, we simply need to add all of these images together.



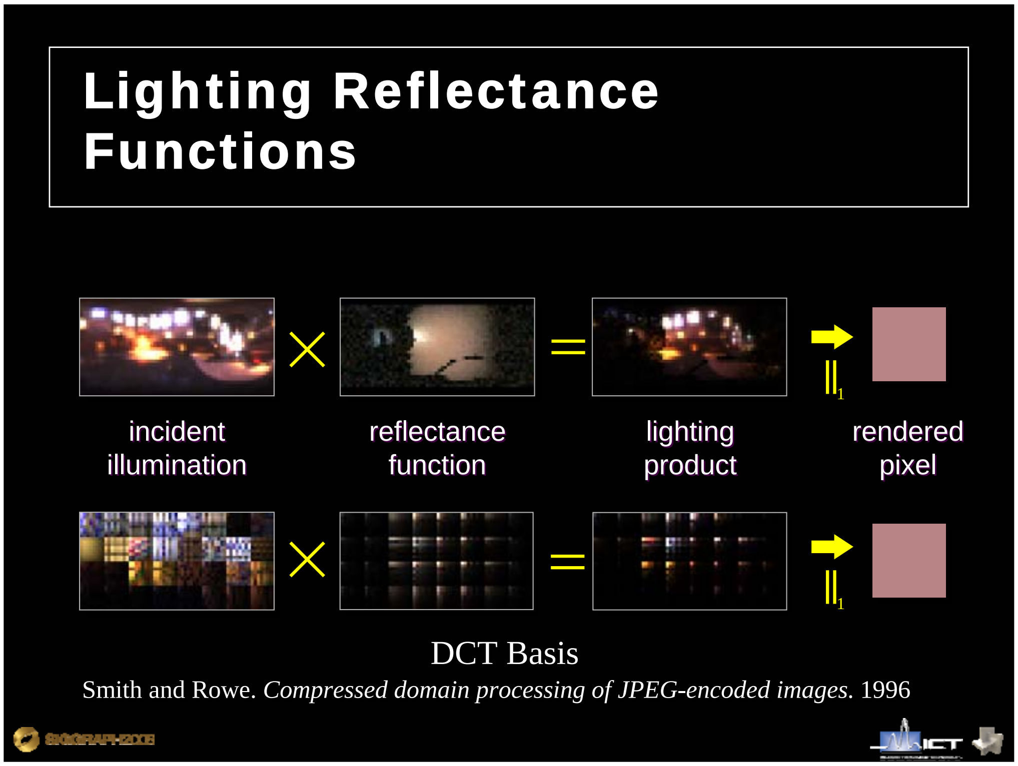
Adding all the images together, since light is additive, yields an image of the person in the novel lighting environment. It's even easy to change the lighting environment!

Here are four other lighting environments from various light probe images.



It is informative to also look at the transpose of the 4D data which we just saw as a 2D grid of faces. If you pick a particular pixel on the face, we have recorded about 2000 pixel values for it according to the incident lighting direction. These can be shown in a latitude-longitude 2D image representation. We call these 2D pixel maps *reflectance functions*, because they encode how a given reflects light from any possible incident direction.

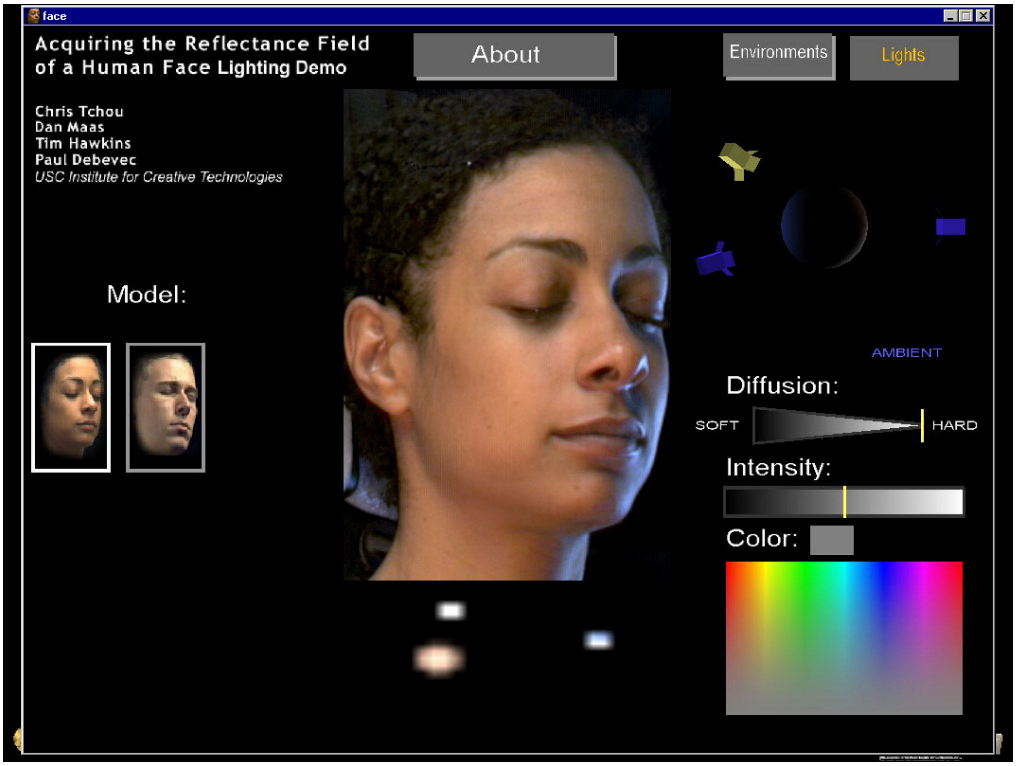
Reflectance functions begin to look like slices of the facial pixel BRDFs, since they include specular lobes and diffuse lobes of reflectance. But they also include non-local reflectance effects such as indirect illumination, self-shadowing, and subsurface scattering. These particular reflectance functions also include some shadowing from the phi-bar and glare from the light source in the back.



We can perform the same lighting calculations in this transposed reflectance function space. The spherical map of incident illumination and the reflectance function yields the color of that pixel illuminated by that lighting environment.

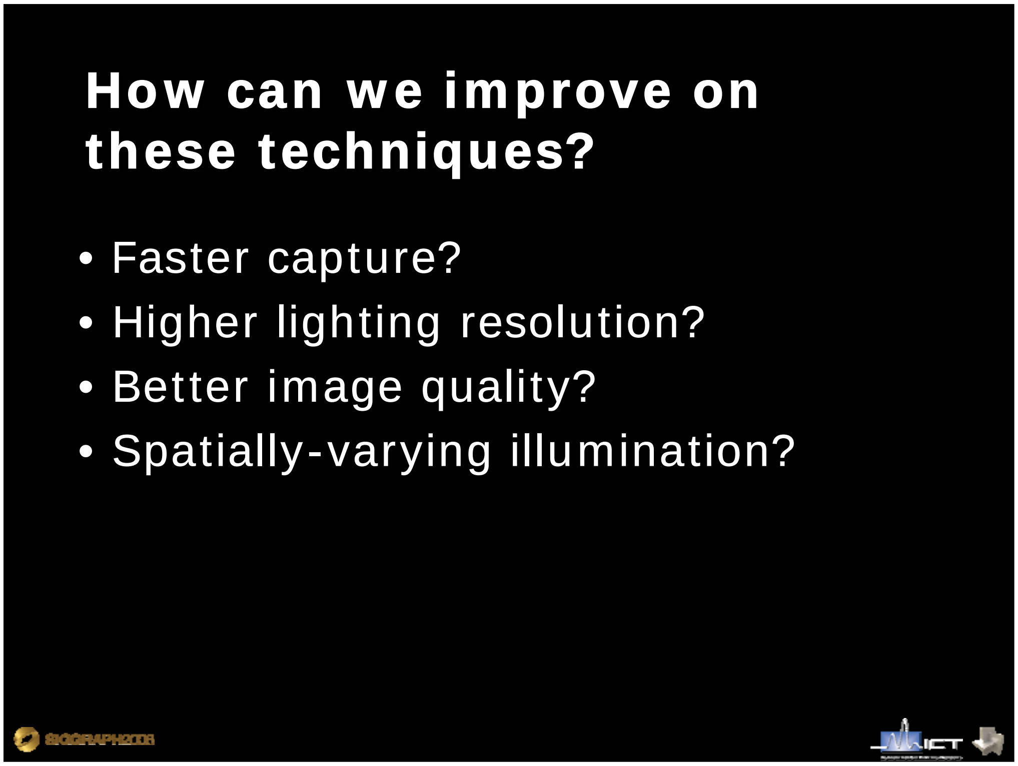
Reflectance functions, even more than regular images, tend to be compressible. In the bottom row you can see here that the reflectance function projected onto the DCT basis concentrates energy in relatively few lighting coefficients. The HDR lighting environment, in contrast, has a lot more frequency content in comparison.

What's particularly cool and useful is that you can still do the relighting process directly on the transformed coefficients and arrive at the same rendered pixel values. That's because the particular transform we are using – the DCT – is an *orthonormal* transform. The techniques of "Precomputed Radiance Transfer" (e.g. Sloan et al SIGGRAPH 2002) all leverage this fact to perform real-time relighting of CG objects essentially based on pre-rendered light stage datasets of the objects.



Performing the relighting in frequency space allows high-resolution datasets to be re-illuminated in real time, such as seen in the real-time face "Facial Reflectance Field Demo" at http://gl.ict.usc.edu/Data/FaceDemo/

A number of publicly-available light stage datasets taken with Light Stage 6 are available on the graphics lab web site.



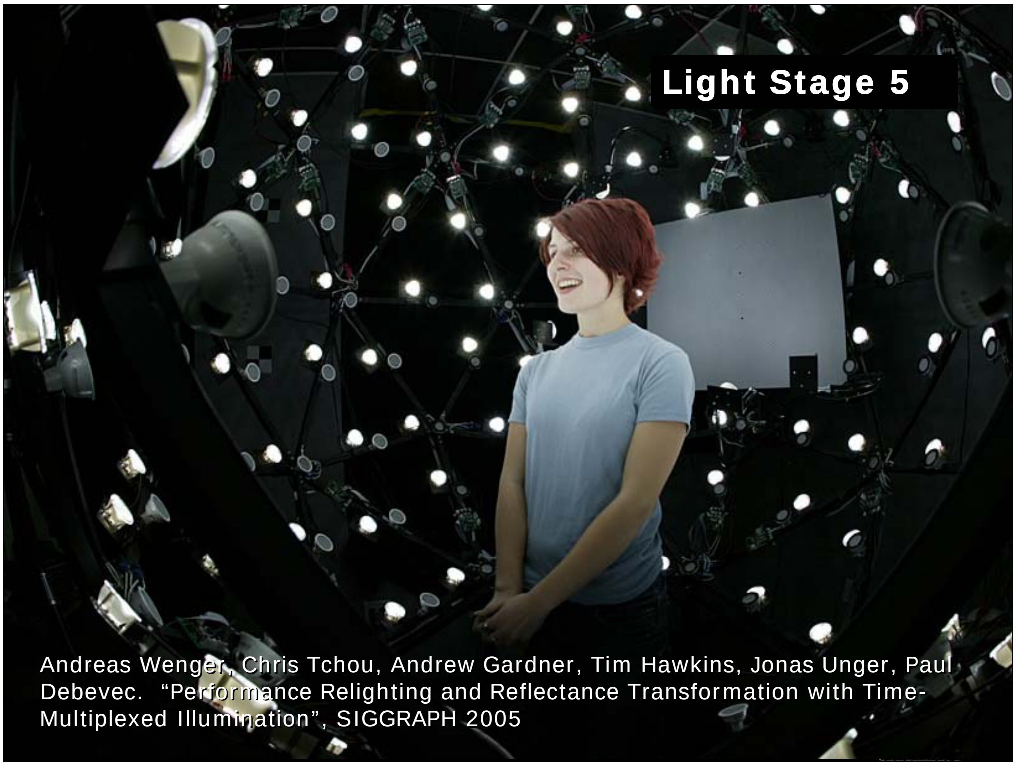
So let's think about how we can do better than what we've seen so far. Can we improve on these techniques to allow for:

•Faster capture?

•Higher lighting resolution?

•Better image quality?

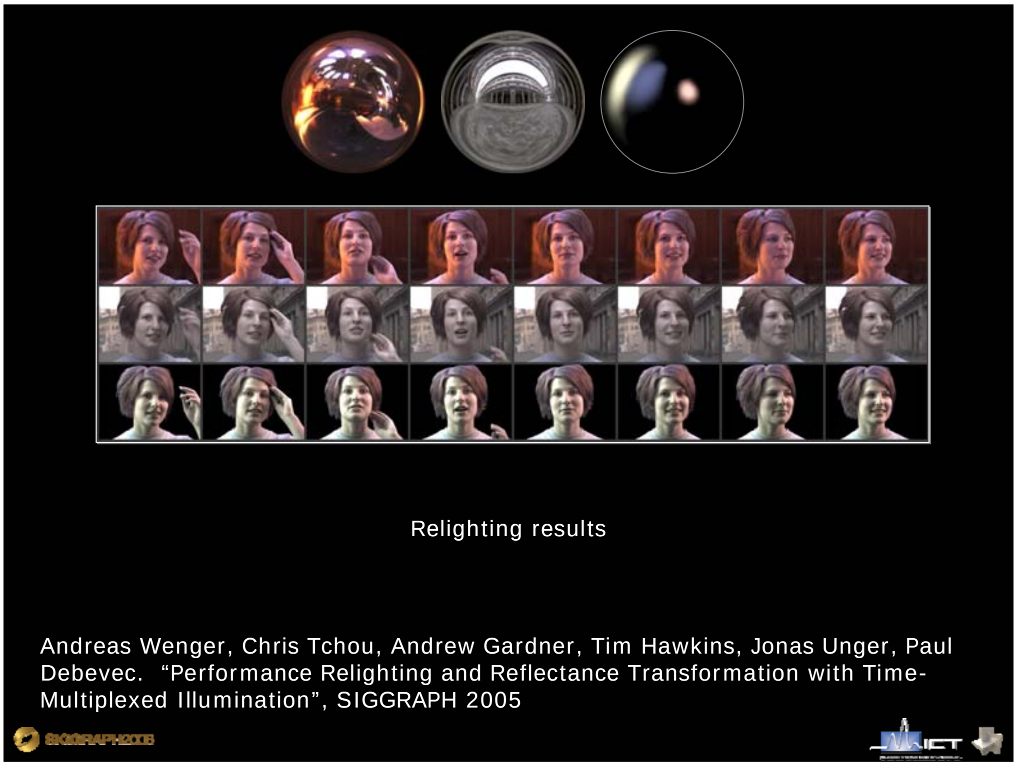
•Spatially-varying illumination?



Faster capture can be done through hardware techniques as it turns out. You just need a light stage with a whole sphere of rapidly-controllable bright LED lights, …

and a high-speed camera, like this Vision Research Phantom v7.1. It can capture images at 800x600 pixel resolution at up to 4800 frames per second.

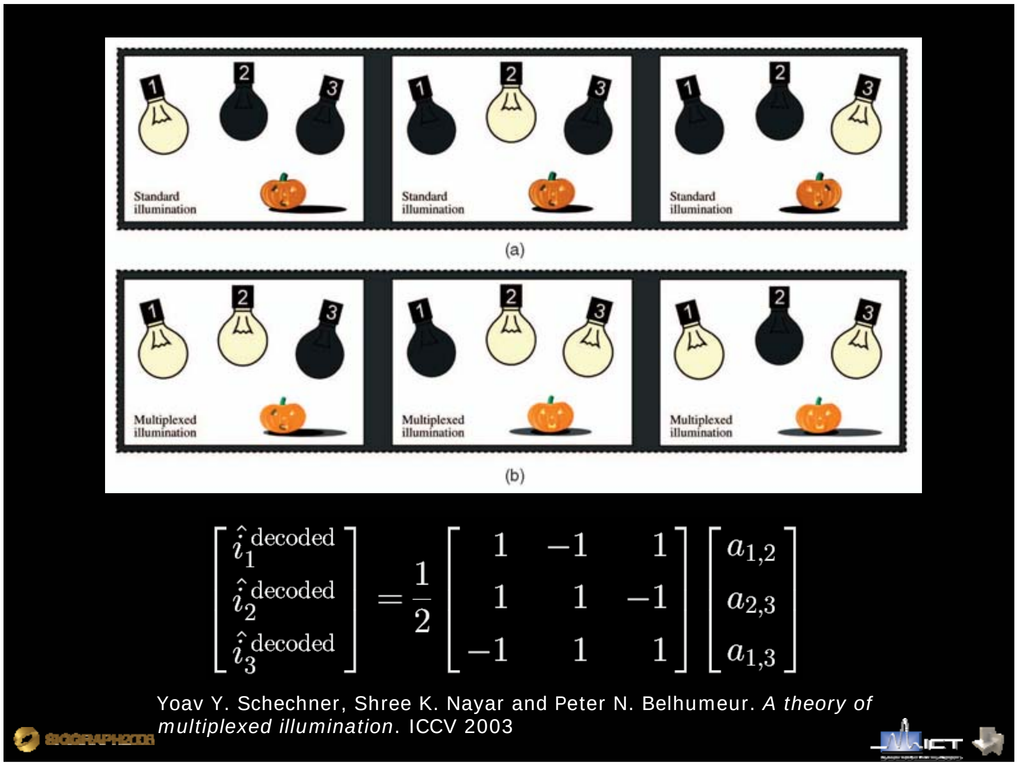
The Light Stage 5 apparatus shown in Figure 3 is a 2m sphere of 156 white LED light sources that surround an actor. The LED lights are controlled by a microcontroller that can change the lighting direction thousands of times per second, fast enough that the illumination appears as a fluttering sphere of light rather than sequential lighting directions. Filming the actor with a synchronized high-speed video camera yields a stream of images of the actor under the repeating sequence of 156 lighting directions, with each complete sequence taking as little as 1/24th of a second of capture time.



With this data, the images can be recombined with image-based relighting to show the actor’s performance, in motion, in any new lighting environment.

To achieve the sharpest results, some motion warping through optical flow is required to give the appearance that each set of 156 images was taken all at the same time.

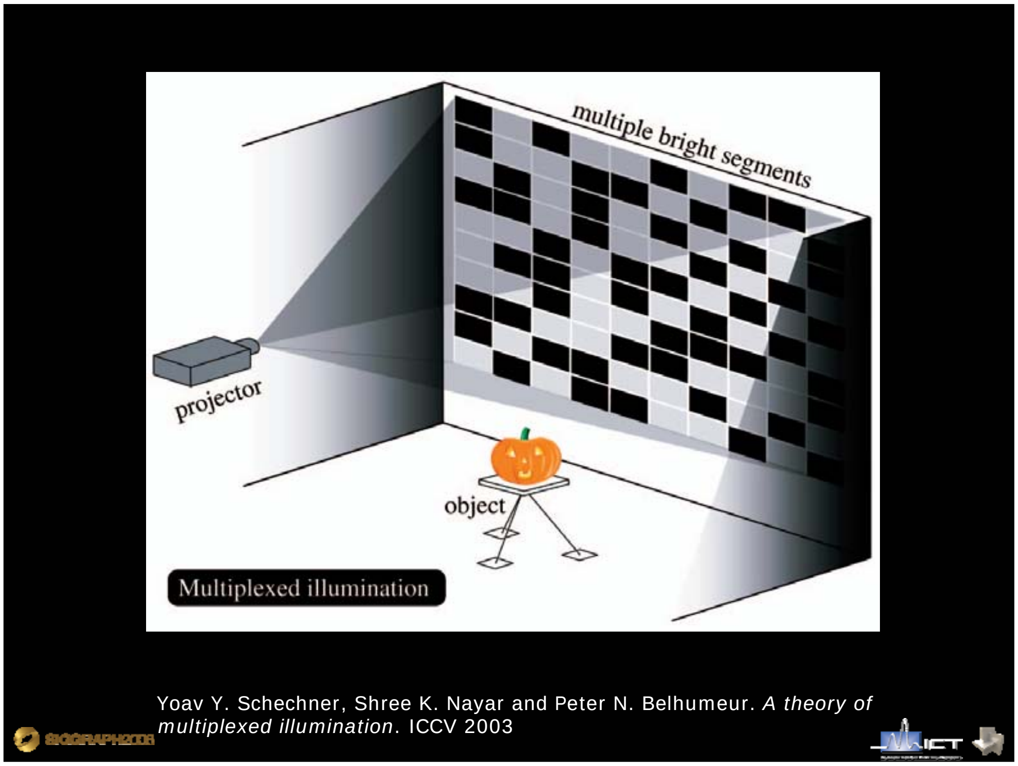
Later in this talk, we’ll discuss other ways of achieving more time-efficient capture, including using gradient illumination patterns, and compressed sensing.



One problem with this kind of high-speed capture is that you become very limited by the amount of light available given that there are such short exposures. The individual lighting direction images from the high-speed camera can actually look quite noisy. Let’s now ask ourselves if we could capture our datasets in a different way which could alleviate this problem, and we’ll look to some work from Yoav Schechner and his colleagues for this.

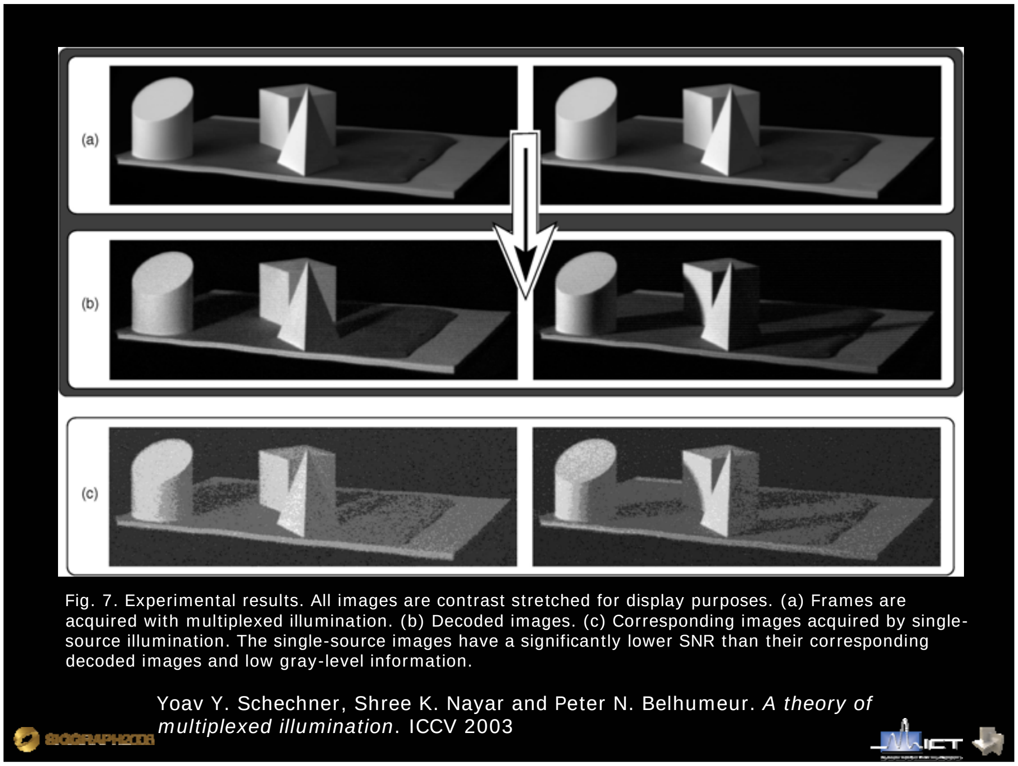
When you capture a light stage dataset, there is nothing requiring you to just turn on one light at a time (as seen for three lights in the first row). If you instead turn different sets of lights which are linearly independent and thus span the same space as the set of single light sources (as seen in the second row, when two lights at a time are turned on). You just need to run the resulting images through an inverse matrix to get back to the images illuminated by single light sources, as seen at the bottom for this small example. Why would you want to do this, other than some fun with linear algebra and a higher electric bill?

Well, as it turns out, the images you get by demultiplexing will generally have a different signal-to-noise ratio than the single-light-source images. Suppose there is additive noise of variance sigma^2 in each pixel of every image taken. Then, the demultiplexed images will have a sigma of (3/4)sigma^2, which is less than the variance in single light source images. (For example, the variance of a1,2 – a2,3 + a1,3 is three times the variance of the original images since Var(a+b)=Var(a)+Var(b) if a and b are uncorrelated, and the variance of half this quantity is one quarter that value since Var(k*a)= k^2 Var(a).



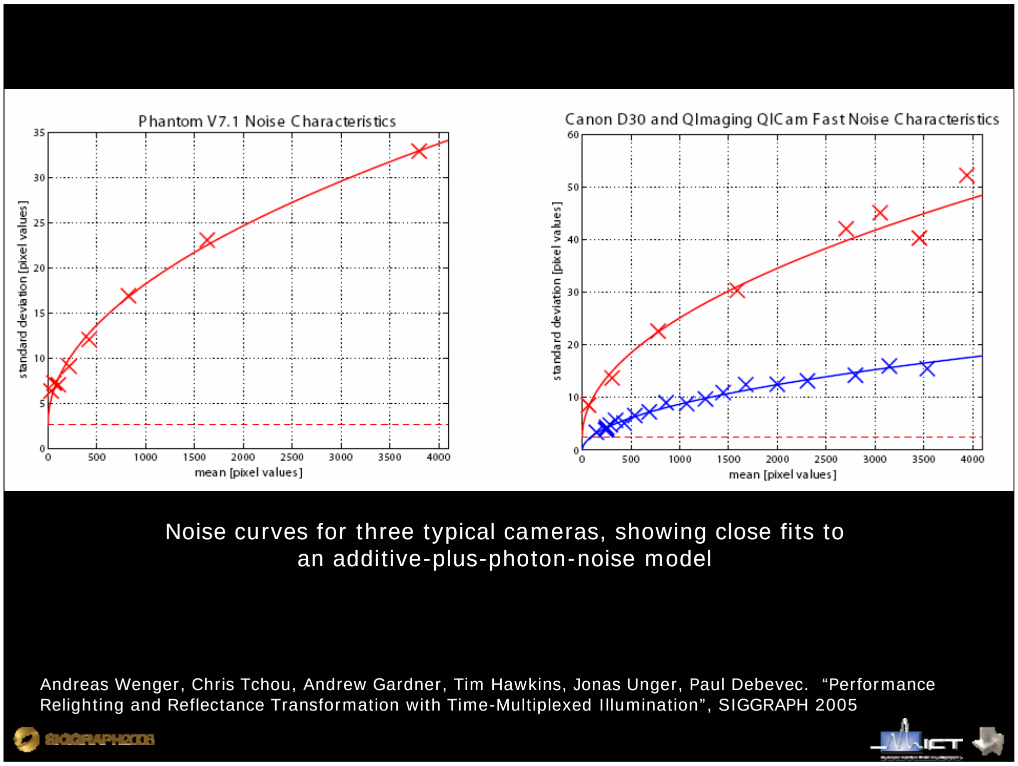
Schechner et al tried this approach with a larger number of light soruces by projecting patterns of rectangles of light onto a wall of a room to act as a set of light sources reflecting back onto a subject (the pumpkin, in this diagram) over a subset of the incident lighting sphere.

In their work they used Hadamard patterns, formed using an S-matrix, to illuminate the scene. There are the same number of Hadamard patterns as there are individual lights, but each Hadamard has just over half of the lights turned on.



The top two images show actual images taken under the Hadamard patterns. The have nice noise characteristics, but they are not the final images we are interested in. Instead, we can . The images are certainly noisier, but still look reasonable.

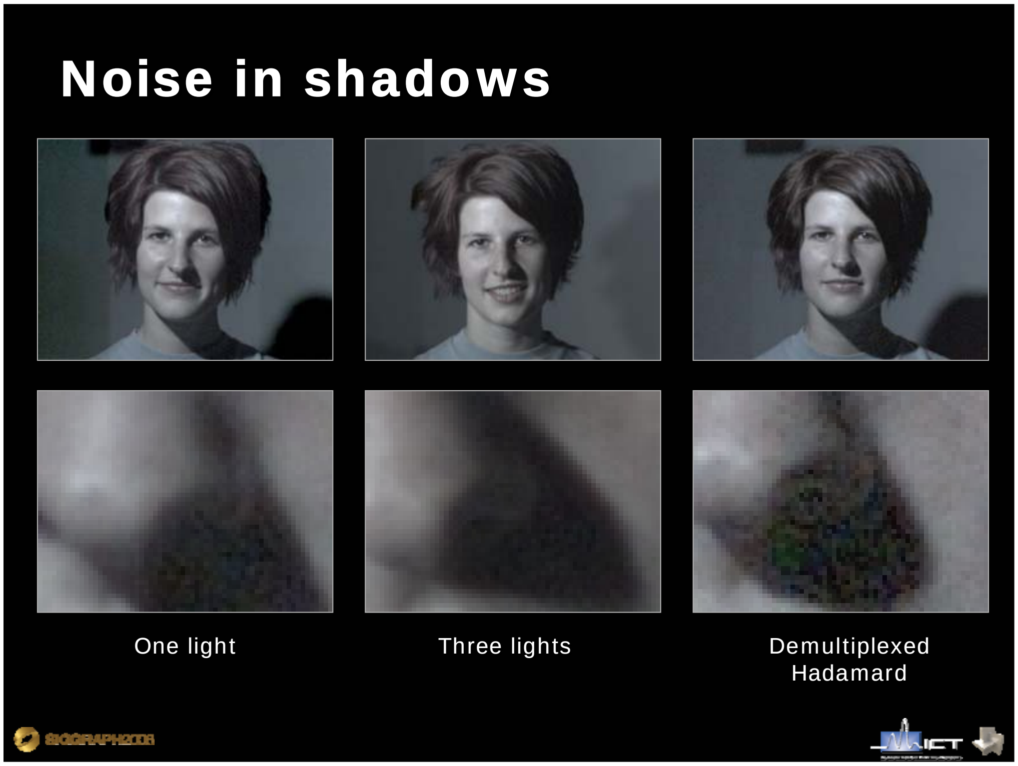
The bottom two images show images taken under single light sources. Since a single light source is pretty dim, the images are quite dark. Here they have brightened considerably in order to show the image, and it's clear the signal is so small that the quantization noise of the camera has considerably degraded the signal. Since quantization noise is additive noise, it can be reduced using Hadamard multiplexing.



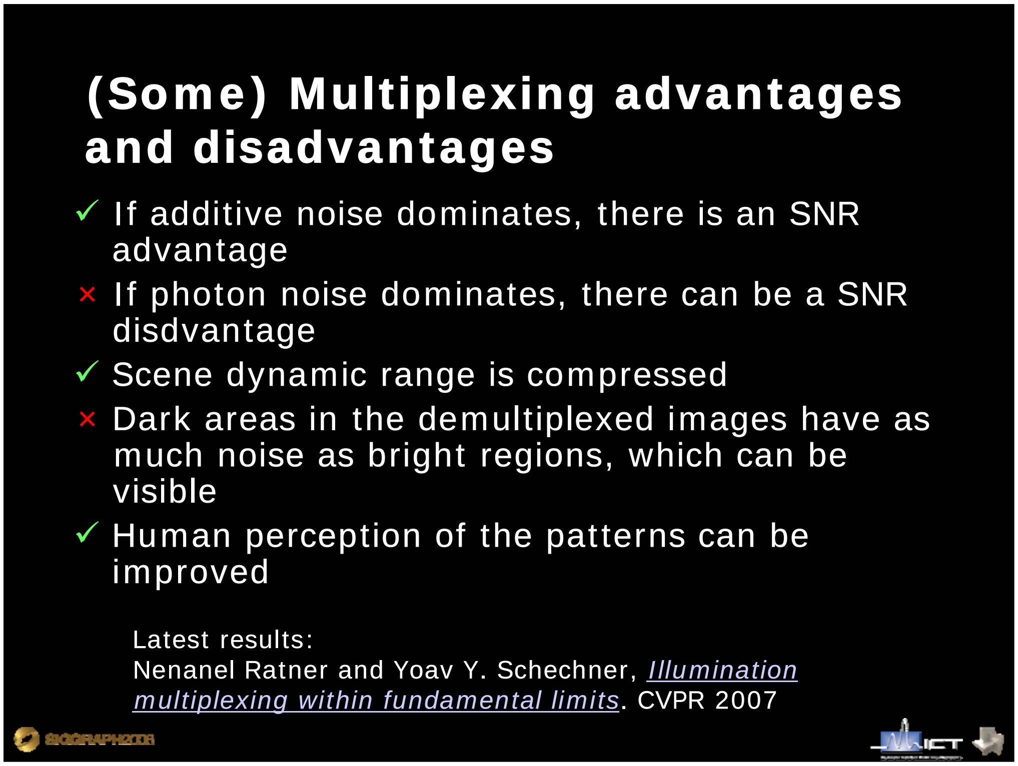
There is an issue which arises when applying Hadamard multiplexing with real images, which is that typical cameras have both additive noise (due to quantization and dark current) plus some amount of photon noise whose variance is proportional to the signal.

These three curves show noise response curves for three cameras we've used in our laboratory. We shot 100 images of uniformly lit patches at various brightness levels, and graphed the variance for a pixel in each patch against the mean pixel value for that patch. The cameras include the high speed camera from the Light Stage 5 project, a Canon D30 still camera, and a cooled QICam machine vision camera. Each curve was well described as a constant amount of additive noise plus photon noise with a standard deviation proportional to the square root of the signal, i.e. sideways parabolas. The cooled QICam had the lease dark current noise of all the camera, almost negligible.

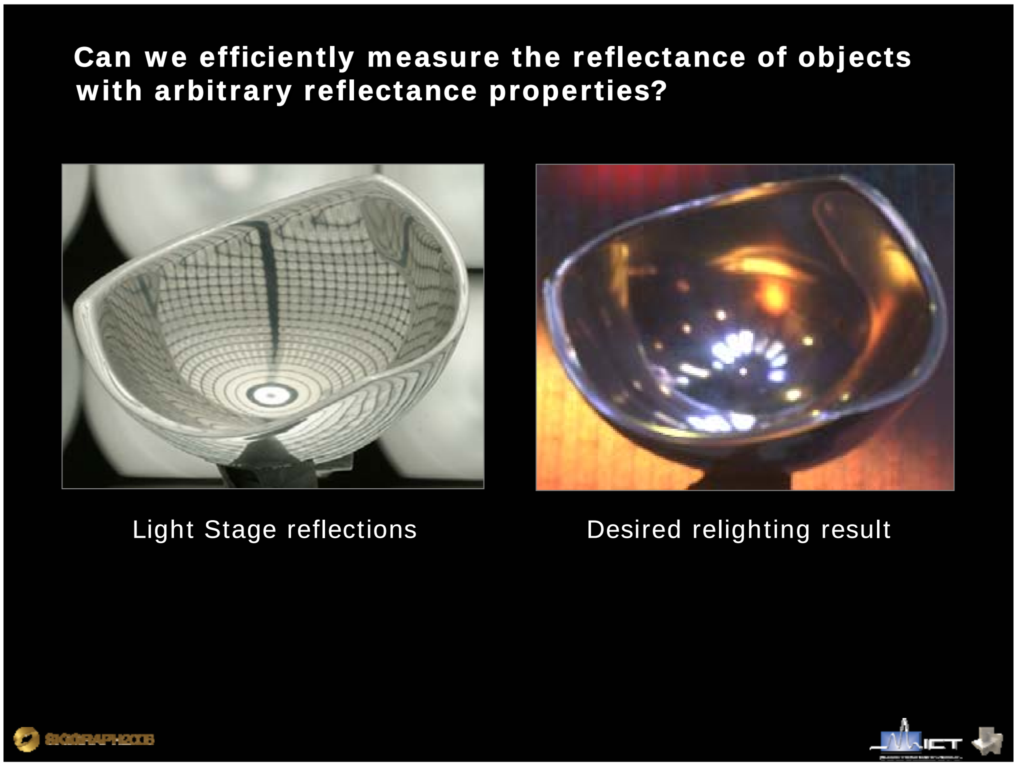
The problem is that when photon noise dominates, there can actually be a multiplexing disadvantage, as bad in theory as doubling the variance of the demultiplexed signals in the worst case.



Another problem is that demultiplexed Hadamard images can have visible noise in shadow regions, since all areas of the images tend to have an equal amount of noise.

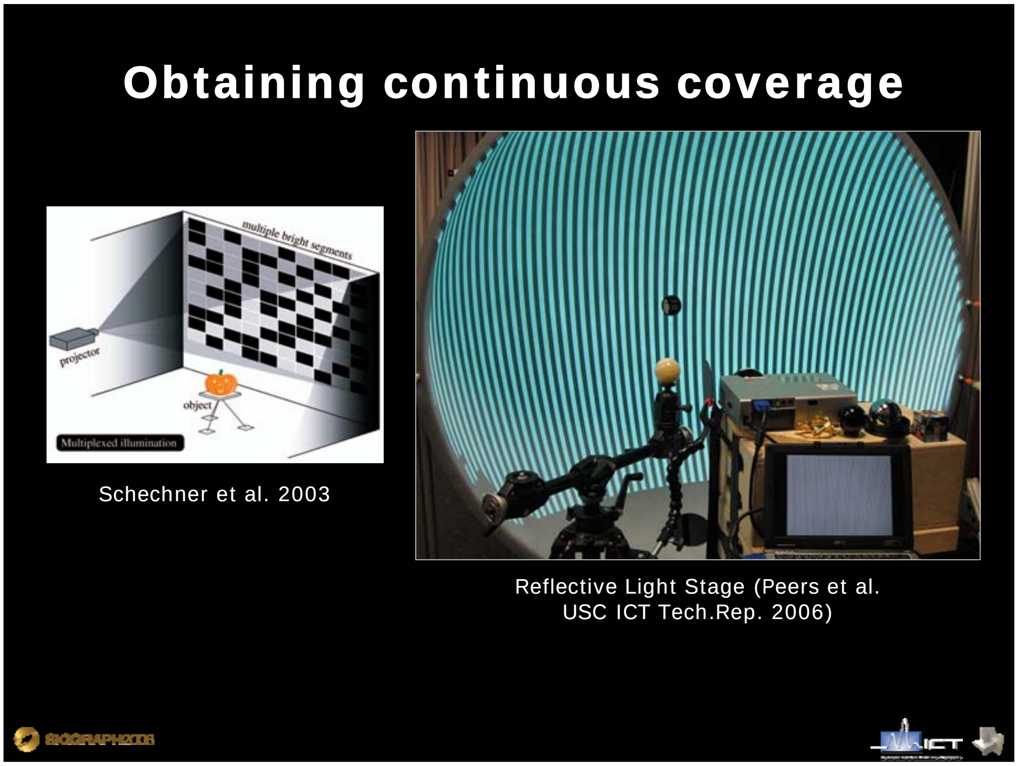


Hadamard patterns are a clear win when single-light images are very underexposed, and quantization or dark current noise dominates. If you are able to expose your images properly, single-lit images may give the most pleasing results due to the photon noise effect shadow noise issue. When we tried Hadamard patterns in Light Stage 5, we actually found that the flashing Hadamard patterns were more comfortable for the subjects than the single-light patterns since the lights blinked well above the rate of perception, bathing the actor in a relatively constant glow. Hadamard patterns also have distinct advantages when a scene includes both diffuse and sharp specular reflections, since the wide-area patterns bring the brightness of the reflections more in line with each other, alleviating problems in capturing the full dynamic range of the scene.

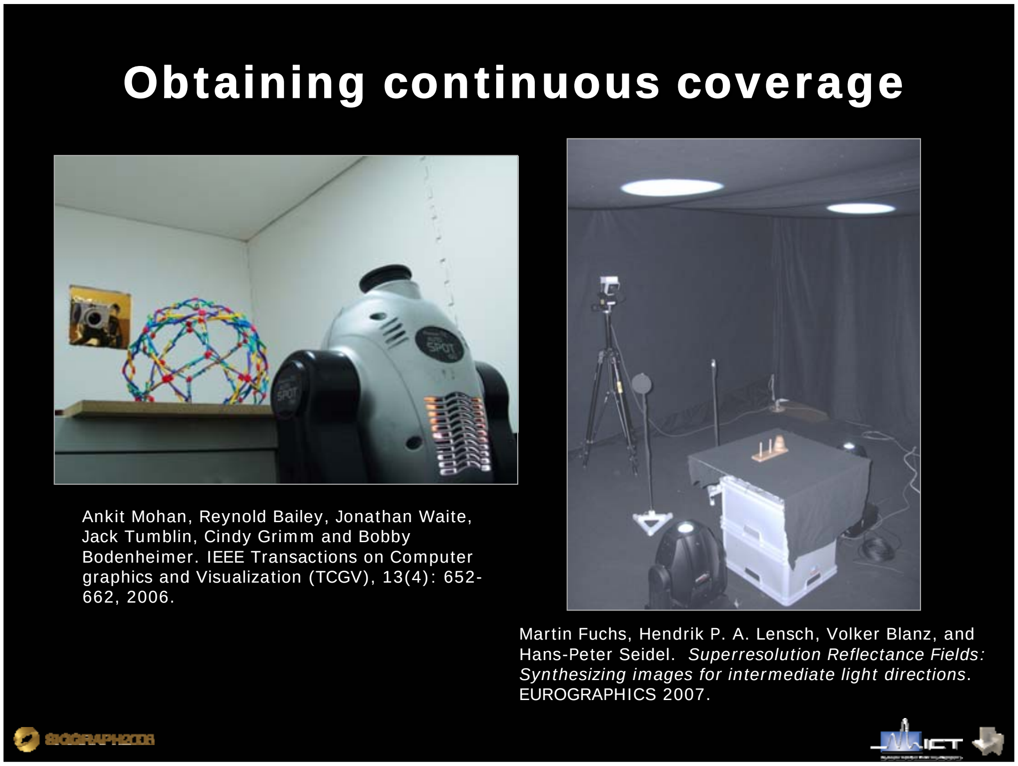


For now, though, let’s think about achieving higher lighting resolution.

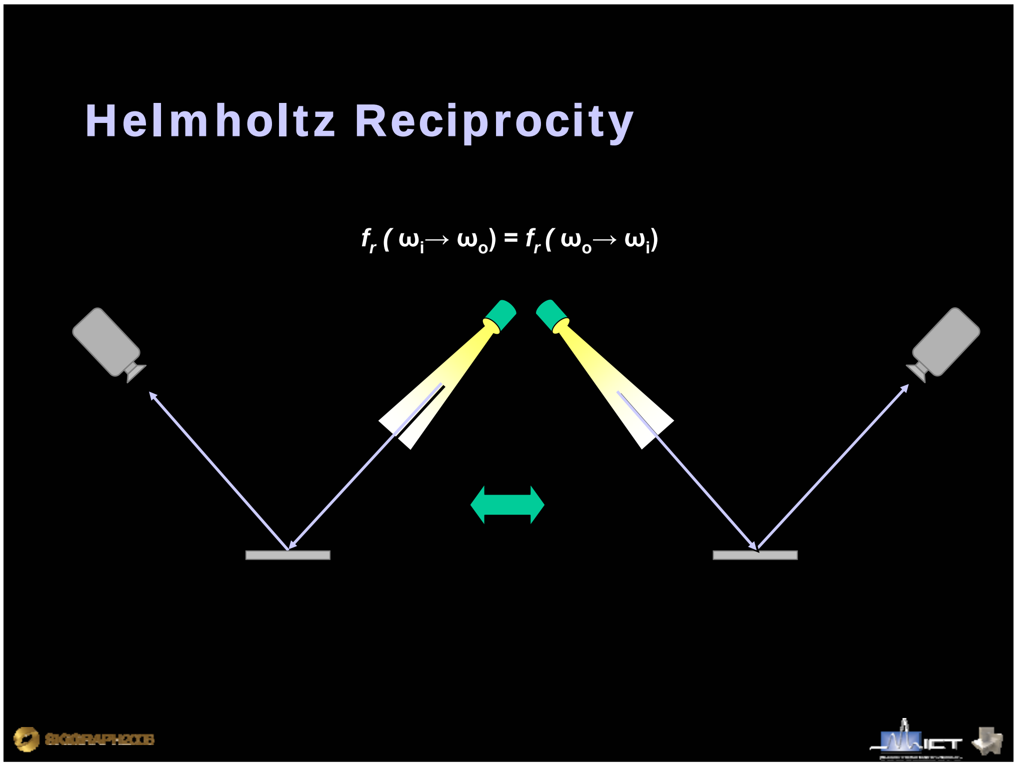
Objects with shiny reflections or translucency can be difficult to capture with light stage techniques, since specular reflections can be very sharp. Obtaining better lighting resolution would be great for these objects.



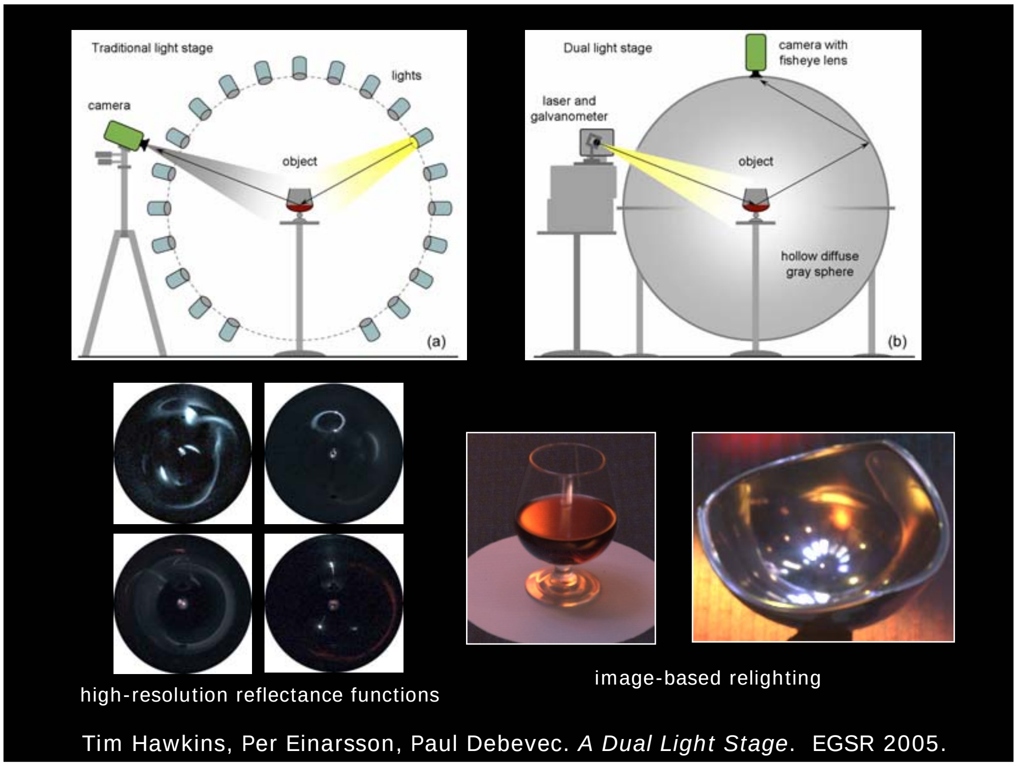
One approach to continuous illumination uses video projectors. We saw how Schechner used a video projector to light up a wall in front of the object. The Reflective Light Stage of Peers et al. lights up a whole hemisphere surrounding an object using an Elumens fisheye video projector. A rough specular painted surface on the inside of the dome increases light efficiency.



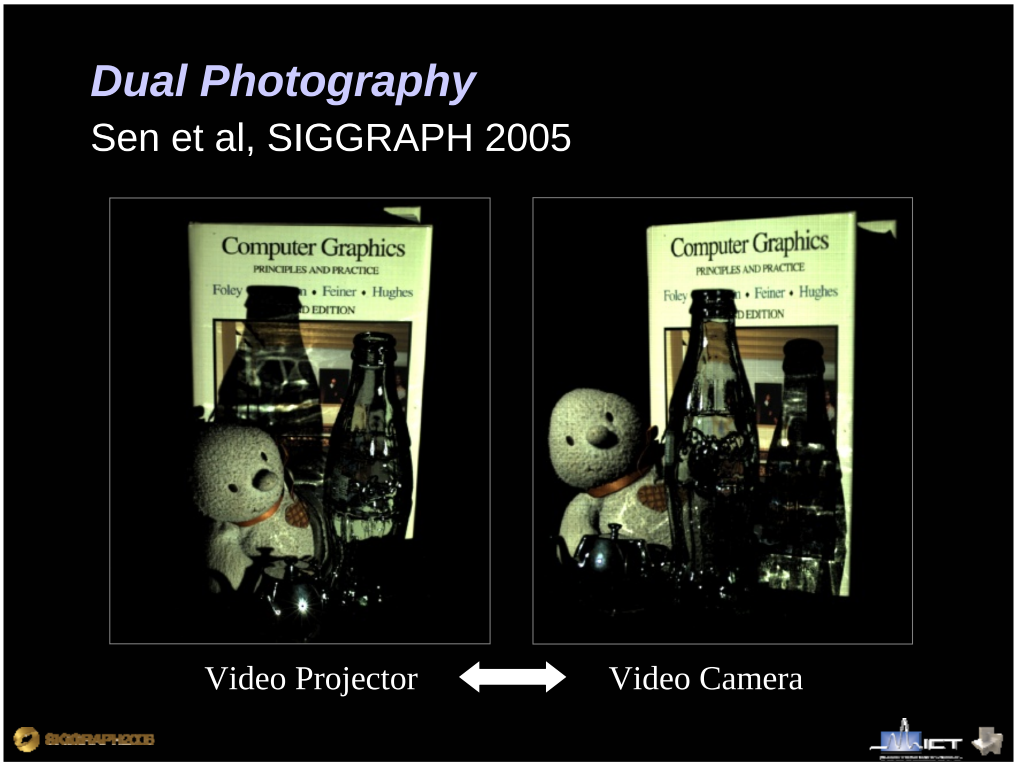
These two projects built light stages out of a single computer-controlled Disco light which projected a spot of illumination onto a projection surface surrounding the object. Mohan et al. At the expense of needed to take longer exposures, Fuchs et al projected onto a room of black felt, which greatly reduced the indirect illumination on the scene. They used the width of the beam to acquire adaptively sampled resolution patterns, and found ways of interpolating between different lighting conditions to given the appearance of super-resolution reflectance fields.



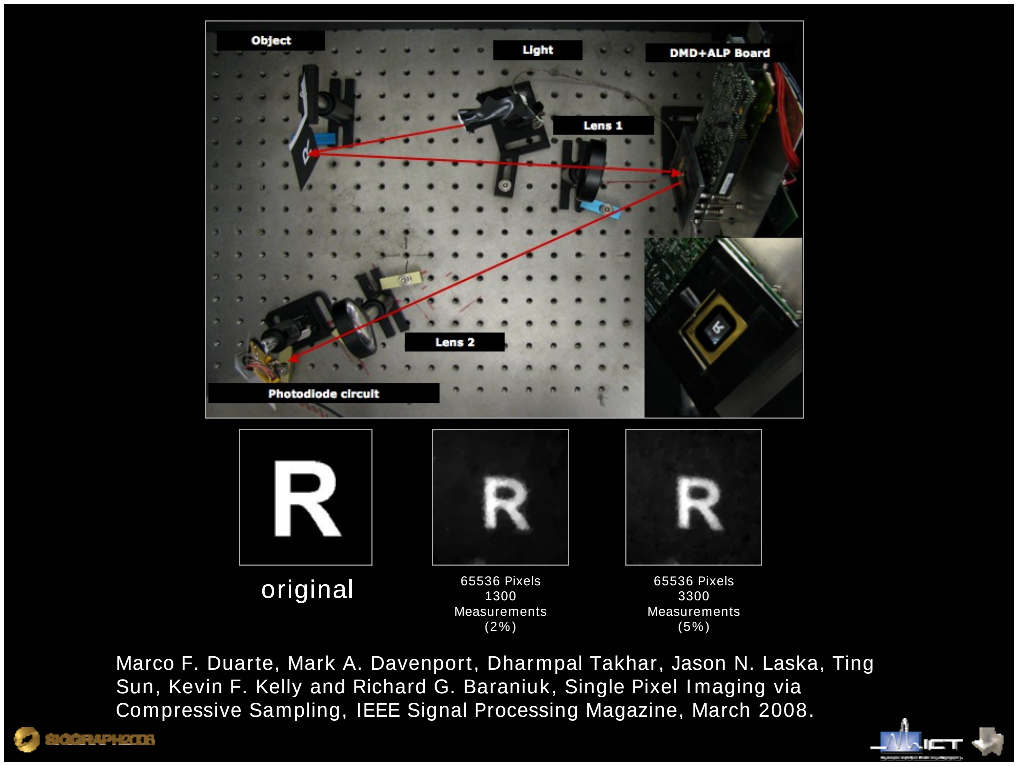
Another technique for achieving high-resolution lighting capture leverages Helmholtz Reciprocity – the condition that light rays are reversible, in that if you switch a sensor and a light emitter in a scene, the same amount of light will still get from one to the other, no matter the complexities of the light path(s).



With the Dual Light Stage, the object is surrounded in a diffusely painted grey sphere. A very bright laser sweeps across the object, and at each point, the laser reflects, refracts, and scatters to form images of each pixel's reflectance function on the inside of the sphere. These complete-sphere images are then recorded by a camera with a fisheye lens at the top of the sphere. The photographed reflectance functions have hundreds of thousands of pixels – enough to see clear reflections in still liquid, and sharp . However, the spatial image resolution is not optimized – the images themselves are relatively low resolution (200x200).



Another project which made use of Helmholz Reciprocity was the Dual Photography project from Stanford. This project showed that an image of an object could be obtained from the position of a video projector just as well as from a camera. Sen et all used a set of adaptive patterns to greatly increase the speed at which the light transport matrix could be measured, and achieved spatially-varying relighting from up to sixteen points of view. However, since the patterns were adaptive, the images could not be taken particularly quickly since processing in between the images was required.



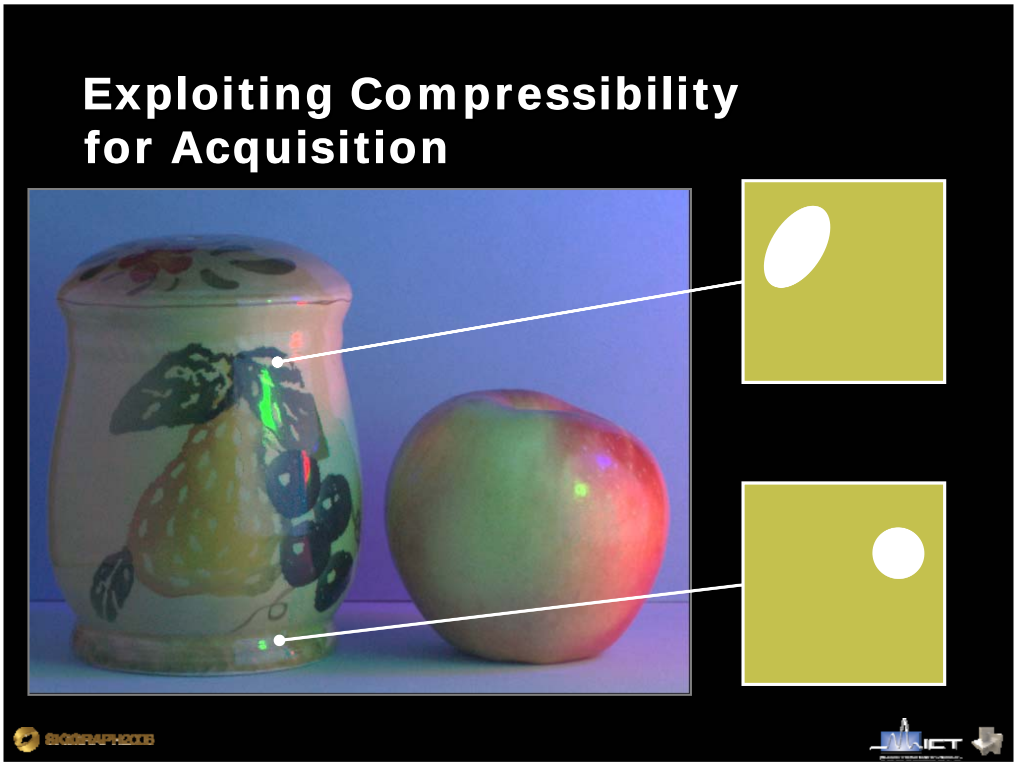
We've heard in the last talk a little about compressive sensing (CS) for novel imaging application. With CS, a compressed full-resolution version of a signal (such as this image of the letter R from Rice University's single-pixel camera project) can be inferred from a much smaller number of *non-adaptive* measurements than from exhaustive capture, as long as it is sparse in some projectable basis. Since CS uses non-adaptive input signals, no online processing is required to obtain the reconstructed results, and the patterns are scene-independent.

Can we apply CS to capturing object reflectance as well s regular images? Of course!

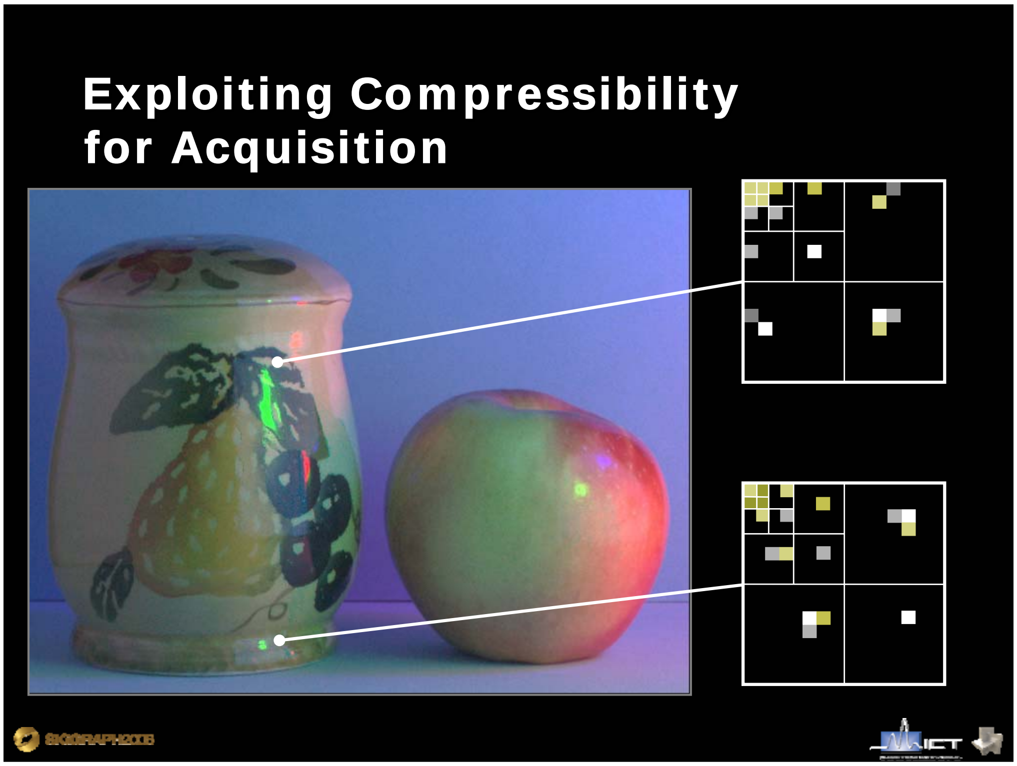
Exploiting Compressibility for Acquisition

The time required to capture a high resolution reflectance field is directly proportional to the number of photographs that need to be acquired. The number of photographs is in turn directly proportional to the lighting resolution. Thus, for high resolution reflectance fields, an impractically large number of photographs need to be recorded.

We will now look at specific properties of reflectance fields, that can help speed up the acquisition process. For this purpose, consider the following scene.

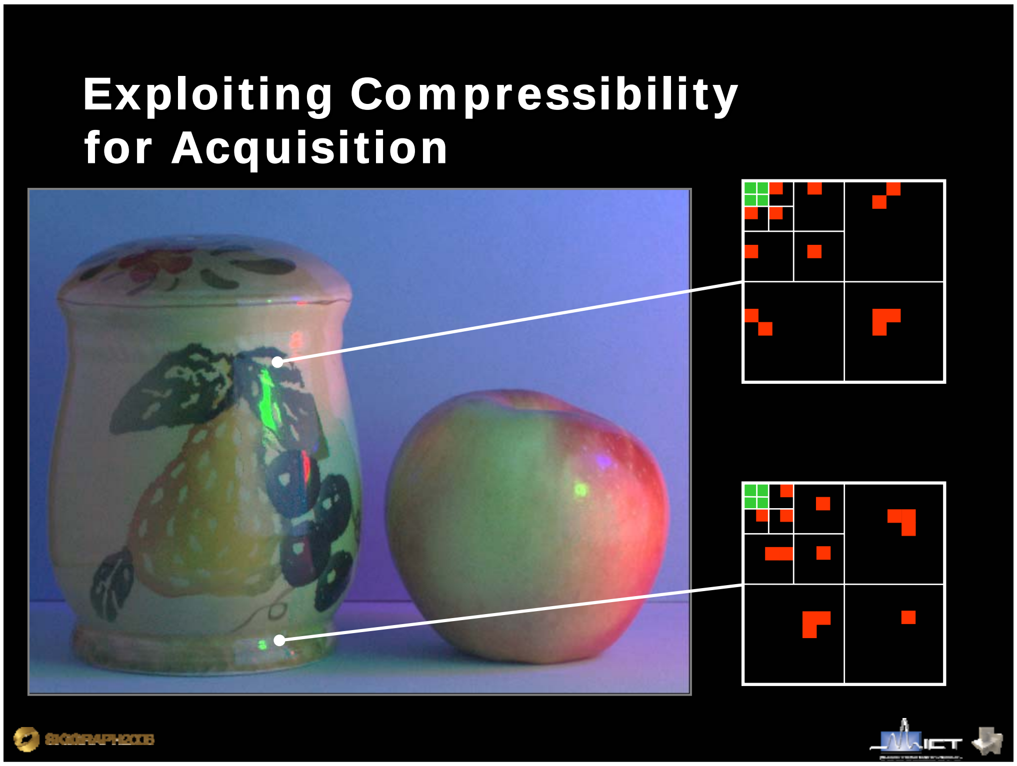


Two randomly selected reflectance functions might look something like this. Both functions are very similar in appearance, and are both relatively simple in content. In order to exploit this apparent simplicity, reflectance functions have often been transformed into a different basis to express this simplicity in a more formal way.



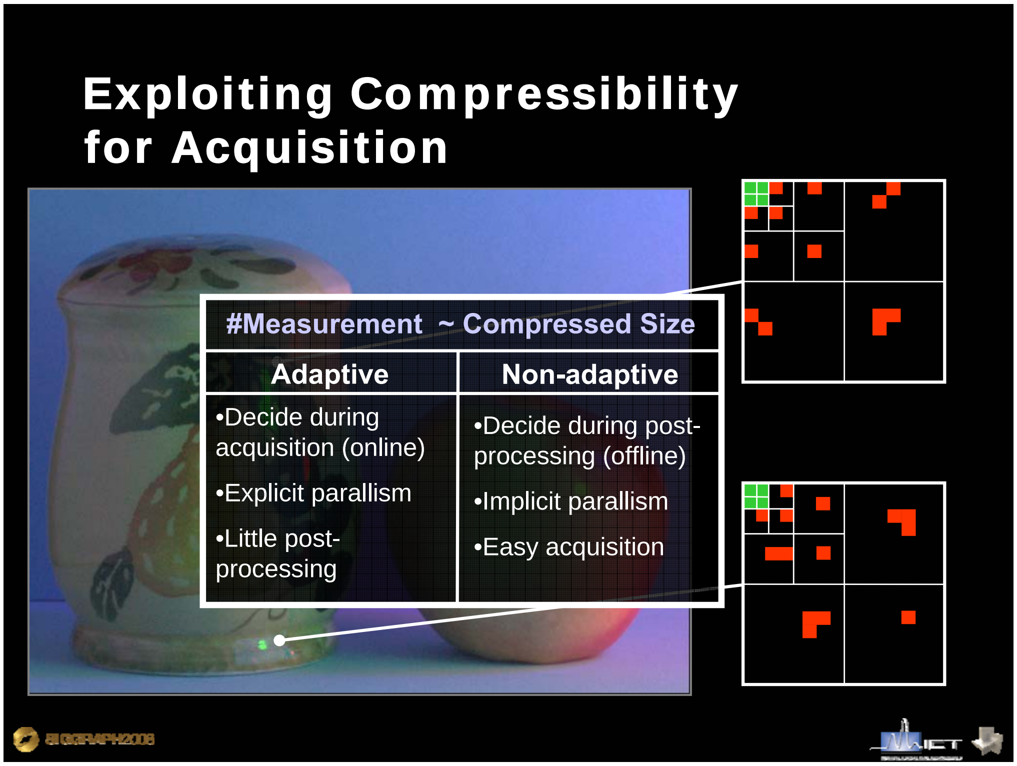
For example, if we convert these functions into a wavelet basis, we could get functions that look like this. First thing to note is that these functions contain many zero or near zero elements. The simplicity in appearance of before is now quantitatively expressed by just a few important (non-zero) coefficient in the new basis. Note, that if we set the near-zero elements to zero, an approximation of the original function is obtained, that is not exactly the same, but very similar. This method, is for example used in compressing images.

So what does it mean to have just a few non-zero coefficient of reflectance function in a specific basis. Well, this means that we only need to measure these coefficients to obtain a good approximation of the reflectance functions. As we have seen before, measuring the response of a specific coefficient of a basis is equal to emitting that basis function onto the scene and observing its response. Thus, by only emitting the basis functions that correspond to non-zero coefficient, we can potentially measure a reflectance field much faster.



However, there is s light problem: we are measuring not just a single reflectance function, but a whole reflectance field, which is a collection of many reflectance functions. Each reflectance function might have just a few non-zero coefficients, which is good, but also that the set of non-zero coefficients for each function is different. So in the worst case, we still have to measure all coefficients, and thus obtain no speed up. For example, in our example here, the green coefficients are shared, and thus emitting the corresponding basis functions yields information gain for both pixels. However, the red marked coefficients are not shared. So when you measure on of these red coefficients, you will only measure additional information for a single pixel, and not gain any information for the other.

From this it is clear that in order to have a fast acquisition, we need to somehow find a way to maximize the information gain for all pixels. The solution is to emit multiple basis functions at the same time, trying to make sure that we know for each functions which basis functions has an effect and which ones don't.



There a two possible solutions. The first one, tries to explicitly coordinate the parallelism of measuring coefficient during the acquisition phase. This are called an adaptive methods.

An adaptive method, uses the information it has of the reflectance field, to schedule new measurements such that information gain is maximized. This requires some processing during acquisition, but usually no after acquisition. Also, the illumination patterns will differ when measuring different scenes.

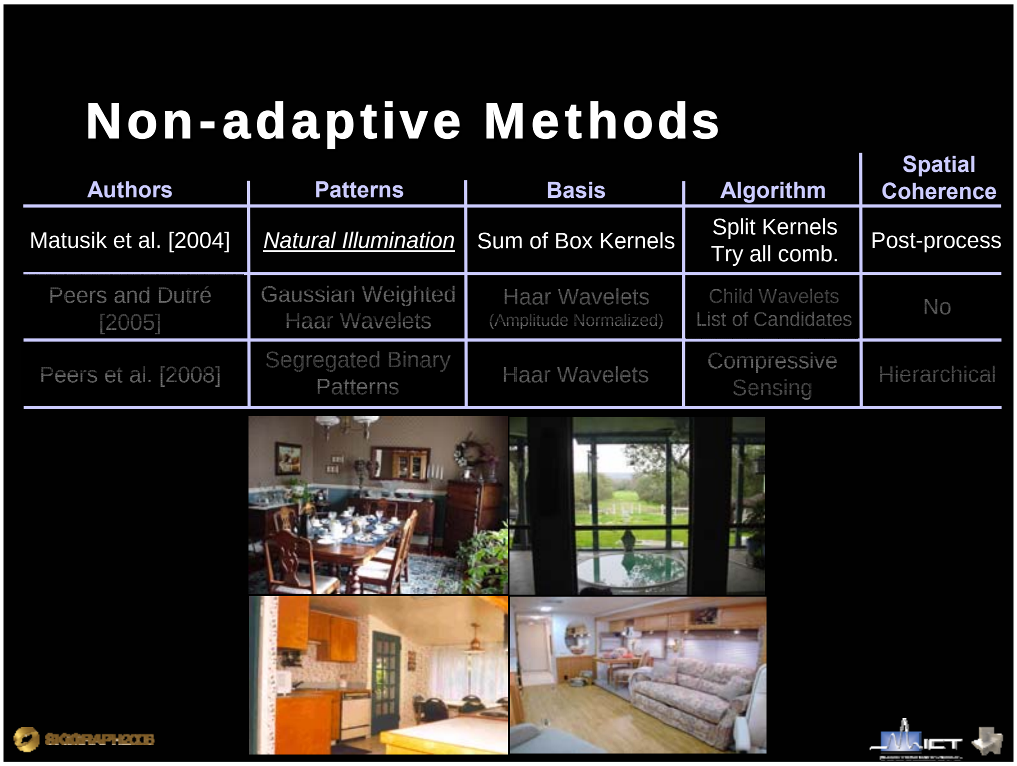
The second method is called non-adaptive. This method always uses the same illumination patterns, which are specially designed to maximize the information gain per measurement. In order words, these measure responses for every basis function (but with different weights) in each measurement. During post-processing, each reflectance functions needs to be inferred in an adaptive fashion from the measurement. In other words, there is an implicit parallelism during acquisition. Non-adaptive methods move the complexity of the system from the measurements to the post-processing.

Both methods have their advantages and disadvantages, and are capable of measuring a reflectance fields in a number of measurements that is proportional to their compressed size. This is to some degree independent of the illumination resolution.

Non-adaptive Methods

Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	Gaussian Weighted Haar Wavelets	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical

We will now discuss three non-adaptive methods briefly. The first method was presented by Matusik et al. In 2004.



This method uses natural illumination as illumination patterns (photographs). These photographs are emitted from a CRT monitor onto the scene. They represent the reflectance field as a sum a box kernel functions which different sizes and positions. During post-processing they split the kernel in each reflectance function that explains the observed responses under the *known* natural illumination best. Furthermore, to improve the results, a spatial correction is performed after post-processing to ensure that neighboring pixels have similar reflectance functions.

Non-adaptive Methods

Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	Gaussian Weighted Haar Wavelets	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical


Relit (24 subdiv.)


Reference Photograph

Here you can see a result of their method. On the right you see a reference photograph of the scene, and on the left the same scene relit using the same illumination and using the computed reflectance field. Each reflectance function is the result of 24 subdivisions.

Non-adaptive Methods

Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	Gaussian Weighted Haar Wavelets	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical

Relit (24 subdiv.)

Reference Photograph

Here you see a different scene. Again, 24 subdivisions were performed to obtain these results. The number of measurements was approximately 200.

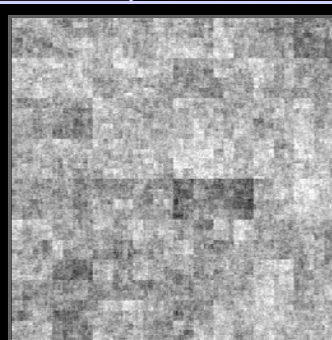
Non-adaptive Methods

Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	Gaussian Weighted Haar Wavelets	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical

In 2005, Peers and Dutré presented a system that uses Wavelet noise as illumination patterns.

Non-adaptive Methods

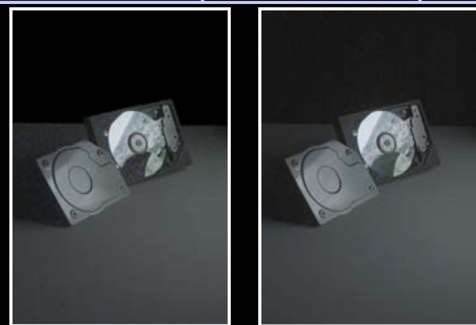
Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	*Gaussian Weighted Haar Wavelets*	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical



Here you can see such a wavelet noise illumination pattern. This allowed them to use a Haar wavelet basis instead of the weighted sum of box kernel functions of Matusik et al. A Haar wavelet basis has the advantage that it is a real basis, and has been study really well in image-compression, and thus a lot of mathematical properties are known. Their algorithm builds a list of candidate coefficients, and at each step adds the best candidate. When adding a coefficient to the solution, it children's coefficients are estimated are added to the list of candidates. No spatial coherence is enforced in their algorithm.

Non-adaptive Methods

Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	Gaussian Weighted Haar Wavelets	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical



Relit (64 coeff.)　Reference Photograph

Here you see a scene lit by a photograph of a stone bridge. On the right a reference photograph, and on the left a relit image. Each reflectance function was reconstructed from 250 photographs, and contained 64 Haar wavelet coefficients.

Non-adaptive Methods

Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	Gaussian Weighted Haar Wavelets	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical


Relit (128 coeff.)

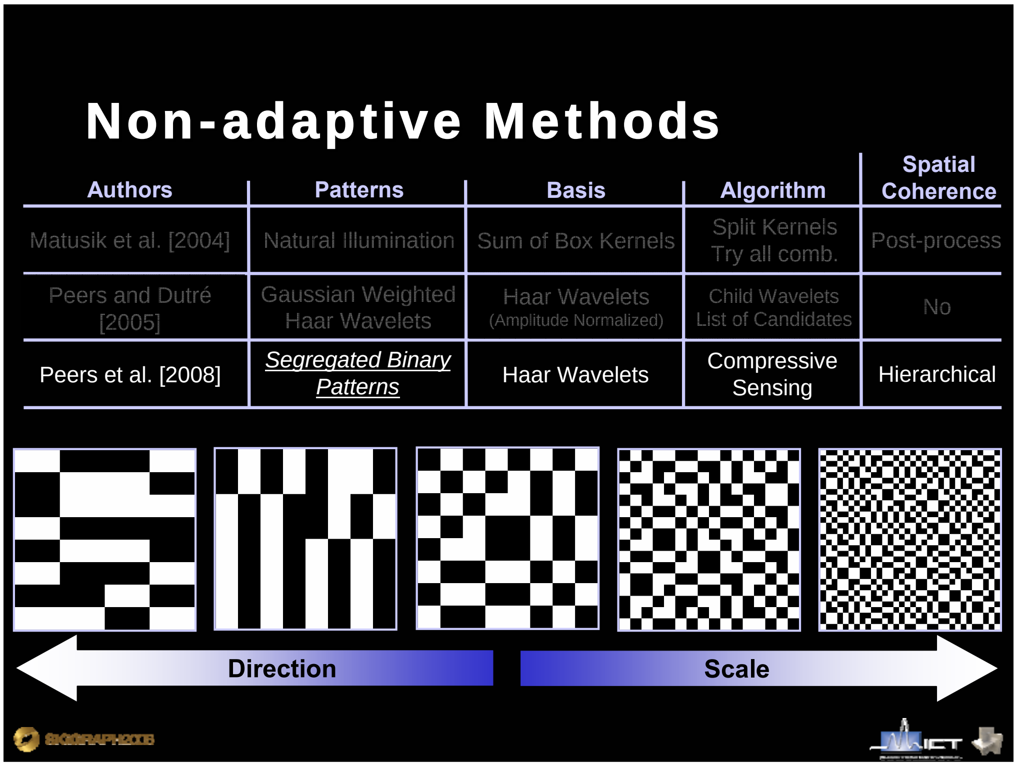

Reference Photograph

Here a different scene is shown. This time inferred from 500 measurements, and each reflectance function is reconstructed using 128 Haar wavelet coefficients.

Non-adaptive Methods

Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	Gaussian Weighted Haar Wavelets	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical

This year, Peers et al., presented a technique that utilizes the mathematical theory of compressive sensing (as explained before). However, directly applying compressive sensing does not work well. For this they made a number of enhancements that make it better suited for measuring reflectance fields.



First they use newly designed measurement patterns called segregated binary patterns. These patterns are binary (black and white), and are segregated in two ways. First there is scale. As you can see on the three patterns on the right. All three look somewhat similar, but have different scales. Second, each scale is further segregated into three groups, shown on the three patterns on the left. As you can see, the first pattern exhibits more horizontal structures, the middle one more vertical, while the third one does not prefer any direction.

By capturing a balanced mix of these patterns, segregated in scale and direction, compressive sensing can be used to reconstruct reflectance fields. Additionally, they exploit the spatial coherence of the reflectance field by using a hierarchical algorithm to infer the reflectance functions.

Non-adaptive Methods

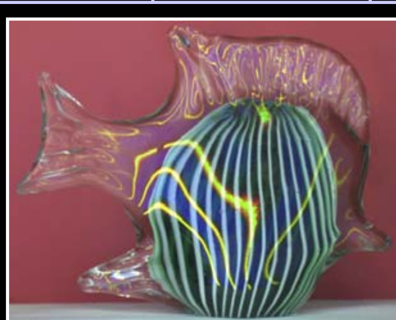
Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	Gaussian Weighted Haar Wavelets	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical

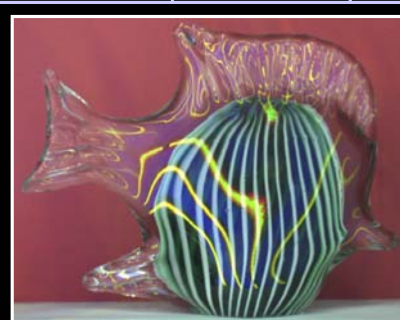
Relit (128 coeff.) Reference Photograph

Here you can see a scene, illuminated by a natural illumination condition (emitted from a CRT monitor on the right). On the right a reference photograph, on the left a relit image. The reflectance field contains 128 Haar wavelet coefficients per reflectance function. Approximately 1000 measurement were performed to obtain this result.

Non-adaptive Methods

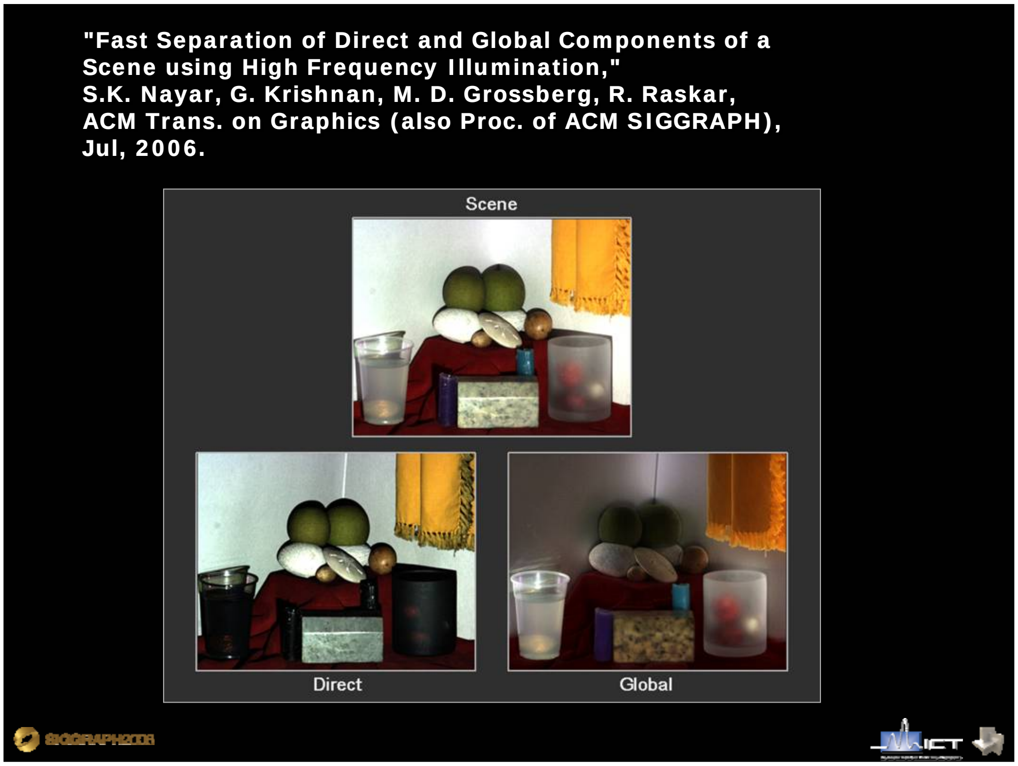
Authors	Patterns	Basis	Algorithm	Spatial Coherence
Matusik et al. [2004]	Natural Illumination	Sum of Box Kernels	Split Kernels Try all comb.	Post-process
Peers and Dutré [2005]	Gaussian Weighted Haar Wavelets	Haar Wavelets (Amplitude Normalized)	Child Wavelets List of Candidates	No
Peers et al. [2008]	Segregated Binary Patterns	Haar Wavelets	Compressive Sensing	Hierarchical


Relit (128 coeff.)

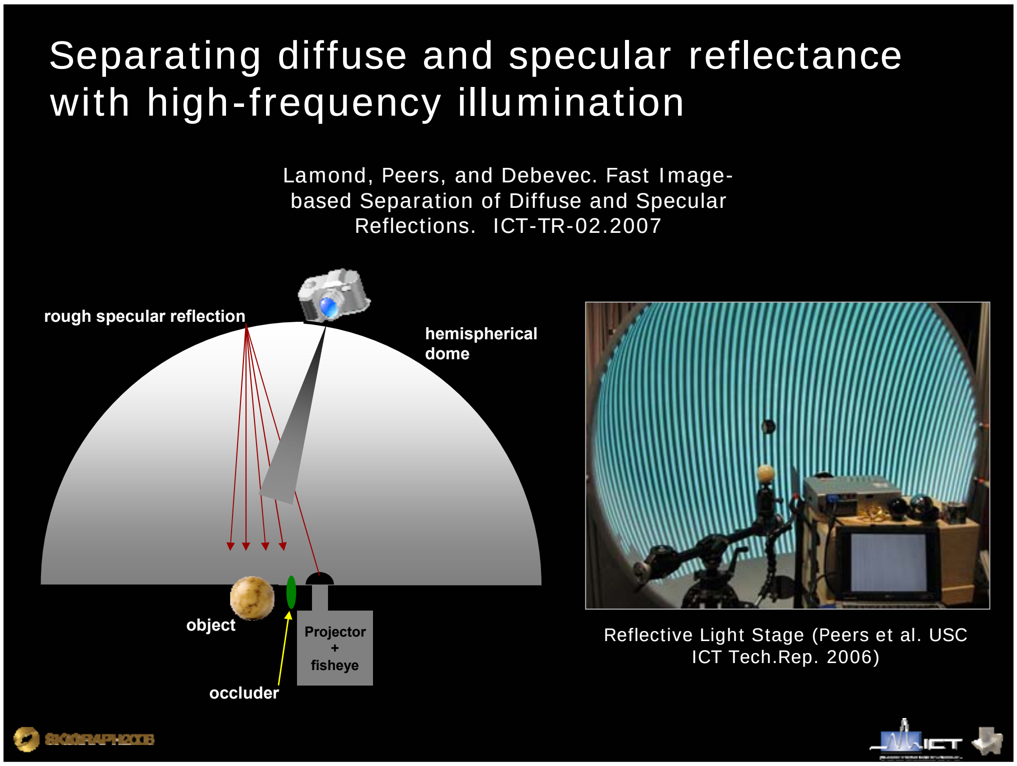

Reference Photograph

Here is a final example. In this case the illumination is emitted from the hemispherical Reflective Light Stage shown earlier. Again, 1000 measurements were performed, and 128 coefficients reconstructed per reflectance functions.

As we will see, it can be easier to capture and analyze reflectance signals if it is possible to separate various reflectance behaviors in the reflectance functions, such as diffuse and specular reflections.

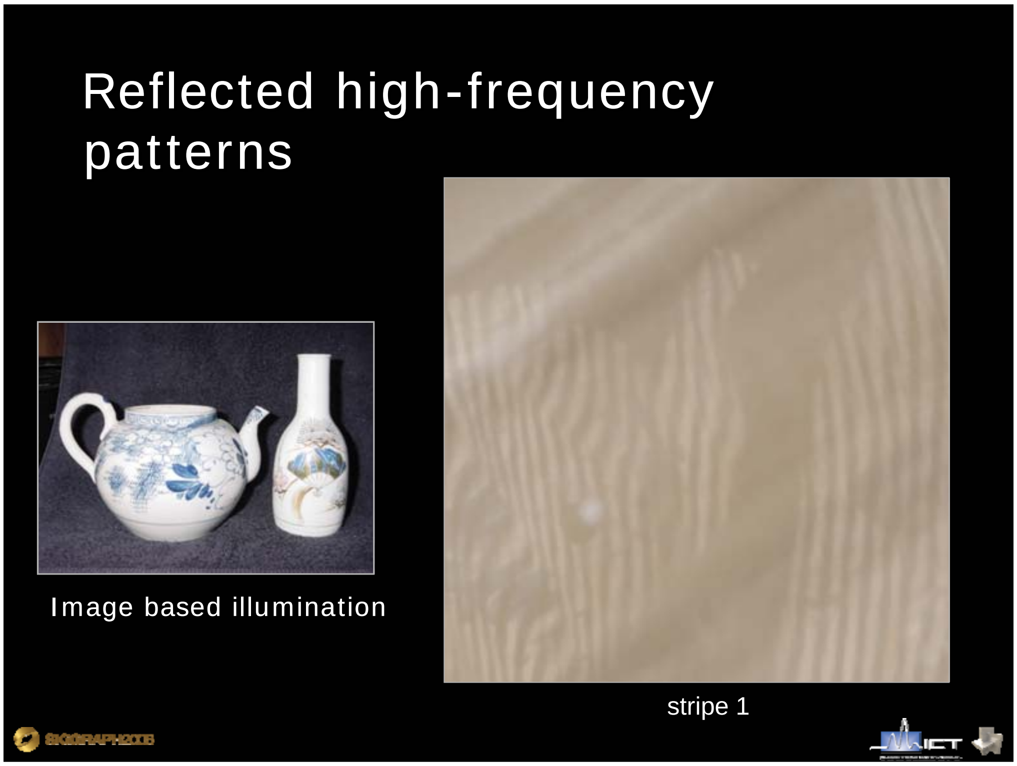


Nayar et al. used high-frequency illumination patterns to quickly separate “direct” and “global” components. Basically, the global components stay the same as you phase-shift high-frequency illumination on the scene, while the direct components appear and disappear at different pixels. Taking the minimum value over a sequence of phase shifts yields the global component, multiplied by the fill ratio of the patterns; the maximum minus the minimum yields the direct component.

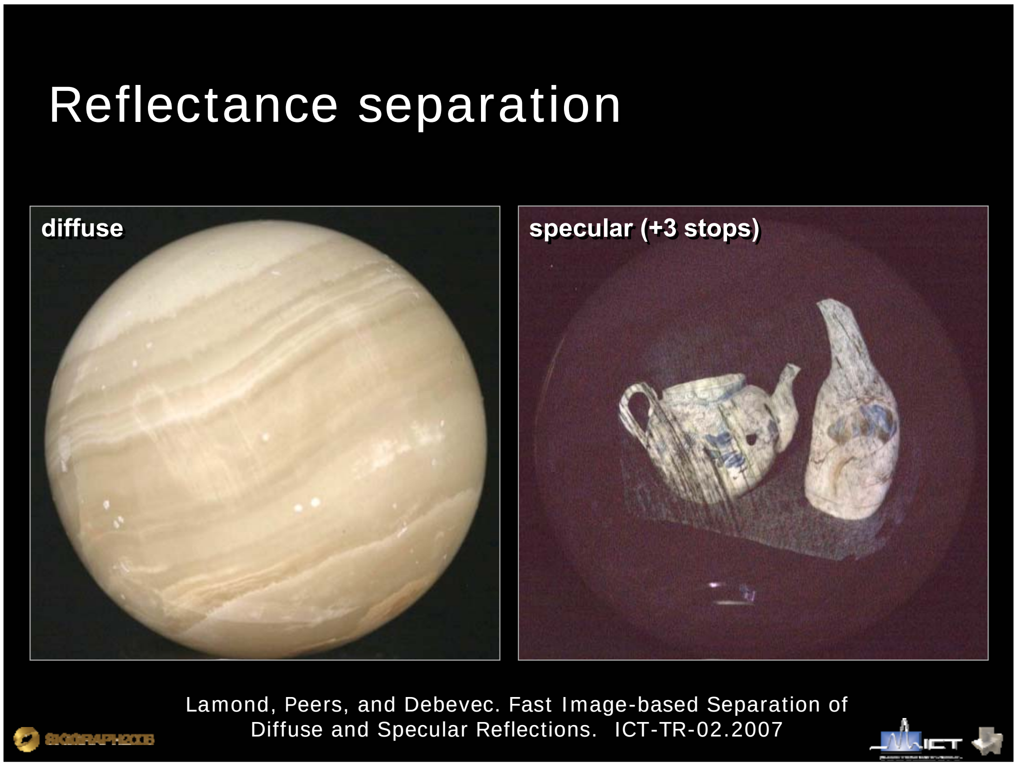


This project uses a similar technique to Nayar et al 2006 to separate diffuse and specular reflections, rather than direct and global components. It makes use of the Reflective Light Stage.

Here is the object whose components we will separate – a marble ball, with both specular and diffuse reflections.

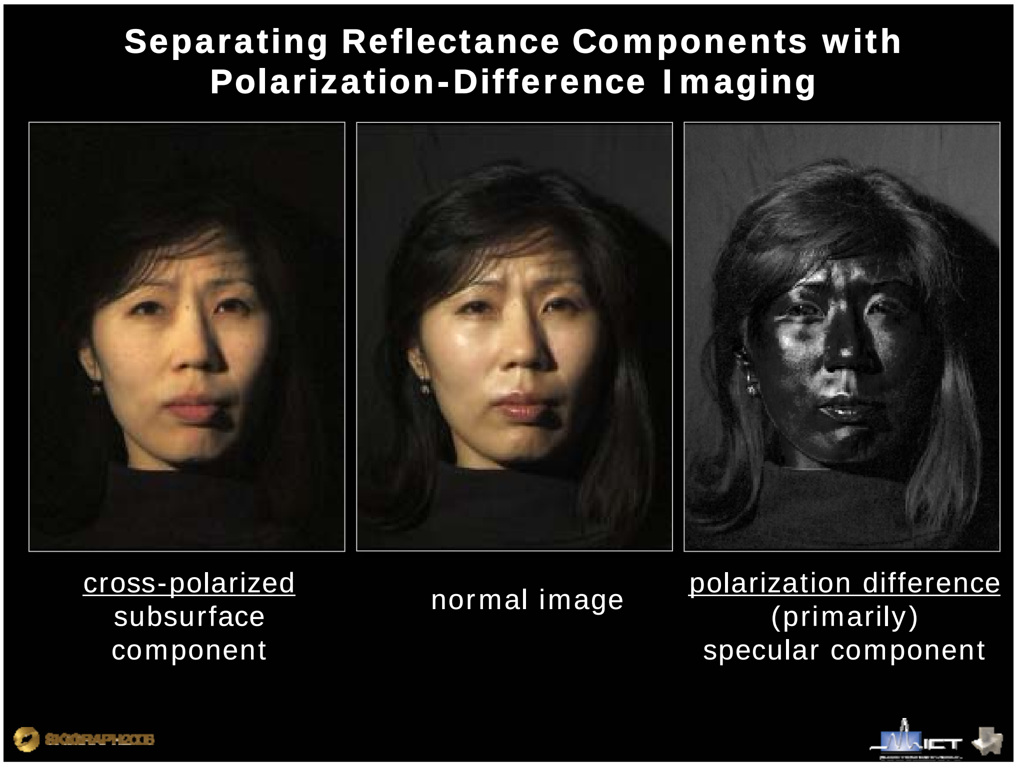


The image on the left is an image we will consider projecting into the dome and observing its reflections back from the object. We won't project the pattern itself, since that would yield both diffuse and specular reflections mixed together. Instead, we will modulate the patter by four phase-shifted high-frequency illumination patterns and project those instead. A detail of the reflection of one of these patterns is shown at the right.

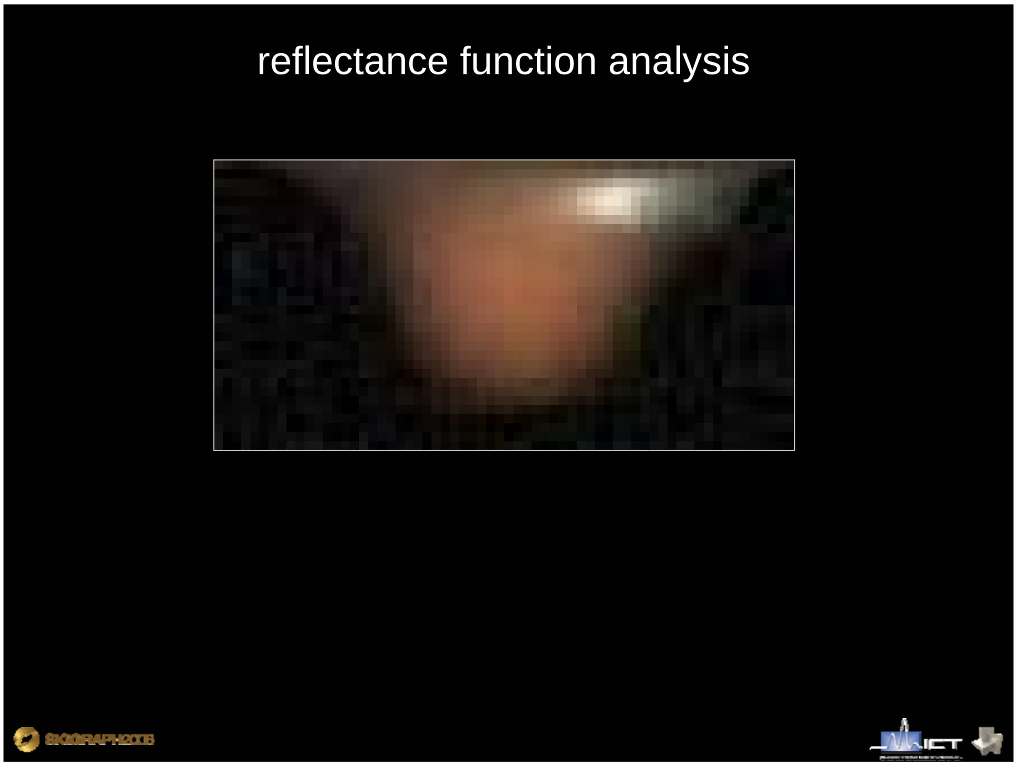


Taking the mimimum pixel value over the four-pattern sequence yields the left image showing the diffuse reflectance of the subject under the image-based illumination environment.

Finding the difference between the maximum and minimum values for each pixel over the sequence yields an estimate of the specular component of the object's reflectance, shown at the right. For this object, this component reveals a clear image of the image-based illumination environment.

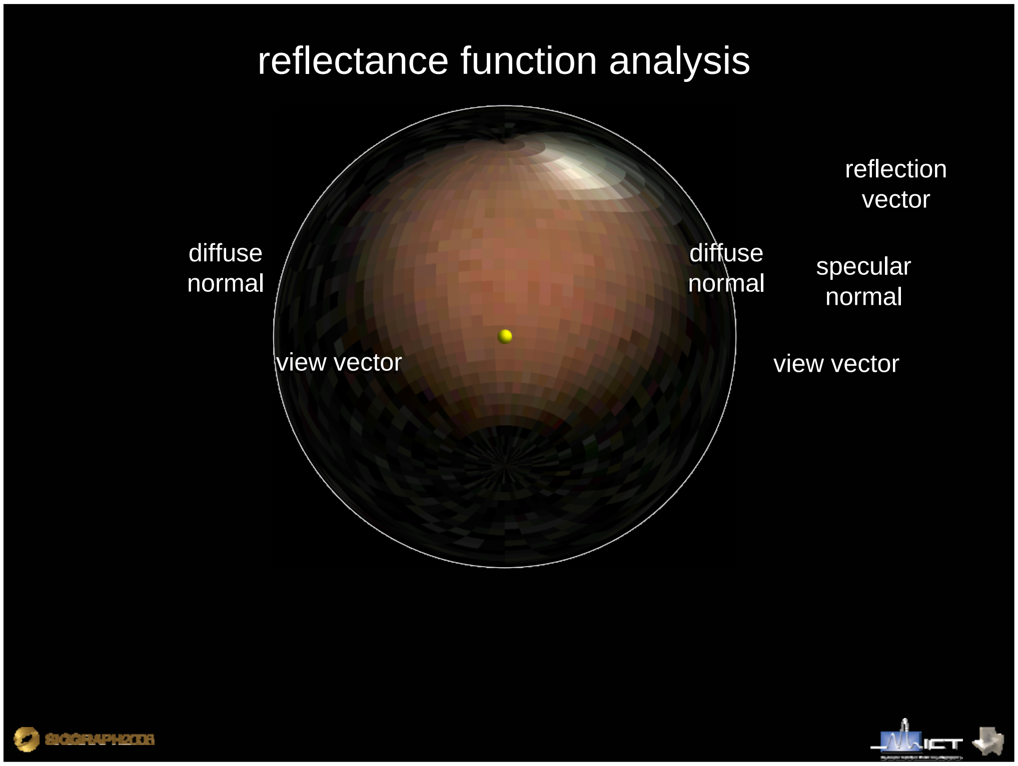


Another useful way to separate diffuse (or subsurface) and specular reflections is through polarization. Placing perpendicular polarizers on a light source and the camera removes specular reflections, which maintain polarization. The subsurface and diffuse components remain since they depolarize the light. Turning one of the polarizers so that it becomes parallel to the other tunes the specular reflections back in. The difference between two such images shows only the light which maintains polarization, which is primarily the specular component of the illumination.

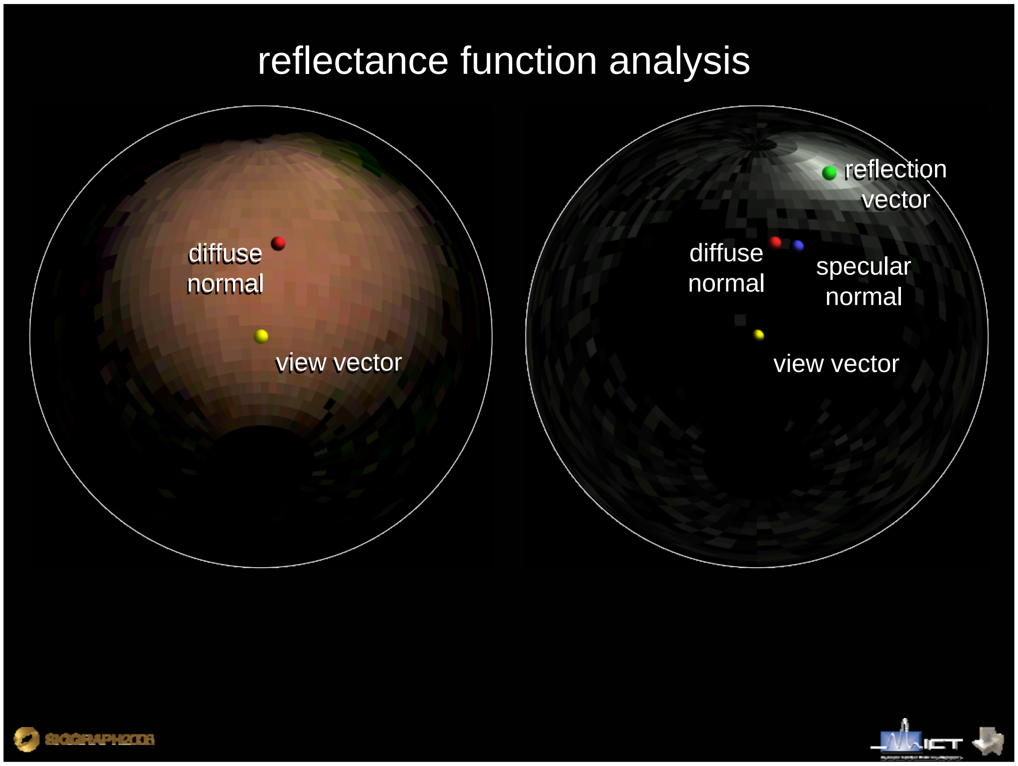


This polarization separation process can be combined with a *gradient illumination* technique to obtain useful estimates of facial surface normals.

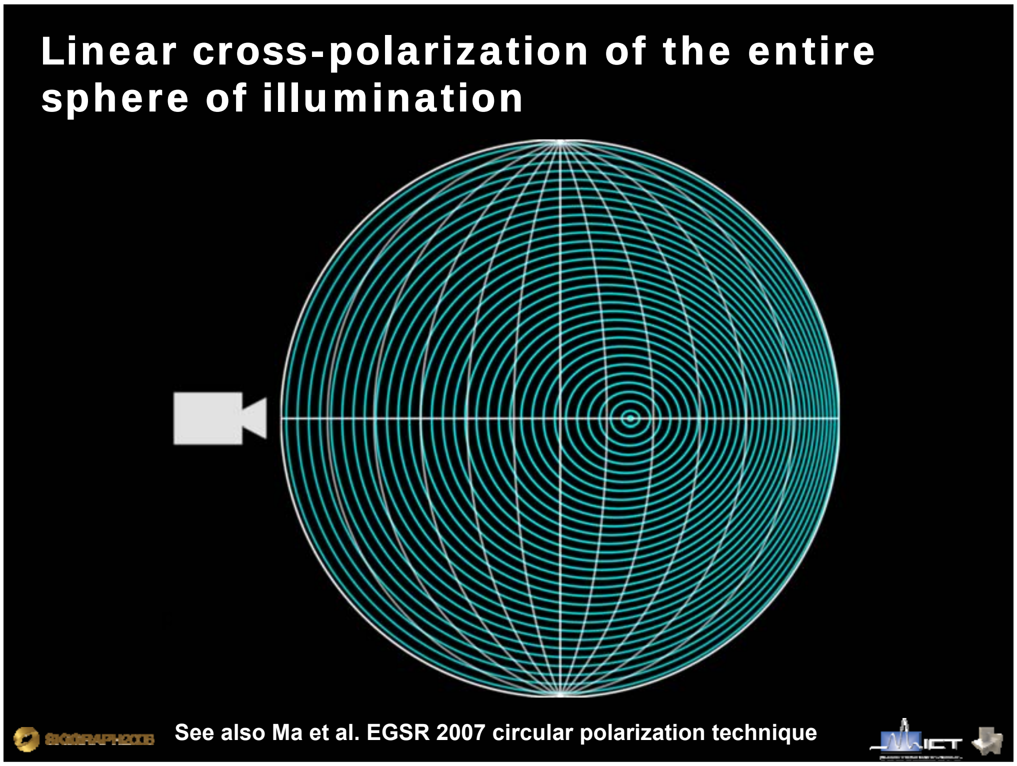
Let’s return to looking at a canonical reflectance function, with a diffuse and a specular component.



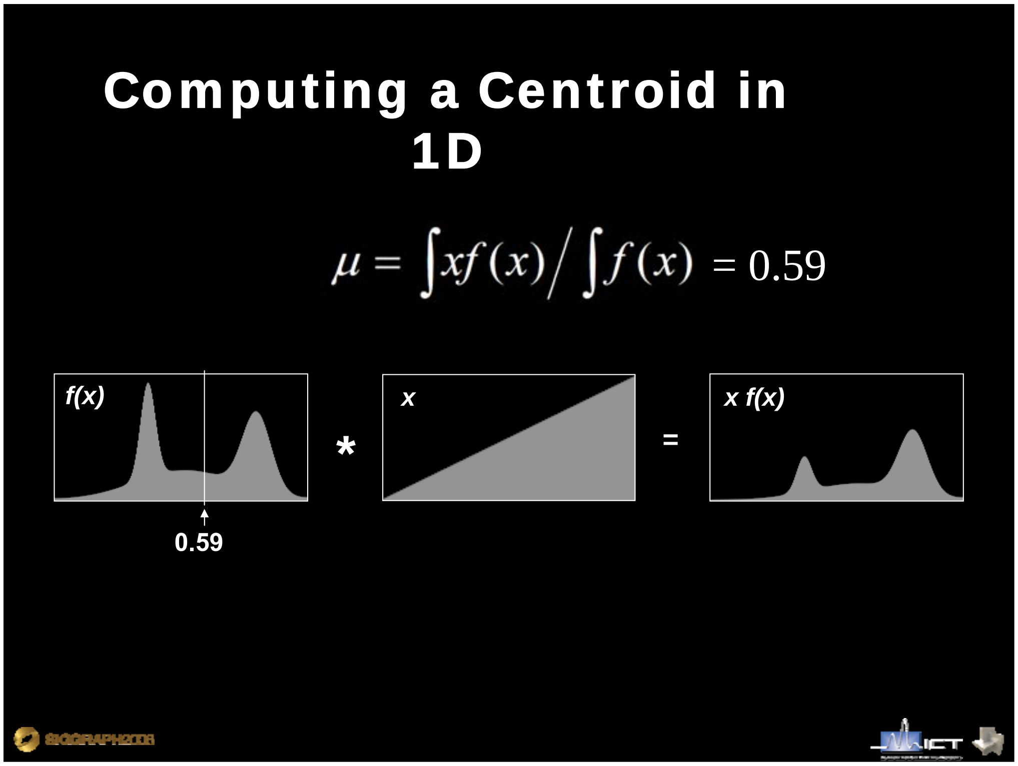
If we wanted to develop a shorthand for the reflectance function, we could say a lot about it if we knew the color and centroid of the diffuse lobe, and the intensity and centroid of the specular lobe.



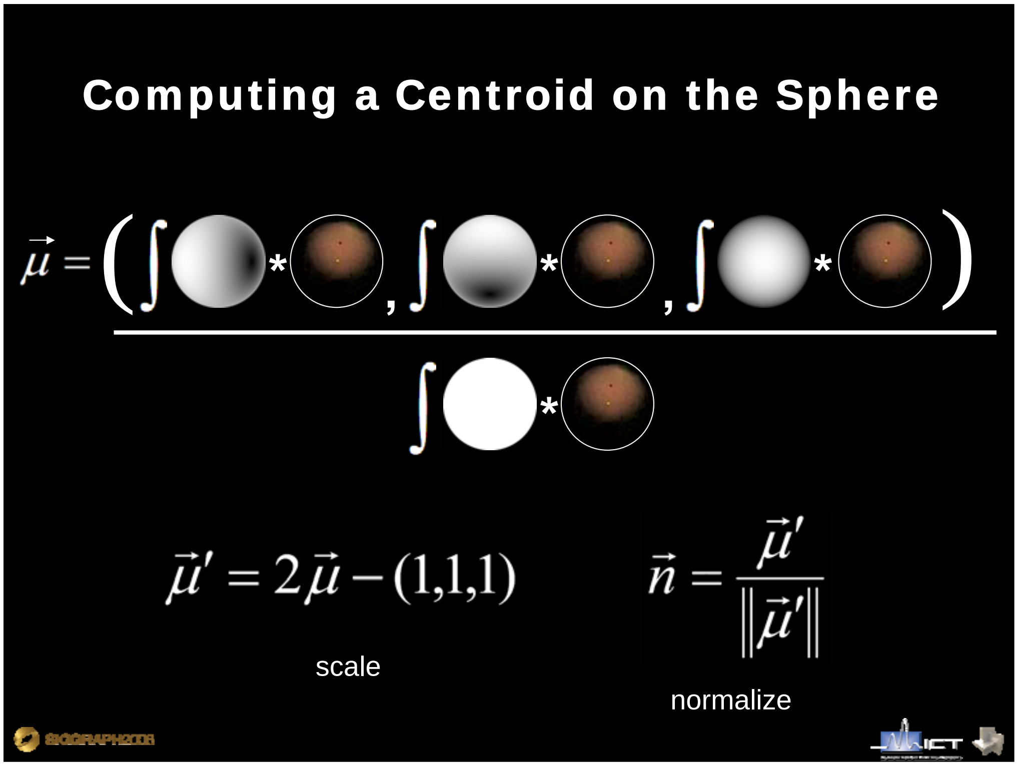
If there were a way to separate the reflectance function into its diffuse and specular components, and then have each function available, this is relatively straightforward. Debevec et al 2000 used a colorspace separation technique to separate the reflectance function seen above. The surface normal as determined from the diffuse reflection is essentially equivalent to the centroid of the diffuse normal. The centroid of the specular reflection actually is an estimate of the reflection vector; the corresponding surface normal the lies halfway back to the view vector. Note that this does not necessarily yield the same normal estimate.



This pattern of linear polarizers from [Ma et al 2007] allows the entire sphere of illumination to be cross-polarized from a particular camera position at the front of the stage. Ignoring a few complications regarding the Fresnel equations and in particular the Brewster angle, this allows spherical reflectance functions to be separated into diffuse and specular components through polarization difference imaging.



Let’s now build a process for measuring the centroid of reflectance functions (based on their diffuse or specular component) using a small number of incident illumination patterns. First let’s simplify the problem to 1D. We can compute the centroid of a 1D function by integrating it against a linear function, or gradient, as seen in this slide. We’ll see in a second how we can perform the same process for a spherical reflectance function to estimate a surface normal or a reflection vector.



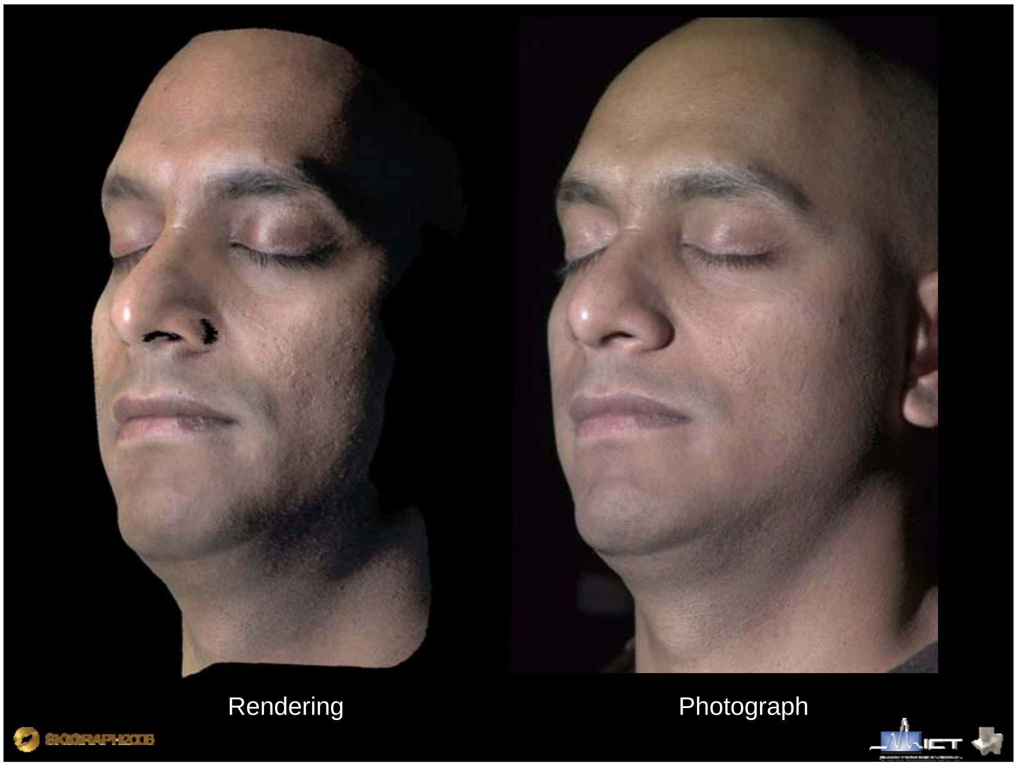
To find the centroid of a spherical reflectance function, such as the diffuse one shown above, we can integrate it against gradients of spherical illumination across the three coordinate axes. As before, we also integrate the total energy of the function in order to normalize the coordinate estimates.

Instead of measuring these functions exhaustively, we note that what we really need to compute the surface normal is the dot product of the reflectance function with each gradient illumination condition. These we can actually measure directly by lighting the subject with the gradient illumination conditions directly, instead of measuring the reflectance functions first. In this way, we will use the illumination to compute the surface normals in the scene!

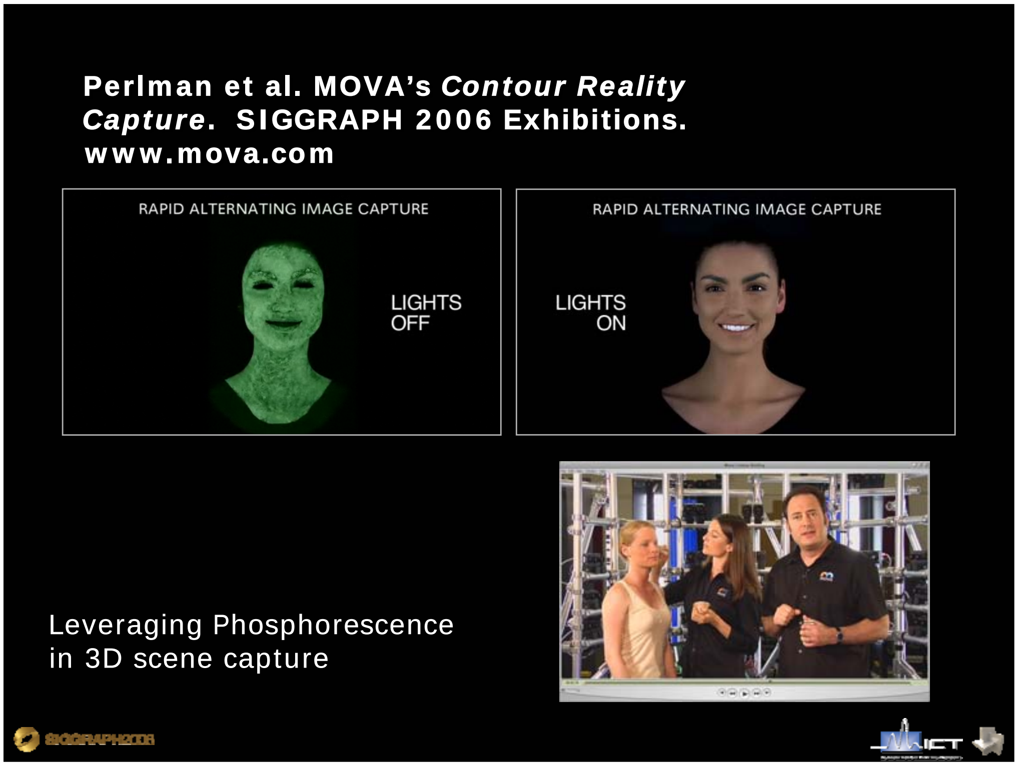
Four images of the scene under the gradient illumination patterns yield estimates of the surface normals from either the diffuse or specular components of the subject's reflectance.

Gradient Illumination techniques are efficient enough that surface normal maps for a face can be recorded in real time with relatively standard camera equipment. Surface normal maps can be used for a variety of normal map relighting tricks.

The specular normal map in particular provides a great deal of high-resolution geometric information about the face, since this measures the reflection of light from . This high-resolution facial scan was created from a lower-resolution structured light scan (accurate to a millimeter or two) increased in resolution (to perhaps 0.1mm) using the specular normal map, as in Ma et al. 2007.

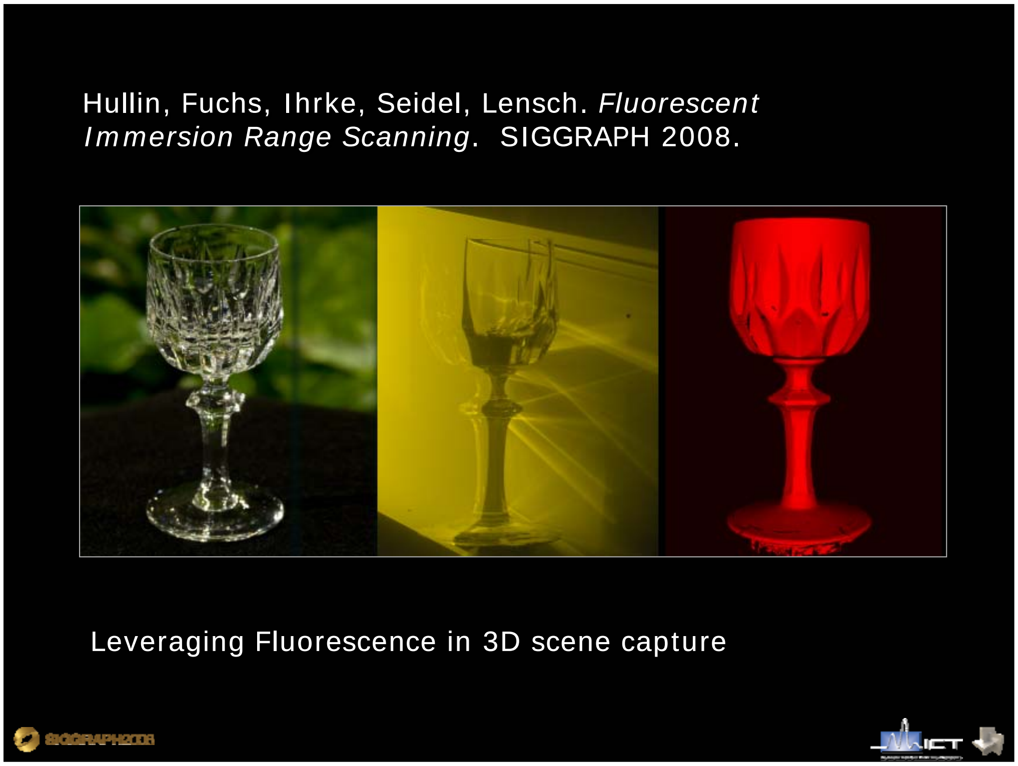


The specular and diffuse surface normals can also be used for hybrid normal map rendering. This rendering of a face on the left, seen from a novel viewpoint and illumination condition, uses surface normals computed from the face's diffuse component to shade the diffuse reflection, and normals computed from the face's specular component to shade the specular reflection. The real-time rendering closely resembled the validation photograph to its right. The skin-like appearance results from differences in the surface normals due to the subsurface scattering properties of the skin.

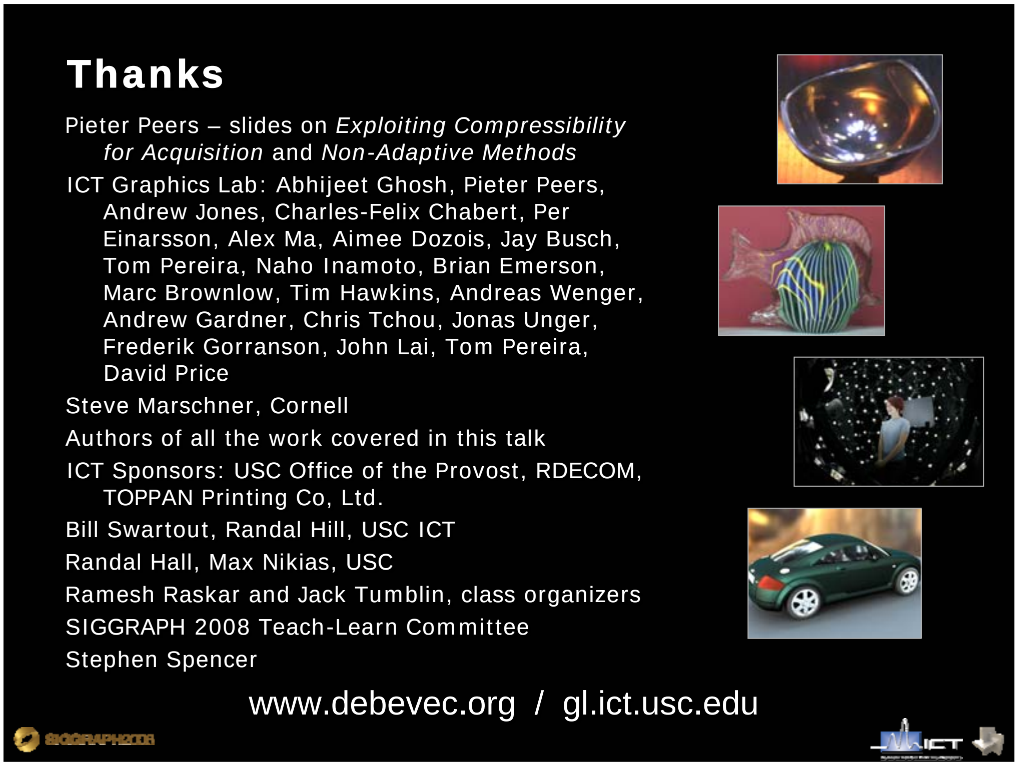


The last two projects I want to mention make clever use of wavelength-altering reflectance properties use of for scene capture, which we haven't yet seen.

MOVA's "Coutour" system leverages fluorescence to capture 3D facial geometry. It uses stereo correspondence from multiple cameras to record the 3D geometry of the face, but it avoids a common problem of lacking sufficient facial texture to match points between the viewpoints. They do this by applying a random pattern of glow-in-the-dark makeup to the performer, and illuminated the subject with a combination of ultraviolet and visible lights which turn on approximately thirty times per second. The cameras record a visible texture image when the lights are on, and during this time the ultraviolet lights charge the phosphorescent makeup. Then, a second image is acquired when the lights are off, observing the makeup glowing on its own. This also alleviates the problem of specular facial reflectance in geometry matching, since the emmissive surface has a relatively even light output over the hemisphere. In this way, natural facial shapes can be captured.



Hullin et al's work being presented at this year's SIGGRAPH conference presents an ingenious technique for 3D scanning otherwise difficult-to-scan objects using ***Fluorescent Immersion Range Scanning.*** In this technique, a laser stripe moves over the object, but object is placed in a liquid with fluorescent dye which glows orange when the green laser stripe travels through it. When the stripe reaches the object, the volume no longer fluoresces, and there is a dark contour which indicates the intersection of that plane of the laser with the object, easily triangulated into 3D. The fluorescence (and a filter on the camera blocking the wavelength of the laser light) removes the effect of multiple scattering, which would be a problem if the laser were traveling through a cloudy liquid.


Thanks
Pieter Peers – slides on *Exploiting Compressibility for Acquisition* and *Non-Adaptive Methods*
ICT Graphics Lab: Abhijeet Ghosh, Pieter Peers, Andrew Jones, Charles-Felix Chabert, Per Einarsson, Alex Ma, Aimee Dozois, Jay Busch, Tom Pereira, Naho Inamoto, Brian Emerson, Marc Brownlow, Tim Hawkins, Andreas Wenger, Andrew Gardner, Chris Tchou, Jonas Unger, Frederik Gorranson, John Lai, Tom Pereira, David Price
Steve Marschner, Cornell
Authors of all the work covered in this talk
ICT Sponsors: USC Office of the Provost, RDECOM, TOPPAN Printing Co, Ltd.
Bill Swartout, Randal Hill, USC ICT
Randal Hall, Max Nikias, USC
Ramesh Raskar and Jack Tumblin, class organizers
SIGGRAPH 2008 Teach-Learn Committee
Stephen Spencer
www.debevec.org / gl.ict.usc.edu

Thank you!

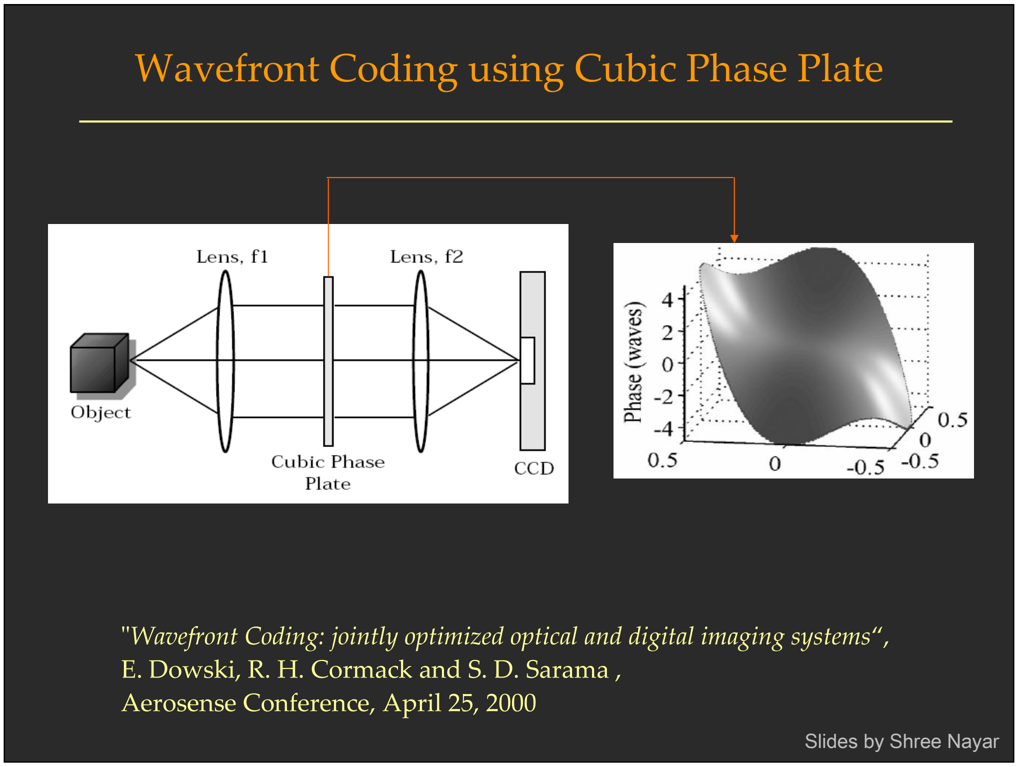
Optics: Computatable Extensions

Ramesh Raskar
MIT Media Lab

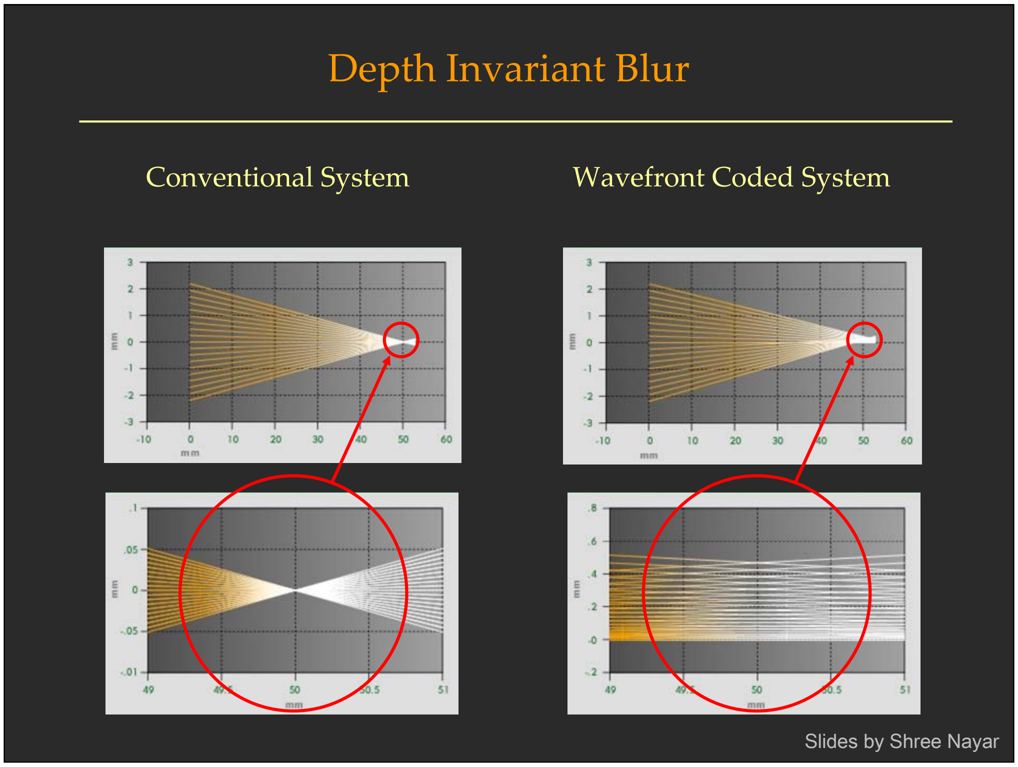
Optics: Computable Extensions

- Unusual Optics
  - Wavefront Coding
  - Folded Optics
  - Nonlinear Optics
- Material (Refractive Index)
  - Graded Index (GRIN) materials
  - Photonics crystals
  - Negative Index
- Imaging
  - Schlieren Imaging
  - Agile Spectrum Imaging
  - Random Lens Imaging
- Animal Eyes
  - What can we learn

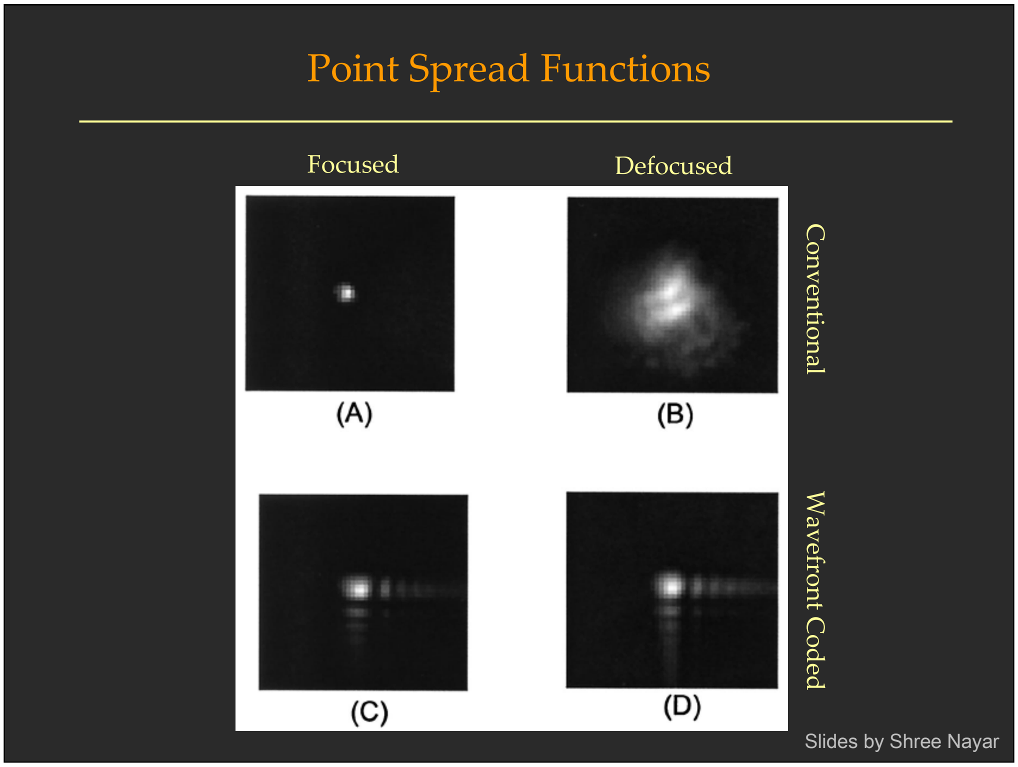
Conventional lenses have a limited depth of field. One can increase the depth of field and reduce the blur by stopping down the aperture. However, this leads to noisy images.



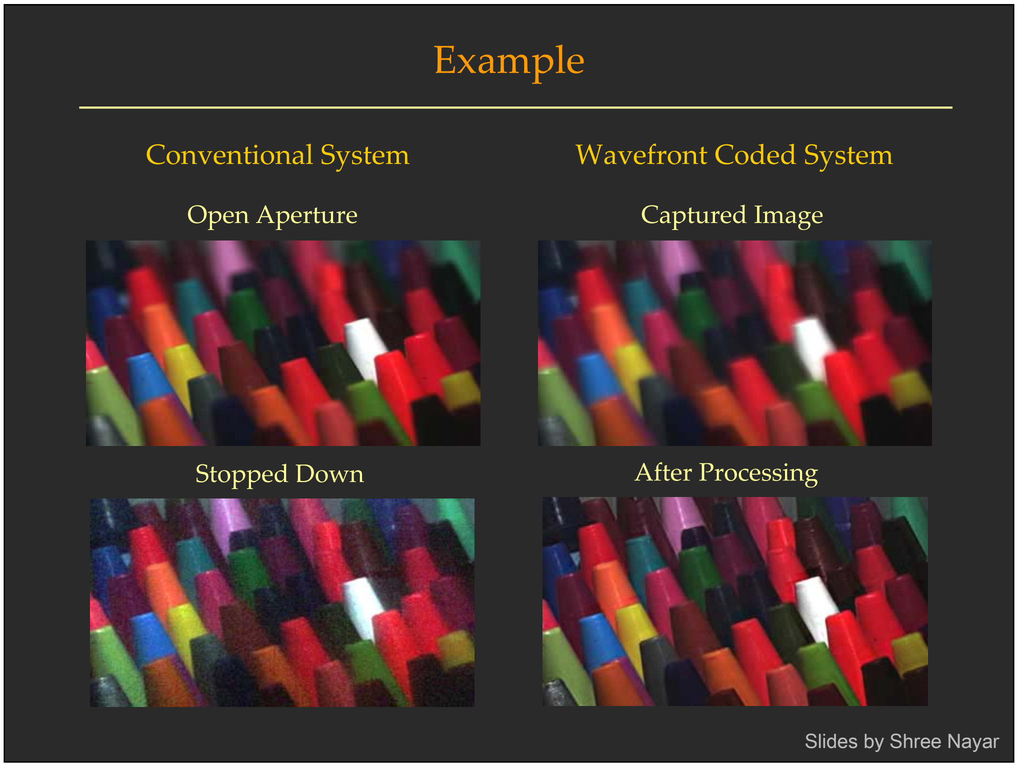
A solution proposed by authors and now commercialized by CDM optics uses a cubic phase plate. The effect of cubic phase plate is equivalent so summing images due to lens position was different planes of focus. Note that the cubic phase plate can be made up of glass of varying thickness OR glass of varying refractive index. The example here shows a total phase difference of just 8 periods.



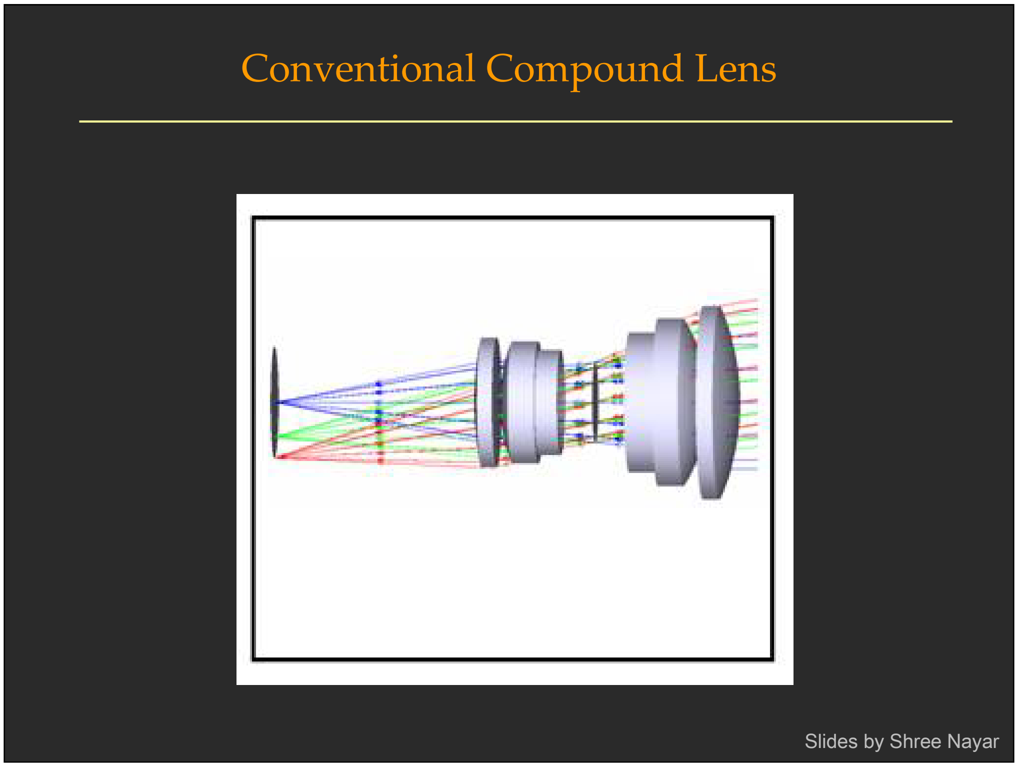
Unlike traditional systems, where you see a conical profile of lightfield for a point in focus, for CDM the profile is more like a twisted cylinder of straws. This makes the point spread function somewhat depth independent.



The point spread function is not a point even when the point is in sharp focus. One can say it is equally worse over a large range of focus.

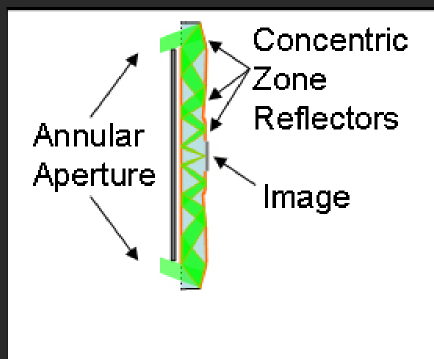


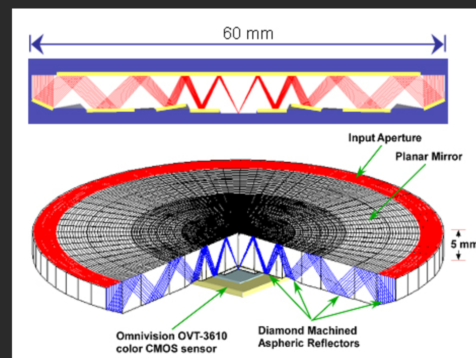
After software decoding one can recover sharp images. These images show good tradeoff between depth of field and image noise.



Traditional lenses are long, the length is close to the focal length in mm.

“Origami Lens”: Thin Folded Optics (2007)



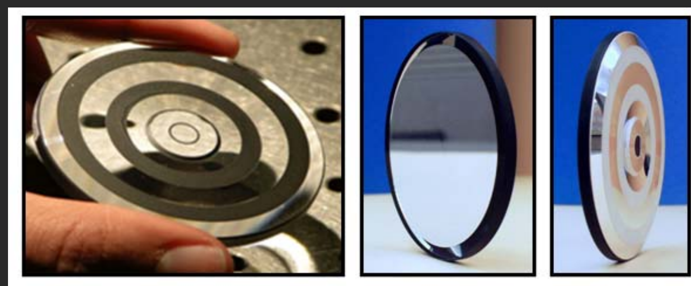


“Ultrathin Cameras Using Annular Folded Optics, “
E. J. **Tremblay**, R. A. Stack, R. L. Morrison, J. E. **Ford**
Applied Optics, 2007 - OSA

Slides by Shree Nayar

New techniques are trying decrease this distance using a folded optics approach. The origami lens uses multiple total internal reflection to propagate the bundle of rays.

Origami Lens

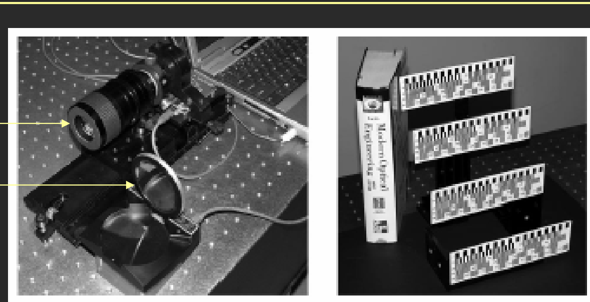


Slides by Shree Nayar

Optical Performance

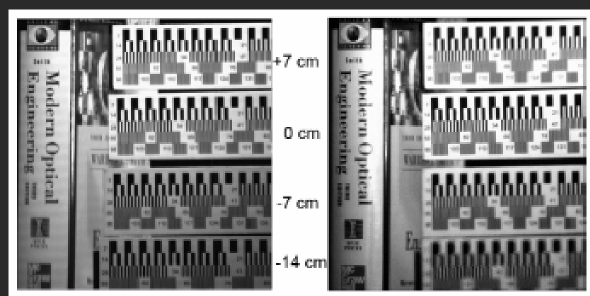
Conventional

Origami



Scene

Conventional
Lens Image

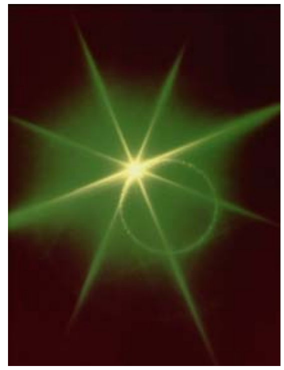


Origami
Lens Image

Slides by Shree Nayar

Nonlinear Optics

- Linear Optics
  - Light is deflected or delayed without wavelength change
  - (a 'linear system')
- High intensity light in non-linear media
  - Think of it as 'self-modifying programs'
  - Mainly lasers and crystals
- Nonlinear optics
  - Change the color of a light beam
  - Change beam's shape in space and time
  - Light intensity refractive index
  - Self-focussing via gaussien high intensity beam
  - Switch and gate telecommunications systems
  - Create shortest events (in femtoseconds, 10^{-15}).



Input: Infrared laser
Output: Green light

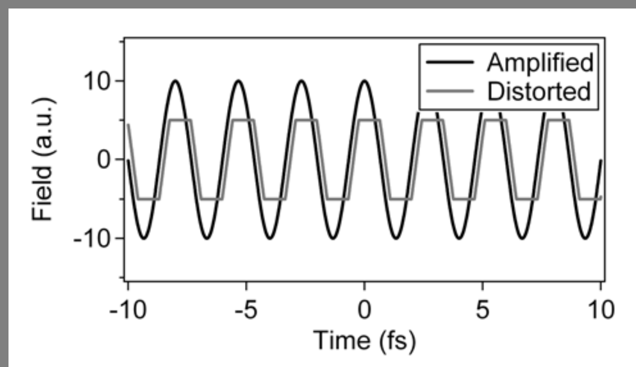
www.physics.gatech.edu/gcuo/UltrafastOptics/3803/OpticsI23NonlinearOptics.ppt

Nonlinear optics has been a rapidly growing field in recent decades. It is based on the study of effects and phenomena related to the interaction of intense coherent light radiation with matter.

Nonlinear optics (NLO) is the branch of optics that describes the behaviour of light in nonlinear media, that is, media in which the dielectric polarization P responds nonlinearly to the electric field E of the light. This nonlinearity is typically only observed at very high light intensities such as those provided by pulsed lasers.

Why do nonlinear effects occur, in general?

Imagine playing music through a cheap amplifier that just can’t quite put out the power necessary to hit the loud notes.



The sharp edges correspond to higher frequencies—harmonics!

www.physics.gatech.edu/gcuo/UltrafastOptics/3803/OpticsI23NonlinearOptics.ppt

Non-linear effects are always present but usually barely noticeable.
However, when the input signal intensity is high, the system does not behave as a linear system. The higher harmonics have progressively smaller amplitude.

Nonlinear Optics

- Interesting Applications
  - Change in refractive index (RI),
    - Original RI: n_0
    - Nonlinear RI: n_2
    - Incident light intensity, I
    - Modified refractive index, $n = n_0 + n_2 * I$
  - Self-focusing
    - *N2* is positive,
    - RI is higher for higher intensity, usually center of beam
    - Create a 'convex' lens,
    - (But collapsing beam on itself till material is damaged)

www.physics.gatech.edu/gcuo/UltrafastOptics/3803/OpticsI23NonlinearOptics.ppt

For Computational Photography, an interesting future direction could be programmable refractive index. We will see more on the next slide.

Self-focusing is induced by the change in refractive index. A medium whose refractive index increases with the electric field intensity acts as a focusing lens for a laser beam. The peak intensity of the self-focused region keeps increasing as the wave travels through the medium, until defocusing effects or medium damage interrupt this process. Since n_2 is positive in most materials, the refractive index becomes larger in the areas where the intensity is higher, usually at the centre of a beam, creating a focusing density profile which potentially leads to the collapse of a beam on itself.[

Self-focusing is often observed when radiation generated by femtosecond lasers propagates through many solids, liquids and gases. Depending on the type of material and on the intensity of the radiation, several mechanisms produce variations in the refractive index which result in self-focusing: a well known example is Kerr-induced self-focusing.

Optical Kerr Effect: Intensity dependent Refractive Index

The refractive index including higher order terms

$$n = \sqrt{1 + \chi^{(1)} + \chi^{(3)} \left|E\right|^2}$$

Now, the usual refractive index (which we'll call n_0) is: $n_0 = \sqrt{1 + \chi^{(1)}}$

So:

$$n \;=\; \sqrt{n_0^2 + \chi^{(3)} \left|E\right|^2} \;=\; n_0 \sqrt{1 + \chi^{(3)} \left|E\right|^2 / n_0^2}$$

Assume that the nonlinear term << n_0:

So:

$$n \;\approx\; n_0 \left[1 + \tfrac{1}{2} \chi^{(3)} \left|E\right|^2 / n_0^2\right] \;\approx\; n_0 + \chi^{(3)} \left|E\right|^2 / 2n_0$$

Usually, we define a "Nonlinear Refractive Index":

$$n_2 \propto \chi^{(3)} / 2n_0$$

$$\boxed{n \approx n_0 + n_2 I}$$

since $I \propto \left|E\right|^2$

Before 1960, optics mathematical equations manifested a linear system of equations, using usual refractive index n0.

The optical Kerr effect, or AC Kerr effect is the case in which the electric field is due to the light itself. This causes a variation in index of refraction which is proportional to the local irradiance of the light. This refractive index variation is responsible for the nonlinear optical effects of self-focusing and self-phase modulation, and is the basis for Kerr-lens modelocking. This effect only becomes significant with very intense beams such as those from lasers.

First Demo: Second-Harmonic Generation

P.A. Franken, et al, Physical Review Letters 7, p. 118 (1961)

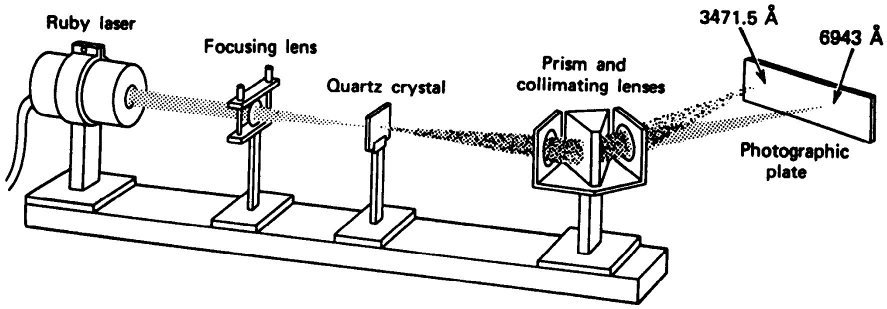


Figure 12.1. Arrangement used in the first experimental demonstration of second-harmonic generation [1]. A ruby-laser beam at $\lambda = 0.694\ \mu m$ is focused on a quartz crystal, causing the generation of a (weak) beam at $\frac{1}{2}\lambda = 0.347\ \mu m$. The two beams are then separated by a prism and detected on a photographic plate.

Input: Ruby laser (695 nm) **Output**: Ruby laser (649nm) + UV (347nm)

www.physics.gatech.edu/gcuo/UltrafastOptics/3803/OpticsI23NonlinearOptics.ppt

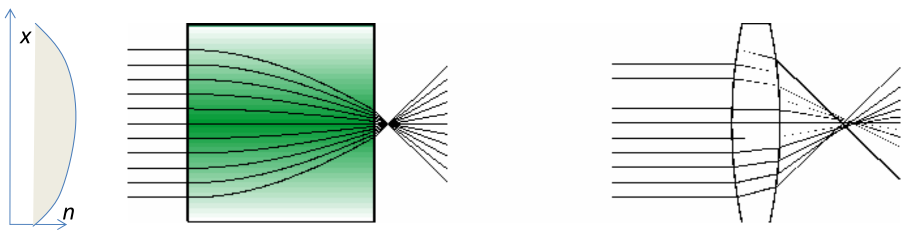
Nonlinear materials like quartz crystal create a second harmonic (twice the frequency, i.e. half the wavelength) when a high intensity laser is incident.
This figure is borrowed from Margaret Murnane’s HHG talk.

Optics: Computable Extensions

- Unusual Optics
  - Wavefront Coding
  - Folded Optics
  - Nonlinear Optics
- Material (Refractive Index)
  - Graded Index (GRIN) materials
  - Photonics crystals
  - Negative Index
- Imaging
  - Schlieren Imaging
  - Agile Spectrum Imaging
  - Random Lens Imaging
- Animal Eyes
  - What can we learn

Gradient Index (GRIN) Optics

Refractive Index
along width



Gradient Index 'Lens'

Continuous change of the refractive index within the optical material

Change in RI is very small, 0.1 or 0.2

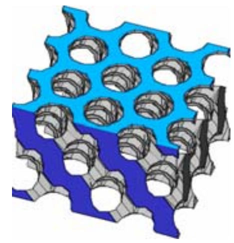
Conventional Convex Lens

Constant refractive index but carefully designed geometric shape

Consider a conventional lens: An incoming light ray is first refracted when it enters the shaped lens surface because of the abrupt change of the refractive index from air to the homogeneous material. It passes the lens material in a direct way until it emerges through the exit surface of the lens where it is refracted again because of the abrupt index change from the lens material to air (see Fig. 1, right). A well-defined surface shape of the lens causes the rays to be focussed on a spot and to create the image. The high precision required for the fabrication of the surfaces of conventional lenses aggrevates the miniaturization of the lenses and raises the costs of production.

GRIN lenses represent an interesting alternative since the lens performance depends on a continuous change of the refractive index within the lens material. Instead of complicated shaped surfaces plane optical surfaces are used. The light rays are continuously bent within the lens until finally they are focussed on a spot. Miniaturized lenses are fabricated down to 0.2 mm in thickness or diameter. The simple geometry allows a very cost-effective production and simplifies the assembly. Varying the lens length implies an enormous flexibility at hand to fit the lens parameters as, e.g., the focal length and working distance. For example, appropriately choosing the lens length causes the image plane to lie directly on the surface plane of the lens so that sources such as optical fibers can be glued directly onto the lens surface.

Photonic Crystals



- Photonic Crystal
  - Nanostructure material with ordered array of holes
  - A lattice of high-RI material embedded within a lower RI
  - High index contrast
  - 2D or 3D periodic structure
- Photonic band gap
  - Highly periodic structures that blocks certain wavelengths
  - (creates a ‘gap’ or notch in wavelength)
- Applications
  - ‘Semiconductors for light’: mimics silicon band gap for electrons
  - Highly selective/rejecting narrow wavelength filters (Bayer Mosaic?)
  - Light efficient LEDs
  - Optical fibers with extreme bandwidth (wavelength multiplexing)
  - Hype: future terahertz CPUs via optical communication on chip

Current explosion in information technology has been derived from our ability to control the flow of electrons in a semiconductor in the most intricate ways. Photonic crystals promise to give us similar control over photons - with even greater flexibility because we have far more control over the properties of photonic crystals than we do over the electronic properties of semiconductors.

Negative Refractive Index

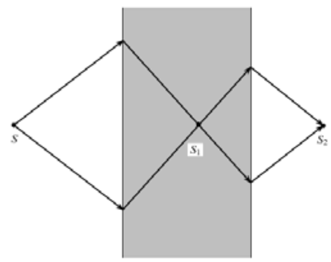
- Left-handed material
  - Refractive index N = +/- sqrt(ε μ)
  - Usual matter, both permittivity ε and permeability μ are positive
  - NIM: both ε and μ are negative
    - Usually over a limited wavelength band
  - Incident and refracted waves are on the same side of the normal to the boundary
  - The electric field, magnetic field and wave vector follow a left-hand rule

- Applications
  - Perfect Lens
    - A planar slab of NIM,
    - Imaging beyond diffraction limits
    - No glare (lens reflection)
  - Sub-wavelength imaging

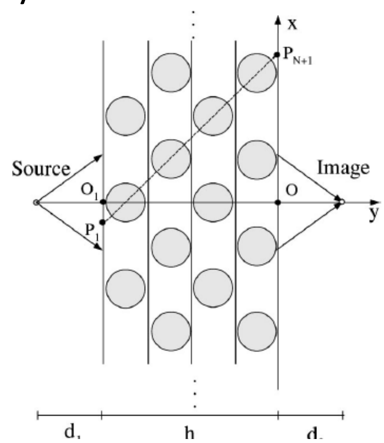
- Implementations
  - 'Metamaterials'
  - So far only in lower frequencies, microwave
  - Array of wires and spring-ring resonators
  - Carefully designed photonic crystal (glass rods in air)

All transparent or translucent materials that we know of possess positive refractive index--a refractive index that is greater than zero. However, is there any fundamental reason that there should not be materials with negative refractive index?

Perfect Lens* using
Photonic Crystal Slab



NIM Slab



Rods in Air

*Only points whose distance from the lens surface does not exceed the thickness of the lens have real images

http://www.ee.duke.edu/~drsmith/nim_pubs/pubs_2004/bliokh_physics_uspekhi_2004.pdf
http://www.ee.duke.edu/~drsmith/nim_pubs/pubs_2004/he_prb_2004.pdf

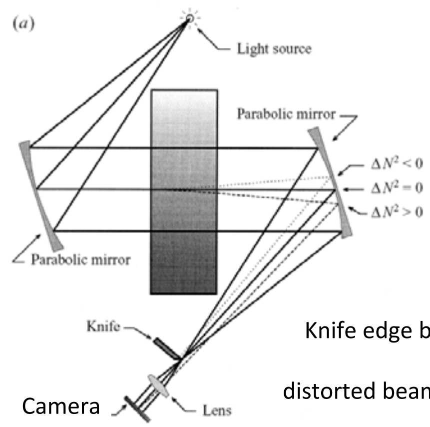
Refraction of a plane wave at the interface of a left-handed medium and a right-handed one looks quite unusual.
The fact that the incident and refracted waves are on the same side of the normal to the boundary enables one to manufacture quite unusual optic elements of left-handed media. For instance, a plane-parallel plate made of a left-handed material works as a collecting lens.

Optics: Computable Extensions

- Unusual Optics
  - Wavefront Coding
  - Folded Optics
  - Nonlinear Optics
- Material (Refractive Index)
  - Graded Index (GRIN) materials
  - Photonics crystals
  - Negative Index
- Imaging
  - Schlieren Imaging
  - Agile Spectrum Imaging
  - Random Lens Imaging
- Animal Eyes
  - What can we learn

Schlieren Photography

- Image of small index of refraction gradients in a gas
- Invisible to human eye (subtle mirage effect)



Changes in the index of refraction of air are made visible by Schlieren Optics. This special optics technique is extremely sensitive to deviations of any kind that cause the light to travel a different path.

In schlieren photography, the collimated light is focused with a lens, and a knife-edge is placed at the focal point, positioned to block about half the light. In flow of uniform density this will simply make the photograph half as bright. However in flow with density variations the distorted beam focuses imperfectly, and parts which have focussed in an area covered by the knife-edge are blocked. The result is a set of lighter and darker patches corresponding to positive and negative fluid density gradients in the direction normal to the knife-edge.

Clearest results are obtained from flows which are largely two-dimensional and not volumetric.


http://www.mne.psu.edu/psgdl/FSSPhotoalbum/index1.htm

Full-Scale Schlieren Image Reveals The Heat Coming off of a Space Heater, Lamp and Person

http://www.mne.psu.edu/psgdl/FSSPhotoalbum/index1.htm

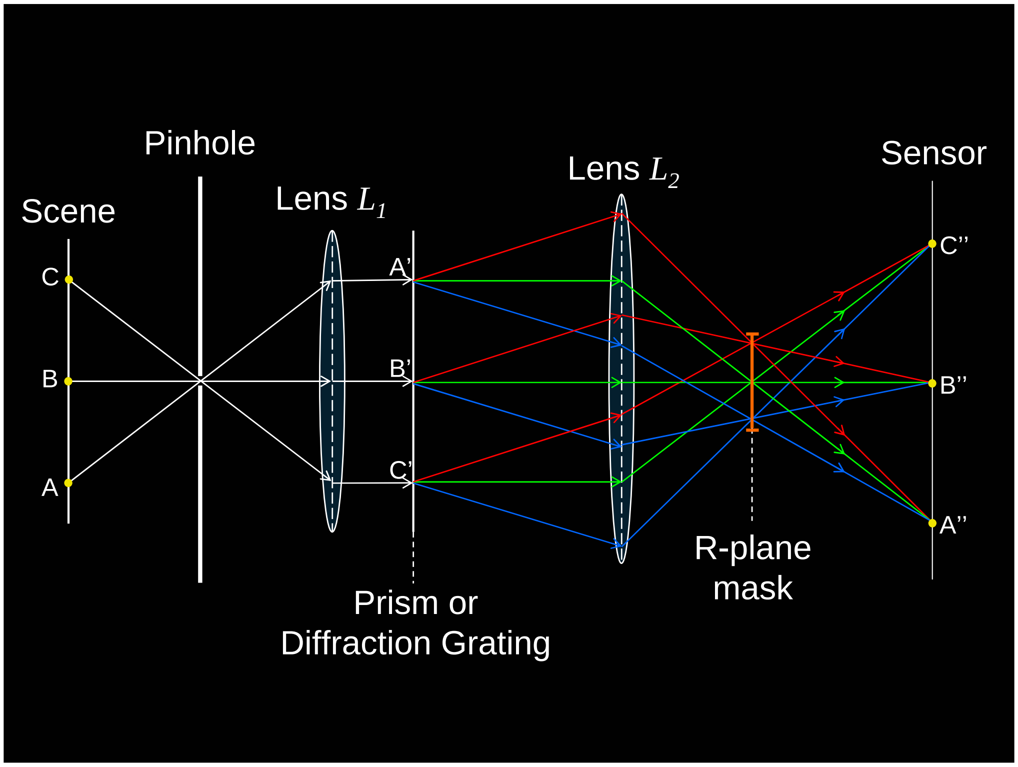
Full-Scale Schlieren Image Reveals A Gas Leak.

For amateur photography, the setup is surprisingly simple and one can record stunning images.

Agile Spectrum Imaging

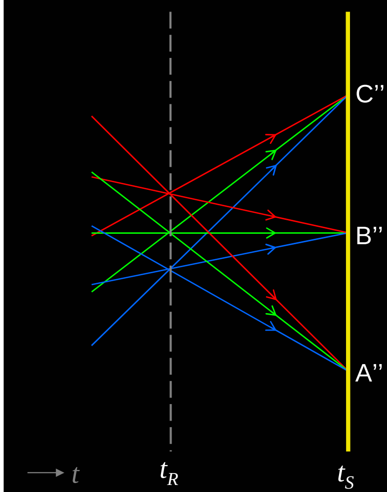
Ankit Mohan, Jack Tumblin
Northwestern University

Ramesh Raskar
MERL / MIT

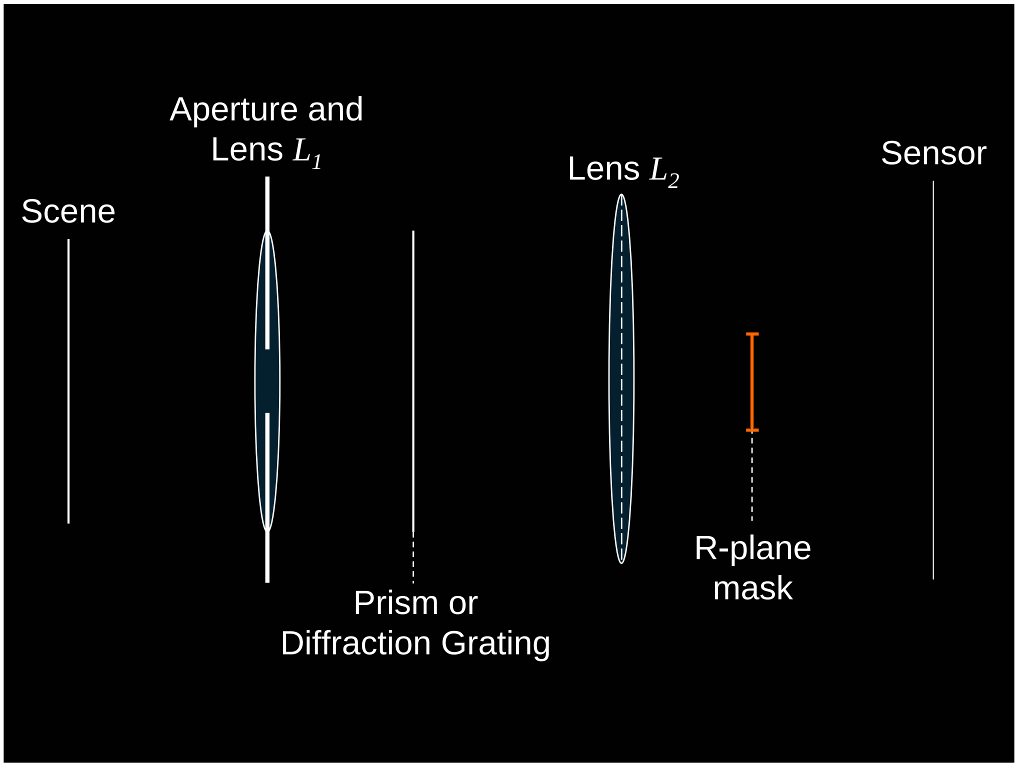


Pinhole
Scene
Lens L_1
Lens L_2
Sensor
C
B
A
A'
B'
C'
C''
B''
A''
R-plane
mask
Prism or
Diffraction Grating

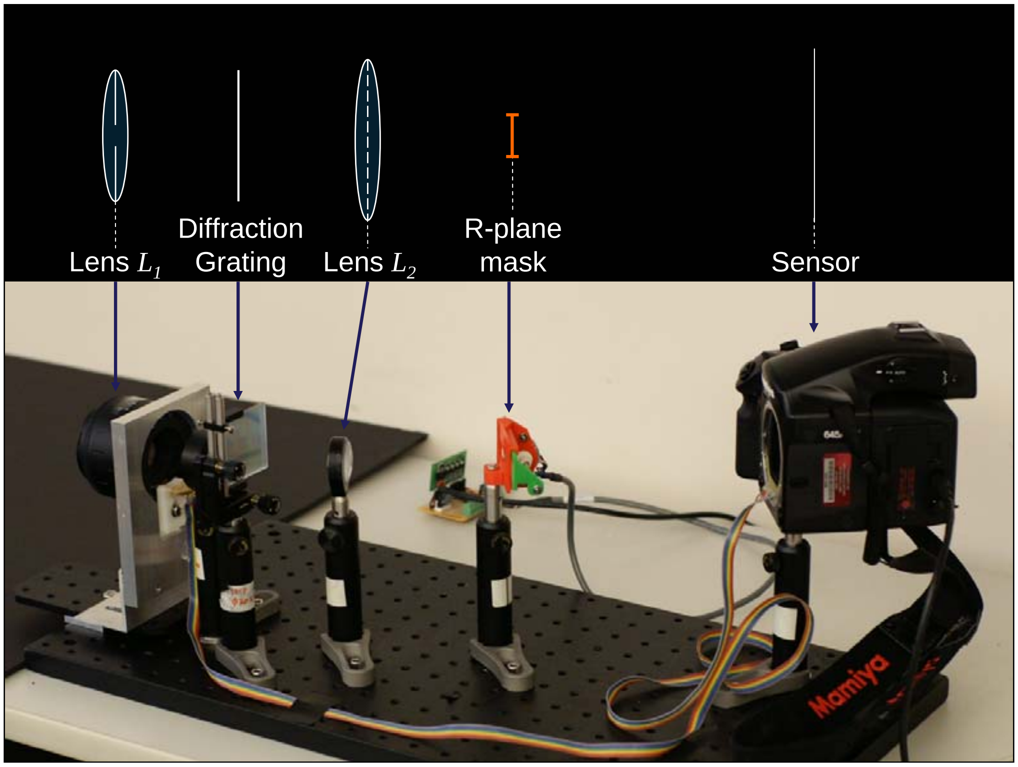
Rainbow plane ($t = t_R$)

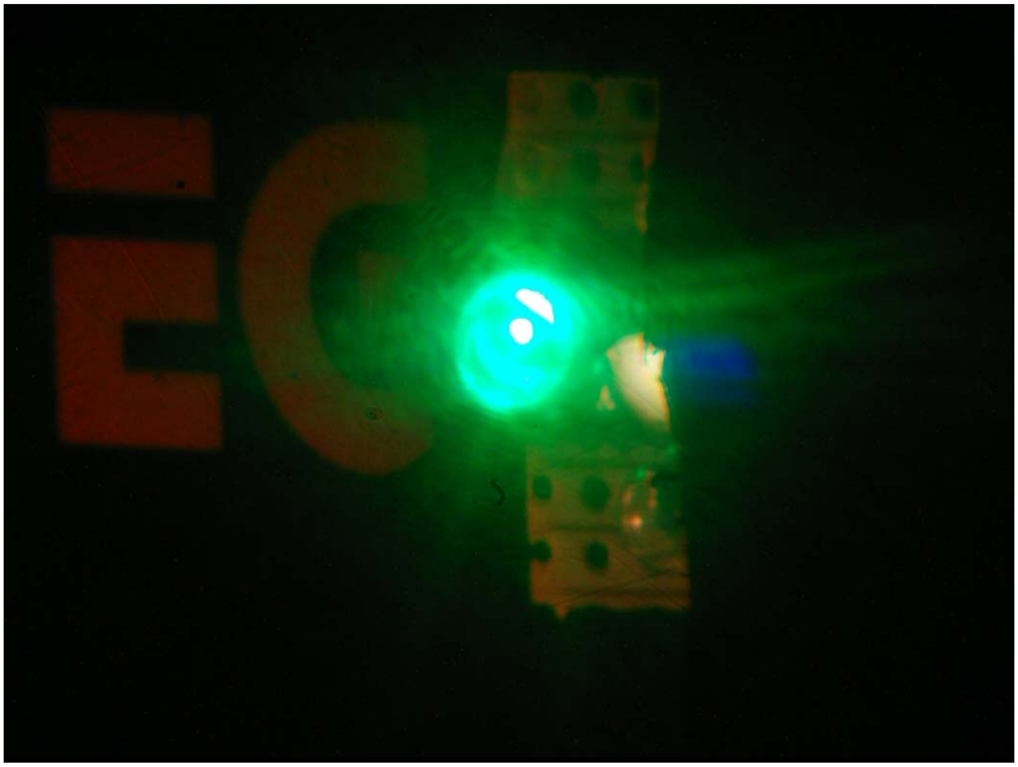


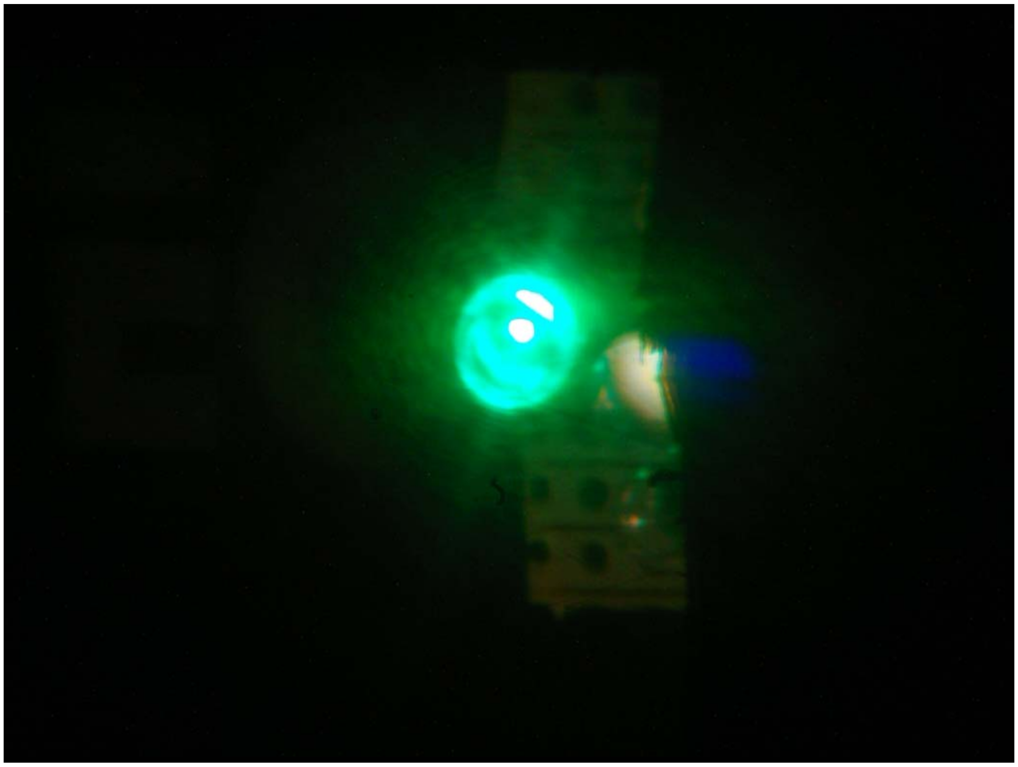
Control the *spectral sensitivity* of the sensor by placing an appropriate *grayscale masks* in the R-plane.

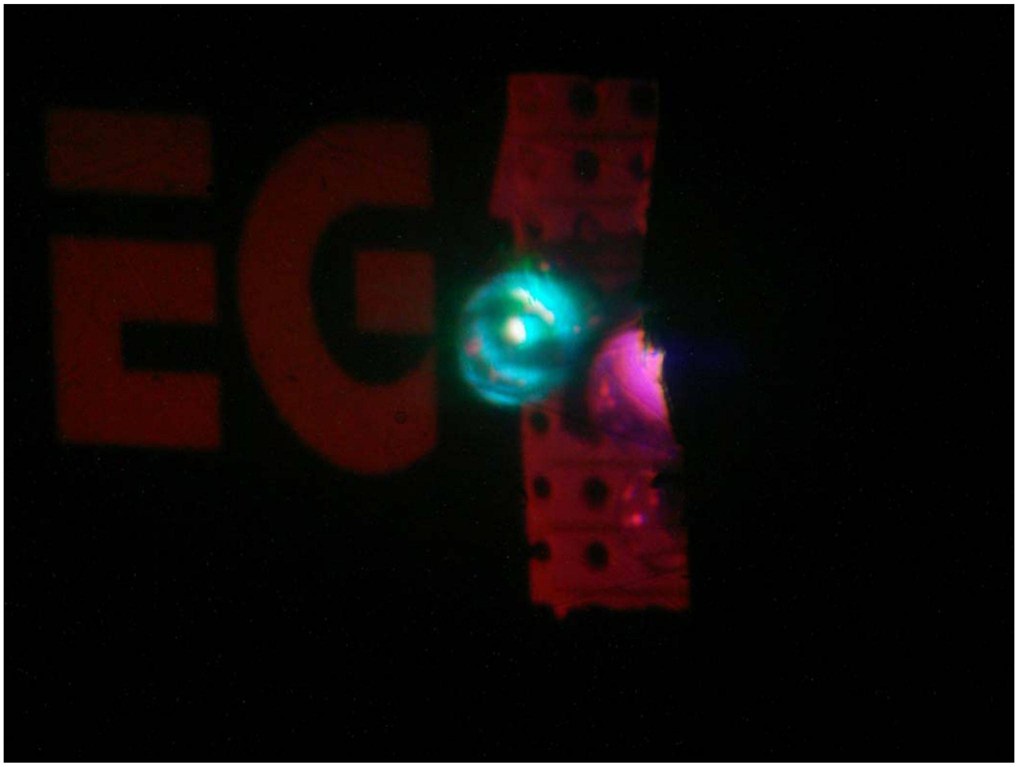


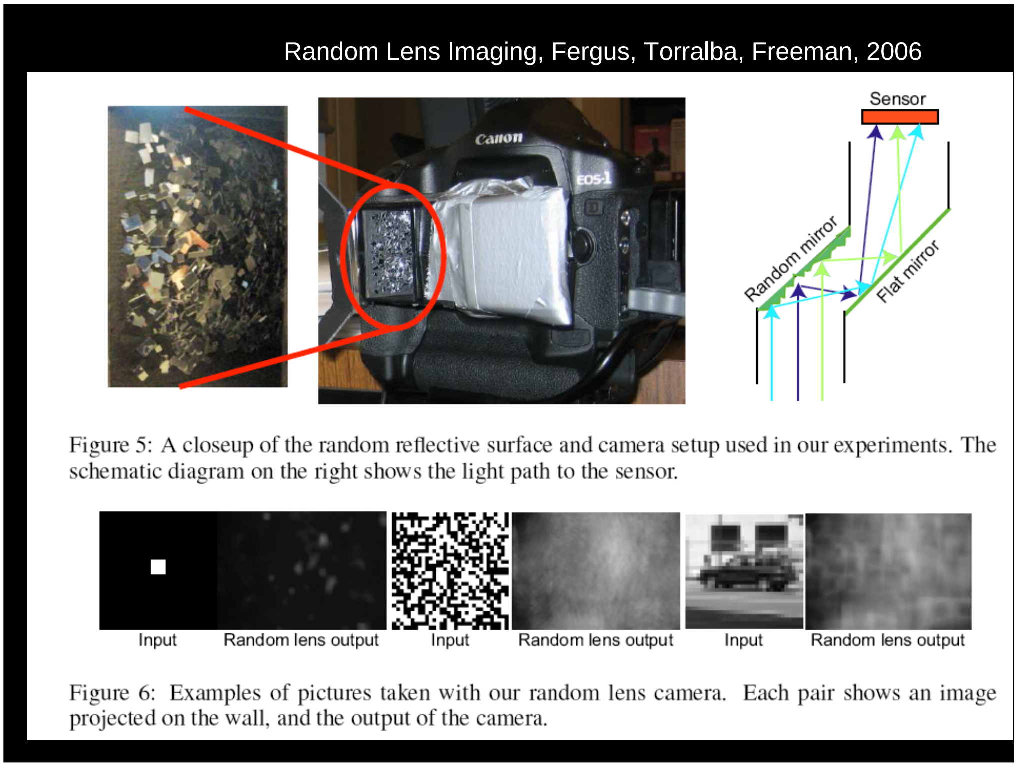
Aperture and
Lens L_1
Lens L_2
Sensor
Scene
R-plane
mask
Prism or
Diffraction Grating


Lens L_1
Diffraction Grating
Lens L_2
R-plane mask
Sensor
Mamiya









What if the input to output relationship of the light rays were randomized? How could you use such a camera? What would
be the benets and the drawbacks of such an approach? The ability to use a camera with those
characteristics would open new regions of the space of possible optical designs, since in many cases
it is easier to randomize light ray directions than to precisely direct them. Authors show applications in super-resolution and depth estimation.

Community and Social Impact

Ramesh Raskar

Community and Social Impact

- Crowdsourcing
  - Object Recognition
  - CMU's captcha-like games
  - MIT's LabelMe
  - Distributed Search
- Cross-sources Image Visualization
  - Google/Virtual Earth problems
  - From street maps to street-level photos to 3D models
- Trust and Privacy
  - Verification and Forensics
  - Privacy-preserving Computation
- Mobile Phones
  - ZoneSurger: social tagging of photos
  - Govt forms in developing counties
- Social/Political Goals

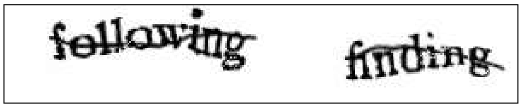
Crowd Sourcing

- Get Help from Crowd for Tasks That Tax Computers
  - Object Recognition/Labeling
- Image-based Applications
  - Fight Spam using CAPTCHAs
  - Completely Automated Turing Test To Tell Computers and Humans Apart
  - reCAPTCHAs
    - For OCR of old text
  - LabelMe
    - Segmentation and object recognition
  - Distributed Search

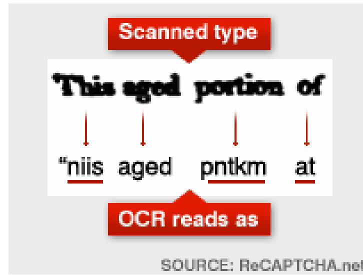
Crowdsourcing is a new online strategy for converting a task difficult for computers and too expensive for a team of dedicated humans, and outsourcing it to an undefined, generally large group of people, in the form of an open call. Thanks to Web2.0 online technologies make participation of thousands of synchronous or asynchronous users possible to solve a talk.

Example: Digitizing Old Books

- Optical Character Recognition for Old Text
  - Poor visibility



Original goal: making segmentation and OCR difficult by adding an angled line



- reCAPTCHA project at CMU
  - Human solves CAPTCHA and solves difficult OCR problems

http://recaptcha.net/ http://en.wikipedia.org/wiki/Captcha
http://news.bbc.co.uk/2/hi/technology/7023627.stm

The CMU team is involved in digitising old books and manuscripts supplied by a non-profit organisation called the Internet Archive, and uses Optical Character Recognition (OCR) software to examine scanned images of texts and turn them into digital text files which can be stored and searched by computers.
But the OCR software is unable to read about one in 10 words, due to the poor quality of the original documents.
Computers cannot read words as easily as humans
The only reliable way to decode them is for a human to examine them individually - a mammoth task since CMU processes thousands of pages of text every month.
To solve this problem the team takes images of the words which the OCR software can't read, and uses them as CAPTCHAs.
These CAPTCHAs, known as reCAPTCHAS, are then distributed to websites around the world to be used in place of conventional CAPTCHAs.
When visitors decipher the reCAPTCHAs to gain access to the web site, the answers - the results of humans examining the images - are sent back to CMU.
Every time an Internet user deciphers a reCAPTCHA, another word from an old book or manuscript is digitised.

LabelMe

Bryan Russell, Antonio Torralba and William T. Freeman at MIT

- Recognition performance increases dramatically when more labeled training data is made available
- Goal: Create a massive high quality database for research on object recognition.
- Multiple users label as many objects and regions as they can within the same image.

http://labelme.csail.mit.edu/

The authors seek to build a large collection of images with ground truth labels to be used for object
detection and recognition research. Such data is useful for supervised learning and quantitative
evaluation. To achieve this, they developed a web-based tool that allows easy image annotation
and instant sharing of such annotations. Using this annotation tool, they have collected a large
dataset that spans many object categories, often containing multiple instances over a wide variety
of images. They quantify the contents of the dataset and compare against existing state of the
art datasets used for object recognition and detection. Also, they show how to extend the dataset
to automatically enhance object labels with WordNet, discover object parts, recover a depth ordering
of objects in a scene, and increase the number of labels using minimal user supervision
and images from the web.

Distributed Patch-wise Image Search

- Example: Steve Fossett's plane, 2007
- Divide and Conquer
  - Hires imagery from DigitalGlobe
  - Amazon's Mechanical Turk splits into small patches
  - Volunteers each review individual patches
  - Report back and aggregate info for professionals

http://www.wired.com/software/webservices/news/2007/09/distributed_search

The search for aviator Steve Fossett, whose plane went missing in Nevada in 2007, in which up to 50,000 people examined high-resolution satellite imagery from DigitalGlobe that was made available via Amazon Mechanical Turk. The helpers are issued squares that represent 278-foot-square pieces of the search area. If they see something worth closer study, participants flag it. Since each square is issued to 10 different people, squares that are flagged by several volunteers are given greater scrutiny.

User Generated Content Visualization

http://phototour.cs.washington.edu/

- Google Map overlayed with geo-tagged photos
- Image-based Mashups

Microsoft Photosynth and U-Washington's Phototourism software takes a large collection of photos of a place or an object, analyzes them for similarities, and then displays the photos in a reconstructed **three-dimensional space**, showing you how each one relates to the next.
New options on Google Maps allows users to post and view populated map with geo-tagged photos provided by Panoramio.

Mobile Photography

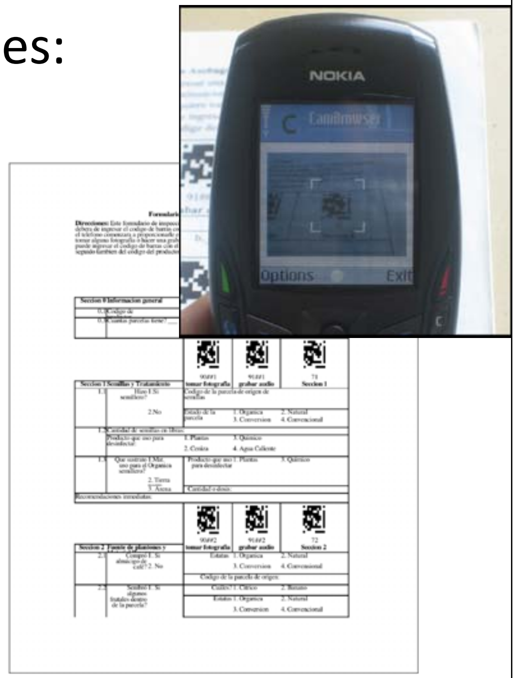
- ZoneSurfer
  - (Yahoo Research)
  - Spatial - social - topical mobile photo browser
  - Mobile window to the world of multimedia
  - Social interface based on Flickr

- Mobile phone-based entrepreneurship
  - Developing countries
  - Many examples: http://nextbillion.mit.edu/

Zurfer from Yahoo Research uses channel metaphor to give users contextual access to media of interest according to key dimensions: spatial, social, and topical. Elements of social interaction and communication aroundcthe photographs are built into the mobile application, to increasecuser engagement. The application utilizes Flickr.comc as an image and social-network data source (From http://yahooresearchberkeley.com/blog/wp-content/uploads/2007/09/sp50a-hwang.pdf)

Developing Countries: CAMForms

- Paper forms with barcodes
- 83-bit 2D codes (including seven bits of error correction)



Parikh (2005)

The camera phone provides an easy interface to fill in and verify government forms. The paper form is printed with 2D bar-codes which are decoded by camera phone and info is transmitted to a central location.

Socio-Political Goals: B'Tselem

From websiteof : Israeli Information Center for Human Rights in the Occupied Territories

Israeli Information Center for Human Rights in the Occupied Territories captures photos of events.
From their website: Goal is to Document and educate Israeli public and policymakers about human rights violations in Occupied Territories.
Second goal is to combat phenomenon of denial prevalent among Israeli public.
They hope to create human rights culture in Israel

Course: Computational Photography Advanced Topics

Module 1: 90 minutes

9:00:	A.1	Introduction and Overview	(Raskar, 15 minutes)
9:15:	A.2	Concepts in Computational Photography	(Tumblin, 15 minutes)
9:30:	A.3	Optics: Computable Extensions	(Raskar, 30 minutes)
10:00:	A.4	Sensor Innovations	(Tumblin, 30 minutes)
10:30:	Q & A		(5 minutes)

10:35: Break: 25 minutes

Module 2: 90 minutes

11:00:	B.1	Illumination As Computing	(Debevec, 25 minutes)
11:25:	B.2	Scene and Performance Capture	(Debevec, 20 minutes)
11:45:	B.3	Image Aggregation & Sensible Extensions	(Tumblin, 20 minutes)
12:05:	B.4	Community and Social Impact	(Raskar, 20 minutes)
12:25:	B.4	Summary and Discussion, Q&A	(All, 10 minutes)

Course Page : http://computationalphotography.org/