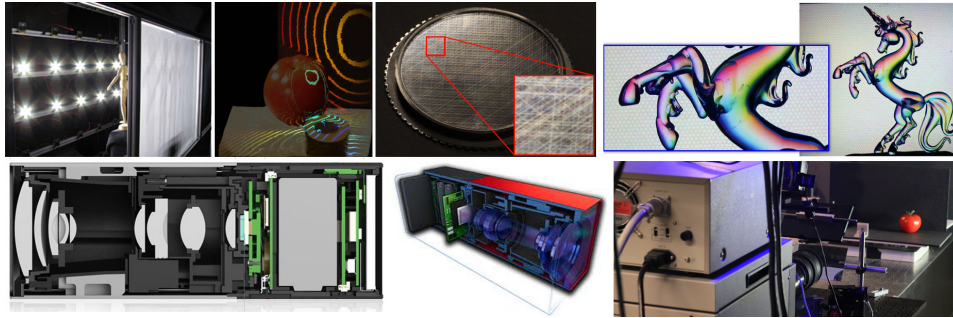

Computational Plenoptic Imaging



SIGGRAPH 2012 Course

Monday, 6 August 2012, 9:00-10:30 am

Los Angeles Convention Center, Room 408B

Gordon Wetzstein
MIT Media Lab
gordonw@media.mit.edu

Ivo Ihrke
Saarland University
ihrke@mpi-inf.mpg.de

Douglas Lanman
MIT Media Lab
dlanman@media.mit.edu

Wolfgang Heidrich
University of British
Columbia
heidrich@cs.ubc.ca

Kurt Akeley
Lytro, Inc.
kakeley@lytro.com

Ramesh Raskar
MIT Media Lab
raskar@media.mit.edu



Abstract and Motivation

Spawned by the introduction of the Lytro light field camera to the consumer market and recent accomplishments in the speed at which light can be captured [Raskar et al. 2011], a new generation of computational cameras is emerging. By exploiting the co-design of camera optics and computational processing, these cameras capture unprecedented details of the plenoptic function — a ray-based model for light that includes the color spectrum as well as spatial, temporal, and directional variation. Although digital light sensors have greatly evolved in the last years, the visual information captured by conventional cameras has remained almost unchanged since the invention of the daguerreotype. All standard CCD and CMOS sensors integrate over the dimensions of the plenoptic function as they convert photons into electrons; in the process all visual information is irreversibly lost, except for a two-dimensional, spatially-varying subset — the common photograph. In this course, we review the plenoptic function and discuss approaches that aim at optically encoding high-dimensional visual information that is then recovered computationally in post-processing.

This course is intended to review the state of the art in joint optical light modulation and computational reconstruction of visual information transcending that captured by traditional photography. In addition to the plenoptic dimensions, we also consider high dynamic range image acquisition as common sensors have a limited dynamic range. In contrast to prior courses on general computational photography [Raskar and Tumblin 2006,2007] and a recent survey on the topic [Wetzstein et al. 2011], this course gives a broad, well-structured, and intuitive overview of all aspects of plenoptic image acquisition and focuses on two recent developments: light field acquisition and ultra-fast cameras. We unveil the secrets behind capturing light at a trillion frames per second and the Lytro camera. Our course serves as a resource for interested parties by providing a categorization of recent research and help in the identification of unexplored areas in the field.

We will discuss all aspects of plenoptic image acquisition in detail. Specifically, we begin by giving an overview of the plenoptic dimensions and show how much of this visual information is irreversibly lost in conventional image acquisition. We proceed by discussing the state of the art in joint optical modulation and computation reconstruction for the acquisition of high dynamic range imagery as well as spectral information. Two parts, focusing on light field acquisition and ultra-fast optics respectively, will unveil the secrets behind imaging techniques that have recently been

featured in the news. We outline other aspects of light that are of interest for various applications and wrap the course up with a short summary, while leaving enough time for questions and a short discussion.

Prerequisites

This introductory-level course has no prerequisites.

Speaker Biographies

Gordon Wetzstein

MIT Media Lab

gordonw@media.mit.edu

<http://web.media.mit.edu/~gordonw>

Gordon Wetzstein is a Postdoctoral Researcher at the MIT Media Lab. His research interests include light field and high dynamic range displays, projector-camera systems, computational optics, computational photography, computer vision, computer graphics, and augmented reality. Gordon received a Diplom in Media System Science with Honors from the Bauhaus-University Weimar in 2006 and a Ph.D. in Computer Science at the University of British Columbia in 2011. His doctoral dissertation focuses on computational light modulation for image acquisition and display. He is co-chairing the first workshop on Computational Cameras and Displays at CVPR 2012, is serving in the general submissions committee at SIGGRAPH 2012, has served on the program committees of IEEE ProCams 2007 and IEEE ISMAR 2010, won a Laval Virtual Award in 2005 for his work on projector-camera systems, and a best paper award for “Hand-Held Schlieren Photography with Light Field Probes” at the International Conference on Computational Photography in 2011, introducing light field probes as computational displays for computer vision and fluid mechanics applications.

Ivo Ihrke

Saarland University / MPI Saarbrücken

ihrke@mpi-inf.mpg.de

<http://giana.mmci.uni-saarland.de/>

Ivo Ihrke is head of the research group “Generalized Image Acquisition and Analysis” within the Cluster of Excellence “Multimodal Computing and Interaction” at Saarland University and Associate Senior Researcher at the MPI Informatik. Prior to joining Saarland University he was a post-doctoral research fellow at the University of British Columbia, Vancouver, Canada, supported by the Alexander von Humboldt-Foundation. He received a MS degree in Scientific Computing from the Royal Institute of Technology (KTH), Stockholm, Sweden (2002) and a PhD in Computer Science from Saarland University (2007). His main research interest

are the modeling of forward and inverse light transport processes and computational algorithms for solving these large scale problems.

Douglas Lanman

MIT Media Lab

dlanman@media.mit.edu

<http://web.media.mit.edu/~dlanman>

Douglas Lanman is a Postdoctoral Associate at the MIT Media Lab. His research is focused on computational imaging and display systems, including light field capture, automultiscopic (glasses-free) 3D displays, and active illumination for 3D reconstruction. He received a B.S. in Applied Physics with Honors from Caltech in 2002 and M.S. and Ph.D. degrees in Electrical Engineering from Brown University in 2006 and 2010, respectively. Prior to joining MIT and Brown, he was an Assistant Research Staff Member at MIT Lincoln Laboratory from 2002 to 2005. Douglas has worked as an intern at Intel, Los Alamos National Laboratory, INRIA Rhône-Alpes, Mitsubishi Electric Research Laboratories (MERL), and the MIT Media Lab. He presented the “Build Your Own 3D Scanner” course at SIGGRAPH 2009 and SIGGRAPH Asia 2009 and the “Build Your Own 3D Display” course at SIGGRAPH 2010, SIGGRAPH 2011, and SIGGRAPH Asia 2010. Douglas is co-chairing the first workshop on Computational Cameras and Displays at CVPR 2012, is serving on the technical paper committee for SIGGRAPH Asia 2012, and is currently writing a book on 3D Photography to be published by CRC Press in Fall 2012.

Wolfgang Heidrich

University of British Columbia

heidrich@cs.ubc.ca

<http://www.cs.ubc.ca/~heidrich>

Professor Wolfgang Heidrich holds the Dolby Research Chair in Computer Science at the University of British Columbia. He received a PhD in Computer Science from the University of Erlangen in 1999, and then worked as a Research Associate in the Computer Graphics Group of the Max-Planck-Institute for Computer Science in Saarbrücken, Germany, before joining UBC in 2000. Heidrich’s research interests lie at the intersection of computer graphics, computer vision, imaging, and optics. In particular, he has worked on High Dynamic Range imaging and display,

image-based modeling, measuring, and rendering, geometry acquisition, GPU-based rendering, and global illumination. Heidrich has written over 100 refereed publications on these subjects and has served on numerous program committees. He was the program co-chair for Graphics Hardware 2002, Graphics Interface 2004, the Eurographics Symposium on Rendering 2006, and ProCams 2011.

Kurt Akeley

Lytro, Inc.

kakeley@lytro.com

http://www.lytro.com/team/kurt_akeley

A pioneer in the field of computer graphics and a founding member of Silicon Graphics (later known as SGI), Kurt has helped develop innovative products like SGIs RealityEngine and the OpenGL graphics system. Hes contributed as well to 3D display technology, NVIDIA GPUs, and the Microsoft Research lab. He is an indefatigable traveler and prefers to be on the move when closer to home as well, often scheduling “walking meetings” while exploring the tranquil neighborhoods near Lytros Mountain View office. Kurt earned his Ph.D. in computer science from Stanford University and his B.E.E. from the University of Delaware. In 2005, he was awarded membership in the National Academy of Engineering. SIGGRAPH honored him with the Computer Graphics Achievement Award in 1995, and hes also been inducted as a Fellow of the Association for Computing Machinery.

Ramesh Raskar

MIT Media Lab

raskar@media.mit.edu

<http://web.media.mit.edu/~raskar>

Ramesh Raskar joined the Media Lab from Mitsubishi Electric Research Laboratories in 2008 as head of the Labs Camera Culture research group. His research interests span the fields of computational light transport, computational photography, inverse problems in imaging and human-computer interaction. Recent projects and inventions include transient imaging to look around a corner, a next generation CAT-Scan machine, imperceptible markers for motion capture (Prakash), long distance barcodes (Bokode), touch+hover 3D interaction displays (BiDi

screen), low-cost eye care devices (Netra,Catra), new theoretical models to augment light fields (ALF) to represent wave phenomena and algebraic rank constraints for 3D displays (HR3D).

He is a recipient of TR100 award from Technology Review, 2004, Global Indus Technovator Award, top 20 Indian technology innovators worldwide, 2003, Alfred P. Sloan Research Fellowship award, 2009 and Darpa Young Faculty award, 2010. Other awards include Marr Prize honorable mention 2009, LAUNCH Health Innovation Award, presented by NASA, USAID, US State Dept and NIKE, 2010, Vodafone Wireless Innovation Award (first place), 2011. He holds over 40 US patents and has received four Mitsubishi Electric Invention Awards. He is currently co-authoring a book on Computational Photography.

Course Outline

10 minutes: Introduction and Overview

Gordon Wetzstein

This part will introduce the speakers, present a motivation of the course, and outline the individual parts. We will also introduce the plenoptic function in this part and give an overview over its dimensions along with a variety of different applications for their acquisition.

10 minutes: High Dynamic Range Imaging

Wolfgang Heidrich

This part will review approaches to high dynamic range (HDR) image acquisition. The focus of this part of the course will be on techniques that exploit joint optical light modulation and computational reconstruction for HDR imaging.

10 minutes: Spectral Imaging

Ivo Ihrke

This part of the course will give a brief overview of multi-spectral and hyper-spectral imaging and then proceed to discuss computational photography techniques for spectral imaging in more detail.

10 minutes: Light Field Acquisition

Douglas Lanman

This part of the course reviews the general concepts behind light fields and discusses a variety of light field acquisition approaches, ranging from single-shot cameras (e.g., Lytro) to more exotic multi-camera and high-dimensional multiplexing approaches. This section also reviews emerging applications for light fields, including 3D displays, human-computer interaction, and medical imaging.

20 minutes: Inside the Lytro Light Field Camera

Kurt Akeley

With the recent launch of the first consumer light field camera by Lytro, computational cameras have started to enter the mass market. In this part of the course, we reveal the secrets behind the Lytro camera, which was listed by TIME Magazine as one of the 50 best inventions of 2011.

20 minutes: Light Capture with a Trillion Frames per Second

Ramesh Raskar

This part of the course will cover the latest developments in ultra-fast image acquisition. We will show how light propagation in free space can be captured at one trillion frames per second. We will show how ultra-fast image acquisition in combination with advanced computational reconstructions allows unprecedented applications such as looking around corners and measuring BRDFs in the wild.

5 minutes: Further Light Properties

Gordon Wetzstein

In this part of the course, we outline the acquisition of light properties that are not directly included in the plenoptic function but related, such as polarization and Schlieren imaging. We will show how joint optical light modulation and computational processing can also be useful for these applications.

5 minutes: Summary and Q & A

All

This part will summarize how computational plenoptic cameras are becoming the future of imaging by exploiting the co-design of display optics and computational processing. We will outline future directions of this emerging field and allow for sufficient time to answer questions and stimulate discussions.

Computational Plenoptic Imaging

SIGGRAPH 2012 Course Notes

Gordon Wetzstein¹

Ivo Ihrke²

Douglas Lanman¹

Wolfgang Heidrich³

Kurt Akeley⁴

Ramesh Raskar¹

¹MIT Media Lab

²Saarland University

³University of British Columbia

⁴Lytro, Inc.

The plenoptic function is a ray-based model for light that includes the color spectrum as well as spatial, temporal, and directional variation. Although digital light sensors have greatly evolved in the last years, one fundamental limitation remains: all standard CCD and CMOS sensors integrate over the dimensions of the plenoptic function as they convert photons into electrons; in the process, all visual information is irreversibly lost, except for a two-dimensional, spatially-varying subset — the common photograph. In this part of the course notes, we provide a review of approaches that optically encode the dimensions of the plenoptic function transcending those captured by traditional photography and reconstruct the recorded information computationally. The literature survey provided in this section is an updated version of a survey paper recently published by some of the authors [Wetzstein et al. 2011a]; a course on the dual topic of computational displays is being taught at SIGGRAPH 2012 as well [Wetzstein et al. 2012].

1 Introduction

Evolution has resulted in the natural development of a variety of highly specialized visual systems among animals. The mantis shrimp retina, for instance, contains 16 different types of photoreceptors [Marshall and Oberwinkler 1999]. The extraordinary anatomy of their eyes not only allows the mantis shrimp to see 12 different color channels, ranging from ultra-violet to infra-red, and distinguish between shades of linear and circular polarization, but it also allows the shrimp to perceive depth using trinocular vision with each eye. Other creatures of the sea, such as cephalopods [Mäthger et al. 2009], are also known to use their ability to perceive polarization for communication and unveiling transparency of their prey. Although the compound eyes found in flying insects have a lower spatial resolution compared to mammalian single-lens eyes, their temporal resolving power is far superior to the human visual system.

Traditionally, cameras have been designed to capture what a single human eye can perceive: a two-dimensional trichromatic image. Inspired by the natural diversity of perceptual systems and fueled by advances of digital camera technology, computational processing, and optical fabrication, image processing has begun to transcend limitations of film-based analog photography. Applications for the computerized acquisition of images with a high spatial, temporal, spectral, and directional resolution are manifold; medical imaging, remote sensing, shape reconstruction, surveillance, and automated fabrication are only a few examples.

The plenoptic function [Adelson and Bergen 1991] provides a ray-based model of light encompassing most properties that are of interest for image acquisition. As illustrated in Figure 1, these include the color spectrum as well as spatial, temporal, and directional light variation. In addition to these more traditional plenoptic dimensions [Adelson and Bergen 1991], we also consider dynamic range a desirable property, as common sensors have a limited dynamic range.

1.1 Computational Photography and Plenoptic Imaging

What makes plenoptic imaging different than general computational photography? Plenoptic imaging considers a subset of computational photography approaches; specifically, those that aim at acquiring the dimensions of the plenoptic function with combined optical light modulation and computational reconstruction. Computational pho-


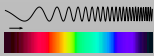
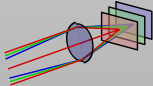
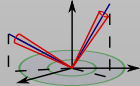
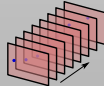



Plenoptic Dimension					
Acquisition Approach	Dynamic Range	Color Spectrum	Space Focal Surfaces	Directions Light Fields	Time
 Single Shot Acquisition	Assorted Pixels Gradient Camera Adaptive DR Imaging	Color Filter Arrays Assorted Pixels Dispersive Optics	Coded Apertures Focal Sweep Field Correction	Plenoptic Cameras w/ Lenses, Masks, or Mirrors Compound Eye Cameras	Assorted Pixels Flutter Shutter Reinterpretable Imager Sensor Motion
 Sequential Image Capture	Exposure Brackets Generalized Mosaics HDR Video	Narrow Band Filters Generalized Mosaicing Agile Spectrum Imaging	Focal Stack Jitter Camera Super-Resolution	Programmable Aperture Camera & Gantry	High-Speed Imaging Temporal Dithering
 Multi-Device Setup	Split Aperture Imaging Optical Splitting Trees	Multi-Camera Arrays Optical Splitting Trees	Multi-Camera Arrays	Multi-Camera Arrays	Multi-Camera Arrays Hybrid Cameras

Figure 1: Taxonomy and overview of plenoptic image acquisition approaches.

tography has grown tremendously in the last years with dozens of published papers per year in a variety of graphics, vision, and optics venues. The dramatic rise in publications in this interdisciplinary field, spanning optics, sensor technology, image processing, and illumination, has made it difficult to encompass all research in a single survey.

We provide a structured review of the subset of research that has recently been shown to be closely-related in terms of optical encoding and especially in terms of reconstruction algorithms [Ihrke et al. 2010a]. Additionally, our report serves as a resource for interested parties by providing a categorization of recent research and is intended to aid in the identification of unexplored areas in the field.

1.2 Overview and Definition of Scope

In this report, we review the state of the art in joint optical light modulation and computational reconstruction approaches for acquiring the dimensions of the plenoptic function. Specifically, we discuss the acquisition of high dynamic range imagery (Section 2), the color spectrum (Section 3), light fields and directional variation (Section 4), spatial super-resolution and focal surfaces (Section 5), as well as high-speed events (Section 6). We also outline the acquisition of light properties that are not directly included in the plenoptic function, but related, such as polarization, phase imaging, and time-of-flight (Section 7) and point the reader to more comprehensive literature on these topics. Conclusions and possible future avenues of research are discussed in Section 8.

Due to the fact that modern, digital acquisition approaches are often closely related to their analog predecessors, we outline these whenever applicable. For each of the plenoptic dimensions we also discuss practical applications of the acquired data. As there is an abundance of work in this field, we focus on imaging techniques that are designed for standard planar sensors. We will only highlight examples of modified sensor hardware for direct capture of plenoptic image information. We do not cover pure image processing techniques, such as tone-reproduction, dynamic range compression and tone-mapping [Reinhard et al. 2010], or the reconstruction of geometry [Ihrke et al. 2010b], BSDFs and reflectance fields.

2 High Dynamic Range Imaging

High dynamic range (HDR) image acquisition has been a very active area of research for more than a decade. With the introduction of the HDR display prototype [Seetzen et al. 2004] and its successor models becoming consumer products today, the demand for high-contrast photographic material is ever increasing. Other applications for high dynamic range imagery include digital photography, physically-based rendering and lighting [Debevec 2002], image

editing, digital cinema, perceptual difference metrics based on absolute luminance [Mantiuk et al. 2005; Mantiuk et al. 2011], virtual reality, and computer games. For a comprehensive overview of HDR imaging, including applications, radiometry, perception, data formats, tone reproduction, and display, the reader is referred to the textbook by Reinhard et al. [Reinhard et al. 2010]. In this section, we provide a detailed and up-to-date list of approaches for the acquisition of high dynamic range imagery.

2.1 Single-Shot Acquisition

According to DxOMark (www.dxomark.com), the latest high-end digital SLR cameras are equipped with CMOS sensors that have a measured dynamic range of up to 13.5 f-stops, which translates to a contrast of 11,000:1. This is comparable to that of color negative films [Reinhard et al. 2010]. In the future, we can expect digital sensors to perform equally well as negative film in terms of dynamic range, but this is not the case for most sensors today.

Specialized sensors that allow high dynamic range content to be captured, have been commercially available for a few years. These include professional movie cameras, such as Grass Valley’s Viper [Valley 2010] or Panavision’s Genesis [Panavision 2010]. The SpheroCam HDR [SpheronVR 2010] is able to capture full spherical 360-degree images with 26 f-stops and 50 megapixels in a single scan. A technology that allows per-pixel exposure control on the sensor, thereby enabling adaptive high dynamic range capture, was introduced by Pixim [Pixim 2010]. This level of control is achieved by including an analog-to-digital converter for each pixel on the sensor.

Capturing image gradients rather than actual pixel intensities was shown to increase the dynamic range of recorded content [Tumblin et al. 2005]. In order to reconstruct intensity values, a computationally expensive Poisson solver needs to be applied to the measured data. While a *Gradient Camera* is an interesting theoretical concept, to the knowledge of the authors this camera has never actually been built.

The maximum intensity that can be resolved with standard ND filter arrays is limited by the lowest transmission of the employed ND filters. Large, completely saturated regions in the sensor image are usually filled with data interpolated from neighboring unsaturated regions [Nayar and Mitsunaga 2000]. An analysis of sensor saturation in multiplexed imaging along with a Fourier-based reconstruction technique that boosts the dynamic range of captured images beyond the previous limits was recently proposed [Wetzstein et al. 2010]. Figure 2 shows an example image that is captured with an ND filter array on the left and a Fourier-based reconstruction of multiplexed data on the right.

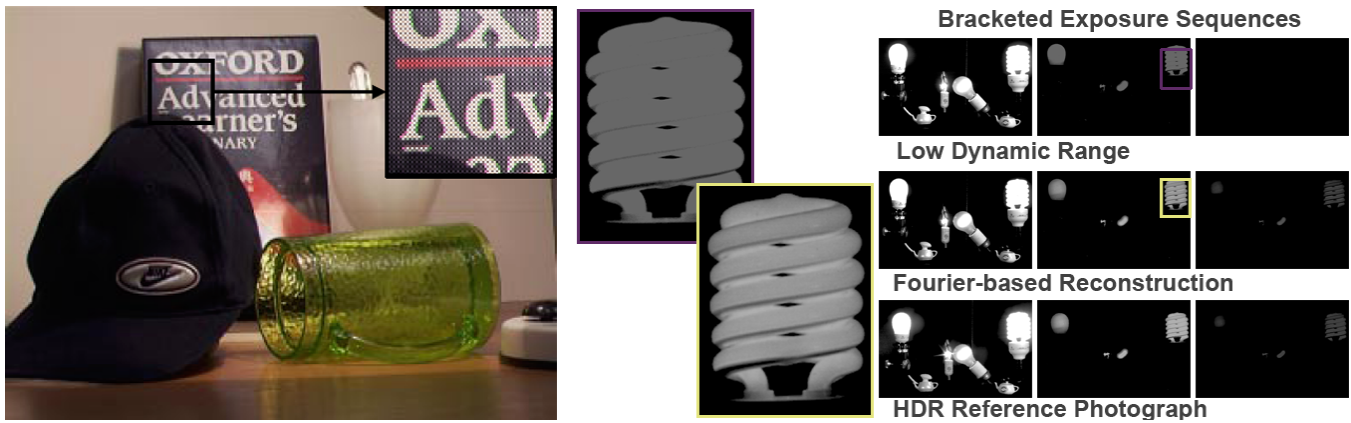


Figure 2: Sensor image captured with an array of ND filters [Nayar and Mitsunaga 2000] (left). Exposure brackets and magnifications for Fourier-based HDR reconstruction from multiplexed sensor images [Wetzstein et al. 2010] (right).

An alternative to mounting a fixed set of ND filters in front of the sensor is an aligned spatial light modulator, such as a digital micromirror device (DMD). This concept was explored as *Programmable Imaging* [Nayar et al. 2004; Nayar

et al. 2006] and allows for adaptive control over the exposure of each pixel. Unfortunately, it is rather difficult to align a DMD with a sensor on a pixel-precise basis, partly due to the required additional relay optics; for procedures to precisely calibrate such a system please consult [Ri et al. 2006]. Although a transmissive spatial light modulator can, alternatively, be mounted near the aperture plane of the camera, as proposed by Nayar and Branzoi [Nayar and Branzoi 2003], this *Adaptive Dynamic Range Imaging* approach only allows lower spatial frequencies in the image to be modulated. The most practical approach to adaptive exposures is a per-pixel control of the readout in software, as implemented by the Pixim camera [Pixim 2010]. This has also been simulated for the specific case of CMOS sensors with rolling shutters [Gu et al. 2010], but only on a per-scanline basis. The next version of the *Frankencamera* [Adams et al. 2010] is planned to provide non-destructive sensor readout for small image regions of interest [Levoy 2010], which would be close to the desired per-pixel exposure control.

Rouf et al. [Rouf et al. 2011] proposed to encode both saturated highlights and low-dynamic range content in a single sensor image using cross-screen filters. Computerized tomographic reconstruction techniques were employed to estimate the saturated regions from glare created by the optical filters.

2.2 Multi-Sensor and Multi-Exposure Techniques

The most straightforward way of acquiring high dynamic range images is to sequentially capture multiple photographs with different exposure times and merge them into a single, high-contrast image [Mann and Picard 1995; Debevec and Malik 1997; Mitsunaga and Nayar 1999; Robertson et al. 1999]. Some of these approaches simultaneously compute the non-linear camera response function from the image sequence [Debevec and Malik 1997; Mitsunaga and Nayar 1999; Robertson et al. 1999]. Extensions to these techniques also allow HDR video [Kang et al. 2003]. Here, successive frames in the video are captured with varying exposure times and aligned using optical flow algorithms. Today, all of these methods are well established and discussed in the textbook by Reinhard et al. [Reinhard et al. 2010].

In addition to capturing multiple exposures, a static filter with varying transmissivity, termed *Generalized Mosaicing* [Schechner and Nayar 2003b], can be mounted in front of the camera but also requires multiple photographs to be captured. Alternatively, the optical path of an imaging device can be divided using prisms [Aggarwal and Ahuja 2004] (*Split Aperture Imaging*) or beam-splitters [McGuire et al. 2007] (*Optical Splitting Trees*), so that multiple sensors capture the same scene with different exposure times. While these approaches allow dynamic content to be recorded, the additional optical elements and sensor hardware make them more expensive and increase the form factor of the device.

2.3 Analysis and Tradeoffs

Given a camera with known response function and dynamic range, Grossberg and Nayar [Grossberg and Nayar 2003] analyze the best possible set of actual exposure values for a low dynamic range (LDR) image sequence used to compute an HDR photograph. By also considering variable ISO settings, Hasinoff et al. [Hasinoff et al. 2010] provide the optimal choice of parameters for HDR acquisition with minimal noise.

3 Spectral Imaging

Imaging of the electromagnetic spectrum comes in a number of flavors. For photographs or movies, the goal is typically to capture the colors perceived by the human visual system. Since the human visual system is based on three types of color sensing cells (the cones), three color bands are sufficient to form a natural color impression. This discovery is usually credited to Maxwell [Maxwell 1860].

In this report we are mainly concerned with methods for capturing the physical properties of light in contrast to their perceptual counterparts that are dealt with in the areas of Applied Perception and Color Sciences. For readers interested in issues of color perception, we refer to standard literature: Wyszeski and Stiles [Wyszecki and Stiles

1982] provide raw data for many perceptual experiments. Fairchild's book [Fairchild 2005] is a higher-level treatise focusing on models for perceptual effects as, for instance, adaptation issues. Hunt's books [Hunt 1991; ?] deal with measurement and reproduction of color for human observers (e.g., in digital imaging, film, print, and television). Reinhard et al. [Reinhard et al. 2008] discuss color imaging from a computer graphics perspective.

In this section we discuss spectral imaging from a radiometric, i.e. physical, perspective. To simplify the discussion we first introduce some terminology as used in this subfield of plenoptic imaging.

3.1 Glossary of Terms

Spectral Radiance is the physical quantity emitted by light sources or reflected by objects. The symbol is L_λ and its unit is $[W/m^2 \cdot sr \cdot nm]$. Spectral radiance is constant along a ray. It is the quantity returned by the plenoptic function.

Spectral Filters selectively attenuate parts of the electromagnetic spectrum. There are two principles of operation, *absorptive* spectral filters remove parts of the spectrum by converting photons into kinetic energy of the atoms constituting the material. Interference-based filters, also referred to as *dichroic* filters, consist of a transparent substrate which is coated with thin layers that selectively reflect light, reinforcing and attenuating different wavelengths in different ways. The number and thicknesses of the layers determine the spectral reflection profile. Absorptive filters have a better angular constancy, but heating may be an issue for narrow-band filters. Interference-based filters have the advantage that the spectral filter curve can be designed within certain limits by choosing the parameters of the coatings. However, the angular variation of these filters is significant. In general, filters are available both for transmission and reflection modes of operation.

Narrow-Band Filters have a small support in the wavelength domain.

Broad-Band Filters have a large support in the wavelength domain. They are also known as *panchromatic* filters.

The **Spectral Response Curve** of a sensor is a function that describes its quantum efficiency with respect to photons of different wavelengths. A higher value means a better response of the sensor to photons of a particular wavelength, i.e. more electrons are freed due to the photo-electric effect.

Color is the perceptual interpretation of a given electro-magnetic spectrum.

The **Gamut** of an imaging or color reproduction system is the range of correctly reproducible colors.

Multi-Spectral Images typically consist of a low number of spectral bands. They often include a near infrared (NIR) band. The bands typically do not form a full spectrum, there can be missing regions [Vagni 2007].

Hyper-Spectral Images contain thirty to several hundred spectral bands which are approximations to the full spectrum [Vagni 2007]. The different spectral bands do not necessarily have the same spatial resolution. In this report, we will use the term multi-spectral to refer both to multi-spectral and hyper-spectral image acquisition methods.

A **Multi-Spectral Data Cube** is a stack of images taken at different wavelength bands.

3.2 Color Imaging

In a limited sense, the most common application of multi-spectral imaging is the acquisition of color images for human observers. In principle, three spectral bands mimicking the human tri-stimulus system, are sufficient to capture color images. This principle was first demonstrated by Maxwell performing color photography by time-sequential acquisition of three images using different band-pass filters, see Figure 3. Display was achieved by superimposed spectrally filtered black-and-white projections using the same filters as used for capture. This acquisition principle was in use for quite some time until practical film-based color photography was invented. One of the

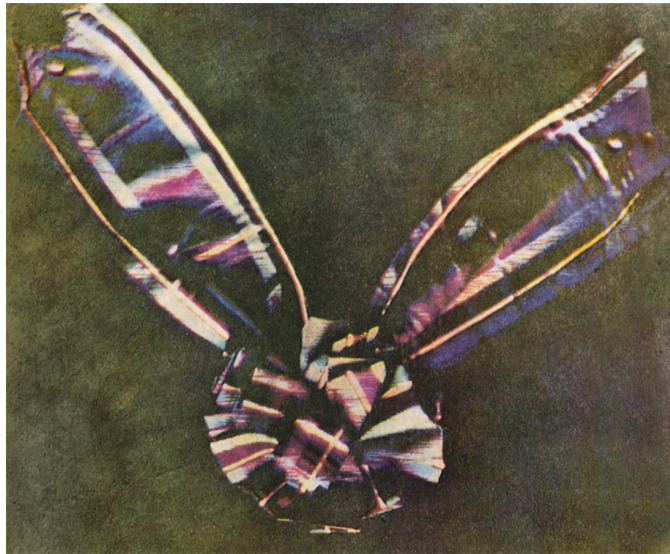


Figure 3: *Tartan Ribbon*, considered to be the world's first color photograph, taken by Thomas Sutton for James Clerk Maxwell in 1861 by successively placing three color filters in front of the camera's main lens and taking three monochromatic photographs (Wikimedia Commons).

earliest collections of color photographs was assembled by the Russian photographer Sergej Mikhailovich Prokudin-Gorskij [Prokudin-Gorskii 1912]. Time-sequential imaging through different filters is still one of the main modes of capturing multi-spectral images (see Sec. 3.3).

In the digital age, color films have been replaced by electronic CMOS or CCD sensors. The two technologies to capture an instantaneous color image are optical splitting trees employing dichroic beam-splitter prisms [Optec 2011], as used in three-CCD cameras, and spatial multiplexing [Narasimhan and Nayar 2005; Ihrke et al. 2010a], trading spatial resolution for color information. The spatially varying spectral filters in multiplexing applications are also known as color filter arrays (CFAs). A different principle, based on volumetric, or layered measurements is employed by the Foveon sensor [Foveon 2010], which captures tri-chromatic images at full spatial resolution.

The most popular spatial multiplexing pattern is the well known Bayer pattern [Bayer 1976]. It is used in most single-sensor digital color cameras. The associated problem of reconstructing a full-resolution color image is generally referred to as demosaicing. An overview of demosaicing techniques is given in [Ramanath et al. 2002; Gunturk et al. 2005; Li et al. 2008b]. Li et al. [Li et al. 2008b] present a classification scheme of demosaicing techniques depending on the prior model being used (explicitly or implicitly) and an evaluation of different classes of algorithms. An interesting result is that the common constant-hue assumption seems to be less valid for modern imagery with a wider gamut and higher dynamic range than the classical Kodak photo CD test set [Eastman Kodak Company], which was scanned from film and has predominantly been used for evaluating demosaicing algorithms. Mostly, demosaicing is evaluated through simulation. However, in a realistic setting, including camera noise, Hiraakawa and Parks [Hiraakawa and Parks 2006] have shown that demosaicing on noisy images performs poorly and that subsequent denoising is affected by demosaicing artifacts. They propose a joint denoising and demosaicing framework that can be used with different demosaicing and denoising algorithms.

In recent years, a large number of alternative CFAs have been explored by camera manufacturers, some of which are already being used in consumer products. Examples and many references to alternative CFA designs can be found in Hiraakawa and Wolfe's work [Hiraakawa and Wolfe 2008]. Traditionally, imaging through CFAs and reconstruction of the signal have been seen as sub-sampling and up-sampling operations, respectively. Recent research in the analysis of these multiplexing patterns has, however, produced a new view of multiplexing as a projection onto a linear subspace of basis functions (the spectral responses of the filters in this case), i.e. of multiplexing as a coding

operation [Lu and Vetterli 2009; Ihrke et al. 2010a]. Correspondingly, in this view, the reconstruction process is seen as a recovery of the subspace, or a decoding of the signal. This view originated in Fourier analysis of color filter arrays [Alleyson et al. 2005], stimulated by the desire to apply digital signal processing methodology to the color multiplexing problem. Being a linear framework, it allows for the optimization of the subspace onto which color is projected [Hirakawa and Wolfe 2007; Hirakawa and Wolfe 2008; Lu and Vetterli 2009]. Practical realizations are alternative CFA designs that suffer from less aliasing than their ad-hoc, heuristically-designed counterparts. While in [Hirakawa and Wolfe 2007; Hirakawa and Wolfe 2008; Lu and Vetterli 2009] a fixed number of primary color response functions are assumed which can be linearly mixed to optimize the CFA, [Parmar and Reeves 2006; Parmar and Reeves 2010] optimize the spectral response functions themselves in order to improve CFA design. More recently, [Sajadi et al. 2011] proposed to use shiftable layers of CFAs; this design allows the color primaries to be switched dynamically and provides an optimal SNR in different lighting conditions.

Generalizing color filter arrays, Narasimhan and Nayar [Narasimhan and Nayar 2005] proposed the *Assorted Pixels* framework, where individual pixels can be modulated by arbitrary plenoptic filters, yielding an image mosaic that has to be interpolated to obtain the full-resolution multi-channel image. In subsequent work [Yasuma et al. 2010], aliasing within the *Assorted Pixels* framework was minimized. Ihrke et al. [Ihrke et al. 2010a] have shown how this (and other) approaches that are tiling the image in a super-pixel fashion can be interpreted as belonging to one group of imaging systems that share common analysis and reconstruction approaches.

In a different application, Wetzstein et al. [Wetzstein et al. 2010] explore CFA designs that lead to dynamic range constraints in Fourier space. They show how an optimized CFA pattern in conjunction with optimization algorithms allow trichromatic high dynamic range images to be captured.

3.3 Multi-Spectral Imaging

As for the other plenoptic dimensions, the three basic approaches of Figure 1, single-shot capture, sequential image acquisition, and multiple device setups are valid alternatives for multi-spectral imaging and have been investigated intensely.

3.3.1 Spectrometers

Traditionally, spectroscopy has been carried out for single rays entering an instrument referred to as spectrometer. It was invented by Joseph von Fraunhofer in 1814 and used to discover the missing lines in the solar spectrum bearing his name. Typically the ray is split into its constituent wavelengths which are displaced spatially. This is achieved by placing either dispersive or diffractive elements into the light path, where the latter come both in transmissive and reflective variants. If dispersion is used to split the ray, typically a prism is employed. The separation of wavelengths is caused by the wavelength-dependent refractive index of the prism material. The function mapping wavelength to refractive index is typically decreasing with increasing wavelength, but usually in a non-linear fashion. Under certain conditions, it can even have an inverted slope (anomalous dispersion) [Hecht 2002]. Diffractive elements are usually gratings, where maxima of the diffraction pattern are spatially shifted according to the grating equation [Hecht 2002]. After the light path is split by some means, the light is brought onto a photo-detector which can, for instance, be a CCD. Here, relative radiance of the individual wavelengths is measured. Spectrometers have to be calibrated in two ways: first, the mapping of wavelengths to pixels has to be determined. This is usually done using light sources with few and very narrow emission lines of known wavelength, the pixel positions of other wavelengths are then interpolated [Gaigalas et al. 2009]. The second step establishes the relative irradiance measured for every wavelength. This is done by measuring a surface of known flat reflectance, for example Spectralon, which is illuminated with a known broad-band spectrum. The relative inhomogeneities imposed by the device are then divided out [Gaigalas et al. 2009]. Spectrometers that are designed to image more than one ray are referred to as *imaging spectrometers*.

3.3.2 Scanning Imaging Spectrometers

Traditional devices are usually based on some form of scanning. Either a full two-dimensional image is acquired with changed band-pass filters, effectively performing a spectral scan, or a pushbroom scan is performed where the two-dimensional CCD images a spatial dimension on one axis of the image and the spectral dimension on the other. The full multi-spectral data cube is then obtained by scanning the remaining spatial dimension.

Spectral Scanning can be performed in a variety of ways. Most of them involve either a filter wheel (e.g., [Wang and Heidrich 2004]) or electronically tunable filters (ETFs). The former method usually employs narrow-band filters such that spectral bands are imaged directly. The disadvantage is a low light throughput. Toyooka and Hayasaka [Toyooka and Hayasaka 1997] present a system based on broad-band filters with computational inversion. Whether or not this is advantageous depends on the camera noise [Ihrke et al. 2010a]. Electronically tunable filters are programmable devices that can exhibit varying filter curves depending on control voltages applied to the device. Several incarnations exist; the most well known include Liquid Crystal Tunable Filters (LCTFs) [cri inc 2009], which are based on a cascade of Lyot-filter stages [Lyot 1944], acousto-optical tunable filters, where an acoustically excited crystal serves as a variable diffraction grating, and interferometer-based systems, where the spectrum is projected into the Fourier basis. In the latter, the spectral scan is performed in a multiplexing manner: by varying the position of the mirror in one arm of an interferometer, for instance a Michelson-type device, different phase shifts are induced for every wavelength. The resulting spectral modulation is in the form of a sinusoid. Thus, effectively, the measurements are performed in the Fourier basis, similar to the Dappled Photography technique [Veeraraghavan et al. 2007] for light fields (see Sec. 4). The spectrogram is obtained by taking an inverse Fourier transform. A good overview of these technologies is given in [Gat 2000].

A more flexible way of programmable wavelength modulation is presented by Mohan et al. [Mohan et al. 2008b]. They modulate the spectrum of an entire image by first diffracting the light and placing an attenuating mask into the “rainbow plane” of the imaging system. However, the authors do not recover multiplexed spectra but only demonstrate modulated spectrum imaging. Usually, the scanning and the imaging process have to be synchronized, i.e. the camera should only take an image when the filter in front of the camera is set to a known value. Schechner and Nayar [Schechner and Nayar 2004] introduce a technique to computationally synchronize video streams taken with a periodically moving spectral filter.

All previously discussed techniques attempt to recover scene spectra passively. An alternative technique using active spectral illumination in a time-sequential manner is presented by Park et al. [Park et al. 2007]. The scene, which can include ambient lighting, is imaged under different additional spectral lighting. The acquired images allow for reasonably accurate per-pixel spectra to be recovered.

Spatial Scanning has been widely employed in satellite-based remote sensing. Two technologies are commonly used: pushbroom and whiskbroom scanning. Whereas pushbroom scanning uses a two-dimensional sensor and can thus recover one spectral and one spatial dimension per position of the satellite, whiskbroom systems employ a one-dimensional sensor, imaging the spectrum of a single point which is then scanned to obtain a full scan-line with a rotating mirror. The main idea is that a static scene can be imaged multiple times using different spectral bands and thus a full multi-spectral data cube can be assembled. A good overview of space-borne remote sensing, and more generally, multi-spectral imaging techniques is given in [Vagni 2007]

In computer vision, a similar concept, called *Generalized Mosaicing*, has been introduced by Schechner and Nayar [Schechner and Nayar 2001]. Here, a spatially varying filter is mounted in front of the main lens, filtering each column of the acquired image differently. By moving the camera and registering the images, a full multi-spectral data cube can be recovered [Schechner and Nayar 2002].

3.3.3 Single-Shot Imaging Spectrometers

To enable the spectral acquisition of fast-moving objects, it is necessary to have single-shot methods available. Indeed, this appears to be the focus of research in recent years. We can differentiate between three major modes of operation. The first is a trade of spatial for spectral resolution. Optical devices are implemented that provide empty space on the sensor which can, with a subsequent dispersion step through which a scene ray is split into its wavelength constituents, be filled with spectral information. The second option is multi-device setups which operate mostly like their spectral scanning counterparts, replacing sequential imaging by additional hardware. The third class of devices employs computational imaging, i.e. computational inversion of an image formation process where spectral information is recorded in a super-imposed manner.

Spatial Multiplexing of the spectrum, in general, uses a dispersive or diffractive element in conjunction with some optics redirecting rays from the scene onto parts of the sensor surrounded by void regions. The void regions are then filled with spectral information. All these techniques take advantage of the high resolution of current digital cameras. Examples using custom manufactured redirecting mirrors include [Harvey et al. 2005; Gao et al. 2009; Gorman et al. 2010]. These techniques achieve imaging of up to 25 spectral bands in real-time and keep the optical axis of the different slices of the multi-spectral data cube constant. Bodkin et al. [Bodkin et al. 2009] and Du et al. [Du et al. 2009] propose a similar concept by using an array of pinholes that limits the rays that can reach the sensor from the scene. The pinholes are arranged such that a prism following in the optical path disperses the spectrum and fills the pixels with spectral information. A different approach is taken by Fletcher-Holmes et al. [Fletcher-Holmes and Harvey 2005]. They are interested in only providing a small “foveal region” in the center of the image with multi-spectral information. For this, the center of the image is probed with fiber optic cables which are fed into a standard spectrometer. Mathews et al. [Mathews 2008] and Horstmeyer et al. [Horstmeyer et al. 2009] describe light field cameras with spectrally filtered sub-images. An issue with this design is the problem of motion parallax induced by the different view points when registering the images (see Sec. 4). In general, this registration problem is difficult and requires knowledge of scene geometry and reflectance which cannot easily be estimated.

Multi-Device Setups are similar in spirit to spectral scanning spectrometers, replacing the scanning process by additional hardware. A straightforward solution recording five spectral bands is presented by Lau et al. [Lau and Yang 2005]. They use a standard multi-video array where different spectral filters are mounted on each camera. The motion-parallax problem mentioned previously is even worse in this case. McGuire et al. [McGuire et al. 2007] discuss optical splitting trees where the individual sensors are aligned such that they share a single optical axis. The design of beam-splitter/filter trees is non-trivial and the authors propose an automatic solution based on optimization. A hybrid camera was proposed by Cao et al. [Cao et al. 2011], where the output of a high resolution RGB camera was combined with that of a prism-based low spatial-resolution, high spectral-resolution camera.

Computational Spectral Imaging aims at trading computational complexity for simplified optical designs. *Computed Tomography Image Spectrometry* (CTIS) was developed by Okamoto and Yamaguchi [Okamoto and Yamaguchi 1991]. They observed that by placing a diffraction grating in the optical path, several spectral copies overlay on the image sensor. Every pixel is measuring a line integral along the spectral axis. Knowing the imaging geometry enables a tomographic reconstruction of the spectra. A drawback to this technique is that not all data can be measured and thus an ill-conditioned problem, similar to limited angle tomography, is encountered. The technique was extended to single shot imaging by Descour et al. [Descour and Dereniak 1995; Descour et al. 1997].

A relatively novel technique is referred to as *Coded Aperture Snapshot Spectral Imaging* (CASSI) [Gehm et al. 2007; Wagadarikar et al. 2008; Wagadarikar et al. 2009]. In a series of papers the authors show how to construct different devices to exploit the compressive sensing paradigm [Candès et al. 2006] which promises to enable higher resolution computational reconstructions with less samples than predicted by the Shannon-Nyquist sampling theorem.

The results presented for both CTIS and CASSI have only been demonstrated for relatively low-resolution, low-quality spectral images. Therefore, these approaches are not yet suitable for high-quality photographic applications.

3.4 Applications

There is a huge amount of applications for multi-spectral imaging and we are just beginning to explore the possibilities in computer graphics and vision. Traditional users of multi-spectral imaging technology are in the fields of astronomy and remote sensing where, for instance, the mapping of vegetation, minerals, water surfaces, and hazardous waste monitoring are of interest. In addition, multi-spectral imaging is used for material discrimination [Du et al. 2009], ophthalmology [Lawlor et al. 2002], the study of combustion dynamics [Hunicz and Piernikarski 2001], cellular dynamics [Kindzelskii et al. 2000], surveillance [Harvey et al. 2000], for deciphering ancient scrolls [Mansfield 2005], flower photography [Rorslett 2008], medicine, agriculture, manufacturing, forensics, and microscopy. It should not be forgotten that the military is an interested party [Vagni 2007].

4 Light Field Acquisition

While the 5D plenoptic function parameterizes all possible images of a general scene, being a slice of constant time and wavelength of the full plenoptic function, it remains problematic to capture its directional variation over a wide field of view. Consider the inside of a ceramic vase; a camera must be inserted inside to capture the intensity and color of light rays traveling entirely within the concavity. However, if the viewer is restricted to move through an isotropic, transparent medium (e.g., air or water) outside of the convex hull of a given object, then the plenoptic function can be measured by translating a digital camera throughout the allowed viewer region. Levoy and Hanrahan [Levoy and Hanrahan 1996] and Gortler et al. [Gortler et al. 1996] realized that, when the viewer is restricted to move outside the convex hull, the 5D plenoptic function possesses one dimension of redundancy: the radiance of a given ray does not change in free space. Thus, in a region free of occluders, the 5D plenoptic function can be expressed as a 4D *light field*.

The concept of a light field predates its introduction in computer graphics. The term itself dates to the work of Gershun [Gershun 1936], who derived closed-form expressions for illumination patterns projected by area light sources. Ashdown [Ashdown 1993] continued this line of research. Moon and Spencer [Moon and Spencer 1981] introduced the equivalent concept of a *photic field* and applied it to topics spanning lighting design, photography, and solar heating. The concept of a light field is similar to epipolar volumes in computer vision [Bolles et al. 1987]. As demonstrated by Halle [Halle 1994], both epipolar volumes and holographic stereograms can be captured by uniform camera translations. The concept of capturing a 4D light field, for example by translating a single camera [Levoy and Hanrahan 1996; Gortler et al. 1996] or by using an array of cameras [Wilburn et al. 2002], is predated by integral photography [Lippmann 1908], parallax panoramagrams [Ives 1903], and holography [Gabor 1948].

This section catalogues existing devices and methods for light field capture, as well as applications enabled by such data sets. Note that a sensor pixel in a conventional camera averages the radiance of light rays impinging over the full hemisphere of incidence angles, producing a 2D projection of the 4D light field. In contrast, light field cameras prevent such averaging by introducing spatio-angular selectivity. Such cameras can be classified into those that primarily rely on multiple sensors or a single sensor augmented by temporal, spatial, or frequency-domain multiplexing.

4.1 Multiple Sensors

As described by Levoy and Hanrahan [Levoy and Hanrahan 1996], a light field can be measured by capturing a set of photographs taken by an array of cameras distributed on a planar surface. Each camera measures the radiance of light rays incident on a single point, defined in the plane of the cameras, for a set of angles determined by the field of view of each camera. Thus, each camera records a 2D slice of the 4D light field. Concatenating these

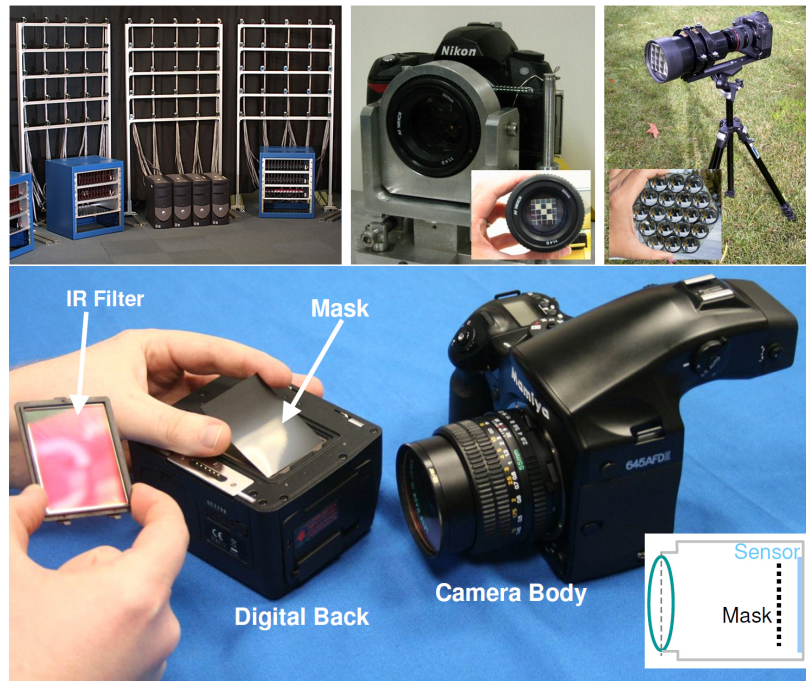


Figure 4: Light field cameras can be categorized by how a 4D light field is encoded in a set of 2D images. Methods include using multiple sensors or a single sensor with temporal, spatial, or frequency-domain multiplexing. (Top, Left) Wilburn et al. [Wilburn et al. 2002] describe a camera array. (Top, Middle) Liang et al. [Liang et al. 2008] achieve temporal multiplexing with a programmable aperture. (Top, Right) Georgiev et al. [Georgiev et al. 2008] capture spatially-multiplexed light fields using an array of lenses and prisms. (Bottom) Raskar et al. [Raskar et al. 2008] capture frequency-multiplexed light fields by placing a heterodyne mask [Veeraraghavan et al. 2007; Veeraraghavan et al. 2008; Lanman et al. 2008] close to the sensor. (Figures reproduced from [Wilburn et al. 2002], [Liang et al. 2008], [Georgiev et al. 2008], and [Raskar et al. 2008].)

slices yields an estimate of the light field. Wilburn et al. [Wilburn et al. 2002; Wilburn et al. 2005] achieve dynamic light field capture using an array of up to 125 digital video cameras (see Figure 4, left). Yang et al. [Yang et al. 2002] propose a similar system using 64 cameras. Nomura et al. [Nomura et al. 2007] create scene collages using up to 20 cameras attached to a flexible plastic sheet, combining the benefits of both multiple sensors and temporal multiplexing. Custom hardware allows accurate calibration and synchronization of the camera arrays. Such designs have several unique properties. Foremost, as demonstrated by Vaish et al. [Vaish et al. 2006], the captured light field can be considered as if it were captured using a single camera with a main lens aperture extending over the region occupied by the cameras. Such large-format cameras can not be practically constructed using refractive optics. Vaish et al. exploit this configuration by applying methods of synthetic aperture imaging to obtain sharp images of objects obscured by thick foliage.

4.2 Temporal Multiplexing

Camera arrays have several significant limitations; foremost, a sparse array of cameras may not provide sufficient light field resolution for certain applications. In addition, the cost and engineering complexity of such systems prohibit their use for many consumer applications. As an alternative, methods using a single image sensor have been developed. For example, Levoy and Hanrahan [Levoy and Hanrahan 1996] propose a direct solution; using a mechanical gantry, a single camera is translated over a spherical or planar surface, constantly reoriented to point towards the object of interest. Alternatively, the object can be mechanically rotated on a computer-controlled turntable. Ihrke et al. [Ihrke et al. 2008] substitute mechanical translation of a camera with rotation of a planar mirror, effectively

creating a time-multiplexed series of virtual cameras. Thus, by distributing the measurements over time, single-sensor light field capture is achieved. Taguchi et al. [Taguchi et al. 2010a] show how capturing multiple images of rotationally-symmetric mirrors from different camera positions allow wide field of view light fields to be captured. Gortler et al. [Gortler et al. 1996] propose a similar solution; the camera is manually translated and computer vision algorithms are used to estimate the light field from such uncontrolled translations. These approaches trace their origins to the method introduced by Chen and Williams [Chen and Williams 1993], which is implemented by QuickTime VR.

The preceding systems capture the light field impinging on surfaces enveloping large regions (e.g., a sphere encompassing the convex hull of a sculpture). In contrast, hand-held light field photography considers capturing the light field passing through the main lens aperture of a conventional camera. Adelson and Wang [Adelson and Wang 1992], Okano et al. [Okano et al. 1999], and Ng et al. [Ng et al. 2005] extend integral photography to spatially multiplex a 4D light field onto a 2D image sensor, as discussed in the following subsection. However, temporal multiplexing can also achieve this goal.

Liang et al. [Liang et al. 2008] propose programmable aperture photography to achieve time-multiplexed light field capture. While Ives [Ives 1903] uses static parallax barriers placed close to the image sensor, Liang et al. use dynamic aperture masks (see Figure 4, middle). For example, consider capturing a sequence of conventional photographs. Between each exposure a pinhole aperture is translated in raster scan order. Each photograph records a pencil of rays passing through a pinhole located at a fixed position in the aperture plane for a range of sensor pixels. Similar to multiple sensor acquisition schemes, each image is a 2D slice of the 4D light field and the sequence can be concatenated to estimate the radiance for an arbitrary light ray passing through the aperture plane. To reduce the necessary exposure time, Liang et al. further apply Hadamard aperture patterns, originally proposed by Schechner and Nayar [Schechner et al. 2007], that are 50% transparent.

The preceding methods all consider conventional cameras with refractive lens elements. Zhang and Chen [Zhang and Chen 2005] propose a lensless light field camera. In their design, a bare sensor is mechanically translated perpendicular to the scene. The values measured by each sensor pixel are recorded for each translation. By the Fourier projection-slice theorem [Ng 2005], the 2D Fourier transform of a given image is equivalent to a 2D slice of the 4D Fourier transform of the light field; the angle of this slice is dependent on the sensor translation. Thus, tomographic reconstruction yields an estimate of the light field using a bare sensor, mechanical translation, and computational reconstruction methods.

4.3 Spatial and Frequency Multiplexing

Time-sequential acquisition reduces the cost and complexity of multiple sensor systems, however it has one significant limitation: dynamic scenes cannot be readily captured. Thus, either a high-speed camera is necessary or alternative means of multiplexing the 4D light field into a 2D image are required. Ives [Ives 1903] and Lippmann [Lippmann 1908] provide two early examples of spatial multiplexing with the introduction of parallax barriers and integral photography, respectively. Such spatial multiplexing allows light field capture of dynamic scenes, but requires a trade-off between the spatial and angular sampling rates. Okano et al. [Okano et al. 1999] and Ng et al. [Ng et al. 2005] describe modern, digital implementations of integral photography, however numerous other spatial multiplexing schemes have emerged.

Instead of affixing an array of microlenses directly to an image sensor, Georgiev et al. [Georgiev et al. 2006] add an external lens attachment with an array of lenses and prisms (see Figure 4, right). Ueda et al. [Ueda et al. 2008a; Ueda et al. 2008b] consider similar external lens arrays; however, in these works, an array of variable focus lenses, implemented using liquid lenses controlled by electrowetting, allow the spatial and angular resolution to be optimized depending on the observed scene.

Rather than using absorbing masks or refractive lens arrays, Unger et al. [Unger et al. 2003], Levoy et al. [Levoy et al. 2004], Lanman et al. [Lanman et al. 2006], and Taguchi et al. [Taguchi et al. 2010b] demonstrate that a single

photograph of an array of tilted, planar mirrors or mirrored spheres produces a spatially-multiplexed estimate of the incident light field. Yang et al. [Yang et al. 2000] demonstrate a large-format, lenslet-based architecture by combining an array of lenses and a flatbed scanner. Related compound imaging systems, producing a spatially-multiplexed light field using arrays of lenses and a single sensor, were proposed by Ogata et al. [Ogata et al. 1994], Tanida et al. [Tanida et al. 2001; Tanida et al. 2003], and Hiura et al. [Hiura et al. 2009].

Spatial multiplexing produces an interlaced array of elemental images within the image formed on the sensor. Veeraraghavan et al. [Veeraraghavan et al. 2007] introduce frequency multiplexing as an alternative method for achieving single-sensor light field capture. The optical heterodyning method proposed by Veeraraghavan et al. encodes the 4D Fourier transform of the light field into different spatio-angular bands of the Fourier transform of the 2D sensor image. Similar in concept to spatial multiplexing, the sensor spectrum contains a uniform array of 2D spectral slices of the 4D light field spectrum. Such frequency-domain multiplexing is achieved by placing non-refractive, light-attenuating masks slightly in front of a conventional sensor (see Figure 4, bottom).

As described by Veeraraghavan et al., masks allowing frequency-domain multiplexing (i.e., heterodyne detection) must have a Fourier transform consisting of an array of impulses (i.e., a 2D Dirac comb). In [Veeraraghavan et al. 2007], a *Sum-of-Sinusoids (SoS) pattern*, consisting of a weighted harmonic series of equal-phase sinusoids, is proposed. As shown in Figure 5, such codes transmit significantly more light than traditional pinhole arrays [Ives 1903]; however, as shown by Lanman et al. [Lanman et al. 2008], these patterns are equivalent to a truncated Fourier series approximation of a pinhole array for high angular sampling rates. In [Lanman et al. 2008], Lanman et al. propose *tiled-broadband patterns*, corresponding to periodic masks with individual tiles exhibiting a broadband Fourier transform. This family includes pinhole arrays, SoS patterns, and the tiled-MURA patterns proposed in that work (see Figure 5). Such patterns produce masks with 50% transmission, enabling shorter exposures than existing methods.

In subsequent work, Veeraraghavan et al. [Veeraraghavan et al. 2008] propose adaptive mask patterns, consisting of aharmonic sinusoids, optimized for the spectral bandwidth of the observed scene. Georgiev et al. [Georgiev et al. 2008] analyze such heterodyne cameras and further propose masks placed external to the camera body. Rather than using a global, frequency-domain decoding scheme, Ihrke et al. [Ihrke et al. 2010a] demonstrate how spatial-domain decoding methods can be extended to frequency-multiplexed light fields.

4.4 Capture Applications

Given the wide variety of light field capture devices, a similarly diverse set of applications is enabled by such high-dimensional representations of light transport. While Kanolt [Kanolt 1918] considers the related concept of a parallax panoramagram to achieve 3D display, light fields have also proven useful for applications spanning computer graphics, digital photography, and 3D reconstruction.

In the field of computer graphics, light fields were introduced to facilitate image-based rendering [Levoy and Hanrahan 1996; Gortler et al. 1996]. In contrast to the conventional computer graphics pipeline, novel 2D images are synthesized by resampling the 4D light field. With sufficient light field resolution, views are synthesized without knowledge of the underlying scene geometry. Subsequent to these works, researchers continued to enhance the fidelity of image-based rendering. For example, a significant limitation of early methods is that illumination cannot be adjusted in synthesized images. This is in stark contrast to the conventional computer graphics pipeline, wherein arbitrary light sources can be supported using ray tracing together with a model of material reflectance properties. Debevec et al. [Debevec et al. 2000] address this limitation by capturing an 8D *reflectance field*. In their system, the 4D light field reflected by an object is measured as a function of the 4D light field incident on the object. Thus, an 8D reflectance field maps variations in the input radiance to variations in the output radiance, allowing image-based rendering to support variation of both viewpoint and illumination.

Light fields parameterize every possible photograph that can be taken outside the convex hull of an object; as a result, they have found widespread application in 3D television, also known as free-viewpoint video. Carranza et

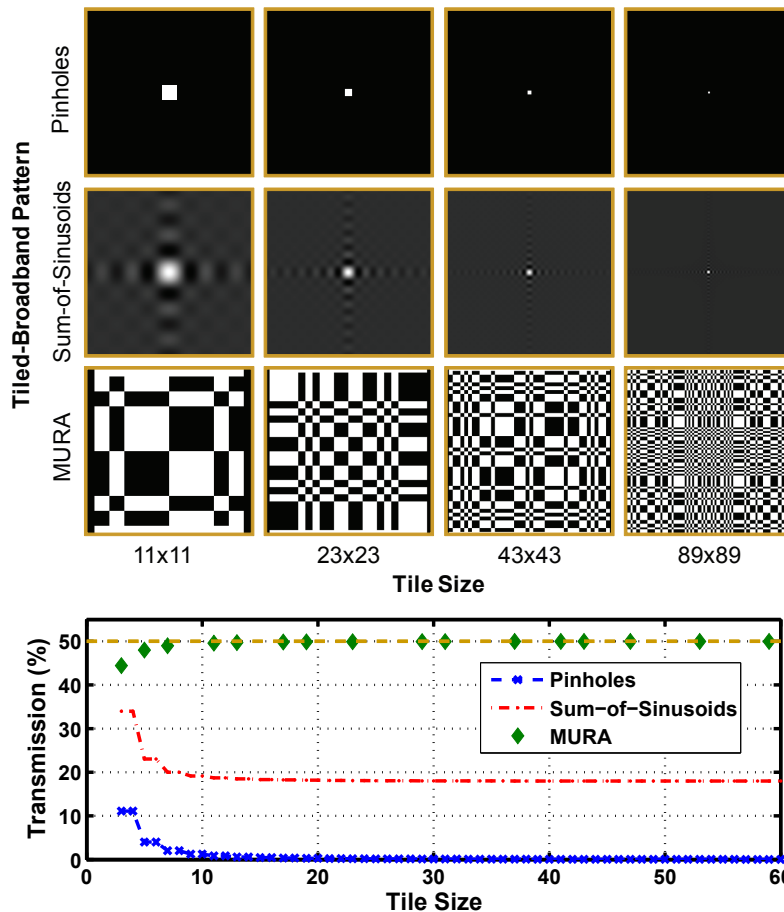


Figure 5: Lanman et al. [Lanman et al. 2008] introduce tiled-broadband patterns for mask-based, frequency-multiplexed light field capture. (Top) Each row, from left to right, shows broadband tiles of increasing spatial dimensions, including: pinholes [Ives 1928], Sum-of-Sinusoids (SoS) [Veeraraghavan et al. 2007], and MURA [Gottesman and Fenimore 1989; Lanman et al. 2008]. (Bottom) The SoS tile converges to 18% transmission, whereas the MURA tile remains near 50%. Note that frequency multiplexing with either SoS or MURA tiles significantly outperforms conventional pinhole arrays in terms of total light transmission and exposure time. (Figures reproduced from [Lanman 2010].)

al. [Carranza et al. 2003] describe a system with an array of cameras surrounding one or more actors. Similar systems have been developed by Matusik et al. [Matusik et al. 2000] and Starck et al. [Starck and Hilton 2008]. Image-based rendering allows arbitrary adjustment of the viewpoint in real-time. Vlasic et al. [Vlasic et al. 2009] further demonstrate 3D reconstruction of human actors from multiple-camera sequences captured under varying illumination conditions.

Light fields, given their similarity to conventional parallax panoramagrams [Ives 1928], have also found application in the design and analysis of 3D displays. Okano et al. [Okano et al. 1999] adapt integral photography to create a 3D television system supporting both multi-view capture and display. Similarly, Matusik and Pfister [Matusik and Pfister 2004] achieve light field capture using an array of 16 cameras and implement light field display using an array of 16 projectors and lenticular screens. Zwicker et al. [Zwicker et al. 2006] develop antialiasing filters for automultiscopic 3D display using a signal processing analysis. Hirsch et al. [Hirsch et al. 2009] develop a *BiDirectional (BiDi) screen*, supporting both conventional 2D image display and real-time 4D light field capture, facilitating mixed multitouch and gesture-based interaction; the device uses a lensless light field capture method, consisting of a tiled-MURA pattern [Lanman et al. 2008] displayed on an LCD panel and a large-format sensor. Recently, Lanman et al. [Lanman

et al. 2010] use an algebraic analysis of light fields to characterize the rank constraints of all dual-layer, attenuation-based light field displays; through this analysis they propose a generalization of conventional parallax barriers, using content-adaptive, time-multiplexed mask pairs to synthesize high-rank light fields with increased brightness and spatial resolution. Wetzstein et al. [Wetzstein et al. 2011b] demonstrate how a stack of attenuating layers, such as spaced transparencies or LCD panels, can be used in combination with computerized tomographic reconstruction to display natural light fields of 3D scenes.

Post-processing of captured light fields can resolve long-standing problems in conventional photography. Ng [Ng 2005] describes efficient algorithms for digital image refocusing, allowing the plane of focus to be adjusted after a photograph has been taken. In addition, Talvala et al. [Talvala et al. 2007] and Raskar et al. [Raskar et al. 2008] demonstrate that high-frequency masks can be combined with light field photography to eliminate artifacts due to glare and multiple scattering of light within camera lenses. Similarly, light field capture can be extended to microscopy and confocal imaging, enabling similar benefits in extended depth of field and reduced scattering [Levoy et al. 2004; Levoy et al. 2006]. Smith et al. [Smith et al. 2009] improve conventional image stabilization algorithms using light fields captured with an array of 25 cameras. As described, most single-sensor acquisition schemes trade increased angular resolution for decreased spatial resolution [Georgiev et al. 2006]; Bishop et al. [Bishop et al. 2009] and Lumsdaine and Georgiev [Lumsdaine and Georgiev 2009] apply priors regarding the statistics of natural images and modified imaging hardware, respectively, to achieve super-resolution light field capture that, in certain conditions, mitigates this resolution loss.

As characterized throughout this report, the plenoptic function of a given scene contains a large degree of redundancy; both the spatial and angular dimensions of light fields of natural scenes are highly correlated. Recent work is exploring the benefits of compressive sensing for light field acquisition. Fergus et al. [Fergus et al. 2006b] introduce *random lens imaging*, wherein a conventional camera lens is replaced with a random arrangement of planar mirrored surfaces, allowing super-resolution and 3D imaging applications. Babacan et al. [Babacan et al. 2009] propose a compressive sensing scheme for light field capture utilizing randomly-coded, non-refractive masks placed in the aperture plane. Ashok and Neifeld [Ashok and Neifeld 2010] propose compressive sensing schemes, again using non-refractive masks, allowing either spatial or angular compressive light field imaging. As observed in that work, future capture methods will likely benefit from joint spatio-angular compressive sensing; however, as discussed later in this report, further redundancies exist among all the plenoptic dimensions, not just the directional variations characterized by light fields.

5 Multiplexing Space and Focal Surfaces

The ability to resolve spatial light variation is an integral part of any imaging system. For the purpose of this report we differentiate between spatial variation on a plane perpendicular to the optical axis and variation along the optical axis inside a camera behind the main lens. The former quantity, transverse light variation, is what all 2D sensors measure. In this section, we discuss approaches for very high-resolution imaging (Sec. 5.1), focal surface curvature correction techniques of the light field inside a camera (Sec. 5.2), and extended depth of field photography (Sec. 5.3).

5.1 Super-Resolution and Gigapixel Imaging

Although exotic camera systems can resolve structures in the order of 100 nm [van Putten et al. 2011], the resolution of standard photographs is usually limited by the physical layout and size of the photosensitive elements, the optical resolution of employed optical elements, and the diffraction limit. Attempts to break these limits, which are significantly larger than 100 nm, are referred to as super-resolution imaging. Such techniques have been of particular interest to the vision community for many years. In most cases a sequence of slightly shifted low-resolution photos is captured and fused into a single high-resolution image. The shifts are usually smaller than the pixel size; an extensive review of such techniques can be found in [Baker and Kanade 2002; Borman and Stevenson 1998]. Sub-pixel precise shifts of low-resolution images can be achieved by mechanical vibrations [LANDOLT et al. 2001; Ben-Ezra

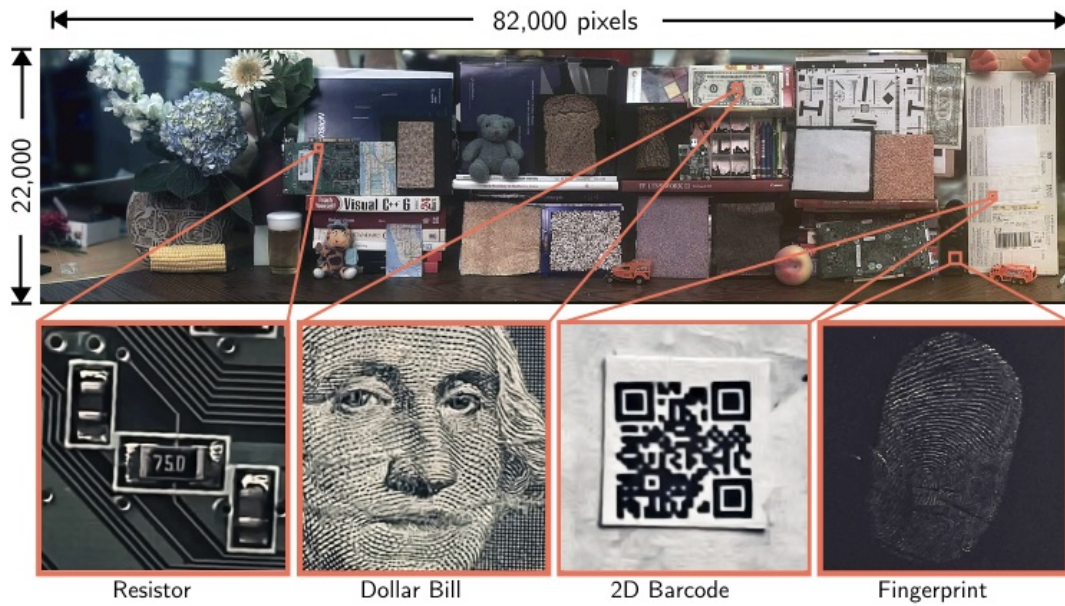


Figure 6: A wide field of view 1.7 gigapixel image captured by Cossairt et al. [Cossairt et al. 2011].

et al. 2005], by coding the camera’s aperture using phase [Ashok and Neifeld 2007] and attenuation [Mohan et al. 2008a] masks, by exploiting object motion in combination with temporally coded apertures [Agrawal and Raskar 2007], or by analyzing multiple frames of a video [Liu and Sun 2011]. For an increased resolution in space and time, successive frames in videos of a single camera [Shahar et al. 2011] or multiple devices [Shechtman et al. 2002; Shechtman et al. 2005] (see Sec. 6.2) can be analyzed instead. All super-resolution approaches require an optimization problem to be solved for the unknown super-resolved image given multiple low-resolution measurements. This is computationally expensive for higher resolutions and is usually an ill-posed problem requiring additional image priors [Baker and Kanade 2002].

Gigapixel imaging is a relatively new field that, similar to super-resolution, aims at capturing very high-resolution imagery. The main difference is that gigapixel imaging approaches generally do not try to beat the limits of sensor resolution, but rather stitch a gigapixel panoramic image together from a set of megapixel images. These can be photographed by mounting a camera on a computer-controlled rotation stage [Kopf et al. 2007], or a high-resolution small-scale sensor that is automatically moved in the image plane of a large-format camera [Ben-Ezra 2011]. Both of these techniques implement the concept of capturing a sequence of images with a single device that are composited into a high-quality photograph. In this case, the parameter that is varied for each image in the sequence is the camera pose. Alternatively, the optics of a camera can be modified, for instance with custom spherical lens elements, to allow a single very high-resolution image to be captured instantaneously with multiple sensors [Cossairt et al. 2011]. An example scene captured with this technique is shown in Figure 6.

5.2 Optical Field Correction

Not only is the actual resolution of digital photographs limited by the pixel count and the diffraction limit, but also by the applied optical elements. Standard spherical lenses have a focal surface that is, unlike most sensors, not actually planar but curved. Significant engineering effort is put into the commercial development of complex lens systems, especially in variable-focus camera objectives, that correct for the resulting image blur at sensor locations away from the optical axis. Several approaches have been proposed to correct for what is usually called field curvature, or more simply lens aberrations, and also phase aberrations caused by atmospheric turbulences in astronomical imaging [Tyson 1991]. These usually integrate secondary optical assemblies into the system, such as fiber optics [Kapany and Hopkins 1957], prisms [Inderhees 1973], lenslet arrays [Hanrahan and Ng 2006; Brady

and Hagen 2009], coded attenuation masks [Pandharkar et al. 2010], or spatial light modulators [Tyson 1991] and oftentimes require computational processing of the measured data.

5.3 Extended Depth of Field Photography

Depth of field (DOF), that is a depth-dependent (de)focus of a pictured scene, plays an important role in photography. Controlled focus and defocus can be useful for highlighting objects of interest, such as people in a portrait where the background is blurred. For most applications, however, all-focused imagery is desirable. Ideally, a photographer should be able to refocus or completely remove all defocus as a post processing step in an image editing software or directly on the camera. While this is one of the main applications for light fields, as discussed in Section 4, in this section we explore alternative focus modulation approaches that do not directly capture the full 4D light field.

Although the depth-dependent size of the point spread function (PSF) or circle-of-confusion is effected by a variety of parameters, the most important ones are the aperture size and the depth expansion of the photographed scene. Larger apertures result in shallower depths of field but allow more light to reach the sensor, thereby decreasing the noise level. While a shallow DOF is often undesirable, it is also unavoidable in many situations where a low noise level is more important.

Removing the DOF blur in a standard photograph is a difficult problem because it requires a deconvolution of the image with a spatially varying PSF. The PSF shape corresponds to that of the camera aperture, which can usually be well approximated with a Gaussian distribution; unfortunately, a deconvolution with a Gaussian is an ill-posed inverse problem, because high frequencies are irreversibly lost in the image capture. Applying natural image priors can improve reconstructions (see e.g., [Levin et al. 2007b]). The spatially-varying PSF size is directly proportional to the depth of the scene, which is in most cases unknown. A common approach to alleviate this problem is to mechanically or optically modify the depth-dependent PSF of the imaging system so that it becomes depth-invariant resulting in a reconstruction that only requires a spatially invariant deconvolution, which is much easier and does not require knowledge of the scene depth.

One family of techniques that only requires a single shot to capture a scene with a depth-invariant PSF is called *Focal Sweep*. Here, the PSF modulation is achieved by moving the object [Häusler 1972] or the sensor [Nagahara et al. 2008] during the exposure time, or by exploiting the wavelength-dependency of the PSF to multiplex multiple focal planes in the scene onto a single sensor image [Cossairt and Nayar 2010].

Alternatively, the apertures of the imaging system can be coded with cubic phase plates [Dowski and Cathey 1995] or other phase masks [Ojeda-Castaneda et al. 2005; Ben-Eliezer et al. 2005; Chi and George 2001], diffusers [Garcia-Guerrero et al. 2007; Cossairt et al. 2010], attenuation patterns [Levin et al. 2007a], polarization filters [Chi et al. 2006], or multi-focal elements [Levin et al. 2009].

All of the above listed focal sweep and coded aperture approaches optically modify the PSF of the optical system for an extended DOF. The captured images usually need to be post-processed, for instance by applying a deconvolution. An analysis of quality criteria of attenuation-based aperture masks for defocus deblurring was presented by Zhou and Nayar [Zhou and Nayar 2009]; this analysis was extended to also consider PSF invertibility [Baek 2010].

Focal Stacks are series of images of the same scene, where the focal plane differs for each photograph in the sequence. A single, focused image can be composited by selecting the best-focused match in the stack for each image region [Pieper and Korpel 1983]. The optimal choice of parameters, including focus and aperture, for the images in a focal stack are well established [Hasinoff et al. 2009; Hasinoff and Kutulakos 2008]. Capturing a focal stack with a large-scale high-resolution camera was implemented by Ben-Ezra [Ben-Ezra 2010]. Kutulakos and Hasinoff [Kutulakos and Hasinoff 2009] proposed to multiplex a focal stack into a single sensor image in a similar fashion as color filter arrays multiplex different color channels into a RAW camera image. However, to the knowledge of the authors, this camera has not yet been built.

Green et al. [Green et al. 2007] split the aperture of a camera using circular mirrors and multiplex the result into



Figure 7: *Multiple frames of a flying bird multiplexed into a single photograph (left). These kinds of photographs were shot with a photographic gun (right) by Etienne-Jules Marey as early as 1882.*

different regions of a single photograph. In principle, this approach captures multiple frames with varying aperture settings at a reduced spatial resolution in a single snapshot.

Other applications for flexible focus imaging include 3D shape reconstruction with shape from (de)focus (see e.g. [Nayar and Nakagawa 1994; Zhou et al. 2009]) or *Confocal Stereo* [Hasinoff and Kutulakos 2006; Hasinoff and Kutulakos 2009], video matting [McGuire et al. 2005], and extended depth of field projection [Grosse et al. 2010].

6 Multiplexing Time

Capturing motion and other forms of movement in photographs has been pursued since the invention of the daguerreotype. Early pioneers in this field include Eadweard Muybridge (e.g. [Muybridge 1957]) and Etienne-Jules Marey (e.g. [Braun 1992]). As illustrated in Figure 7, much of the early work on picturing time focused on the study of anatomy and locomotion of animals and humans; photographic apparatuses were usually custom built at that time (Fig. 7, right). In this section, we discuss two classes of techniques for picturing motion: image capture at temporal resolutions that are significantly lower (Sec. 6.1) or higher (Sec. 6.2) than the resolving capabilities of the human visual system and approaches for joint optical and computational motion deblurring (Sec. 6.3).

6.1 Time Lapse Photography

Photographing scenes at very low temporal sampling rates is usually referred to as time lapse photography. Technically, time lapses can simply be acquired by taking multiple photographs from the same or a very close camera position at larger time intervals and assembling them in a video. In order to avoid temporal aliasing, or in simpler terms provide naturally looking motion, the exposure times should ideally be as long as the interval between successive shots. Timothy Allen, photographer for the BBC, provides a very informative tutorial on time lapse photography on his website [Allen 2010]. The BBC has produced a number of astounding time lapse videos, including many scenes in their Planet Earth and Life series.

6.2 High-Speed Imaging

Analog high-speed film cameras have been developed throughout the last century. A variety of technologies exist that expose film at very high speeds including mechanical movement through temporally registered pins and rotating prisms or mirrors. For a detailed discussion of the history of high-speed photography, applications, and the state of the art about nine years ago, the reader is referred to the book by Ray [Ray 2002].

Single Sensor Approaches

Today, high-speed digital cameras are commercially available. Examples are the Phantom Flex by Vision Research [Research 2010], which can capture up to 2,570 frames per second (fps) at HD resolution, and the FAST-CAM SA5 by Photron, which captures 7,500 fps at megapixel resolution or up to one million frames per second at a reduced resolution (64×16 pixels) [Photron 2010]; both cameras employ CMOS sensor technology. A modified CCD is used in the HyperVision HPV-2 by Shimadzu [Shimadzu 2010], which operates at one million fps for an image resolution of 312×260 pixels. The Dynamic Photomechanics Laboratory at the University of Rhode Island (mcise.uri.edu/dpml/facilities.html) houses an IMACON 468-MkII digital camera operating at 200 million fps, but exact specifications of that camera are unknown to the authors. With the introduction of Casio's Exilim camera series (exilim.casio.com), which records low resolution videos at up to 1,000 fps, high-speed cameras have entered the consumer market.

An alternative to high-speed sensors is provided by *Assorted Pixels* [Narasimhan and Nayar 2005], where spatial resolution is traded for temporal resolution by measuring spatially interleaved, temporally staggered exposures on a sensor. This approach is very similar to what standard color filter arrays do to acquire color information (see Sec. 3.2). While this concept was initially only theoretical, it has recently been implemented by aligning a digital micromirror device (DMD) with a CCD sensor [Bub et al. 2010]. Alternatively, the sensor readout could be controlled on a per-pixel basis, as for instance provided by non-destructive sensor readout (e.g., [Semiconductor 2010]). Reddy et al. [Reddy et al. 2011] built an LCOS-based camera prototype that modulates the exposure of each pixel randomly throughout the exposure time. In combination with a non-linear sparse reconstruction algorithm, the 25 fps prototype has been shown to capture imagery with up to 200 frames per second without loss of spatial resolution by exploiting sparsity in the spatio-temporal volume. Coded rolling shutters [Gu et al. 2010] have the potential to implement this concept on a per-scanline basis.

Agrawal et al. [Agrawal et al. 2010b] demonstrated how a pinhole in the aperture plane of a camera, which moves throughout the exposure time, allows the captured data to be adaptively re-interpreted. For this purpose, temporal light variation is directly encoded in the different views of the light field that is simultaneously acquired with a Sum-of-Sinusoids (SoS) attenuation-mask (see Sec. 4.3) in a single shot. Temporal variation and different viewpoints cannot be separated in this approach.

Multiple Devices

Rather than photographing a scene with a single high-speed camera, multiple synchronized devices can be used. One of the most popular movie scenes that shows high-speed motion captured by a camera array is the bullet time effect in *The Matrix*. Here, a rig of digital SLR cameras, arranged along a virtual camera path, captures a scene at precisely controlled time steps so that a virtual, high-speed camera can be simulated that moves along the predefined path.

The direct capture of high-speed events with camera arrays was scientifically discussed by Wilburn et al. [Wilburn et al. 2004; Wilburn et al. 2005]. In this approach, the exposure windows of the cameras are slightly staggered so that a high-speed video can be composed by merging the data of the individual cameras. Shechtman et al. [Shechtman et al. 2002; Shechtman et al. 2005] proposed to combine the output of multiple low-resolution video cameras for space-time super-resolution. Coded exposures have been shown to optimize temporal super-resolution from multi-camera arrays [Agrawal et al. 2010a] by alleviating the ill-posedness of the reconstruction. As required for spatial super-resolution (see Sec. 5.1), temporal super-resolution requires computationally expensive post-processing of the measured data.

High-Speed Illumination

High-speed imagery can also be acquired by utilizing high-speed illumination. Harold ‘Doc’ Edgerton [Project 2009] created this field by inventing electronic strobes and using them to depict very fast motions in a similar fashion as Eadweard Muybridge and Etienne-Jules Marey had done with more primitive, mechanical technologies decades before him. Today, high-speed illumination, in an attosecond time scale, is more conveniently achieved with lasers rather than stroboscopes [Baker et al. 2006; Ray 2002].

Stroboscopic illumination can be used to compensate for rolling shutter effects and synchronize an array of consumer cameras [Bradley et al. 2009]. Narasimhan et al. [Narasimhan et al. 2008] exploited the high-speed temporal dithering patterns of DLP-based illumination for a variety of vision problems, including photometric stereo and range imaging. Coded strobing, by either illumination or controlled sensor readout, in combination with reconstructions developed in the compressive sensing community, allows high-speed periodic events to be acquired [Veeraraghavan et al. 2011]. Another high-speed imaging approach that is inspired by compressive sensing was proposed by Gupta et al. [Gupta et al. 2010]. Here, a 3D spatio-temporal volume is adaptively encoded with a fixed voxel budget. This approach encodes fast motions with a high temporal, but lower spatial resolution, while the spatial resolution in static parts of the scene is maximized. A co-axial projector-camera pair was used to simulate controllable per-pixel exposures.

Exotic Ultra High-Speed Imaging

Other imaging devices that capture ultra high-speed events are streak cameras. Rather than recording standard 2D photographs, these devices capture 2D images that encode spatial information in one dimension and temporal light variation in the other. These systems are usually combined with pulsed laser illumination and operate at temporal resolutions of about one hundred femtoseconds [Hamamatsu 2010], corresponding to a framerate of ten trillion fps. Another exotic ultra high-speed imager is the STREAM camera [Goda et al. 2009], which optically converts a 2D image into a serial time-domain waveform that is recorded with a single high-speed photodiode at 6.1 million fps.

6.3 Motion Deblurring

Motion deblurring has been an active area of research over the last few decades. It is well known that deblurring is an ill-posed problem, which is why many algorithms apply regularizers [Richardson 1972; Lucy 1974] or natural image statistics (e.g., [Fergus et al. 2006a; Levin et al. 2007b]) to solve the problem robustly. Usually, high-frequency spatio-temporal image information is irreversibly lost in the image formation because a standard shutter along with the sensor integration time create a temporal rect-filter which has many zero-crossings in the Fourier domain. In this section, we review coded image acquisition techniques that optically modify the motion PSF so that the reconstruction becomes a well-posed problem.

Coded Single Capture Approaches

One of the earliest approaches of coded temporal sampling was introduced by Raskar et al. [Raskar et al. 2006] as the *Fluttered Shutter*. The motion PSF in a single sensor image was modified by mounting a programmable liquid crystal element in front of the camera lens and modulating its transmission over the exposure time with optimized binary codes. These codes were designed to preserve high temporal frequencies, so that the required image deconvolution becomes well-posed. Optimized codes and an algorithm for the problem of combined motion deblurring and spatial super-resolution of moving objects with coded exposures were analyzed by Agrawal and Raskar [Agrawal and Raskar 2007]. Both approaches use programmable, attenuation-based shutters to apply the codes, thereby sacrificing light transmission, and require a manual rectification of object motion in the captured images. Optimality criteria for motion PSF invertibility were extended to also allow high-quality PSF estimation [Agrawal and Xu 2009]; this automates the motion rectification step.

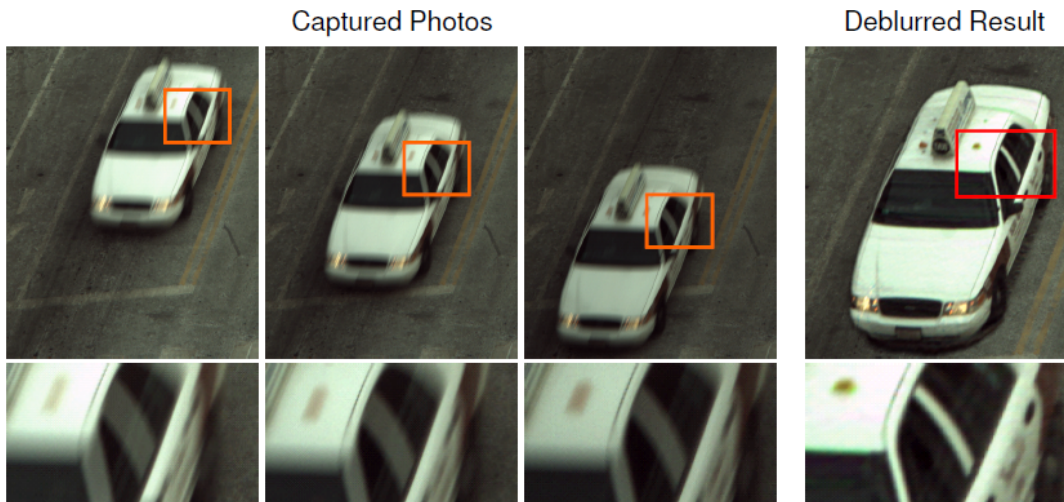


Figure 8: By varying the exposure time for successive frames in a video (left), multi-image deblurring (right) can be made invertible [Agrawal et al. 2009].

Inspired by approaches that create depth-invariant point spread functions (see Sec. 5.3), Levin et al. [Levin et al. 2008] showed how one-dimensional parabolic sensor motion during the exposure time can achieve a motion-invariant PSF along the line of sensor motion. Compared to attenuation-coded temporal exposures, this method does not sacrifice light transmission but requires prior knowledge of object motion and only works along one spatial direction. The optimal tradeoffs for single image deblurring from either attenuation-coded exposures or sensor motion, in terms of signal-to-noise ratio of the reconstructions, were analyzed by Agrawal and Raskar [Agrawal and Raskar 2009].

Image Sequences or Multiple Cameras

Synthetic Shutter Speed Imaging [Telleen et al. 2007] combines multiple sharp but noisy images captured with short exposure times. The resulting image has a lower noise level; motion blur is reduced by aligning all images before they are fused.

A *Hybrid Camera* for motion deblurring, consisting of a rig of two cameras, was introduced by Ben-Ezra and Nayar [Ben-Ezra and Nayar 2003; Ben-Ezra and Nayar 2004]. One of the cameras captures the scene at a high temporal, but low spatial resolution; the output of this camera is used to estimate the motion PSF, which in turn is used to deblur the high-quality image captured by the other camera. Improvements of reconstructions for hybrid cameras have recently been presented [Tai et al. 2010]. Hybrid camera architectures also provide the opportunity to simultaneously deblur captured images and reconstruct a high-resolution depth map of the photographed scene [Li et al. 2008a].

Motion blur in a video can be synthetically removed by applying super-resolution techniques to multiple successive frames [Bascle et al. 1996]. Agrawal et al. [Agrawal et al. 2009] showed that improved results can be achieved by modulating the exposure times for successive frames in video sequences so that a reconstruction from multiple images becomes a well-posed problem. An example for this is shown in Figure 8.

7 Acquisition of Further Light Properties

In this section, we review acquisition approaches for light properties that are not considered dimensions of the plenoptic function, but are closely related in terms of capture or application.

7.1 Polarization

Polarization is an inherent property of the wave nature of light [Collett 2005], which is why we treat it separately from the plenoptic function. Generally, polarization describes the oscillation of a wave traveling through space in the transverse plane, perpendicular to the direction of propagation. Linear polarization refers to transverse oscillation along a line, whereas spherical or elliptical polarization describe corresponding oscillation trajectories.

Although some animals, including mantis shrimp [Marshall and Oberwinkler 1999], cephalopods (squid, octopus, cuttlefish) [Mäthger et al. 2009], and insects [Wehner 1976], are reported to have photoreceptors that are sensitive to polarization, standard solid state sensors are not. The most straightforward way of capturing this information is by taking multiple photographs of a scene with different polarizing filters mounted in front of the camera lens. These filters are standard practice in photography to reduce specular reflections, increase the contrast of outdoor images, and improve the appearance of vegetation. Alternatively, this kind of information can be captured using polarization filter arrays [Schechner and Nayar 2003a] which, similar to generalized mosaics [Schechner and Nayar 2005], require multiple photographs to be captured. Recently, polarized illumination [Ghosh et al. 2010] has been shown to have the potential to acquire all Stokes parameters necessary to describe polarization.

Applications for the acquisition of polarized light include image dehazing [Schechner et al. 2001; Schechner et al. 2003; Namer and Schechner 2005], improved underwater vision [Schechner and Karpel 2004; Schechner and Karpel 2005], specular highlight removal [Wolff and Boulton 1991; Nayar et al. 1993; Müller 1996; Umeyama and Godin 2004], shape [Miyazaki et al. 2003; Miyazaki et al. 2004; Atkinson and Hancock 2005] and BRDF [Atkinson and Hancock 2008] estimation, material classification [Chen and Wolff 1998], light source separation [Cula et al. 2007], surface normal acquisition [Ma et al. 2007], surface normal and refractive index estimation [Sadjadi 2007; Ghosh et al. 2010], separation of transparent layers [Schechner et al. 1999], and optical communication [Siddiqui and Zhou 1991; Schönfelder 2003; Yao 2008].

7.2 Phase Imaging

A variety of techniques has been proposed to visualize and quantify phase retardation in transparent microscopic organisms [Murphy 2001]. Many of these phase-contrast imaging approaches, such as Zernike phase contrast and differential interference contrast (DIC), require coherent illumination and are qualitative rather than quantitative. This implies that changes in phase or refractive events are encoded as intensity variations in captured images, but remain indistinguishable from the intensity variations caused by absorption in the medium. Quantitative approaches exist [Barone-Nugent et al. 2002], but require multiple images, are subject to a paraxial approximation, and are limited to orthographic cameras.

Schlieren and shadowgraph photography are alternative, non-intrusive imaging methods for dynamically changing refractive index fields. These techniques have been developed in the fluid imaging community over the past century, with substantial improvements in the 1940s. An extensive overview of different optical setups and the historic evolution of Schlieren and Shadowgraph imaging can be found in the book by Settles [Settles 2001]. As illustrated in Figure 9, recently proposed applications of Schlieren imaging include the tomographic reconstruction of transparent gas flows using a camera array [Atcheson et al. 2008] and the capture of refractive events with 4D light field probes [Wetzstein et al. 2011d].

7.3 LIDAR and Time-of-Flight

LIDAR (LIght Detection and Ranging) [Wandinger 2005] is a technology that measures the time of a laser pulse from transmission to detection of the reflected signal. It is similar to radar, but uses different wavelengths of the electromagnetic spectrum, typically in the infrared range. Combining such a pulsed laser with optical scanners and a positioning system such as GPS allows very precise depth or range maps to be captured, even from airplanes. Overlaying range data with standard photographs provides a powerful tool for aerial surveying, forestry, oceanography,

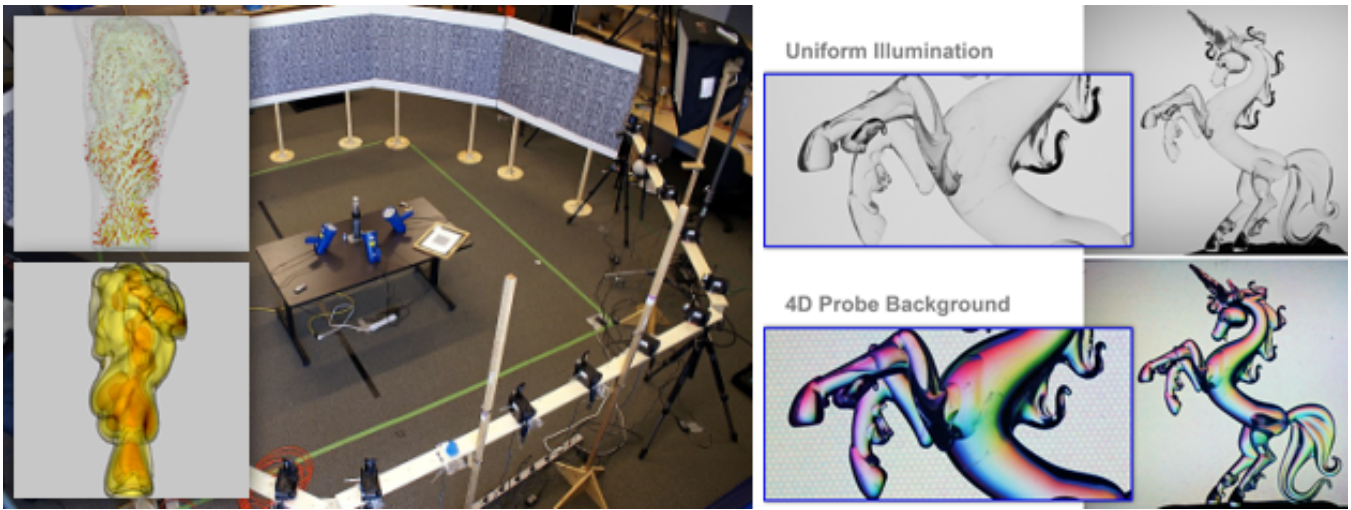


Figure 9: Schlieren imaging for tomographic gas reconstruction (left) and capture of refractive events using light field probes (right). (Figures reproduced from [Atcheson et al. 2008] and [Wetzstein et al. 2011d].)

agriculture, and geology. Flash LIDAR [Lange and Seitz 2001] or time-of-flight cameras [Kolb et al. 2010] capture a photograph and a range map simultaneously for all pixels. Although spatial resolution of the range data is often poor, these cameras usually capture at video rates.

Streak cameras operate in the picosecond [Campillo and Shapiro 1983] or even attosecond [Itatani et al. 2002] range and usually capture 2D images, where one dimension is spatial light variation and the other dimension is time-of-flight. These cameras have recently been used to reveal scene information outside the line of sight of a camera, literally behind corners [Kirmani et al. 2009; Pandharkar et al. 2011].

8 Discussion and Conclusions

In summary, we have presented a review of approaches to plenoptic image acquisition. We have used an intuitive categorization based on plenoptic dimensions and hardware setups for the acquisition. Alternative categorizations may be convenient for the discussion of the more general field of computational photography [Raskar and Tumblin 2009]. The increasingly growing number of publications in this field is one of the main motivations for this state of the art report, which focuses specifically on joint optical encoding and computational reconstruction approaches for the acquisition of the plenoptic function.

Based on the literature reviewed in this report, we make the following observations:

- most of the discussed approaches either assume that some plenoptic dimensions are constant, such as time in sequential image capture, or otherwise restricted, for instance spatially band-limited in single sensor interleaved capture; these assumptions result in fixed plenoptic resolution tradeoffs,
- however, there are strong correlations between the dimensions of the plenoptic function; so far these are almost exclusively exploited in color demosaicing,
- therefore, natural image statistics that can be used as priors in computational image reconstruction and incorporate all plenoptic dimensions with their correlations are desirable; so are sophisticated reconstruction techniques employing these.

It has recently been shown that all approaches for interleaved plenoptic sampling on a single sensor, including spatial [Narasimhan and Nayar 2005] and Fourier multiplexing [Veeraraghavan et al. 2007; Lanman et al. 2008] methods, can be cast into a common reconstruction framework [Ihrke et al. 2010a]. While the exploitation of

correlations between plenoptic dimensions, for example spatial and spectral light variation, is common practice for imaging with color filter arrays and subsequent demosaicing, there is significant potential to develop similar techniques for demosaicing other multiplexed plenoptic information, for instance light fields [Levin and Durand 2010].

Priors for the correlations between plenoptic dimensions can be very useful for plenoptic super-resolution or generally more sophisticated reconstructions. These could, for instance, be derived from plenoptic image databases [Wetzstein et al. 2011c]; we show examples of such data in Figures 10, 11, and 12.

Another promising avenue of future research is adaptive imaging. Precise control of the sampled plenoptic information is the key for flexible and adaptive reconstruction. An intuitive next step for sophisticated imaging with respect to temporal light variation and dynamic range is pixel-precise, non-destructive sensor readout. In the future, however, it is desirable to being able to control the optical modulation of all plenoptic dimensions.

While most of the reviewed approaches make fixed plenoptic resolution tradeoffs, some already show a glimpse of the potential of adaptive re-interpretation of captured data [Agrawal et al. 2010b]. Ideas from the compressive sensing community (see e.g., [Candès et al. 2006]) have also started to play an important role in adaptive plenoptic imaging [Gupta et al. 2010; Veeraraghavan et al. 2011]. In these approaches optical coding is combined with content adaptive reconstructions that can dynamically trade higher-dimensional resolution in post-processing to best represent the recorded data.

Picturing space, color, time, directions, and other light properties has been of great interest to science and art alike for centuries. With the emergence of digital light sensing technology and computational processing power, many new and exciting ways to acquire some of the visual richness surrounding us have been presented. We have, however, only begun to realize how new technologies allow us to transcend the way evolution has shaped visual perception for different creatures on this planet.

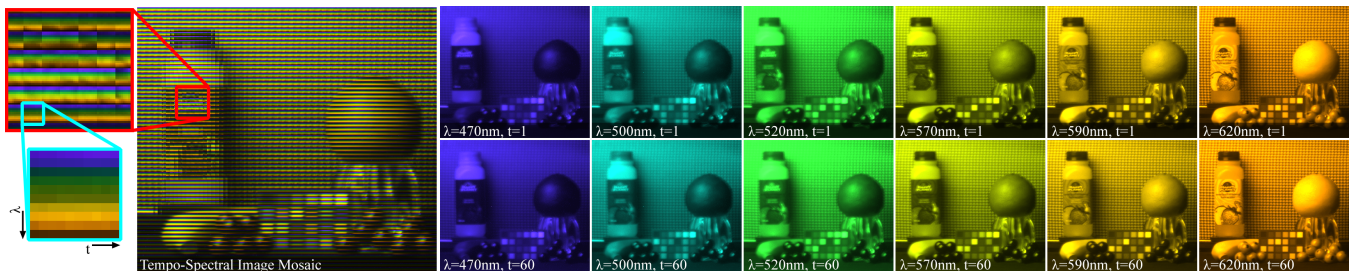


Figure 10: Dataset containing multi-spectral variation and controlled object movement in the scene. Left: image mosaic illustrating the correlation between the spectral and temporal plenoptic dimension of natural scenes. Right: six spectral, color-coded slices of the dataset with two temporal snapshots each. We recorded these datasets using a custom multi-spectral camera that consists of collimating optics, a liquid crystal tunable filter, and a USB machine vision camera.

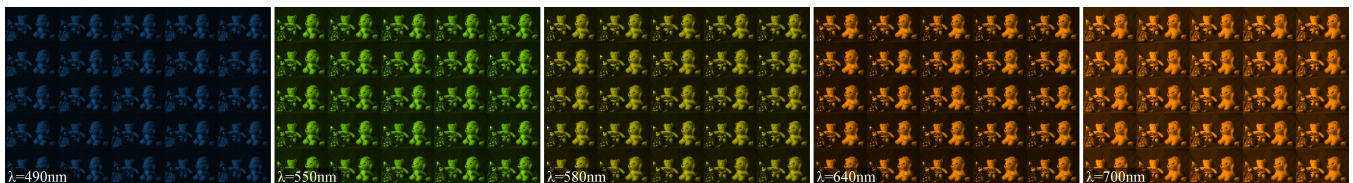


Figure 11: Five color channels of a multi-spectral light field with 5×5 viewpoints and 10 color channels for each viewpoint. This dataset was captured by mounting our multi-spectral camera on an X-Y translation stage and sequentially capturing the spectral bands for each viewpoint.

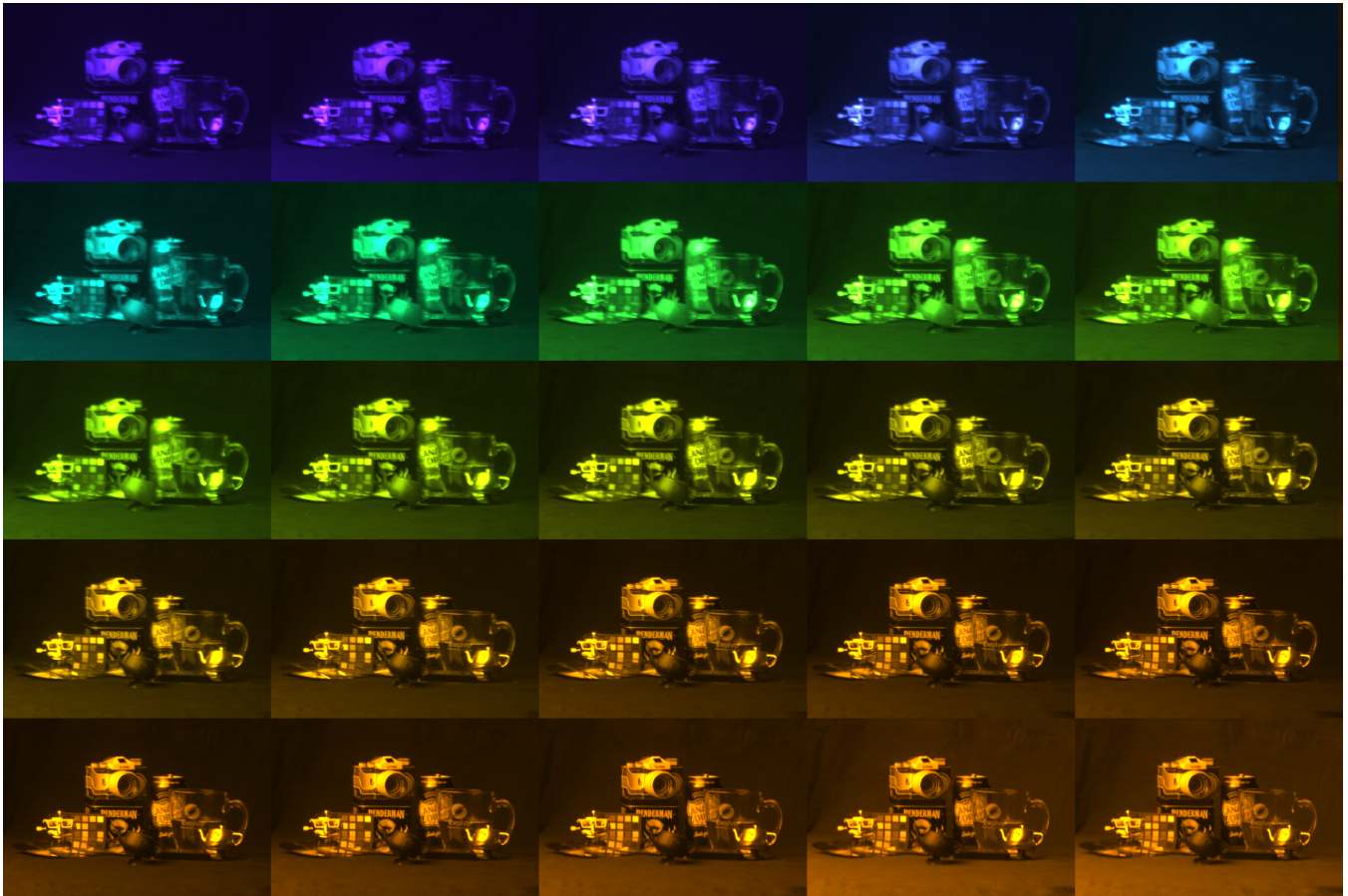


Figure 12: Another multi-spectral light field dataset with 15×15 viewpoints and 23 narrow-band color channels for each viewpoint. The spectral channels range from 460 nm to 680 nm in 10 nm increments. Only 5×5 viewpoints are shown in this mosaic and each of those is color-coded. The photographed scene includes a variety of illumination effects including diffraction, refraction, inter-reflections, and specularities.

References

- ADAMS, A., TALVALA, E.-V., PARK, S. H., JACOBS, D. E., AJDIN, B., GELFAND, N., DOLSON, J., VAQUERO, D., BAEK, J., TICO, M., LENSCH, H. P. A., MATUSIK, W., PULLI, K., HOROWITZ, M., AND LEVOY, M. 2010. The Frankencamera: an Experimental Platform for Computational Photography. *ACM Trans. Graph. (SIGGRAPH)* 29, 29:1–29:12.
- ADELSON, E. H., AND BERGEN, J. R. 1991. The Plenoptic Function and the Elements of Early Vision. In *Computational Models of Visual Processing*, MIT Press, 3–20.
- ADELSON, E., AND WANG, J. 1992. Single Lens Stereo with a Plenoptic Camera. *IEEE Trans. PAMI* 14, 2, 99–106.
- AGGARWAL, M., AND AHUJA, N. 2004. Split Aperture Imaging for High Dynamic Range. *Int. J. Comp. Vis.* 58, 1, 7–17.
- AGRAWAL, A., AND RASKAR, R. 2007. Resolving Objects at Higher Resolution from a Single Motion-Blurred Image. In *Proc. IEEE CVPR*, 1–8.
- AGRAWAL, A., AND RASKAR, R. 2009. Optimal Single Image Capture for Motion Deblurring. In *Proc. IEEE CVPR*, 1–8.
- AGRAWAL, A., AND XU, Y. 2009. Coded Exposure Deblurring: Optimized Codes for PSF Estimation and Invertibility. In *Proc. IEEE CVPR*, 1–8.
- AGRAWAL, A., XU, Y., AND RASKAR, R. 2009. Invertible Motion Blur in Video. *ACM Trans. Graph. (Siggraph)* 28, 3, 95.
- AGRAWAL, A., GUPTA, M., VEERARAGHAVAN, A., AND NARASIMHAN, S. 2010. Optimal Coded Sampling for Temporal Super-Resolution. In *Proc. IEEE CVPR*, 374–380.
- AGRAWAL, A., VEERARAGHAVAN, A., AND RASKAR, R. 2010. Reinterpretable Imager: Towards Variable Post-Capture Space, Angle and Time Resolution in Photography. In *Proc. Eurographics*, 1–10.
- ALLEN, T., 2010. Time Lapse Tutorial. <http://timothyallen.blogs.bbcearth.com/2009/02/24/time-lapse-photography/>.
- ALLEYSON, D., SÜSTRUNK, S., AND HÉRAULT, J. 2005. Linear Demosaicing inspired by the Human Visual System. *IEEE Trans. Im. Proc.* 14, 4, 439–449.
- ASHDOWN, I. 1993. Near-field photometry: A new approach. *Journal of the Illuminating Engineering Society* 22, 1, 163–180.
- ASHOK, A., AND NEIFELD, M. A. 2007. Pseudorandom Phase Masks for Superresolution Imaging from Subpixel Shifting. *Applied Optics* 46, 12, 2256–2268.
- ASHOK, A., AND NEIFELD, M. A. 2010. Compressive Light Field Imaging. In *Proc. SPIE 7690*, 76900Q.
- ATCHESON, B., IHRKE, I., HEIDRICH, W., TEVS, A., BRADLEY, D., MAGNOR, M., AND SEIDEL, H. 2008. Time-resolved 3D Capture of Non-Stationary Gas Flows. *ACM Trans. Graph. (SIGGRAPH Asia)* 27, 5, 132.
- ATKINSON, G., AND HANCOCK, E. 2005. Multi-view surface reconstruction using polarization. In *Proc. ICCV*, vol. 1, 309–316.
- ATKINSON, G. A., AND HANCOCK, E. R. 2008. Two-dimensional BRDF Estimation from Polarisation. *Comput. Vis. Image Underst.* 111, 2, 126–141.
- BABACAN, S. D., ANSORGE, R., LUESSI, M., MOLINA, R., AND KATSAGGELOS, A. K. 2009. Compressive Sensing of Light Fields. In *Proc. ICIP*, 2313–2316.

- BAEK, J. 2010. Transfer Efficiency and Depth Invariance in Computational Cameras. In *Proc. ICCP*, 1–8.
- BAKER, S., AND KANADE, T. 2002. Limits on Super-Resolution and How to Break Them. *IEEE Trans. PAMI* 24, 1167–1183.
- BAKER, S., ROBINSON, J. S., HAWORTH, C. A., TENG, H., SMITH, R. A., CHIRILA, C. C., LEIN, M., TISCH, J. W. G., AND MARANGOS, J. P. 2006. Probing Proton Dynamics in Molecules on an Attosecond Time Scale. *Science* 312, 5772, 424–427.
- BARONE-NUGENT, E. D., BARTY, A., AND NUGENT, K. A. 2002. Quantitative Phase-Amplitude Microscopy I: Optical Microscopy. *Journal of Microscopy* 206, 3, 194–203.
- BASCLE, B., BLAKE, A., AND ZISSERMAN, A. 1996. Motion Deblurring and Super-resolution from an Image Sequence. In *Proc. ECCV*, 573–582.
- BAYER, B. E., 1976. Color imaging array. US Patent 3,971,065.
- BEN-ELIEZER, E., MAROM, E., KONFORTI, N., AND ZALEVSKY, Z. 2005. Experimental Realization of an Imaging System with an Extended Depth of Field. *Appl. Opt.* 44, 11, 2792–2798.
- BEN-EZRA, M., AND NAYAR, S. 2003. Motion Deblurring using Hybrid Imaging. In *Proc. IEEE CVPR*, 657–664.
- BEN-EZRA, M., AND NAYAR, S. 2004. Motion-based Motion Deblurring. *IEEE Trans. PAMI* 26, 6, 689–698.
- BEN-EZRA, M., ZOMET, A., AND NAYAR, S. 2005. Video Superresolution using Controlled Subpixel Detector Shifts. *IEEE Trans. PAMI* 27, 6, 977–987.
- BEN-EZRA, M. 2010. High Resolution Large Format Tile-Scan Camera. In *Proc. IEEE ICCP*, 1–8.
- BEN-EZRA, M. 2011. A Digital Gigapixel Large-Format Tile-Scan Camera. *IEEE Computer Graphics and Applications* 31, 49–61.
- BISHOP, T., ZANETTI, S., AND FAVARO, P. 2009. Light-Field Superresolution. In *Proc. ICCP*, 1–9.
- BODKIN, A., SHEINIS, A., NORTON, A., DALY, J., BEAVEN, S., AND WEINHEIMER, J. 2009. Snapshot Hyperspectral Imaging - the Hyperspectral Array Camera. In *Proc. SPIE* 7334, 1–11.
- BOLLES, R. C., BAKER, H. H., AND MARIMONT, D. H. 1987. Epipolar-plane image analysis: An approach to determining structure from motion. *IJCV* 1, 1, 7–55.
- BORMAN, S., AND STEVENSON, R. 1998. Super-resolution from image sequences - A review. In *Proc. Symposium on Circuits and Systems*, 374–378.
- BRADLEY, D., ATCHESON, B., IHRKE, I., AND HEIDRICH, W. 2009. Synchronization and Rolling Shutter Compensation for Consumer Video Camera Arrays. In *Proc. ProCams*, 1–8.
- BRADY, D. J., AND HAGEN, N. 2009. Multiscale Lens Design. *Optics Express* 17, 13, 10659–10674.
- BRAUN, M. 1992. *Picturing Time: The Work of Etienne-Jules Marey (1830-1904)*. The University of Chicago Press.
- BUB, G., TECZA, M., HELMES, M., LEE, P., AND KOHL, P. 2010. Temporal Pixel Multiplexing for Simultaneous High-Speed, High-Resolution Imaging. *Nature Methods* 7, 209–211.
- CAMPILLO, A., AND SHAPIRO, S. 1983. Picosecond Streak Camera Fluorometry-A Review. *Journal of Quantum Electronics* 19, 4, 585 – 603.
- CANDÈS, E., ROMBERG, J., AND TAO, T. 2006. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Information Theory* 52, 2, 489–509.

- CAO, X., TONG, X., DAI, Q., AND LIN, S. 2011. High Resolution Multispectral Video Capture with a Hybrid Camera System. In *Proc. IEEE CVPR*, 1–8.
- CARRANZA, J., THEOBALT, C., MAGNOR, M. A., AND SEIDEL, H.-P. 2003. Free-viewpoint video of human actors. *ACM Transactions on Graphics (TOG)* 22, 3, 569–577.
- CHEN, S. E., AND WILLIAMS, L. 1993. View interpolation for image synthesis. In *Proc. ACM SIGGRAPH*, 279–288.
- CHEN, H., AND WOLFF, L. B. 1998. Polarization Phase-Based Method For Material Classification In Computer Vision. *IJCV* 28, 1, 73–83.
- CHI, W., AND GEORGE, N. 2001. Electronic Imaging using a Logarithmic Asphere. *Optics Letters* 26, 12, 875–877.
- CHI, W., CHU, K., AND GEORGE, N. 2006. Polarization Coded Aperture. *Optics Express* 14, 15, 6634–6642.
- COLLETT, E. 2005. *Field Guide to Polarization*. SPIE Press.
- COSSAIRT, O., AND NAYAR, S. K. 2010. Spectral Focal Sweep: Extended Depth of Field from Chromatic Aberrations. In *Proc. ICCP*, 1–8.
- COSSAIRT, O., ZHOU, C., AND NAYAR, S. K. 2010. Diffusion Coded Photography for Extended Depth of Field. *ACM Trans. Graph. (Siggraph)* 29, 3, 31.
- COSSAIRT, O., MIAU, D., AND NAYAR, S. K. 2011. Gigapixel Computational Imaging. In *Proc. ICCP*.
- CRI INC, 2009. VariSpec Liquid Crystal Tunable Filters. www.channelsystems.ca/Attachments/VariSpec/VariSpec-Brochure.pdf.
- CULA, O. G., DANA, K. J., PAI, D. K., AND WANG, D. 2007. Polarization Multiplexing and Demultiplexing for Appearance-Based Modeling. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 2, 362–367.
- DEBEVEC, P. E., AND MALIK, J. 1997. Recovering High Dynamic Range Radiance Maps from Photographs. In *Proc. ACM Siggraph*, 369–378.
- DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the Reflectance Field of a Human Face. In *Proc. ACM SIGGRAPH*, 145–156.
- DEBEVEC, P. 2002. Image-Based Lighting. *IEEE Computer Graphics and Applications*, 26–34.
- DESCOUR, M., AND DERENIAK, E. 1995. Computed-tomography Imaging Spectrometer: Experimental Calibration and Reconstruction Results. *Applied Optics* 34, 22, 4817–4826.
- DESCOUR, M. E., C.E.VOLIN, , DERENIAK, E., AND K.J.THOME. 1997. Demonstration of a High-Speed Nonscanning Imaging Spectrometer. *Optics Letters* 22, 16, 1271–1273.
- DOWSKI, E., AND CATHEY, T. 1995. Extended Depth of Field through Wave-Front Coding. *Applied Optics* 34, 11, 1859–1866.
- DU, H., TONG, X., CAO, X., AND LIN, S. 2009. A Prism-Based System for Multispectral Video Acquisition. In *Proc. IEEE ICCV*, 175–182.
- EASTMAN KODAK COMPANY. PhotoCD PCD0992. <http://r0k.us/graphics/kodak>.
- FAIRCHILD, M. D. 2005. *Color Appearance Models*. John Wiley and Sons.
- FERGUS, R., SINGH, B., HERTZMANN, A., ROWEIS, S. T., AND FREEMAN, W. T. 2006. Removing Camera Shake from a Single Photograph. *ACM Trans. Graph.* 25, 787–794.
- FERGUS, R., TORRALBA, A., AND FREEMAN, W. T. 2006. Random Lens Imaging. Tech. Rep. MIT-CSAIL-TR-2006-058, National Bureau of Standards.

- FLETCHER-HOLMES, D. W., AND HARVEY, A. R. 2005. Real-Time Imaging with a Hyperspectral Fovea. *J. Opt. A: Pure Appl. Opt.* 7, S298–S302.
- FOVEON, 2010. X3 Technology. www.foveon.com.
- GABOR, D. 1948. A new microscopic principle. *Nature*, 777–778.
- GAIGALAS, A. K., WANG, L., HE, H.-J., AND DEROSE, P. 2009. Procedures for Wavelength Calibration and Spectral Response Correction of CCD Array Spectrometers. *Journal of Research of the National Institute of Standards and Technology* 114, 4, 215–228.
- GAO, L., KESTER, R. T., AND TKACZYK, T. S. 2009. Compact Image Slicing Spectrometer (ISS) for hyperspectral fluorescence microscopy. *Optics Express* 17, 15, 12293–12308.
- GARCIA-GUERRERO, E. E., MENDEZ, E. R., AND LESKOVA, H. M. 2007. Design and Fabrication of Random Phase Diffusers for Extending the Depth of Focus. *Optics Express* 15, 3, 910–923.
- GAT, N. 2000. Imaging Spectroscopy Using Tunable Filters: A Review. In *Proc. SPIE 4056*, 50–64.
- GEHM, M. E., JOHN, R., BRADY, D. J., WILLETT, R. M., AND SCHULZ, T. J. 2007. Single-Shot Compressive Spectral Imaging with a Dual-Disperser Architecture. *Optics Express* 15, 21, 14013–14027.
- GEORGIEV, T., ZHENG, C., NAYAR, S., CURLESS, B., SALESIN, D., AND INTWALA, C. 2006. Spatio-angular resolution trade-offs in integral photography. In *Proc. EGSR*, 263–272.
- GEORGIEV, T., INTWALA, C., BABACAN, S., AND LUMSDAINE, A. 2008. Unified Frequency Domain Analysis of Lightfield Cameras. In *Proc. ECCV*, 224–237.
- GERSHUN, A. 1936. The light field. *Journal of Mathematics and Physics XVIII*, 51–151. Translated by P. Moon and G. Timoshenko.
- GHOSH, A., CHEN, T., PEERS, P., WILSON, C. A., AND DEBEVEC, P. 2010. Circularly polarized spherical illumination reflectometry. *ACM Trans. Graph. (Siggraph Asia)* 27, 5, 134.
- GODA, K., TSIA, K. K., AND JALALI, B. 2009. Serial Time-Encoded Amplified Imaging for Real-Time Observation of Fast Dynamic Phenomena. *Nature*, 458, 1145–1149.
- GORMAN, A., FLETCHER-HOLMES, D. W., AND HARVEY, A. R. 2010. Generalization of the Lyot Filter and its Application to Snapshot Spectral Imaging. *Optics Express* 18, 6, 5602–5608.
- GORTLER, S., GRZESZCZUK, R., SZELINSKI, R., AND COHEN, M. 1996. The Lumigraph. In *Proc. ACM Siggraph*, 43–54.
- GOTTESMAN, S. R., AND FENIMORE, E. E. 1989. New family of binary arrays for coded aperture imaging. *Applied Optics* 28, 20, 4344–4352.
- GREEN, P., SUN, W., MATUSIK, W., AND DURAND, F. 2007. Multi-Aperture Photography. In *Proc. ACM Siggraph*, 68.
- GROSSBERG, M. D., AND NAYAR, S. K. 2003. High Dynamic Range from Multiple Images: Which Exposures to Combine. In *Proc. ICCV Workshop CPMCV*.
- GROSSE, M., WETZSTEIN, G., GRUNDHÖFER, A., AND BIMBER, O. 2010. Coded Aperture Projection. *ACM Trans. Graph.* 29, 22:1–22:12.
- GU, J., HITOMI, Y., MITSUNAGA, T., AND NAYAR, S. K. 2010. Coded Rolling Shutter Photography: Flexible Space-Time Sampling. In *Proc. IEEE ICCP*, 1–8.

- GUNTURK, B., GLOTZBACH, J., ALTUNBASAK, Y., SCHAFFER, R., AND MERSEREAU, R. 2005. Demosaicking: Color Filter Array Interpolation in Single-Chip Digital Cameras. *IEEE Signal Processing* 22, 1, 44–54.
- GUPTA, M., AGRAWAL, A., VEERARAGHAVAN, A., AND NARASIMHAN, S. G. 2010. Flexible Voxels for Motion-Aware Videography. In *Proc. ECCV*, 100–114.
- HALLE, M. W. 1994. Holographic stereograms as discrete imaging systems. In *SPIE Practical Holography*, 73–84.
- HAMAMATSU, 2010. Streak Systems. <http://sales.hamamatsu.com/en/products/system-division/ultra-fast/streak-systems.php>.
- HANRAHAN, P., AND NG, R. 2006. Digital Correction of Lens Aberrations in Light Field Photography. In *International Optical Design Conference*, 1–3.
- HARVEY, A. R., BEALE, J., GREENAWAY, A. H., HANLON, T. J., AND WILLIAMS, J. 2000. Technology Options for Imaging Spectrometry Imaging Spectrometry. In *Proc. SPIE 4132*, 13–24.
- HARVEY, A. R., FLETCHER-HOLMES, D. W., AND GORMAN, A. 2005. Spectral Imaging in a Snapshot. In *Proc. SPIE 5694*, 1–10.
- HASINOFF, S. W., AND KUTULAKOS, K. N. 2006. Confocal Stereo. In *Proc. ECCV*, 620–634.
- HASINOFF, S. W., AND KUTULAKOS, K. N. 2008. Light-Efficient Photography. In *Proc. of ECCV*, 45–59.
- HASINOFF, S. W., AND KUTULAKOS, K. N. 2009. Confocal Stereo. *Int. J. Comp. Vis.* 81, 1, 82–104.
- HASINOFF, S. W., KUTULAKOS, K. N., DURAND, F., AND FREEMAN, W. T. 2009. Time-Constrained Photography. In *Proc. of ICCV*, 333–340.
- HASINOFF, S. W., DURAND, F., AND FREEMAN, W. T. 2010. Noise-Optimal Capture for High Dynamic Range Photography. In *Proc. IEEE CVPR*, 1–8.
- HÄUSLER, G. 1972. A Method to Increase the Depth of Focus by Two Step Image Processing. *Optics Communications* 6, 1, 38–42.
- HECHT, E. 2002. *Optics, fourth edition*. Addison Wesley.
- HIRAKAWA, K., AND PARKS, T. W. 2006. Joint Demosaicing and Denoising. *IEEE Trans. Im. Proc.* 15, 8, 2146–2157.
- HIRAKAWA, K., AND WOLFE, P. 2007. Spatio-Spectral Color Filter Array Design for Enhanced Image Fidelity. In *Proc. ICIP*, II – 81–II – 84.
- HIRAKAWA, K., AND WOLFE, P. 2008. Spatio-Spectral Color Filter Array Design for Optimal Image Recovery. *IEEE Trans. Im. Proc.* 17, 10, 1876–1890.
- HIRSCH, M., LANMAN, D., HOLTZMAN, H., AND RASKAR, R. 2009. BiDi Screen: a Thin, Depth-Sensing LCD for 3D Interaction using Light Fields. In *ACM Trans. Graph. (SIGGRAPH Asia)*, 1–9.
- HIURA, S., MOHAN, A., AND RASKAR, R. 2009. Krill-eye : Superposition Compound Eye for Wide-Angle Imaging via GRIN Lenses. In *Proc. OMNIVIS*, 1–8.
- HORSTMAYER, R., EULISS, G., ATHALE, R., AND LEVOY, M. 2009. Flexible Multimodal Camera Using a Light Field Architecture. In *Proc. ICCP*, 1–8.
- HUNICZ, J., AND PIERNIKARSKI, D. 2001. Investigation of Combustion in a Gasoline Engine using Spectrophotometric Methods. In *Proc. SPIE 4516*, 307–314.
- HUNT, R. W. G. 1991. *Measuring Color, 3rd ed.* Fountain Press.

- IHRKE, I., STICH, T., GOTTSCHLICH, H., MAGNOR, M., AND SEIDEL, H. 2008. Fast incident light field acquisition and rendering. In *Proc. of WSCG*, 177–184.
- IHRKE, I., WETZSTEIN, G., AND HEIDRICH, W. 2010. A Theory of Plenoptic Multiplexing. In *Proc. IEEE CVPR*, 1–8.
- IHRKE, I., KUTULAKOS, K. N., LENSCH, H. P. A., MAGNOR, M., AND HEIDRICH, W. 2010. Transparent and Specular Object Reconstruction. *Computer Graphics Forum* 29, 8, 2400–2426.
- INDERHEES, J., 1973. Optical field curvature corrector. US patent 3,720,454.
- ITATANI, J., QUR, F., YUDIN, G. L., IVANOV, M. Y., KRAUSZ, F., AND CORKUM, P. B. 2002. Attosecond Streak Camera. *Physical Review Letters* 88, 17, 173903.
- IVES, H., 1903. Parallax Stereogram and Process of Making Same. US patent 725,567.
- IVES, H. 1928. Camera for Making Parallax Panoramagrams. *J. Opt. Soc. Amer.* 17, 435–439.
- KANG, S. B., UYTENDAELE, M., WINDER, S., AND SZELISKI, R. 2003. High Dynamic Range Video. In *Proc. ACM Siggraph*, 319–325.
- KANOLT, C. W., 1918. Parallax Panoramagrams. US patent 1,260,682.
- KAPANY, N. S., AND HOPKINS, R. E. 1957. Fiber Optics. Part III. Field Flatteners. *JOSA* 47, 7, 594–595.
- KINDZELSKII, A. L., YANG, Z. Y., NABEL, G. J., TODD, R. F., AND PETTY, H. R. 2000. Ebola Virus Secretory Glycoprotein (sGP) Diminishes Fc Gamma RIIIB-to-CR3 Proximity on Neutrophils. *J. Immunol.* 164, 953–958.
- KIRMANI, A., HUTCHISON, T., DAVIS, J., AND RASKAR, R. 2009. Looking Around the corner using Transient Imaging. In *Proc. ICCV*, 1–8.
- KOLB, A., BARTH, E., KOCH, R., AND LARSEN, R. 2010. Time-of-Flight Cameras in Computer Graphics. *Computer Graphics Forum* 29, 1, 141–159.
- KOPF, J., UYTENDAELE, M., DEUSSEN, O., AND COHEN, M. F. 2007. Capturing and Viewing Gigapixel Images. *ACM Trans. on Graph. (SIGGRAPH)* 26, 3.
- KUTULAKOS, K. N., AND HASINOFF, S. W. 2009. Focal Stack Photography: High-Performance Photography with a Conventional Camera. In *Proc. IAPR Conference on Machine Vision Applications*, 332–337.
- LANDOLT, O., MITROS, A., AND KOCH, C. 2001. Visual Sensor with Resolution Enhancement by Mechanical Vibrations. In *Proc. Advanced Research in VLSI*, 249–264.
- LANGE, R., AND SEITZ, P. 2001. Solid-State Time-of-Flight Range Camera. *Journal of Quantum Electronics* 37, 3, 390–397.
- LANMAN, D., WACHS, M., TAUBIN, G., AND CUKIERMAN, F. 2006. Spherical Catadioptric Arrays: Construction, Multi-View Geometry, and Calibration. In *Proc. 3DPVT*, 81–88.
- LANMAN, D., RASKAR, R., AGRAWAL, A., AND TAUBIN, G. 2008. Shield Fields: Modeling and Capturing 3D Occluders. *ACM Trans. Graph. (Siggraph Asia)* 27, 5, 131.
- LANMAN, D., HIRSCH, M., KIM, Y., AND RASKAR, R. 2010. Content-Adaptive Parallax Barriers: Optimizing Dual-Layer 3D Displays using Low-Rank Light Field Factorization. *ACM Trans. Graph. (Siggraph Asia)* 28, 5, 1–10.
- LANMAN, D. 2010. *Mask-based Light Field Capture and Display*. PhD thesis, Brown University, School of Engineering.

- LAU, D. L., AND YANG, R. 2005. Real-Time Multispectral Color Video Synthesis using an Array of Commodity Cameras. *Real-Time Imaging 11*, 2, 109–116.
- LAWLOR, J., FLETCHER-HOLMES, D. W., HARVEY, A. R., AND MCNAUGHT, A. I. 2002. In Vivo Hyperspectral Imaging of Human Retina and Optic Disc. *Invest. Ophthalmol. Vis. Sci.* 43, 4350.
- LEVIN, A., AND DURAND, F. 2010. Linear View Synthesis Using a Dimensionality Gap Light Field Prior. In *Proc. IEEE CVPR*, 1–8.
- LEVIN, A., FERGUS, R., DURAND, F., AND FREEMAN, W. 2007. Image and Depth from a Conventional Camera with a Coded Aperture. *ACM Trans. Graph. (Siggraph)* 26, 3, 70.
- LEVIN, A., FERGUS, R., DURAND, F., AND FREEMAN, W. T., 2007. Deconvolution using Natural Image Priors. groups.csail.mit.edu/graphics/CodedAperture/SparseDeconv-LevinEtAl07.pdf.
- LEVIN, A., SAND, P., CHO, T. S., DURAND, F., AND FREEMAN, W. T. 2008. Motion-invariant photography. *ACM Trans. Graph. (Siggraph)* 27, 3, 71.
- LEVIN, A., HASINOFF, S. W., GREEN, P., DURAND, F., AND FREEMAN, W. T. 2009. 4D Frequency Analysis of Computational Cameras for Depth of Field Extension. *ACM Trans. Graph. (Siggraph)* 28, 3, 97.
- LEVOY, M., AND HANRAHAN, P. 1996. Light Field Rendering. In *Proc. ACM Siggraph*, 31–42.
- LEVOY, M., CHEN, B., VAISH, V., HOROWITZ, M., MCDOWALL, I., AND BOLAS, M. 2004. Synthetic Aperture Confocal Imaging. *ACM Trans. Graph. (SIGGRAPH)* 23, 3, 825–834.
- LEVOY, M., NG, R., ADAMS, A., FOOTER, M., AND HOROWITZ, M. 2006. Light Field Microscopy. *ACM Trans. Graph. (Siggraph)* 25, 3, 924–934.
- LEVOY, M., 2010. Computational Photography and the Stanford Frankencamera. Technical Talk. www.graphics.stanford.edu/talks/.
- LI, F., YU, J., AND CHAI, J. 2008. A Hybrid Camera for Motion Deblurring and Depth Map Super-Resolution. In *Proc. IEEE CVPR*, 1–8.
- LI, X., GUNTURK, B., AND ZHANG, L. 2008. Image Demosaicing: a Systematic Survey. In *SPIE Conf. on Visual Comm. and Image Proc.*, 68221J–68221J–15.
- LIANG, C.-K., LIN, T.-H., WONG, B.-Y., LIU, C., AND CHEN, H. H. 2008. Programmable Aperture Photography: Multiplexed Light Field Acquisition. *ACM Trans. Graph. (Siggraph)* 27, 3, 1–10.
- LIPPMANN, G. 1908. La Photographie Intégrale. *Academie des Sciences* 146, 446–451.
- LIU, C., AND SUN, D. 2011. A Bayesian Approach to Adaptive Video Super Resolution. In *Proc. IEEE CVPR*, 1–8.
- LU, Y. M., AND VETTERLI, M. 2009. Optimal Color Filter Array Design: Quantitative Conditions and an Efficient Search Procedure. In *Proc. SPIE 7250*, 1–8.
- LUCY, L. B. 1974. An iterative technique for the rectification of observed distributions. *The Astronomical Journal* 79, 745–754.
- LUMSDAINE, A., AND GEORGIEV, T. 2009. The Focused Plenoptic Camera. In *Proc. ICCP*, 1–8.
- LYOT, B. 1944. Le Filtre Monochromatique Polarisant et ses Applications en Physique Solaire. *Annales d'Astrophysique* 7, 31.
- MA, W.-C., HAWKINS, T., PEERS, P., CHABERT, C.-F., WEISS, M., AND DEBEVEC, P. 2007. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Proc. EGSR*, 183–194.

- MANN, S., AND PICARD, R. W. 1995. Being 'Undigital' with Digital Cameras: Extending Dynamic Range by Combining Differently Exposed Pictures. In *Proc. IS&T*, 442–448.
- MANSFIELD, C. L., 2005. Seeing into the past. 'www.nasa.gov/vision/earth/technologies/scrolls.html.
- MANTIUK, R., DALY, S., MYSKOWSKI, K., AND SEIDEL, H.-P. 2005. Predicting Visible Differences in High Dynamic Range Images - Model and its Calibration. In *Electronic Imaging*, B. E. Rogowitz, T. N. Pappas, and S. J. Daly, Eds., vol. 5666, 204–214.
- MANTIUK, R., KIM, K. J., REMPEL, A., AND HEIDRICH, W. 2011. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Trans. Graph. (Siggraph)* 30, 3, 1–12.
- MARSHALL, J., AND OBERWINKLER, J. 1999. Ultraviolet Vision: the Colourful World of the Mantis Shrimp. *Nature* 401, 6756, 873–874.
- MATHEWS, S. A. 2008. Design and Fabrication of a low-cost, Multispectral Imaging System. *Applied Optics* 47, 28, 71–76.
- MÄTHGER, L. M., SHASHAR, N., AND HANLON, R. T. 2009. Do Cephalopods Communicate using Polarized Light Reflections from their Skin? *Journal of Experimental Biology* 212, 2133–2140.
- MATUSIK, W., AND PFISTER, H. 2004. 3d tv: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes. *ACM Transactions on Graphics* 23, 814–824.
- MATUSIK, W., BUEHLER, C., RASKAR, R., GORTLER, S. J., AND MCMILLAN, L. 2000. Image-based visual hulls. In *ACM SIGGRAPH*, 369–374.
- MAXWELL, J. C. 1860. On the Theory of Compound Colours, and the Relations of the Colours of the Spectrum. *Phil. Trans. R. Soc. Lond.* 150, 57–84.
- MCGUIRE, M., MATUSIK, W., PFISTER, H., HUGHES, J. F., AND DURAND, F. 2005. Defocus Video Matting. *ACM Trans. Graph. (SIGGRAPH)* 24, 3, 567–576.
- MCGUIRE, M., MATUSIK, W., PFISTER, H., CHEN, B., HUGHES, J. F., AND NAYAR, S. K. 2007. Optical Splitting Trees for High-Precision Monocular Imaging. *IEEE Comput. Graph. & Appl.* 27, 2, 32–42.
- MITSUNAGA, T., AND NAYAR, S. K. 1999. Radiometric Self Calibration. In *Proc. IEEE CVPR*, 374–380.
- MIYAZAKI, D., TAN, R. T., HARA, K., AND IKEUCHI, K. 2003. Polarization-based inverse rendering from a single view. In *Proc. ICCV*, 982–998.
- MIYAZAKI, D., KAGESAWA, M., AND IKEUCHI, K. 2004. Transparent Surface Modelling from a Pair of Polarization Images. *IEEE Trans. PAMI* 26, 1, 73–82.
- MOHAN, A., HUANG, X., RASKAR, R., AND TUMBLIN, J. 2008. Sensing Increased Image Resolution Using Aperture Masks. In *Proc. IEEE CVPR*, 1–8.
- MOHAN, A., RASKAR, R., AND TUMBLIN, J. 2008. Agile Spectrum Imaging: Programmable Wavelength Modulation for Cameras and Projectors. *Computer Graphics Forum (Eurographics)* 27, 2, 709–717.
- MOON, P., AND SPENCER, D. E. 1981. *The Photoc Field*. MIT Press.
- MÜLLER, V. 1996. Elimination of Specular Surface-Reflectance Using Polarized and Unpolarized Light. In *Proc. IEEE ECCV*, 625–635.
- MURPHY, D. B. 2001. *Fundamentals of Light Microscopy and Electronic Imaging*. Wiley-Liss.
- MUYBRIDGE, E. 1957. *Animals in Motion*. first ed. Dover Publications, Chapman and Hall 1899.

- NAGAHARA, H., KUTHIRUMMAL, S., ZHOU, C., AND NAYAR, S. 2008. Flexible Depth of Field Photography. In *Proc. ECCV*, 60–73.
- NAMER, E., AND SCHECHNER, Y. Y. 2005. Advanced Visibility Improvement Based on Polarization Filtered Images. In *Proc. SPIE* 5888, 36–45.
- NARASIMHAN, S., AND NAYAR, S. 2005. Enhancing Resolution along Multiple Imaging Dimensions using Assorted Pixels. *IEEE Trans. PAMI* 27, 4, 518–530.
- NARASIMHAN, S. G., KOPPAL, S. J., AND YAMAZAKI, S. 2008. Temporal Dithering of Illumination for Fast Active Vision. In *Proc. ECCV*, 830–844.
- NAYAR, S., AND BRANZOI, V. 2003. Adaptive Dynamic Range Imaging: Optical Control of Pixel Exposures over Space and Time. In *Proc. IEEE ICCV*, vol. 2, 1168–1175.
- NAYAR, S., AND MITSUNAGA, T. 2000. High Dynamic Range Imaging: Spatially Varying Pixel Exposures. In *Proc. IEEE CVPR*, vol. 1, 472–479.
- NAYAR, S. K., AND NAKAGAWA, Y. 1994. Shape from Focus. *IEEE Trans. PAMI* 16, 8, 824–831.
- NAYAR, S., FANG, X.-S., AND BOULT, T. 1993. Removal of Specularities using Color and Polarization. In *Proc. IEEE CVPR*, 583–590.
- NAYAR, S., BRANZOI, V., AND BOULT, T. 2004. Programmable Imaging using a Digital Micromirror Array. In *Proc. IEEE CVPR*, vol. I, 436–443.
- NAYAR, S. K., BRANZOI, V., AND BOULT, T. E. 2006. Programmable Imaging: Towards a Flexible Camera. *IJCV* 70, 1, 7–22.
- NG, R., LEVOY, M., BRÉDIF, M., DUVAL, G., HOROWITZ, M., AND HANRAHAN, P. 2005. Light field photography with a hand-held plenoptic camera. Tech. Rep. Computer Science CSTR 2005-02, Stanford University.
- NG, R. 2005. Fourier Slice Photography. *ACM Trans. Graph. (Siggraph)* 24, 3, 735–744.
- NOMURA, Y., ZHANG, L., AND NAYAR, S. 2007. Scene Collages and Flexible Camera Arrays. In *Proc. EGSR*, 1–12.
- OGATA, S., ISHIDA, J., AND SASANO, T. 1994. Optical Sensor Array in an Artificial Compound Eye. *Optical Engineering* 33, 11, 3649–3655.
- OJEDA-CASTANEDA, J., LANDGRAVE, J. E. A., AND ESCAMILLA, H. M. 2005. Annular Phase-Only Mask for High Focal Depth. *Optics Letters* 30, 13, 1647–1649.
- OKAMOTO, T., AND YAMAGUCHI, I. 1991. Simultaneous Acquisition of Spectral Image Information. *Optics Letters* 16, 16, 1277–1279.
- OKANO, F., ARAI, J., HOSHINO, H., AND YUYAMA, I. 1999. Three-Dimensional Video System Based on Integral Photography. *Optical Engineering* 38, 6, 1072–1077.
- OPTEC, 2011. Separation prism technical data, Jan. www.alt-vision.com/color-prisms-tech.data.htm.
- PANAVISION, 2010. Genesis. www.panavision.com.
- PANDHARKAR, R., KIRMANI, A., AND RASKAR, R. 2010. Lens Aberration Correction Using Locally Optimal Mask Based Low Cost Light Field Cameras. In *Proc. OSA Imaging Systems*, 1–3.
- PANDHARKAR, R., VELTEN, A., BARDAGJY, A., RASKAR, R., BAWENDI, M., KIRMANI, A., AND LAWSON, E. 2011. Estimating Motion and Size of Moving Non-Line-of-Sight Objects in Cluttered Environments. In *Proc. ICCV CVPR*, 1–8.

- PARK, J.-I., LEE, M.-H., GROSSBERG, M. D., AND NAYAR, S. K. 2007. Multispectral Imaging Using Multiplexed Illumination. In *Proc. IEEE ICCV*, 1–8.
- PARMAR, M., AND REEVES, S. J. 2006. Selection of Optimal Spectral Sensitivity Functions for Color Filter Arrays. In *Proc. of ICIP*, 1005–1008.
- PARMAR, M., AND REEVES, S. J. 2010. Selection of Optimal Spectral Sensitivity Functions for Color Filter Arrays. *IEEE Trans. Im. Proc.* 19, 12 (Dec), 3190–3203.
- PHOTRON, 2010. FASTCAM SA5. www.photron.com/datasheet/FASTCAM.SA5.pdf.
- PIEPER, R. J., AND KORPEL, A. 1983. Image Processing for Extended Depth of Field. *Applied Optics* 22, 10, 1449–1453.
- PIXIM, 2010. Digital Pixel System. www.pixim.com.
- PROJECT, E. D. C., 2009. Harold 'Doc' Edgerton. www.edgerton-digital-collections.org/techniques/high-speed-photography.
- PROKUDIN-GORSKII, S. M., 1912. The Prokudin-Gorskii Photographic Records Recreated. www.loc.gov/exhibits/empire/.
- RAMANATH, R., SNYDER, W., BILBRO, G., AND SANDER, W. 2002. Demosaicking Methods for Bayer Color Arrays. *Journal of Electronic Imaging* 11, 3, 306–315.
- RASKAR, R., AND TUMBLIN, J. 2009. *Computational Photography: Mastering New Techniques for Lenses, Lighting, and Sensors*. A. K. Peters.
- RASKAR, R., AGRAWAL, A., AND TUMBLIN, J. 2006. Coded Exposure Photography: Motion Deblurring using Fluttered Shutter. *ACM Trans. Graph. (Siggraph)* 25, 3, 795–804.
- RASKAR, R., AGRAWAL, A., WILSON, C. A., AND VEERARAGHAVAN, A. 2008. Glare Aware Photography: 4D Ray Sampling for Reducing Glare Effects of Camera Lenses. *ACM Trans. Graph. (Siggraph)* 27, 3, 56.
- RAY, S. F. 2002. *High Speed Photography and Photonics*. SPIE Press.
- REDDY, D., VEERARAGHAVAN, A., AND CHELLAPPA, R. 2011. P2C2: Programmable Pixel Compressive Camera for High Speed Imaging. In *Proc. IEEE CVPR*, 1–8.
- REINHARD, E., KHAN, E. A., AKYÜZ, A. O., AND JOHNSON, G. M. 2008. *Color Imaging*. A K Peters Ltd.
- REINHARD, E., WARD, G., DEBEVEC, P., PATTANAIK, S., HEIDRICH, W., AND MYSZKOWSKI, K. 2010. *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann Publishers.
- RESEARCH, V., 2010. Phantom Flex. www.visionresearch.com/Products/High-Speed-Cameras/Phantom-Flex/.
- RI, S., FUJIGAKI, M., MATUI, T., AND MORIMOTO, Y. 2006. Accurate pixel-to-pixel correspondence adjustment in a digital micromirror device camera by using the phase-shifting moiré method. *Applied optics* 45, 27, 6940–6946.
- RICHARDSON, H. W. 1972. Bayesian-Based Iterative Method of Image Restoration. *JOSA* 62, 1, 55–59.
- ROBERTSON, M. A., BORMAN, S., AND STEVENSON, R. L. 1999. Estimation-Theoretic Approach to Dynamic Range Enhancement Using Multiple Exposures. *Journal of Electronic Imaging* 12, 2003.
- RORSLETT, B., 2008. Uv flower photographs. www.naturfotograf.com/index2.html.
- ROUF, M., MANTIUK, R., HEIDRICH, W., TRENTACOSTE, M., AND LAU, C. 2011. Glare Encoding of High Dynamic Range Images. In *Proc. IEEE CVPR*, 1–8.

- SADJADI, F. 2007. Extraction of Surface Normal and Index of Refraction using a Pair of Passive Infrared Polarimetric Sensors. In *Proc. IEEE CVPR*, 1–5.
- SAJADI, B., MAJUMDER, A., HIWADA, K., MAKI, A., AND RASKAR, R. 2011. Switchable Primaries Using Shiftable Layers of Color Filter Arrays. *ACM Trans. Graph. (Siggraph)* 30, 3, 1–10.
- SCHECHNER, Y. Y., AND KARPEL, N. 2004. Clear Underwater Vision. In *Proc. IEEE CVPR*, 536–543.
- SCHECHNER, Y. Y., AND KARPEL, N. 2005. Recovery of Underwater Visibility and Structure by Polarization Analysis. *IEEE Journal of Oceanic Engineering* 30, 3, 570–587.
- SCHECHNER, Y., AND NAYAR, S. 2001. Generalized Mosaicing. In *Proc. IEEE ICCV*, vol. 1, 17–24.
- SCHECHNER, Y., AND NAYAR, S. 2002. Generalized Mosaicing: Wide Field of View Multispectral Imaging. *IEEE Trans. PAMI* 24, 10, 1334–1348.
- SCHECHNER, Y., AND NAYAR, S. K. 2003. Polarization Mosaicking: High dynamic Range and Polarization Imaging in a Wide Field of View. In *Proc. SPIE 5158*, 93–102.
- SCHECHNER, Y., AND NAYAR, S. 2003. Generalized Mosaicing: High Dynamic Range in a Wide Field of View. *IJCV* 53, 3, 245–267.
- SCHECHNER, Y., AND NAYAR, S. 2004. Uncontrolled Modulation Imaging. In *Proc. IEEE CVPR*, 197–204.
- SCHECHNER, Y., AND NAYAR, S. 2005. Generalized Mosaicing: Polarization Panorama. *IEEE Trans. PAMI* 27, 4, 631–636.
- SCHECHNER, Y., SHAMIR, J., AND KIRYATI, N. 1999. Polarization-Based Decorrelation of Transparent Layers: The Inclination Angle of an Invisible Surface. In *Proc. ICCV*, 814–819.
- SCHECHNER, Y., NARASIMHAN, S. G., AND NAYAR, S. K. 2001. Instant Dehazing of Images using Polarization. In *Proc. IEEE CVPR*, 325–332.
- SCHECHNER, Y., NARASIMHAN, S. G., AND NAYAR, S. K. 2003. Polarization-Based Vision through Haze. *Applied Optics* 42, 3, 511–525.
- SCHECHNER, Y., NAYAR, S., AND BELHUMEUR, P. 2007. Multiplexing for Optimal Lighting. *IEEE Trans. PAMI* 29, 8, 1339–1354.
- SCHÖNFELDER, T., 2003. Polarization Division Multiplexing in Optical Data Transmission Systems. US Patent 6,580,535.
- SEETZEN, H., HEIDRICH, W., STUERZLINGER, W., WARD, G., WHITEHEAD, L., TRENTACOSTE, M., GHOSH, A., AND VOROZCOVS, A. 2004. High Dynamic Range Display Systems. *ACM Trans. on Graph. (SIGGRAPH 2004)* 23, 3, 760–768.
- SEMICONDUCTOR, C., 2010. LUPA Image Sensors. www.cypress.com/?id=206.
- SETTLES, G. 2001. *Schlieren & Shadowgraph Techniques*. Springer.
- SHAHAR, O., FAKTOR, A., AND IRANI, M. 2011. Space-Time Super-Resolution from a Single Video. In *Proc. IEEE CVPR*, 1–8.
- SHECHTMAN, E., CASPI, Y., AND IRANI, M. 2002. Increasing Space-Time Resolution in Video. In *Proc. ECCV*, 753–768.
- SHECHTMAN, E., CASPI, Y., AND IRANI, M. 2005. Space-Time Super-Resolution. *IEEE Trans. PAMI* 27, 4, 531–545.

- SHIMADZU, 2010. HyperVision HPV-2. www.shimadzu.com/products/test/hsvc/oh80jt000001d6t.html.
- SIDDIQUI, A. S., AND ZHOU, J. 1991. Two-Channel Optical Fiber Transmission Using Polarization Division Multiplexing. *Journal of Optical Communications* 12, 2, 47–49.
- SMITH, B. M., ZHANG, L., JIN, H., AND AGARWALA, A. 2009. Light Field Video Stabilization. In *Proc. of ICCV*, 1–8.
- SPHERON VR, 2010. SpheroCam HDR. www.spheron.com.
- STARCK, J., AND HILTON, A. 2008. Model-based human shape reconstruction from multiple views. *Computer Vision and Image Understanding (CVIU)* 111, 2, 179–194.
- TAGUCHI, Y., AGRAWAL, A., RAMALINGAM, S., AND VEERARAGHAVAN, A. 2010. Axial Light Fields for Curved Mirrors: Reflect Your Perspective, Widen Your View. In *Proc. IEEE CVPR*, 1–8.
- TAGUCHI, Y., AGRAWAL, A., VEERARAGHAVAN, A., RAMALINGAM, S., AND RASKAR, R. 2010. Axial-Cones: Modeling Spherical Catadioptric Cameras for Wide-Angle Light Field Rendering. *ACM Trans. Graph.* 29, 172:1–172:8.
- TAI, Y., HAO, D., BROWN, M. S., AND LIN, S. 2010. Correction of Spatially Varying Image and Video Motion Blur using a Hybrid Camera. *IEEE Trans. PAMI* 32, 6, 1012–1028.
- TALVALA, E.-V., ADAMS, A., HOROWITZ, M., AND LEVOY, M. 2007. Veiling Glare in High Dynamic Range Imaging. *ACM Trans. Graph. (Siggraph)* 26, 3, 37.
- TANIDA, J., KUMAGAI, T., YAMADA, K., MIYATAKE, S., ISHIDA, K., MORIMOTO, T., KONDOU, N., MIYAZAKI, D., AND ICHIOKA, Y. 2001. Thin Observation Module by Bound Optics (TOMBO): Concept and Experimental Verification. *Applied Optics* 40, 11, 1806–1813.
- TANIDA, J., SHOGENJI, R., KITAMURA, Y., YAMADA, K., MIYAMOTO, M., AND MIYATAKE, S. 2003. Color Imaging with an Integrated Compound Imaging System. *Optics Express* 11, 18, 2109–2117.
- TELLEEN, J., SULLIVAN, A., YEE, J., WANG, O., GUNAWARDANE, P., COLLINS, I., AND DAVIS, J. 2007. Synthetic Shutter Speed Imaging. *Computer Graphics Forum (Eurographics)* 26, 3, 591–598.
- TOYOOKA, S., AND HAYASAKA, N. 1997. Two-Dimensional Spectral Analysis using Broad-Band Filters. *Optical Communications* 137 (Apr), 22–26.
- TUMBLIN, J., AGRAWAL, A., AND RASKAR, R. 2005. Why I want a Gradient Camera. In *Proc. IEEE CVPR*, 103–110.
- TYSON, R. K. 1991. *Principles of Adaptive Optics*. Academic Press.
- UEDA, K., KOIKE, T., TAKAHASHI, K., AND NAEMURA, T. 2008. Adaptive Integral Photography Imaging with Variable-Focus Lens Array. In *Proc SPIE: Stereoscopic Displays and Applications XIX*, 68031A–9.
- UEDA, K., LEE, D., KOIKE, T., TAKAHASHI, K., AND NAEMURA, T., 2008. Multi-Focal Compound Eye: Liquid Lens Array for Computational Photography. ACM SIGGRAPH New Tech Demo.
- UMEYAMA, S., AND GODIN, G. 2004. Separation of Diffuse and Specular Components of Surface Reflection by Use of Polarization and Statistical Analysis of Images. *IEEE Trans. PAMI* 26, 5, 639–647.
- UNGER, J., WENGER, A., HAWKINS, T., GARDNER, A., AND DEBEVEC, P. 2003. Capturing and Rendering with Incident Light Fields. In *Proc. EGSR*, 141–149.
- VAGNI, F. 2007. Survey of Hyperspectral and Multispectral Imaging Technologies. Tech. Rep. TR-SET-065-P3, NATO Research and Technology.

- VAISH, V., SZELISKI, R., ZITNICK, C. L., KANG, S. B., AND LEVOY, M. 2006. Reconstructing Occluded Surfaces using Synthetic Apertures: Stereo, Focus and Robust Measures. In *Proc. IEEE CVPR*, 23–31.
- VALLEY, G., 2010. Viper FilmStream Camera. www.grassvalley.com.
- VAN PUTTEN, E., AKBULUT, D., BERTOLOTTI, J., VOS, W., LAGENDIJK, A., AND MOSK, A. 2011. Scattering Lens Resolves Sub-100 nm Structures with Visible Light. *Physical Review Letters* 106, 19, 1–4.
- VEERARAGHAVAN, A., RASKAR, R., AGRAWAL, A., MOHAN, A., AND TUMBLIN, J. 2007. Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocussing. *ACM Trans. Graph. (Siggraph)* 26, 3, 69.
- VEERARAGHAVAN, A., RASKAR, R., AGRAWAL, A., CHELLAPPA, R., MOHAN, A., AND TUMBLIN, J. 2008. Non-Refractive Modulators for Encoding and Capturing Scene Appearance and Depth. In *Proc. IEEE CVPR*, 1–8.
- VEERARAGHAVAN, A., REDDY, D., AND RASKAR, R. 2011. Coded Strobing Photography: Compressive Sensing of High Speed Periodic Videos. *IEEE Trans. PAMI*, to appear.
- VLASIC, D., PEERS, P., BARAN, I., DEBEVEC, P., POPOVIĆ, J., RUSINKIEWICZ, S., AND MATUSIK, W. 2009. Dynamic Shape Capture using Multi-View Photometric Stereo. In *ACM Trans. Graph. (SIGGRAPH Asia)*, 1–11.
- WAGADARIKAR, A., PITSIANIS, N., SUN, X., AND BRADY, D. 2008. Spectral Image Estimation for Coded Aperture Snapshot Spectral Imagers. In *Proc. SPIE 7076*, 707602.
- WAGADARIKAR, A., PITSIANIS, N., SUN, X., AND BRADY, D. 2009. Video Rate Spectral Imaging using a Coded Aperture Snapshot Spectral Imager. *Optics Express* 17, 8, 6368–6388.
- WANDINGER, U. 2005. *Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere*. Springer.
- WANG, S., AND HEIDRICH, W. 2004. The Design of an Inexpensive Very High Resolution Scan Camera System. *Computer Graphics Forum (Eurographics)* 23, 10, 441–450.
- WEHNER, R. 1976. Polarized-Light Navigation by Insects. *Scientific American* 235, 106115.
- WETZSTEIN, G., IHRKE, I., AND HEIDRICH, W. 2010. Sensor Saturation in Fourier Multiplexed Imaging. In *Proc. IEEE CVPR*, 1–8.
- WETZSTEIN, G., IHRKE, I., LANMAN, D., AND HEIDRICH, W. 2011. Computational Plenoptic Imaging. *Computer Graphics Forum* 30, 23972426.
- WETZSTEIN, G., LANMAN, D., HEIDRICH, W., AND RASKAR, R. 2011. Layered 3D: Tomographic Image Synthesis for Attenuation-based Light Field and High Dynamic Range Displays. *ACM Trans. Graph. (Siggraph)*.
- WETZSTEIN, G., IHRKE, I., GUKOV, A., AND HEIDRICH, W. 2011. Towards a Database of High-dimensional Plenoptic Images. In *Proc. ICCP (Poster)*.
- WETZSTEIN, G., RASKAR, R., AND HEIDRICH, W. 2011. Hand-Held Schlieren Photography with Light Field Probes. In *Proc. ICCP*, 1–8.
- WETZSTEIN, G., LANMAN, D., HIRSCH, M., AND GUTIERREZ, D. 2012. Computational Displays. In *ACM SIGGRAPH 2012 Courses*.
- WILBURN, B., SMULSKI, M., LEE, K., AND HOROWITZ, M. A. 2002. The Light Field Video Camera. In *SPIE Electronic Imaging*, 29–36.
- WILBURN, B., JOSHI, N., VAISH, V., LEVOY, M., AND HOROWITZ, M. 2004. High Speed Video Using a Dense Array of Cameras. In *Proc. IEEE CVPR*, 1–8.

- WILBURN, B., JOSHI, N., VAISH, V., TALVALA, E.-V., ANTUNEZ, E., BARTH, A., ADAMS, A., HOROWITZ, M., AND LEVOY, M. 2005. High Performance Imaging using Large Camera Arrays. *ACM Trans. Graph. (Siggraph)* 24, 3, 765–776.
- WOLFF, L. B., AND BOULT, T. E. 1991. Constraining Object Features using a Polarization Reflectance Model. *IEEE Trans. PAMI* 13, 7, 635–657.
- WYSZECKI, G., AND STILES, W. 1982. *Color Science*. John Wiley and Sons, Inc.
- YANG, J., LEE, C., ISAKSEN, A., AND MCMILLAN, L., 2000. A Low-Cost Portable Light Field Capture Device. ACM SIGGRAPH Technical Sketch.
- YANG, J. C., EVERETT, M., BUEHLER, C., AND MCMILLAN, L. 2002. A Real-Time Distributed Light Field Camera. In *Proc. EGSR*, 77–86.
- YAO, S., 2008. Optical Communications Based on Optical Polarization Multiplexing and Demultiplexing. US Patent 7,343,100.
- YASUMA, F., MITSUNAGA, T., ISO, D., AND NAYAR, S. K. 2010. Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum. *IEEE Trans. Im. Proc.* 99.
- ZHANG, C., AND CHEN, T. 2005. Light Field Capturing with Lensless Cameras. In *Proc. ICIP*, III – 792–5.
- ZHOU, C., AND NAYAR, S. 2009. What are Good Apertures for Defocus Deblurring? In *Proc. ICCP*, 1–8.
- ZHOU, C., LIN, S., AND NAYAR, S. K. 2009. Coded Aperture Pairs for Depth from Defocus. In *Proc. ICCV*.
- ZWICKER, M., MATUSIK, W., DURAND, F., AND PFISTER, H. 2006. Antialiasing for automultiscopic 3D displays. In *Eurographics Symposium on Rendering*.

Computational Plenoptic Imaging

Gordon Wetzstein¹
Wolfgang Heidrich³

Ivo Ihrke²
Kurt Akeley⁴

Douglas Lanman¹
Ramesh Raskar¹

¹MIT Media Lab

²Saarland University

³University of British Columbia

⁴Lytro, Inc.

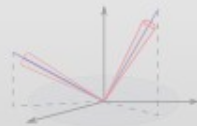
II. High Dynamic Range Imaging



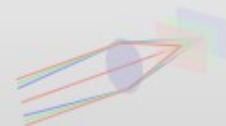
Dynamic Range



Color Spectrum



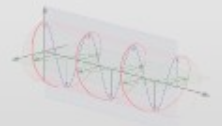
Directions | Light Fields



Space | Focal Surfaces



Time



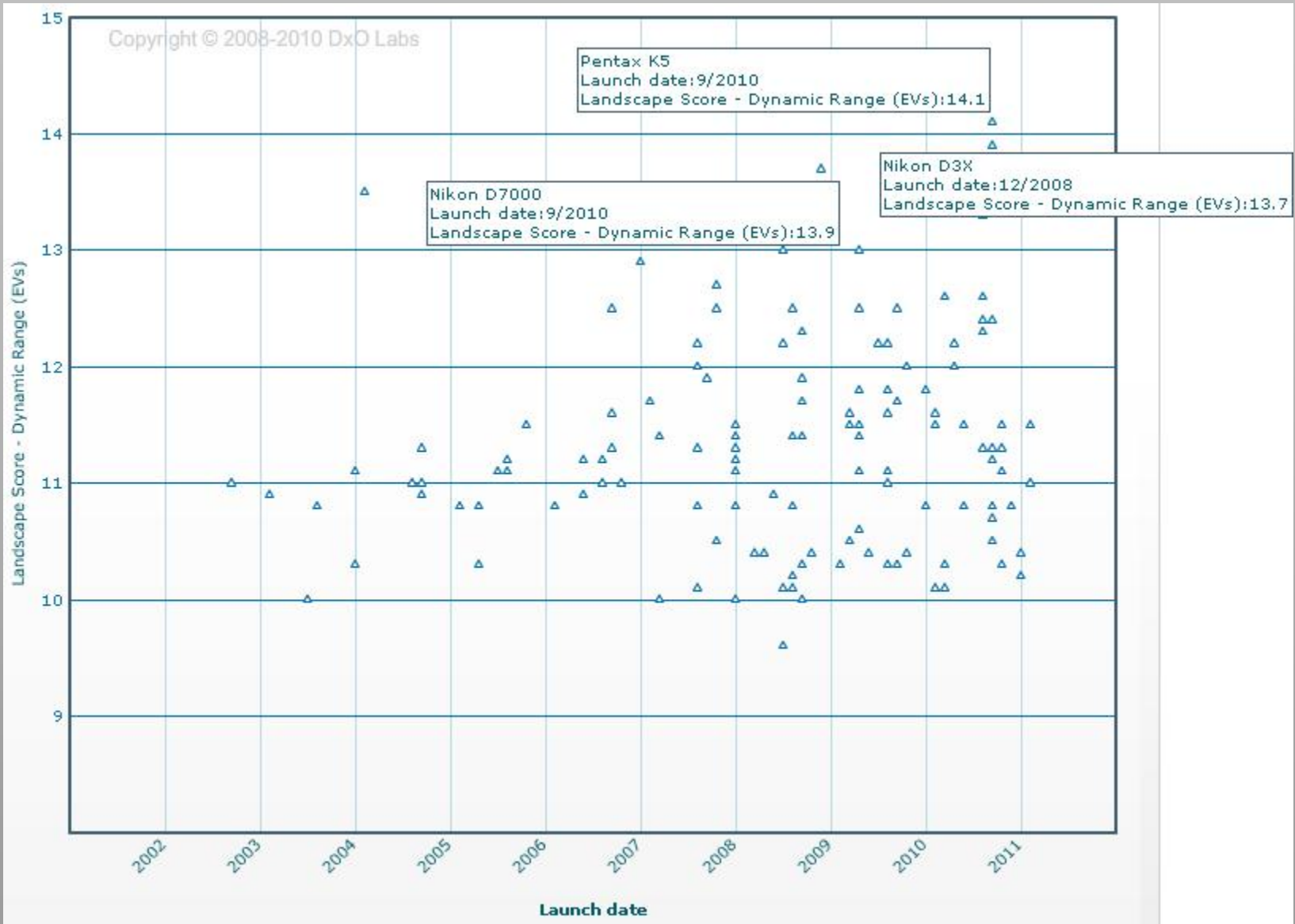
Further Properties

What is HDR?

http://en.wikipedia.org/wiki/High-dynamic-range_imaging

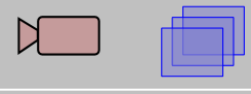


Dynamic Range of Standard Sensors



13.5 EVs or f-stops = contrast 11,000:1 = color negative

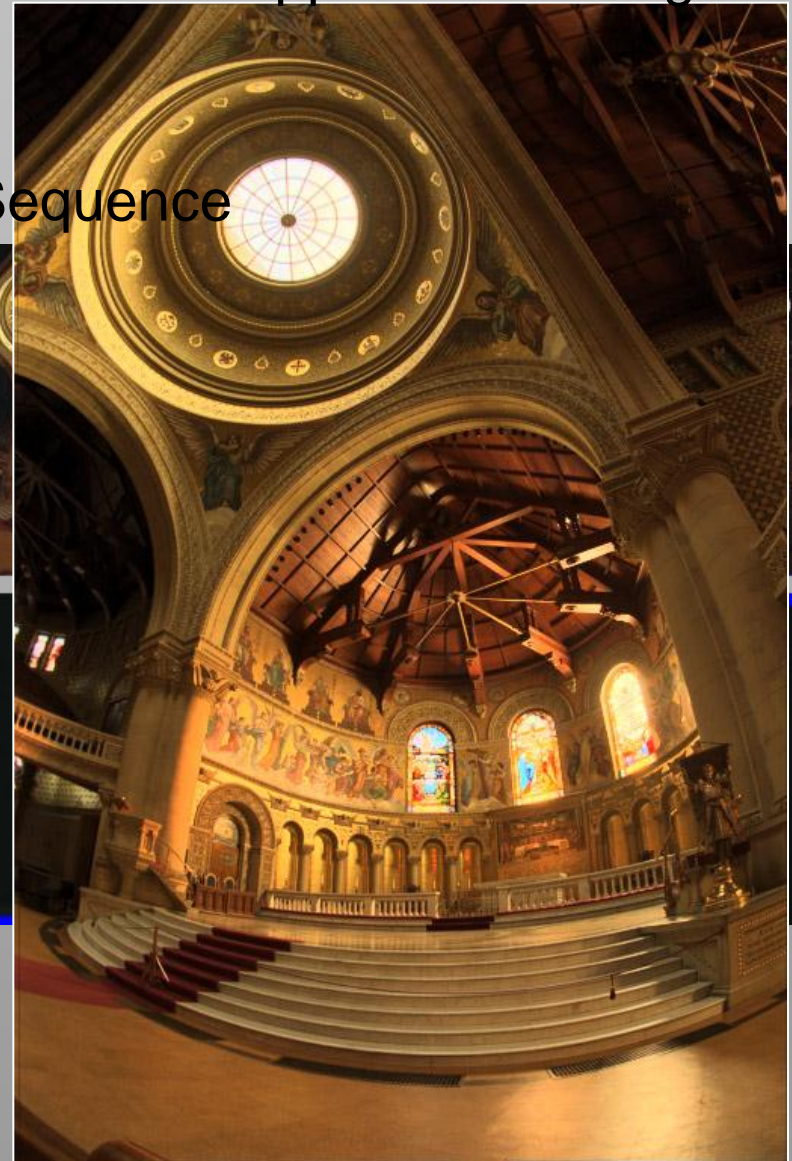
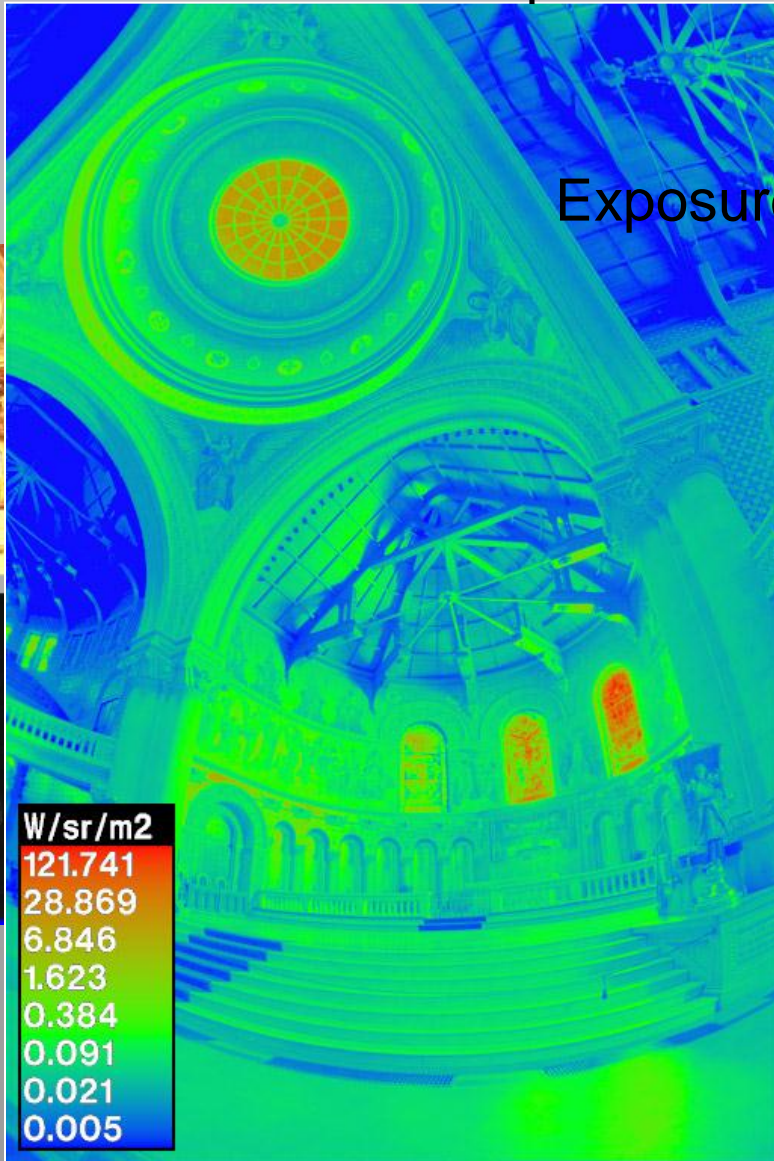
HDR Acquisition – Exposure Brackets



Radiance Map

Tonemapped HDR Image

Exposure Sequence



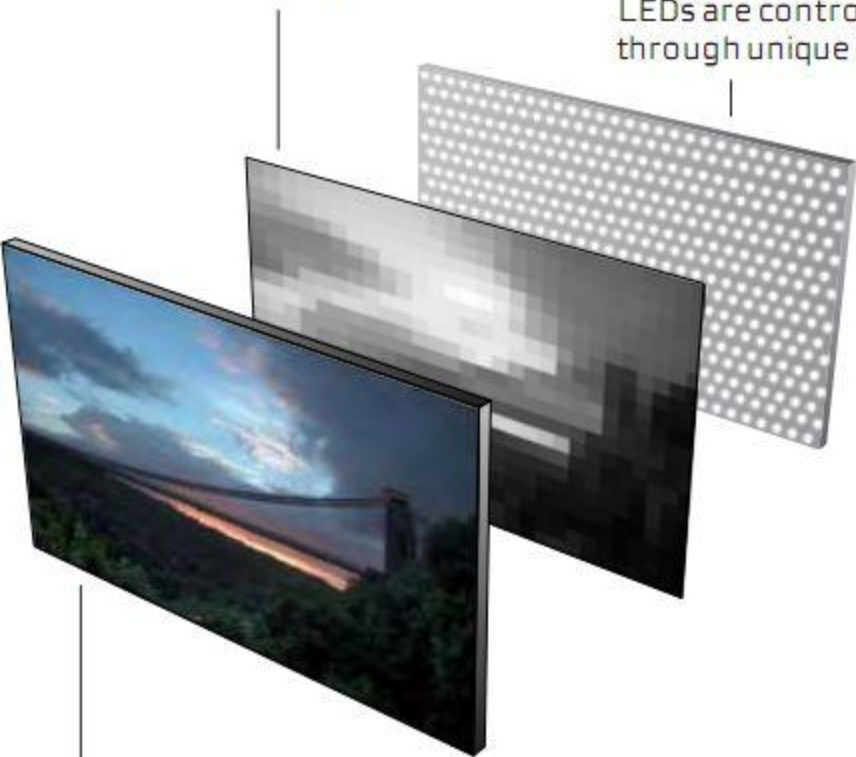
[Debevec & Malik 97]

HDR Display

Backlight Simulation

LED Dynamic Backlight

LEDs are controlled through unique signals.



LCD

Provides color, resolution, and contrast. Contrast and image created by combining LED and LCD images.



- 47" TFT LCD, LED backlight
- aspect ratio 16:9
- resolution 1920 x 1080
- contrast >1,000,000:1
- brightness 4,000 cd/m²

Images courtesy Dolby

Image Based Lighting



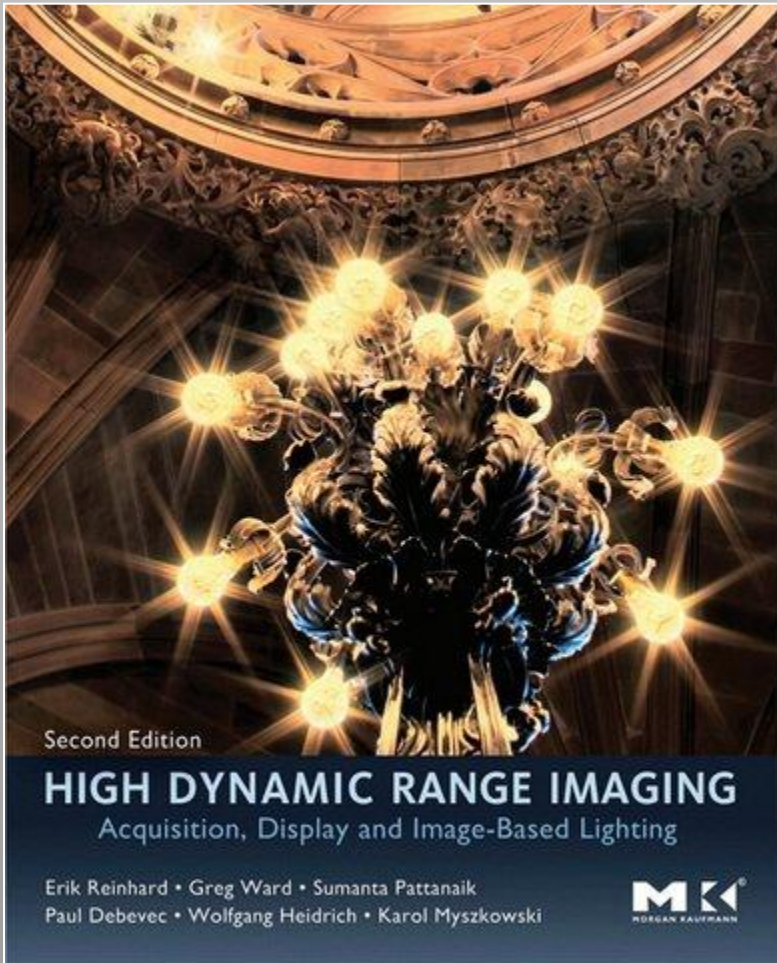
Image-Based Lighting

Rendered with *RADIANCE*



Slides by Paul Debevec

Textbook

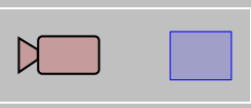


- HDR image / video encoding
- capture, display, tone reproduction
- visible difference predictors
- image based lighting, etc.

We cover some specialized acquisition approaches here.

II.I Single-Shot Acquisition

HDR Cameras



Grass Valley Viper



Spheron



Panavision Genesis



HDR camera and CG
by Spheron VR
www.spheron.com

Per-Pixel Exposure Control

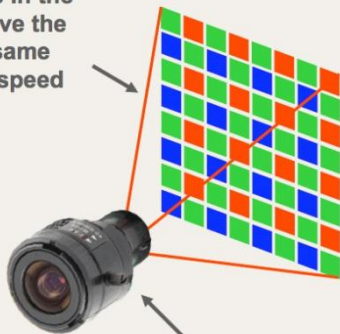


www.pixim.com

Analog CCD Security Camera



All pixels in the array have the exact same shutter speed



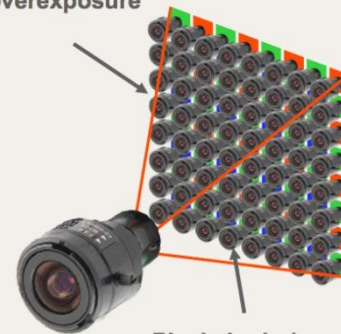
One aperture

- Every pixel in a CCD camera receives exactly the same exposure
- Result
 - Bright areas are overexposed
 - Dark areas are underexposed

Pixim's Digital Pixel System® Technology



Pixels in bright areas automatically adjust to eliminate overexposure



Pixels in dark areas automatically adjust to eliminate underexposure

- Only all-digital solution
- Every pixel automatically adjusts to produce an optimal exposure
- Its like having over 400,000 self-adjusting cameras inside – one for every pixel
- Every Pixel Tells a Story



no pixim



with pixim



no pixim

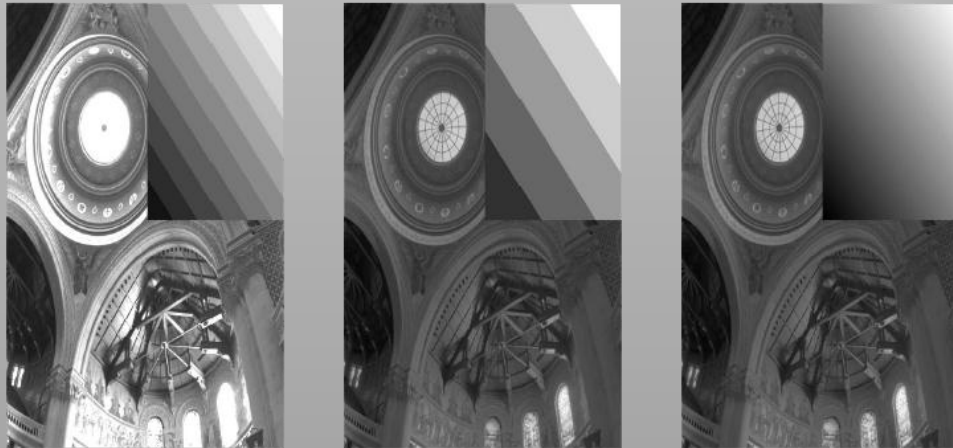


with pixim

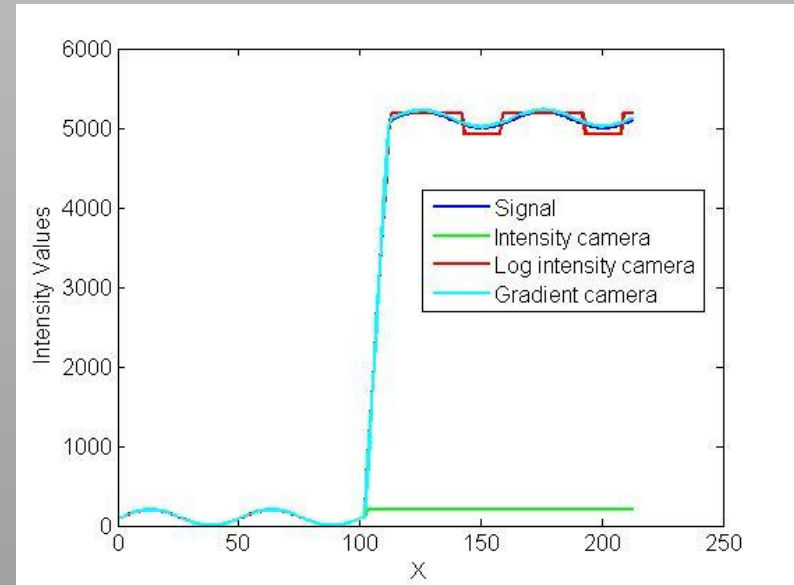
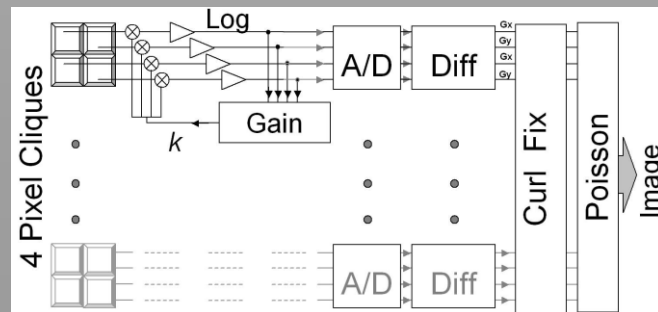
Gradient Camera



- Measure log-gradient
- Reconstruct intensity with Poisson solver

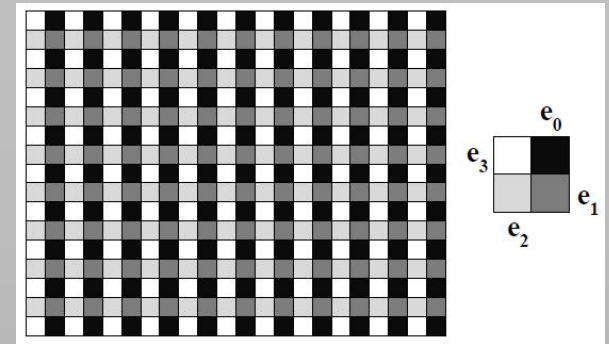


Intensity Log-Intensity Log-Gradient



[Tumblin et al. 05]

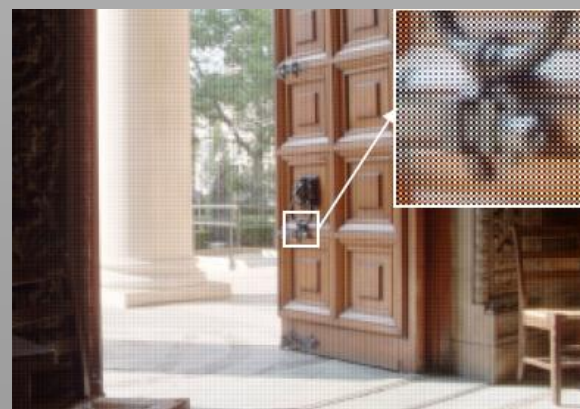
Assorted Pixels



- ND filter arrays
- Less flexible and costly than Pixim



Conventional Camera



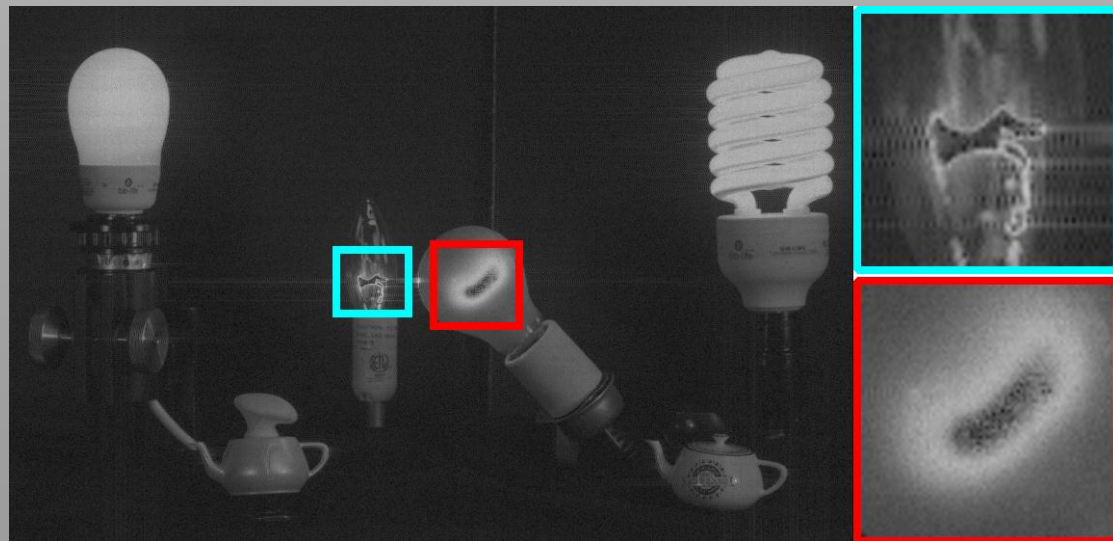
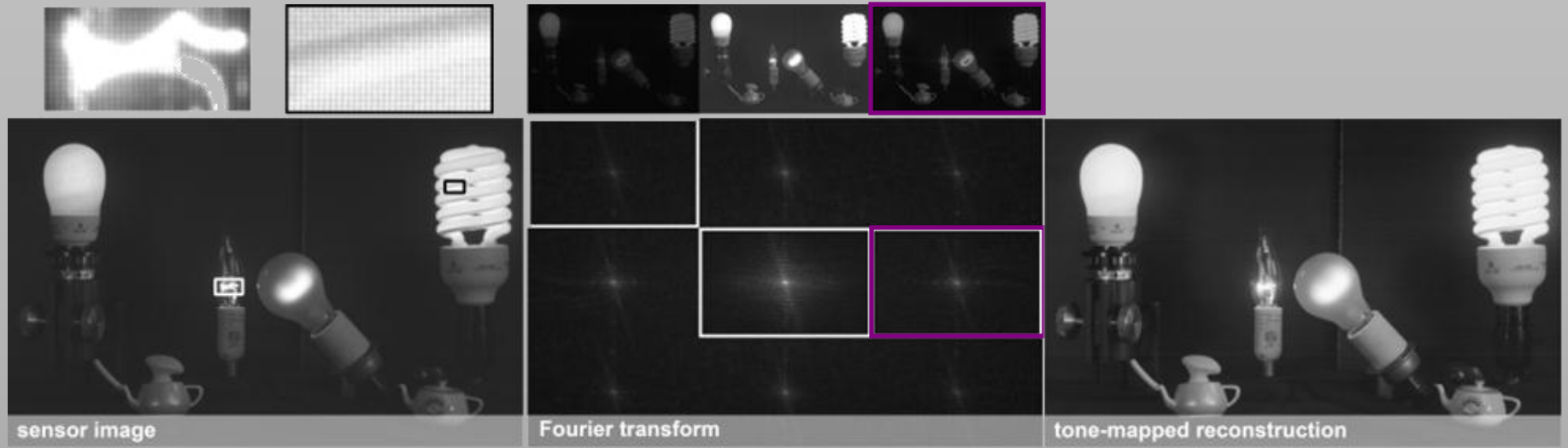
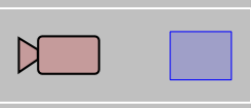
SVE Camera



SVE Reconstruction

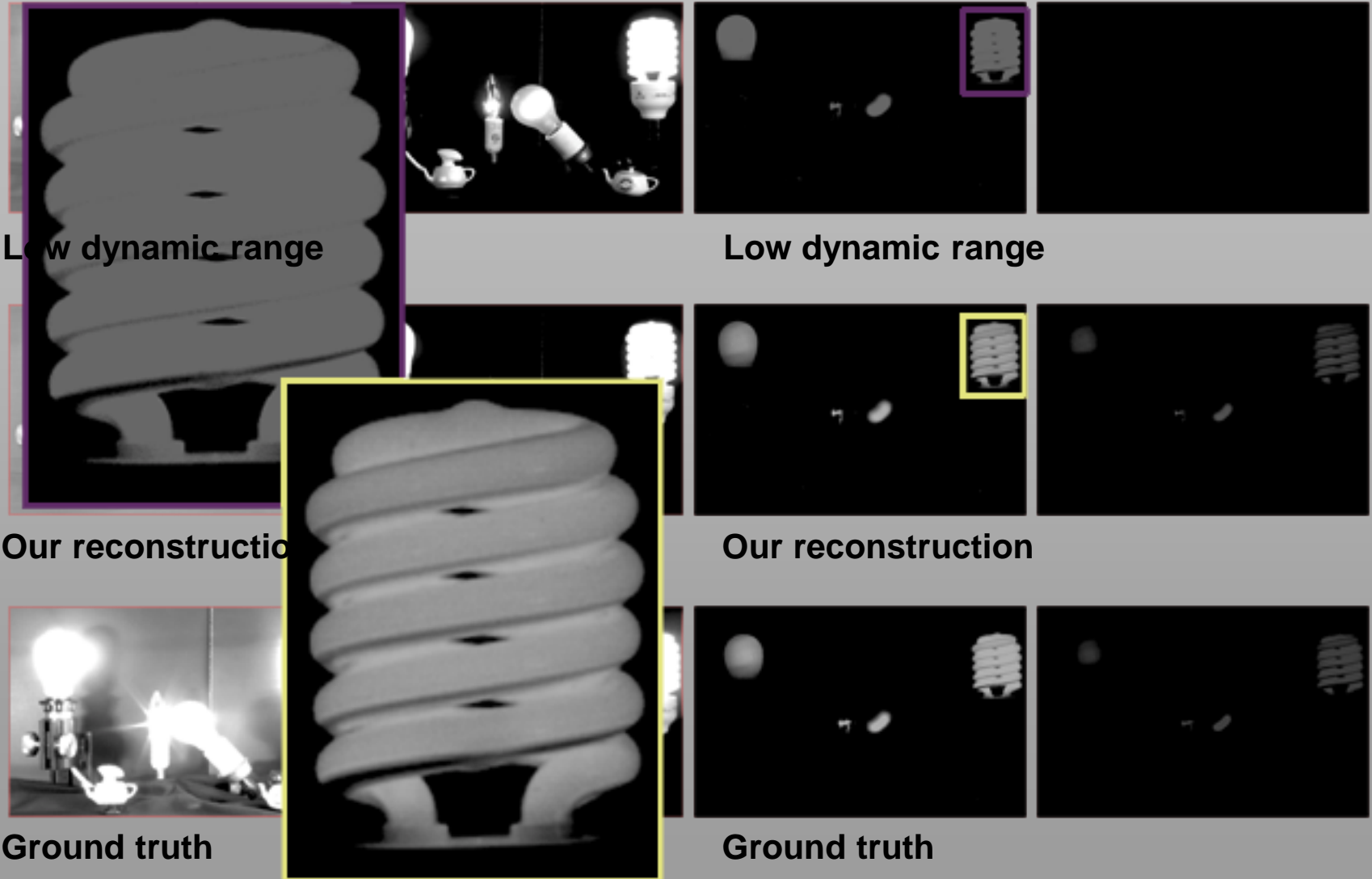
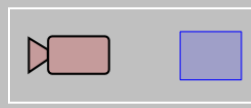
[Nayar and Mitsunaga 00, Narasimhan and Nayar 05]

Fourier-based Analysis and Reconstruction



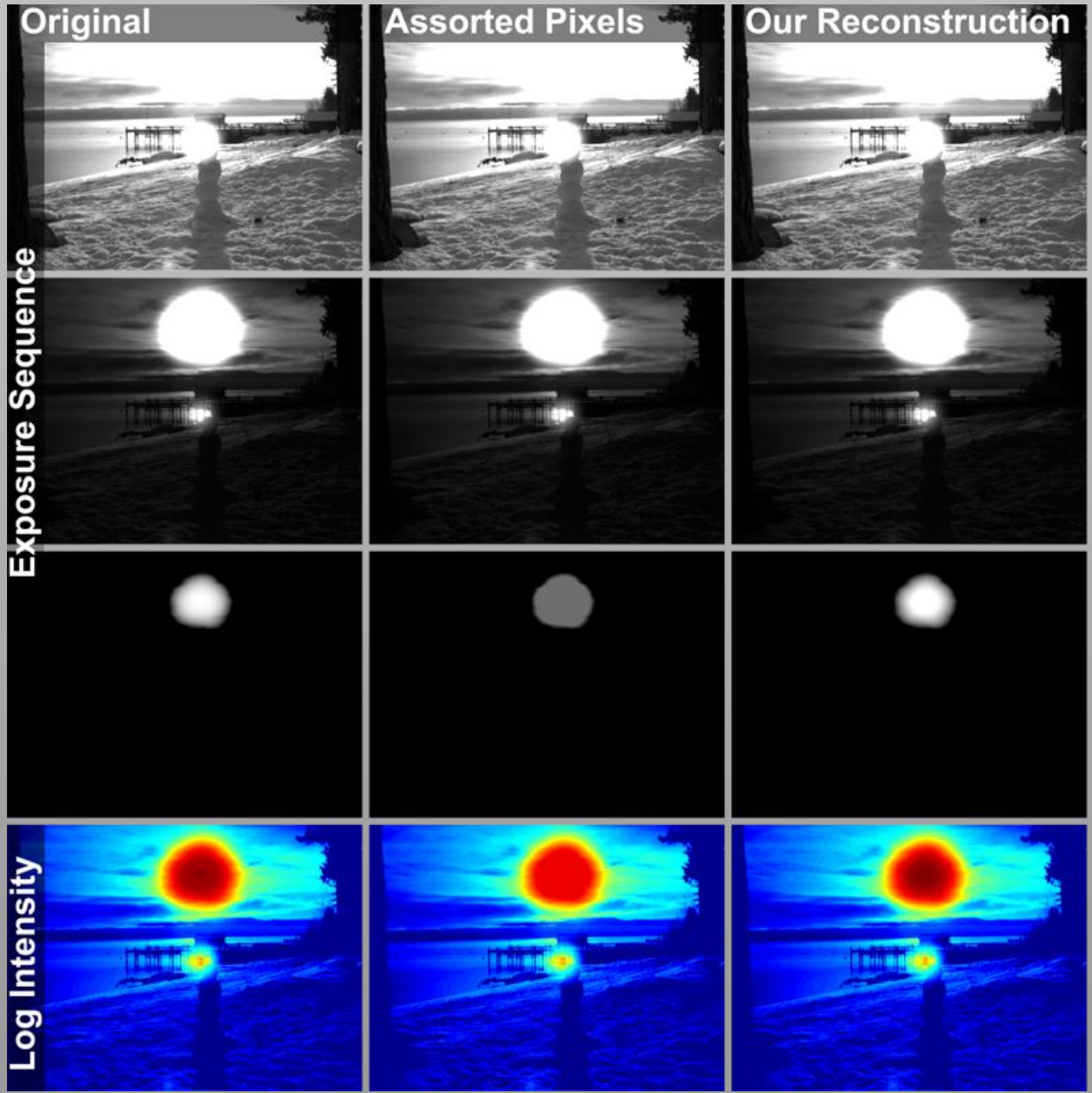
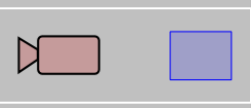
[Wetzstein et al. 10]

Fourier-based Analysis and Reconstruction



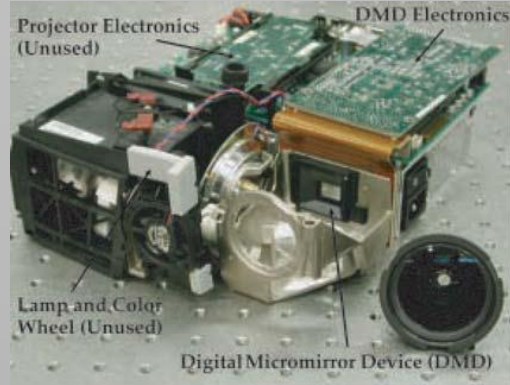
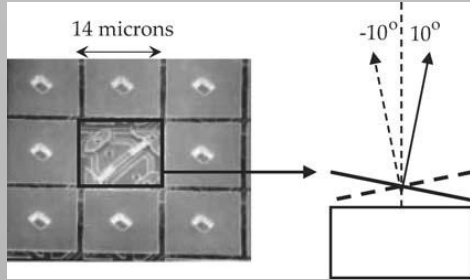
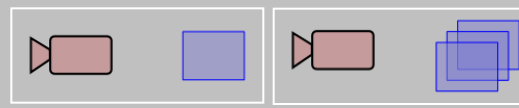
[Wetzstein et al. 10]

Fourier-based Analysis and Reconstruction



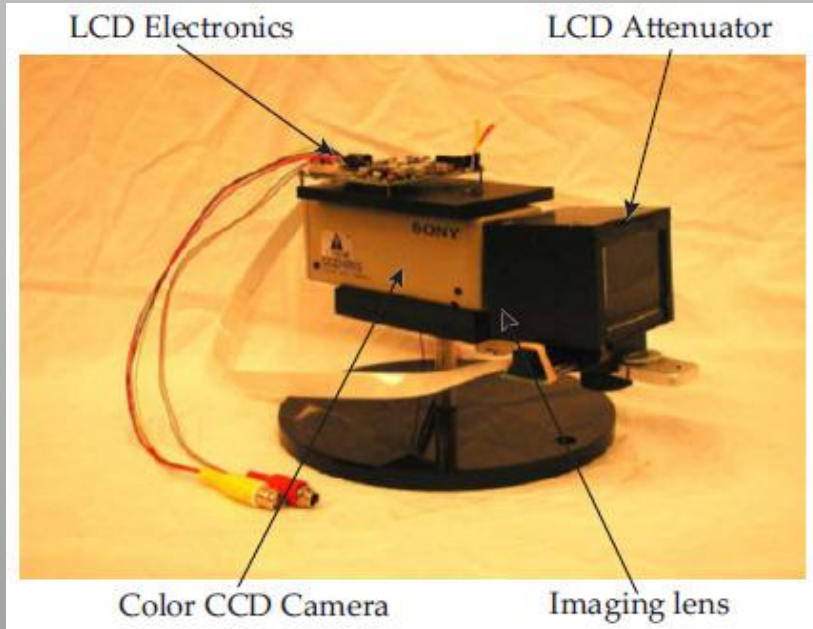
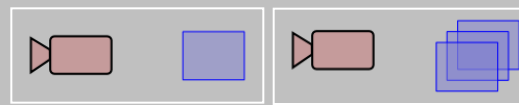
[Wetzstein et al. 10]

Programmable Imaging



[Nayar & Branzoi 06]

Adaptive Dynamic Range Imaging



[Nayar and Branzoi 03]

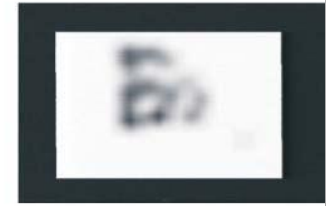
Conventional Camera
(without ADR)



Camera with Adaptive
Dynamic Range (ADR)



Transmittance Function
(LCD Input)



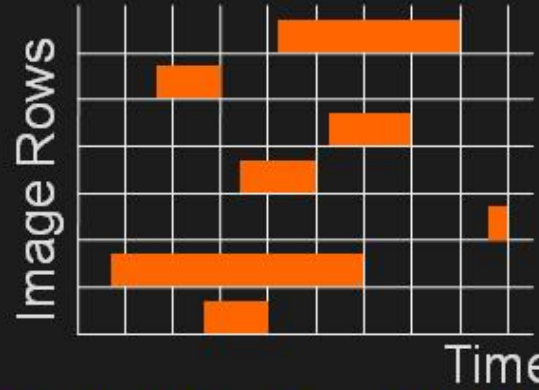
Coded Rolling Shutter



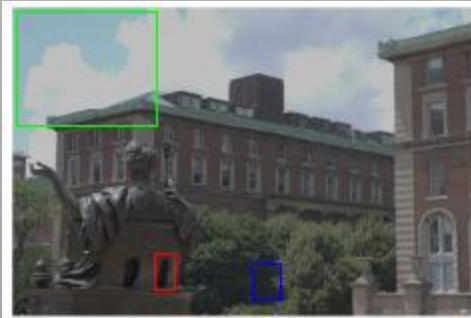
CMOS Image Sensor



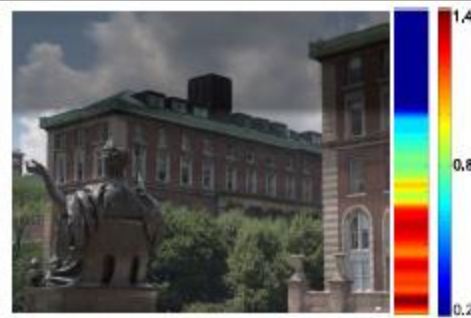
Rolling Shutter (CMOS)



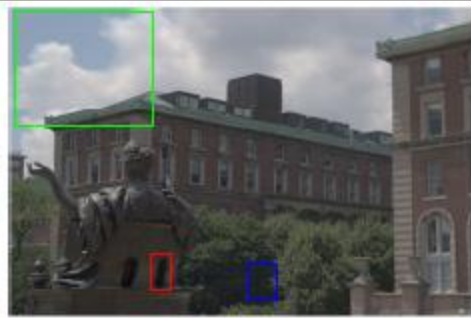
Coded Rolling Shutter



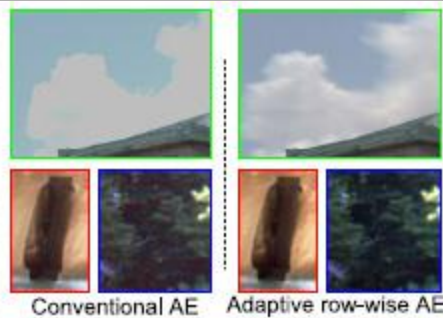
(a) Conventional AE



(b) Input: $I(x, y)$ and $\Delta t_e(y)$



(c) Output: adaptive row-wise AE



(d) Insets of (a) and (c)



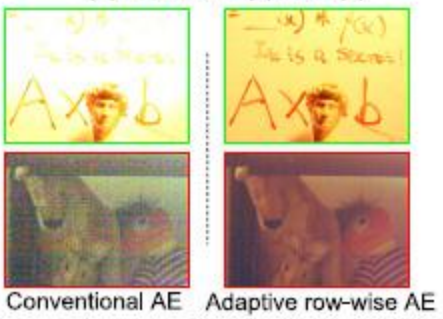
(a) Conventional AE



(b) Input: $I(x, y)$ and $\Delta t_e(y)$

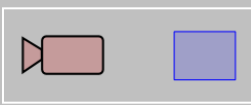


(c) Output: adaptive row-wise AE



(d) Insets of (a) and (c)

Frankencamera

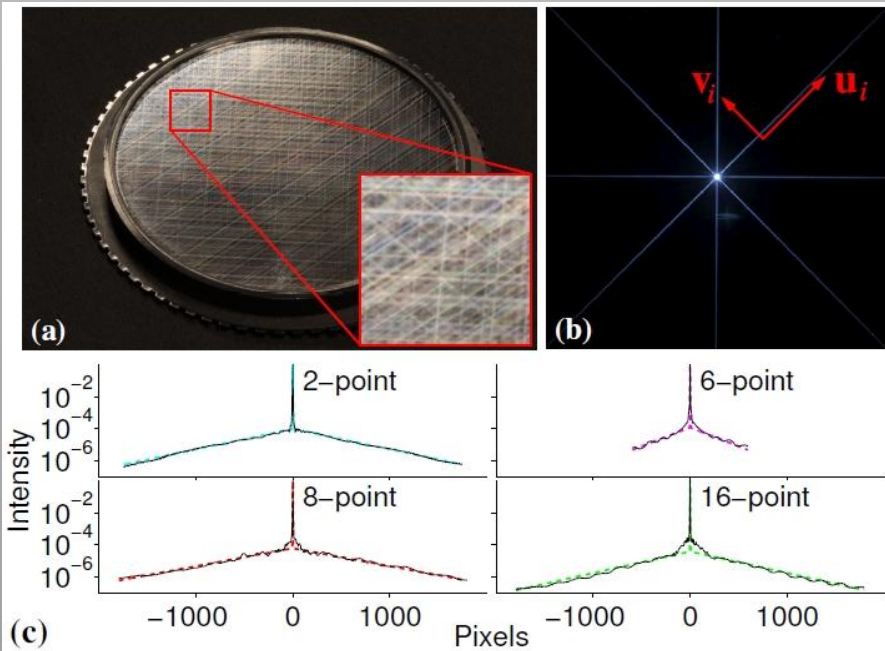


- Version 3.0 is supposed to have non-destructive ROI readout

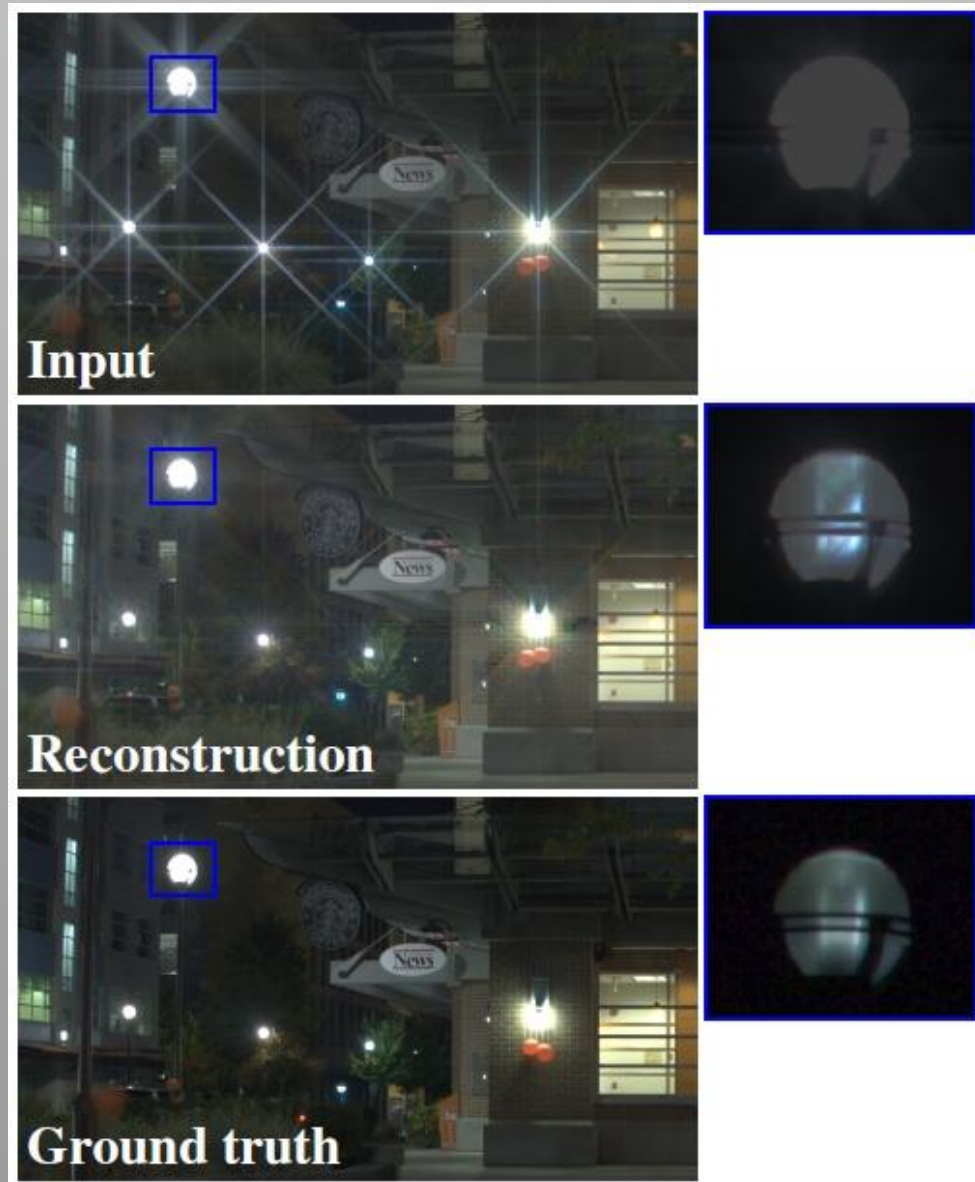


[Adams et al. 10, Levoy 10]

Tomographic HDR from Star Filter

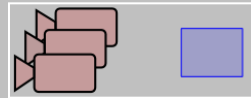


[Rouf et al. 11]

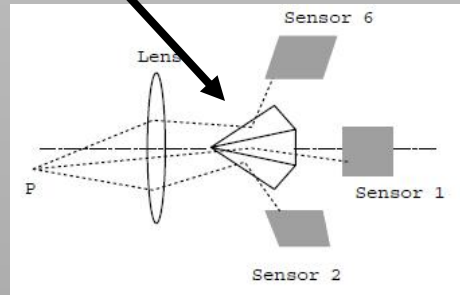
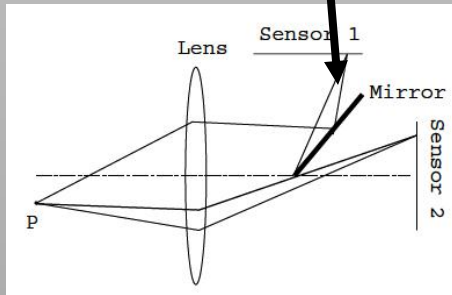
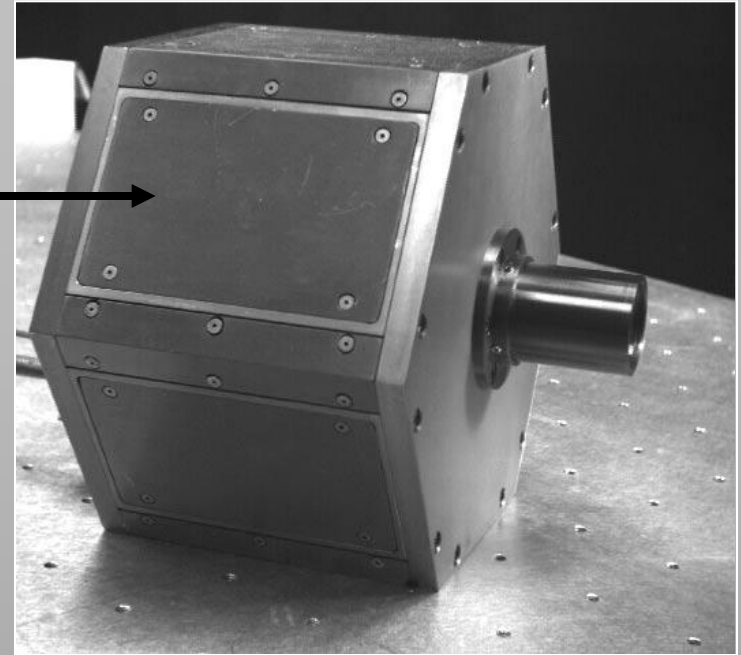


II.II Multi-Device Techniques

Split Aperture Imaging

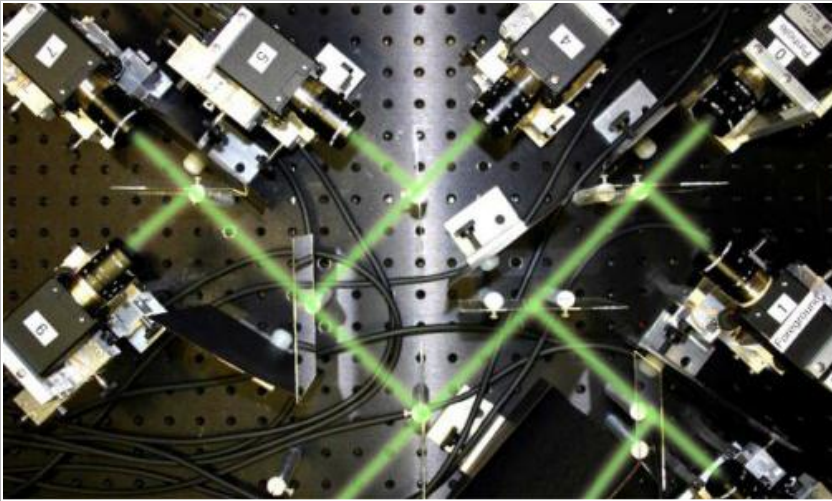
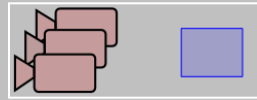


- Multiple sensors
- Mirror pyramid in aperture

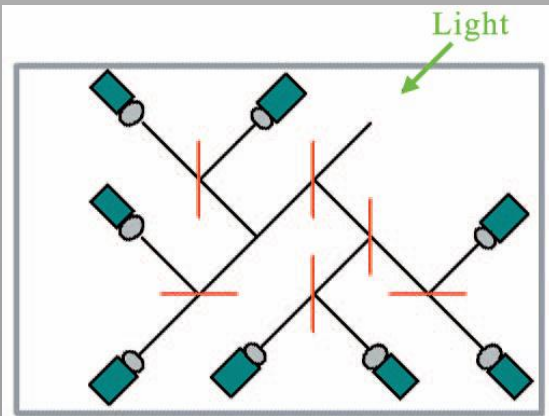


[Aggrawal & Ahuja 04]

Optical Splitting Trees



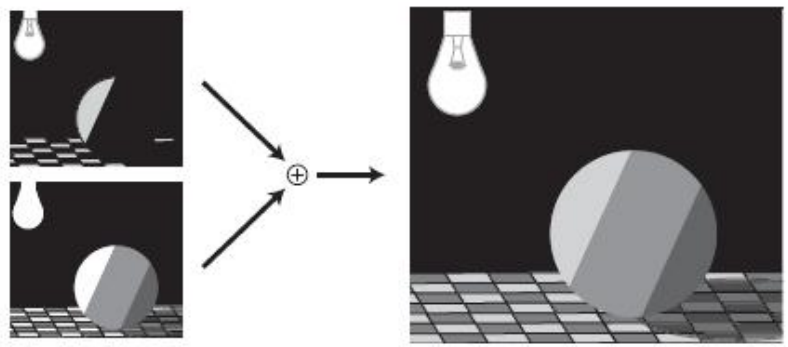
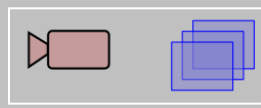
Exposure Sequence & Tonemapped Image



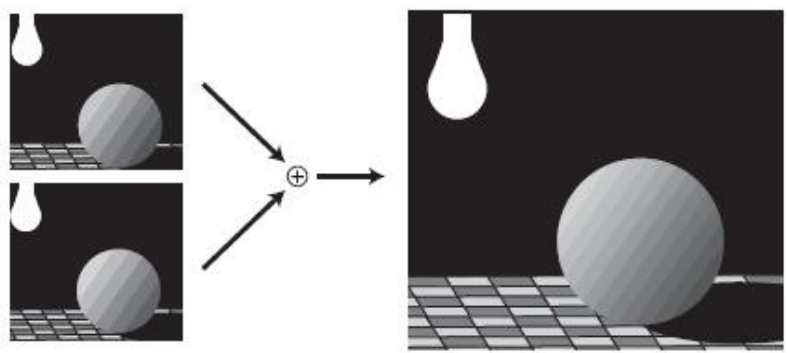
[McGuire et al. 07]

II.III Analysis and Tradeoffs

Which Exposures to Combine?



(a) Small and large exposures combine to capture a high dynamic range

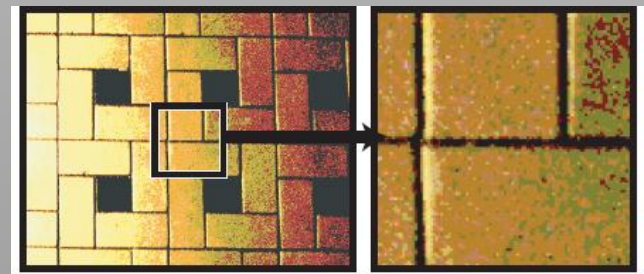


(b) Similar exposures combine to capture subtle variations

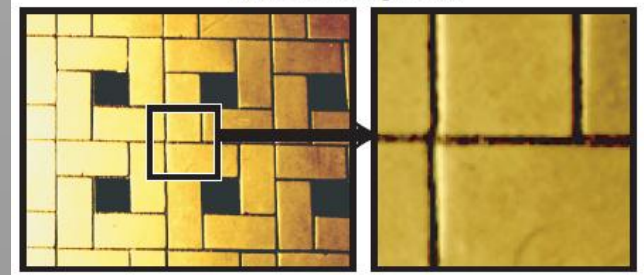
Effective Camera	Number of Exposures	Exposures					
		1	2	3	4	5	6
Linear	2	1	1.003				
	3	1	1.003	2.985			
	4	1	1.015	1.019	3.672		
	5	1	1.003	1.007	1.166	4.975	
	6	1	1.003	1.007	1.031	1.078	5.636
	Gamma = 1/2	2	1	3.094			
Gamma = 1/2	3	1	1.003	5.146			
	4	1	1.003	3.019	11.23		
	5	1	1.003	1.006	5.049	18.56	
	6	1	1.003	1.006	2.866	8.564	33.7
	Constant Contrast (log)	2	1	20.24			
Constant Contrast (log)	3	1	9.91	88.38			
	4	1	4.689	37.23	280		
	5	1	4.831	29.18	144.9	763.5	
	6	1	3.979	16.01	64.22	305.4	1130

Desired Response	
Constant Contrast	
Linear	Dynamic Range
Gamma 1/2	High 1:16,000
	Medium 1:1,000
	Low 1:256

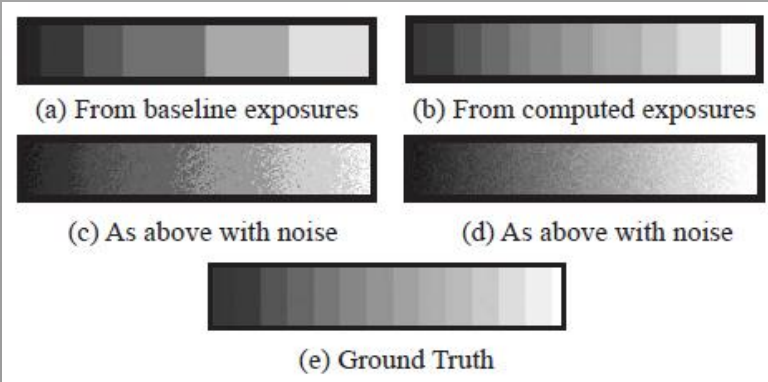
[Grossberg & Nayar 03]



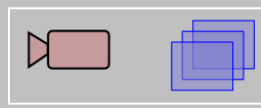
(a) Baseline Exposures



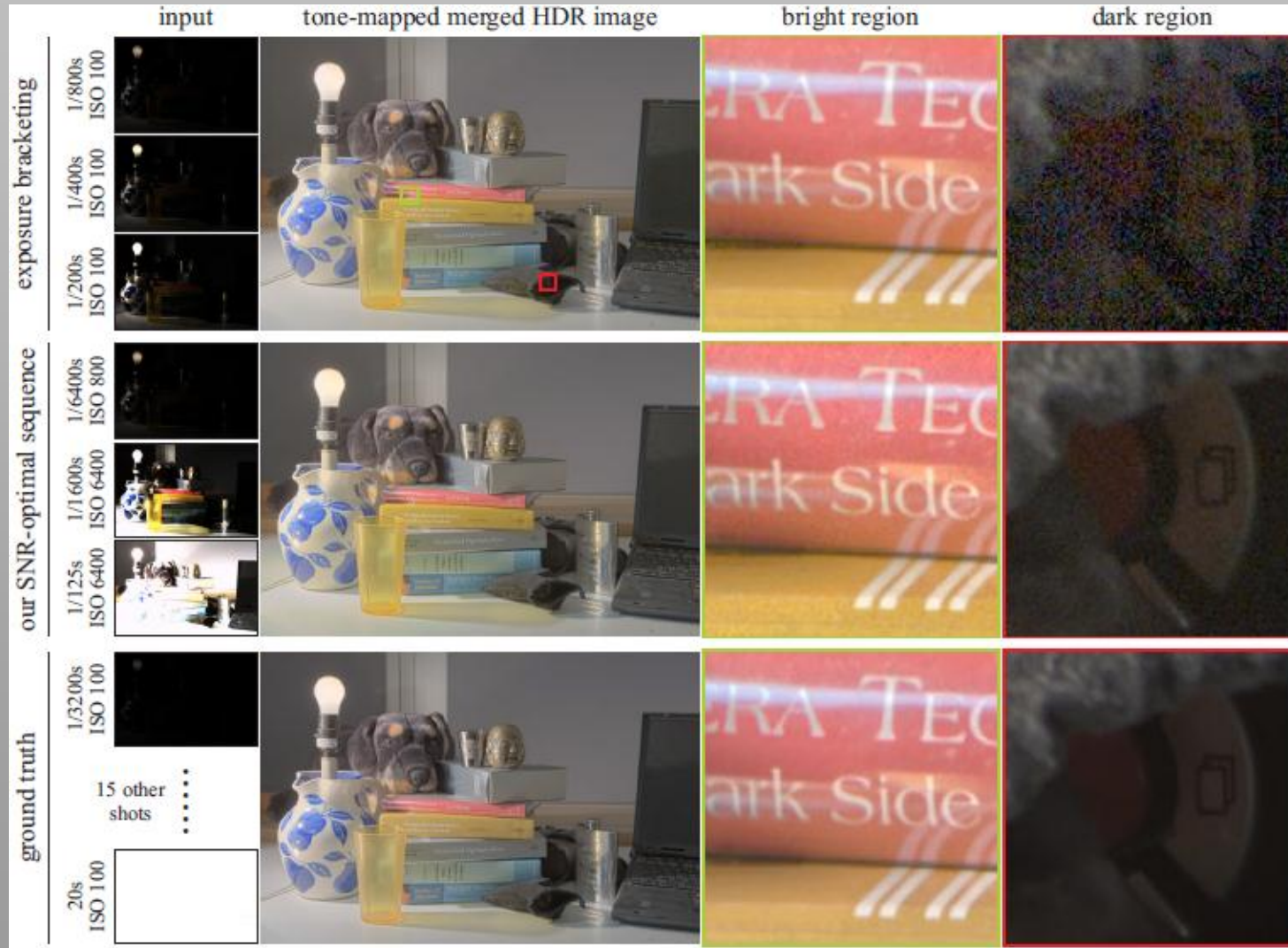
(b) Computed Exposures



Noise Optimal HDR Capture



Vary exposure time and ISO for improved noise in HDR



[Hasinoff et al. 10]

Next: Spectral Imaging

Computational Plenoptic Imaging

Gordon Wetzstein¹
Wolfgang Heidrich³

Ivo Ihrke²
Kurt Akeley⁴

Douglas Lanman¹
Ramesh Raskar¹

¹MIT Media Lab

²Saarland University

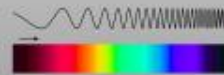
³University of British Columbia

⁴Lytro, Inc.

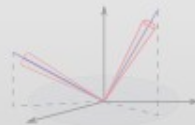
III. Spectral Imaging



Dynamic Range



Color Spectrum



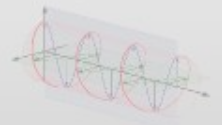
Directions | Light Fields



Space | Focal Surfaces



Time



Further Properties

THE ELECTROMAGNETIC SPECTRUM

ABSORPTION SPECTRA:

HYDROGEN:

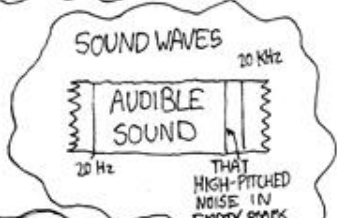


HELIUM:

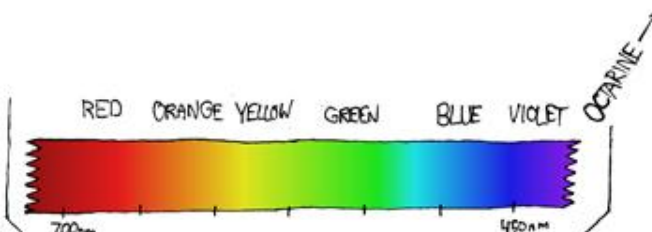
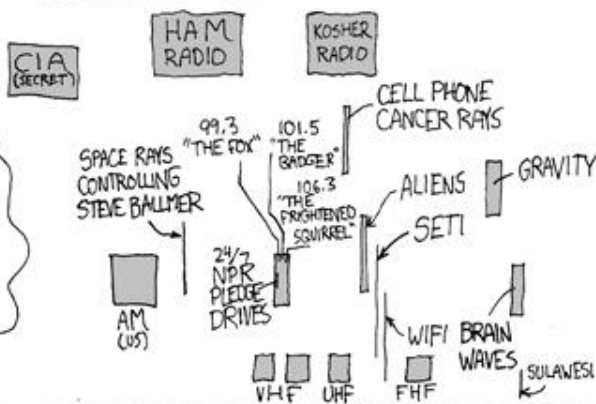


THESE WAVES TRAVEL THROUGH THE ELECTROMAGNETIC FIELD. THEY WERE FORMERLY CARRIED BY THE AETHER, WHICH WAS DECOMMISSIONED IN 1977 DUE TO BUDGET CUTS.

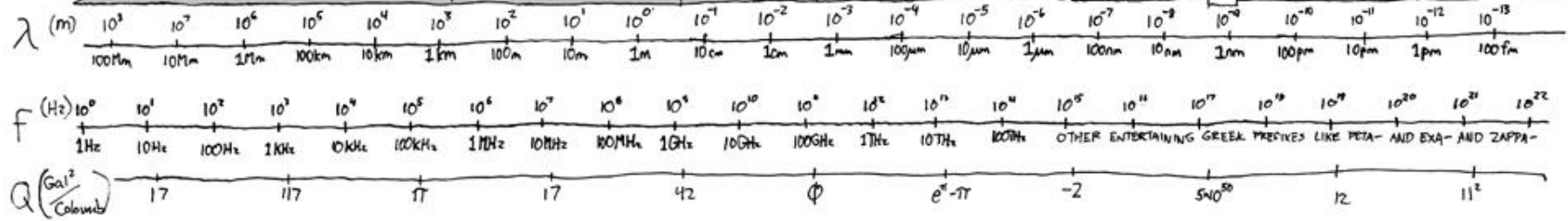
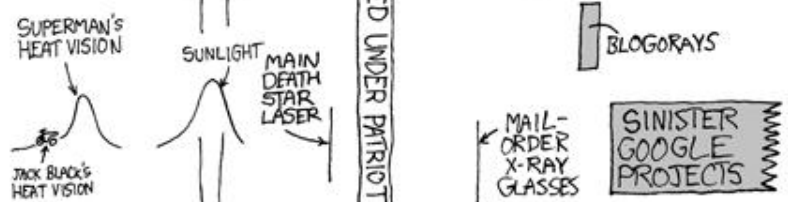
OTHER WAVES:



SHOUTING CAR DEALERSHIP COMMERCIALS

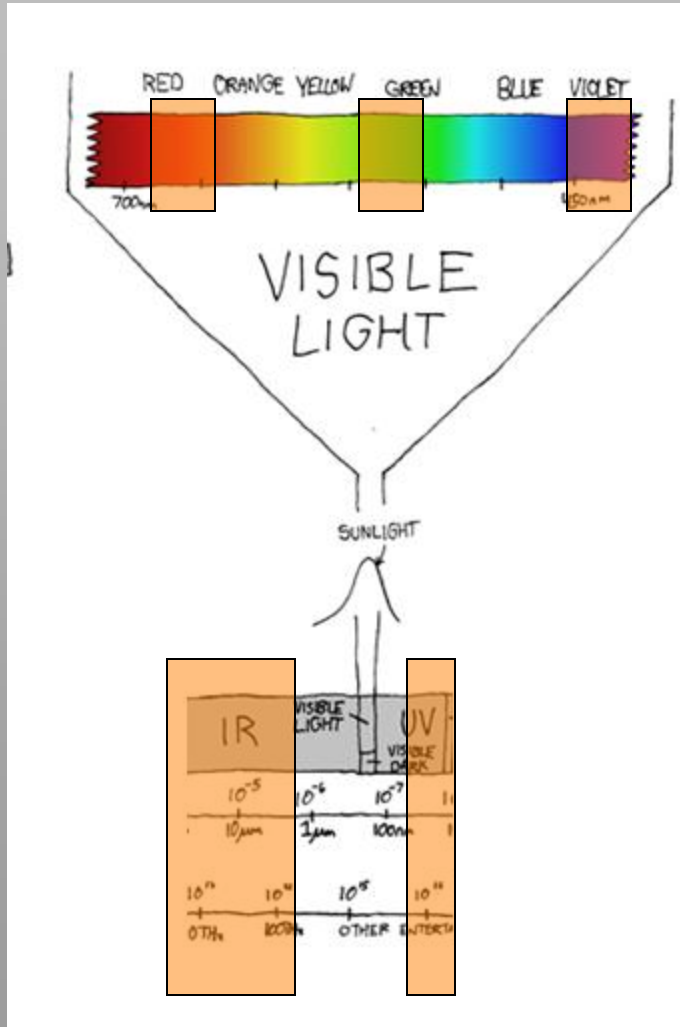


VISIBLE LIGHT

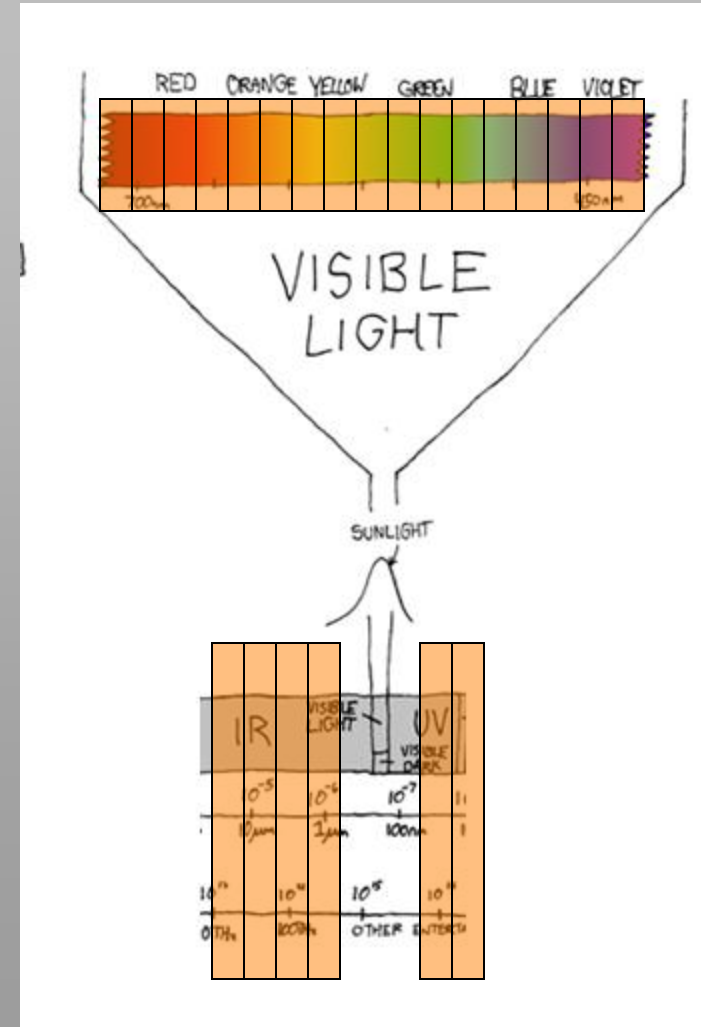


Multi-Spectral vs. Hyper-Spectral

multi-spectral



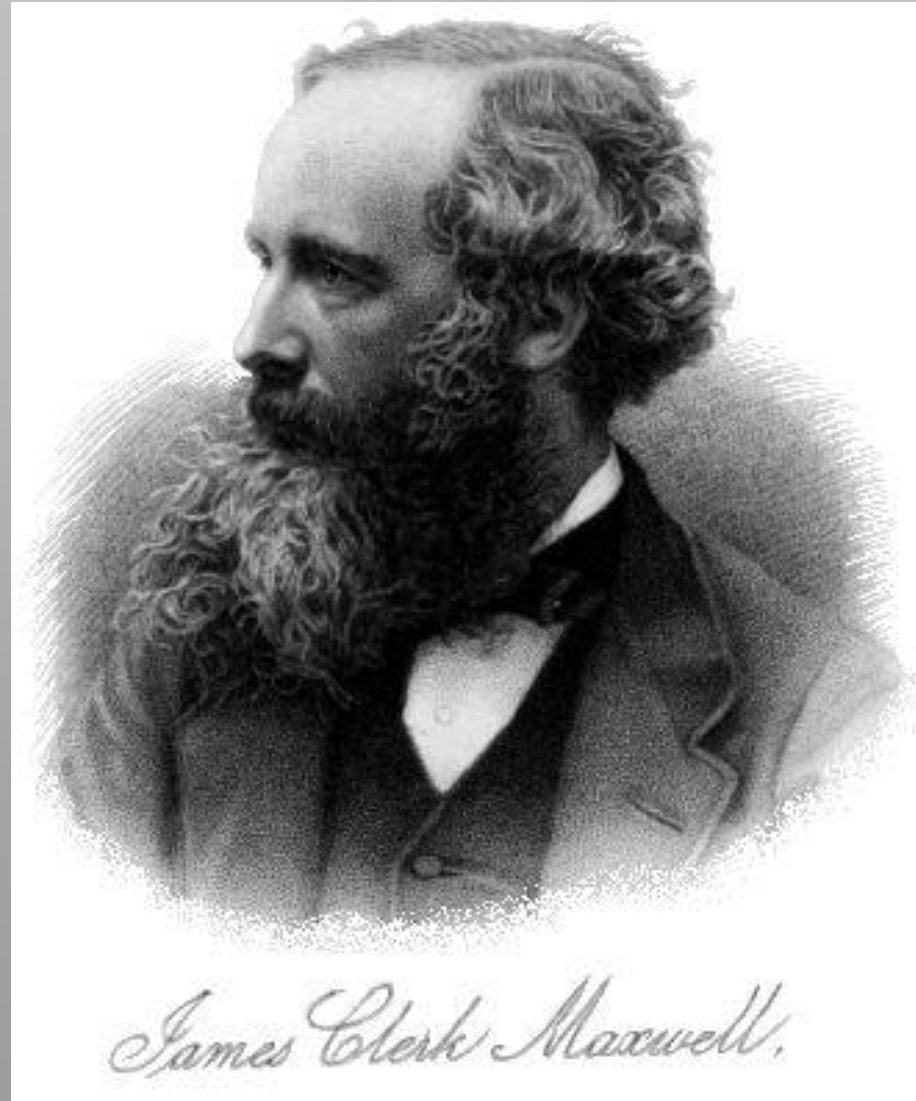
hyper-spectral



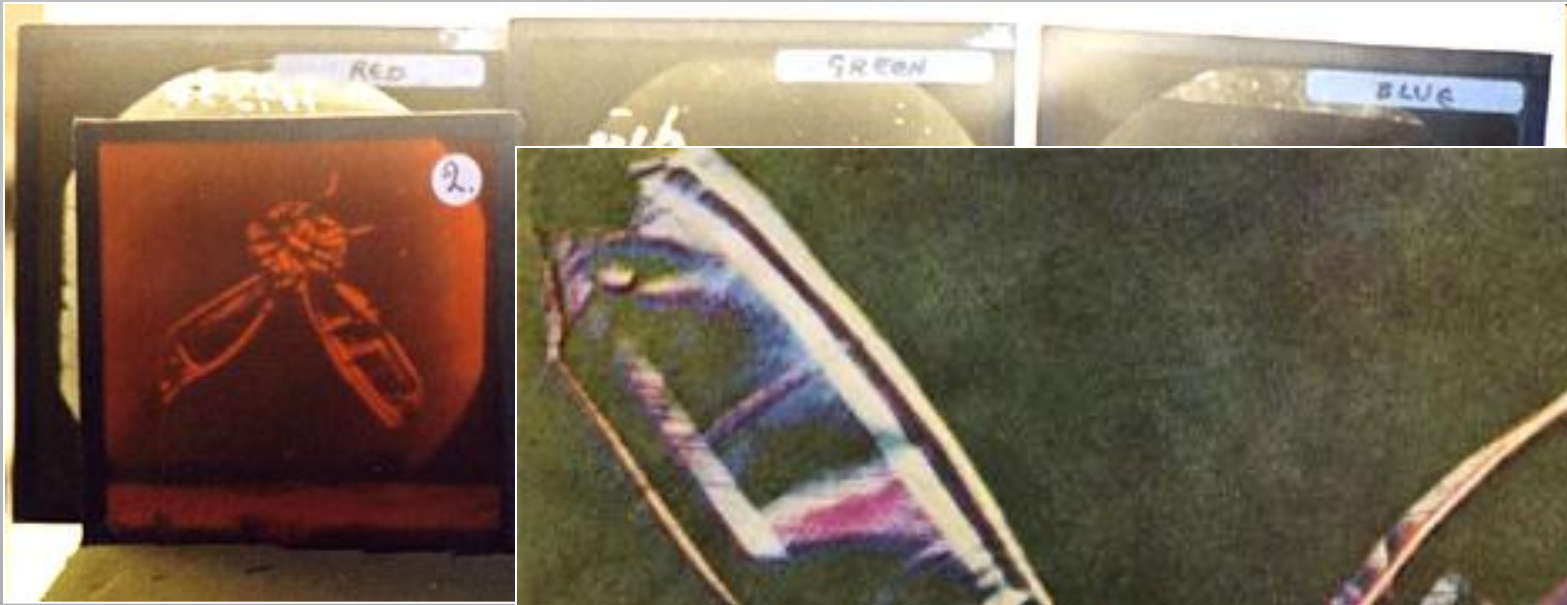
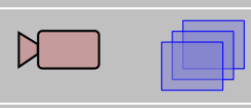
Color Filters

James Clerk Maxwell

1831 - 1879



1861 – first color photograph



© Copyright: For permission to repro



THE LATE THOMAS SUTTON, B.A. (CANTAB.)

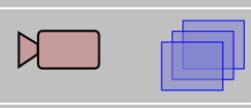
FROM A PHOTOGRAPH BY MR. A. L. HENDERSON.

Examples - Prokudin-Gorskij

- Self-portrait 1915



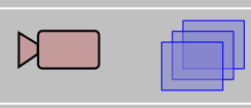
Examples - Prokudin-Gorskij



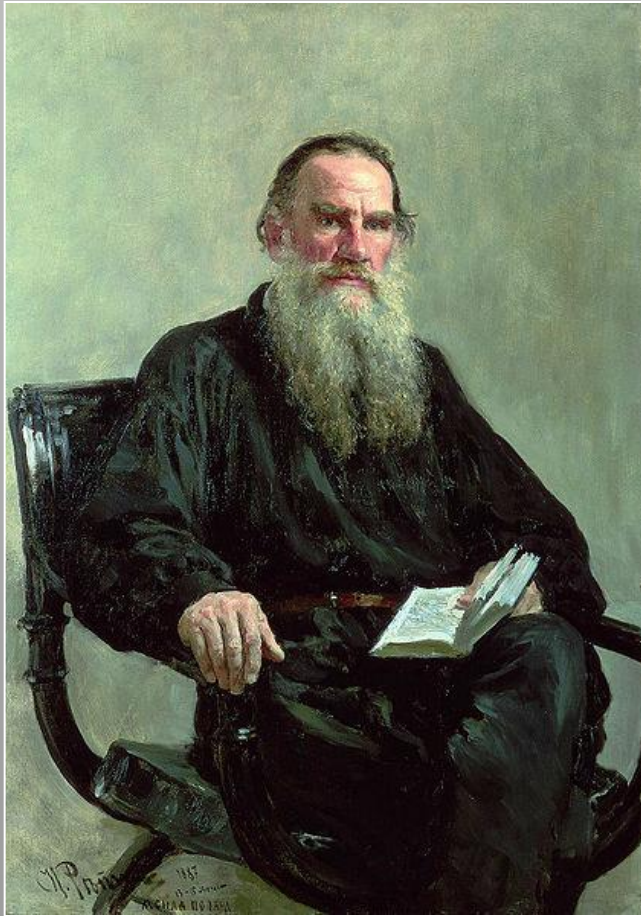
Photograph 1910, Emir of Bukhara , Prokudin-Gorskii



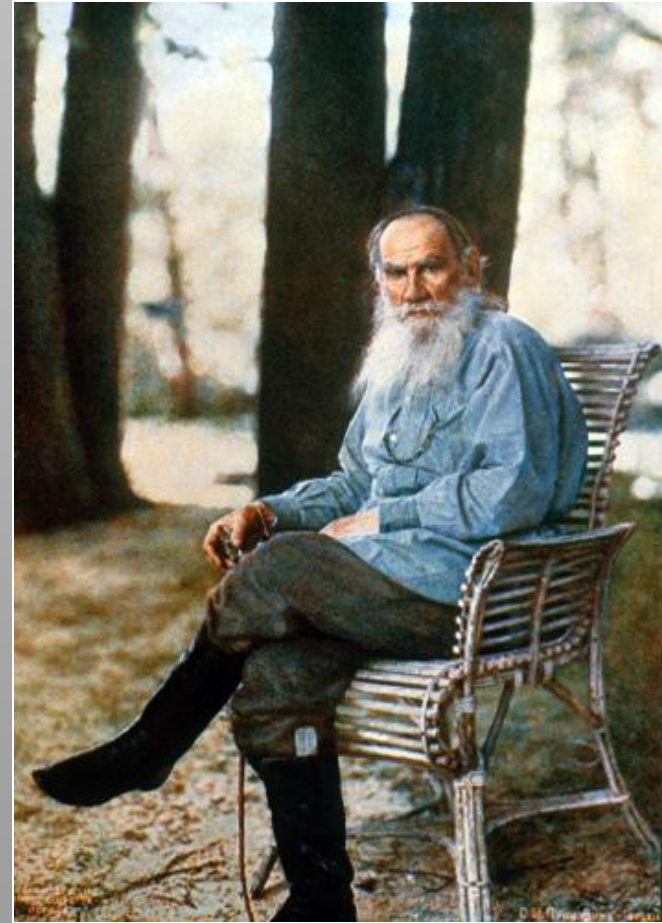
Examples - Prokudin Gorski - Lew Tolstoy



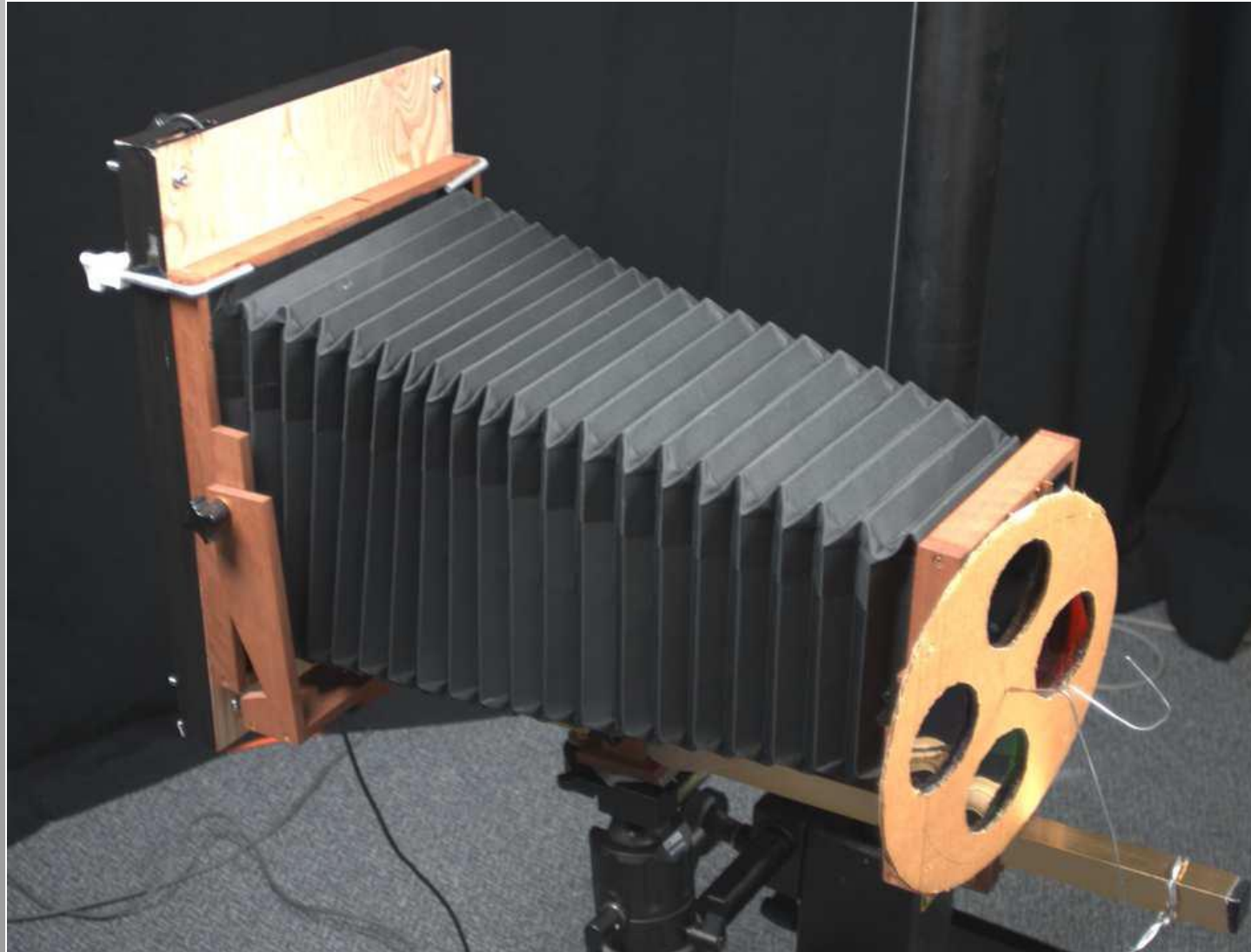
1887 painting, Ilya Repin



1910 photograph, Sergey Prokudin-Gorskii

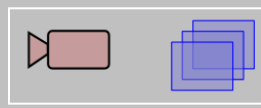


Color Wheel

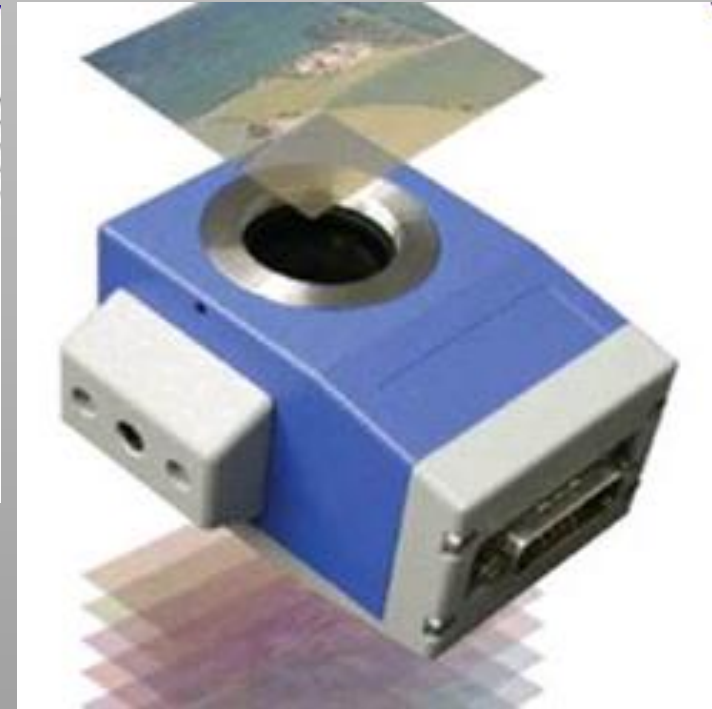
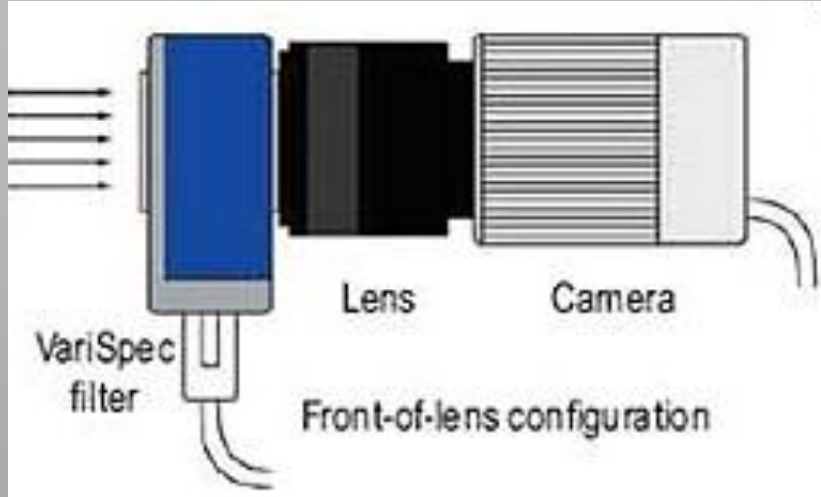


[Wang and Heidrich 2004]

Liquid Crystal Tunable Filter (LCTF)



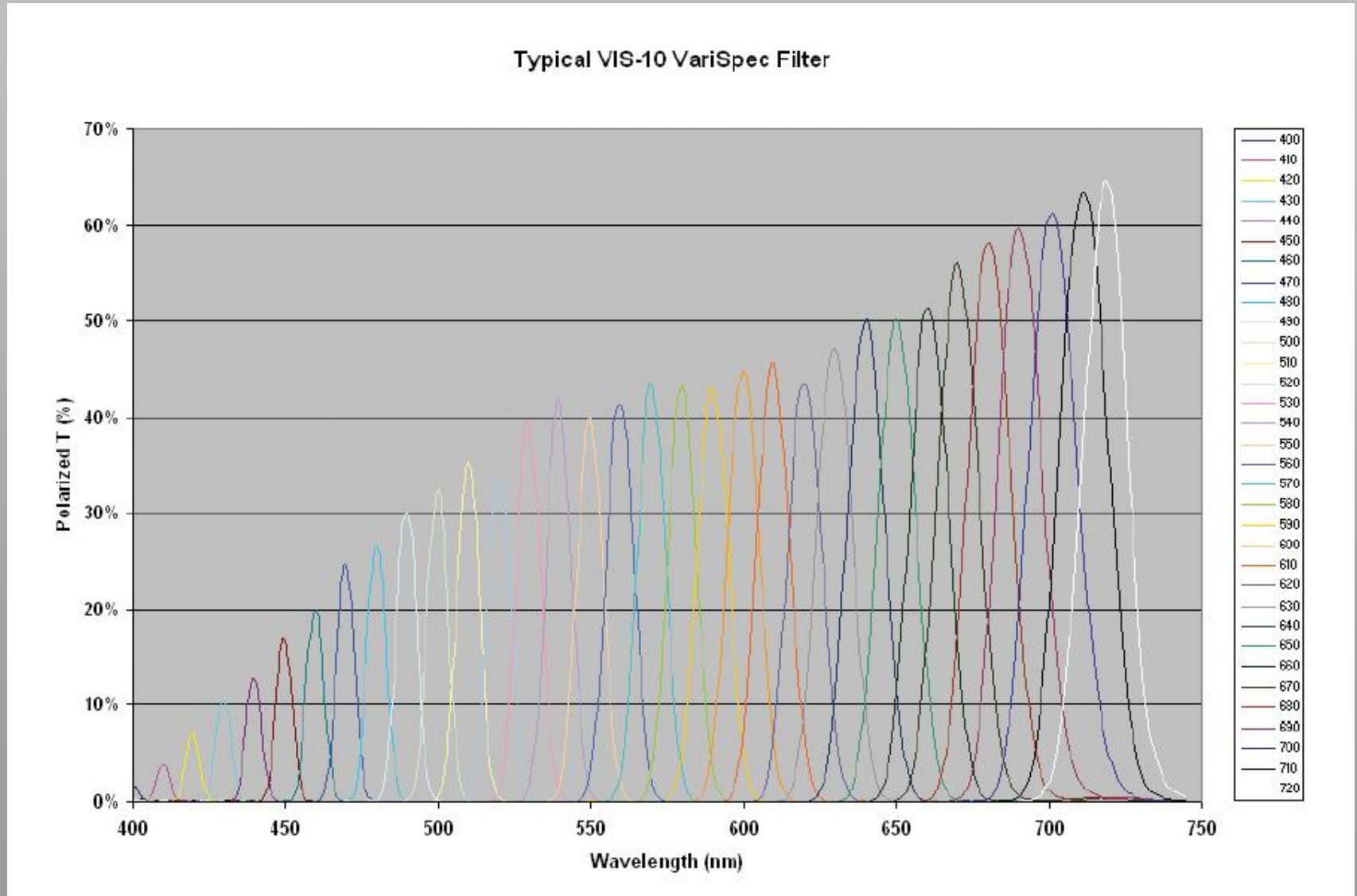
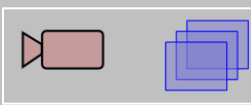
- Computer controllable spectral filter



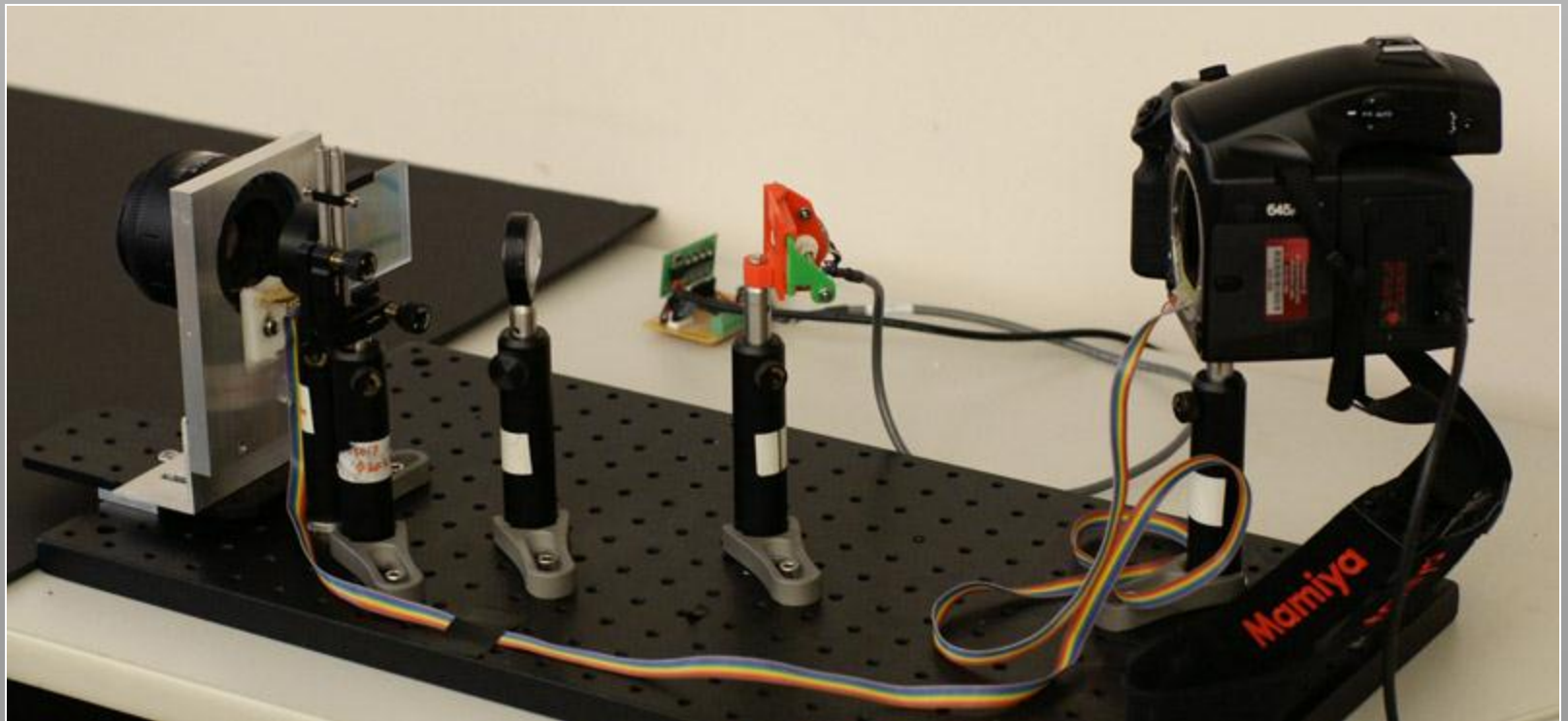
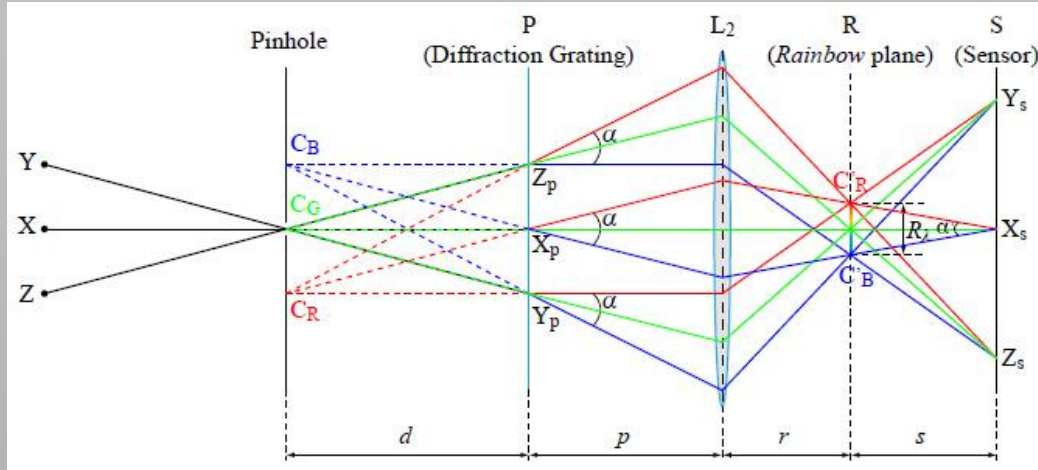
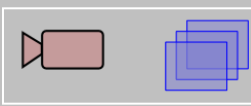
VariSpec LCTF



VariSpec spectral curves



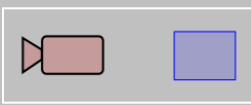
Agile Spectrum Imaging



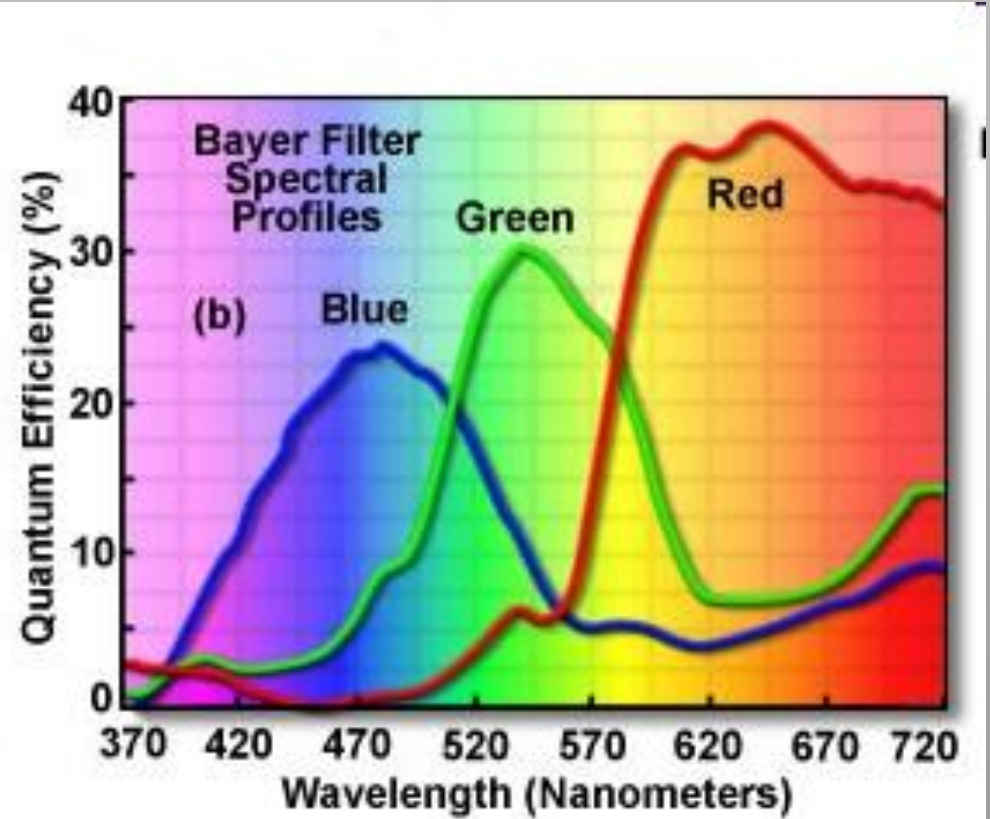
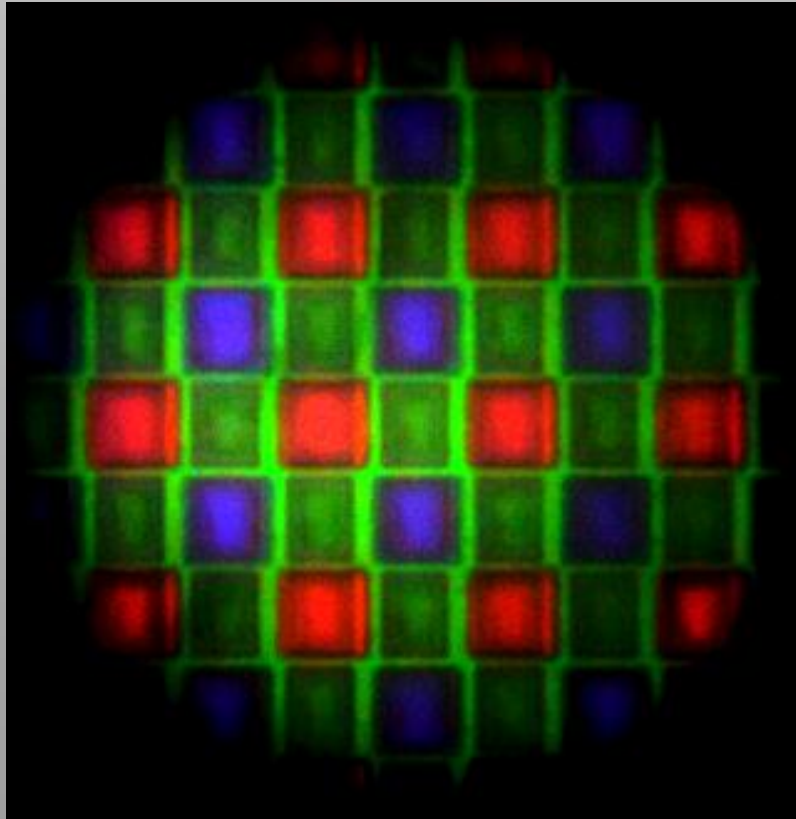
[Mohan et al. 08]

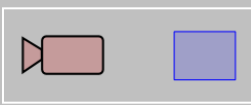
Color Filter Arrays

Color Filter Arrays

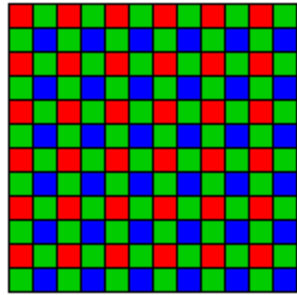


microscopic image of CCD (courtesy of Kevin Collins)

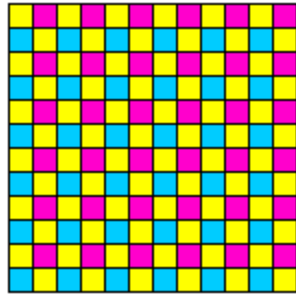




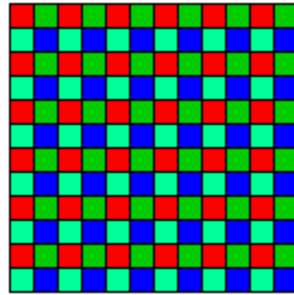
- alternative CFA designs



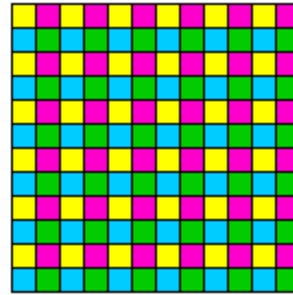
standard
Bayer



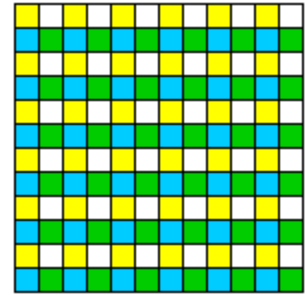
subtractive
primaries
e.g. Kodak
DCS 620x



RGB/Emerald
e.g. Sony
DCS F828



subtractive
primaries +
green
some video
cameras for
increased light
sensitivity



yellow/cyan/
green/white
e.g. JVC

Color Filter Arrays

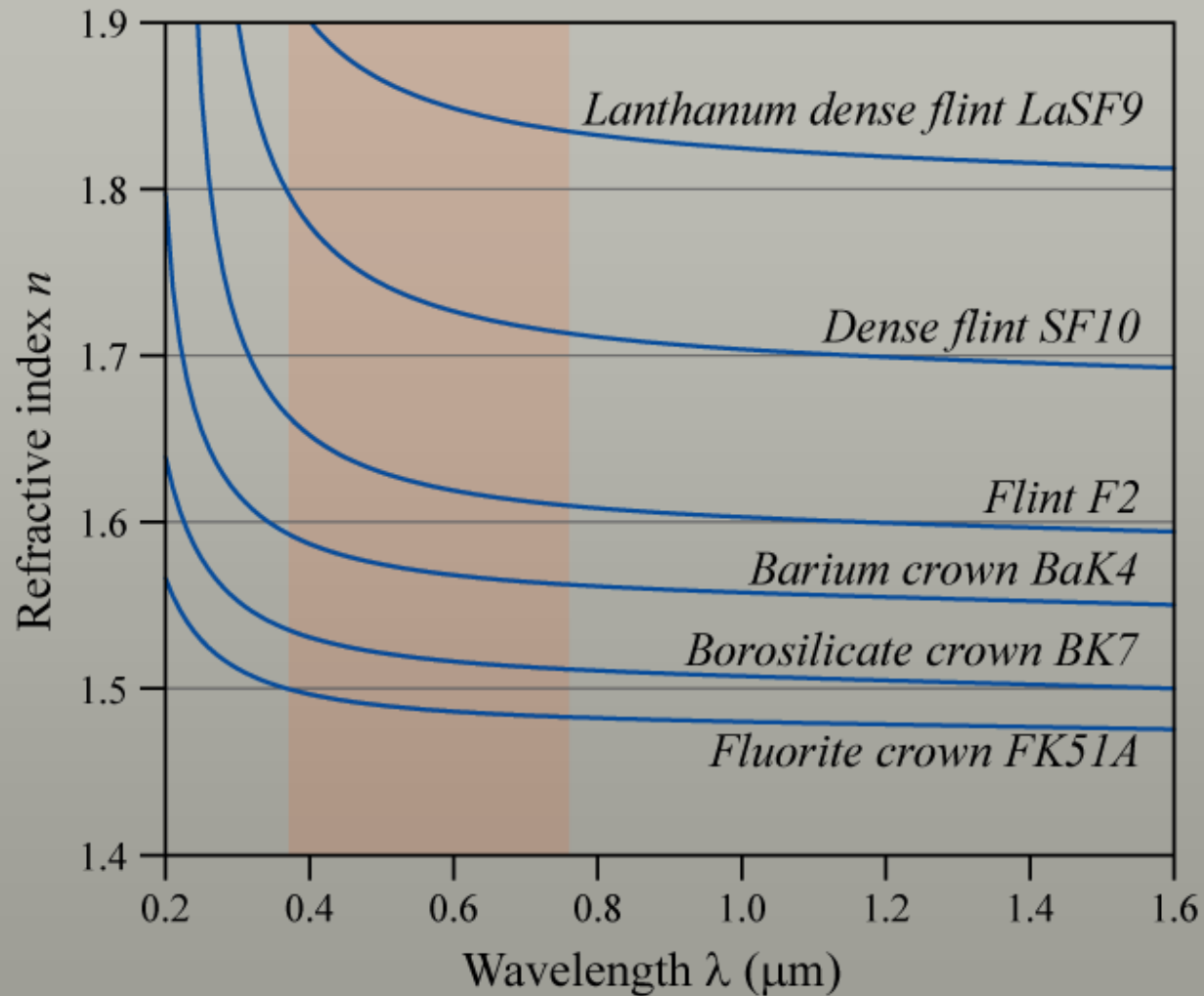
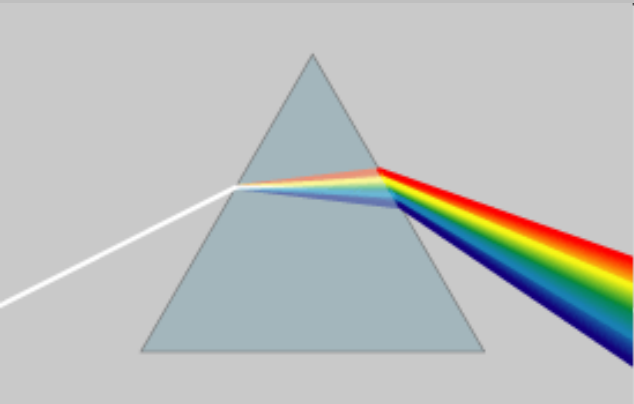
- Lots of work on optimized spectral transmissions & layout of CFAs – see STAR
- Siggraph '11 paper on dynamically switchable CFAs coming up...
- Assorted Pixels generalizes the concept

Single Pixel Spectrometers



Rev. Wimmer 1857

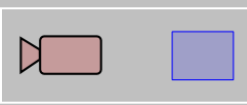
Principle of Operation - Dispersion



disadvantage:

- dispersion relation is nonlinear
- spatial position of wavelengths on screen must be calibrated

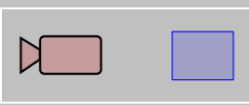
Prism-Based Systems



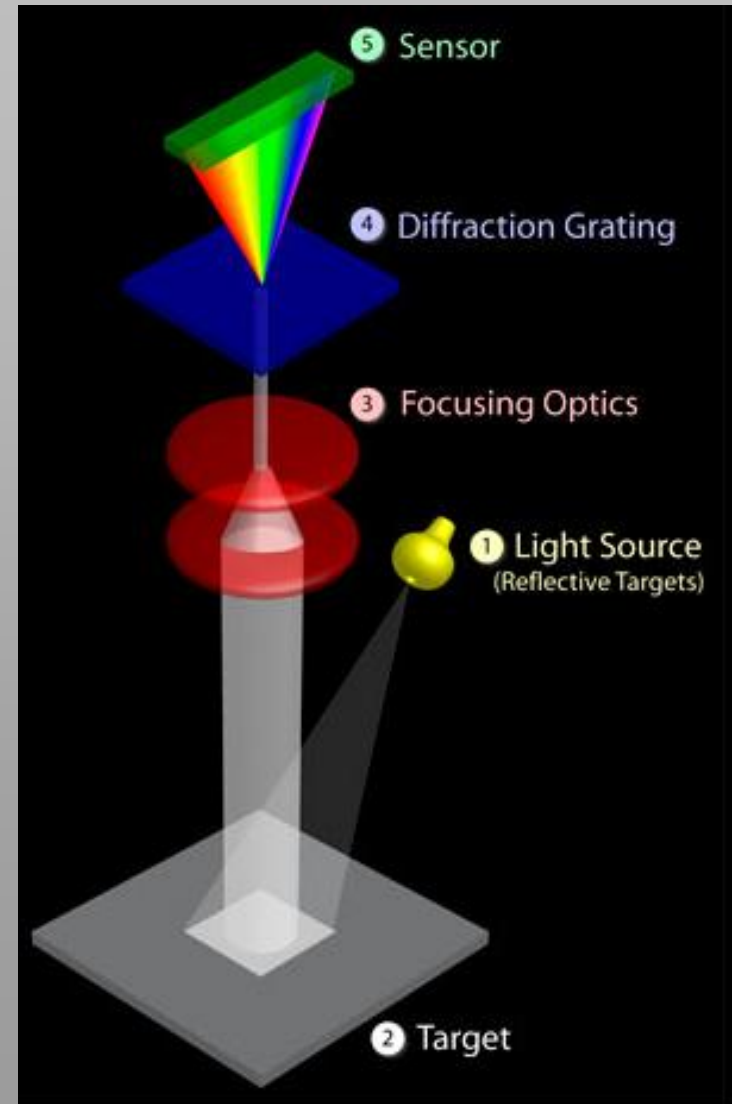
- Vintage Prism-Based Spectrometer



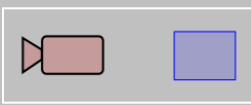
Spectrometer (Diffraction-Based)



- Spectrometer calibration (all types)
 1. mapping pixel – wavelength
 2. relative intensity of wavelength



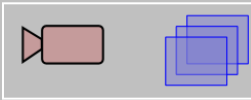
Diffraction Grating



- Center and first order diffraction
- At center, no diffraction
- For higher orders, diffraction is taking place (wavelength dependent)



Scanning Spectrometers



whiskbroom scanning

- one-dimensional sensor
- one spectral dimension
- mirror scanning of first spatial dimension
- second spatial dimension by sensor motion

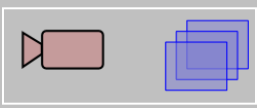
pushbroom scanning

- two-dimensional sensor
- one spatial dimension
- one spectral dimension
- second spatial dimension by sensor motion

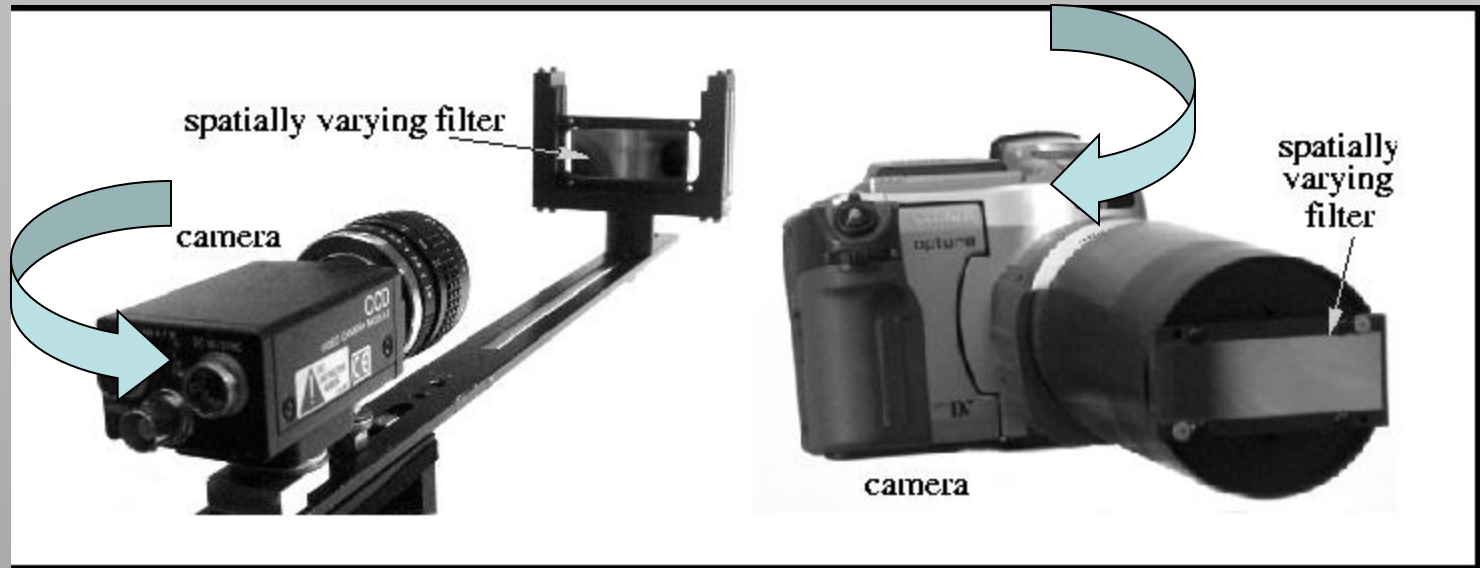
5

[Optoiq]

Spatial Scanning



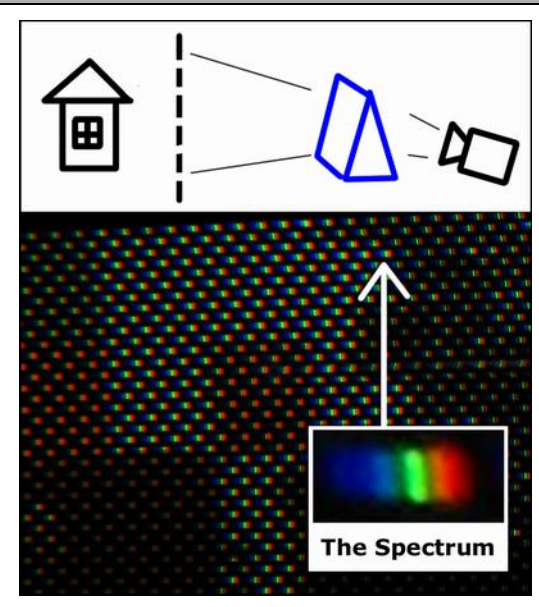
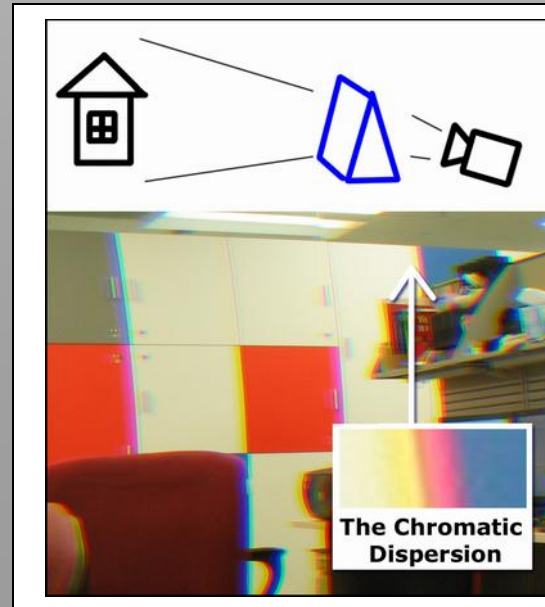
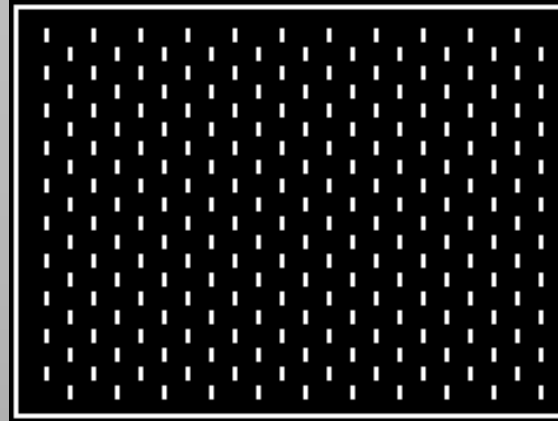
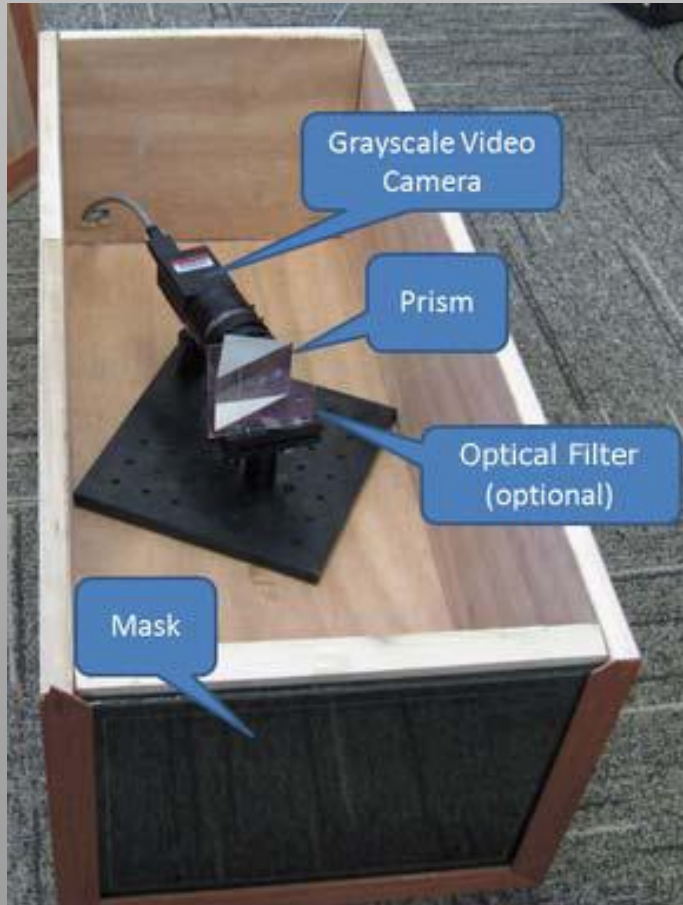
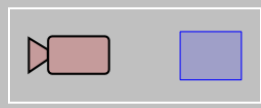
- Generalized Mosaics [Schechner & Nayar]



- linear filter
- each pixel column filtered differently
- rotational motion & registration to assemble image stack

Snapshot Imaging Spectrometers

Multiplexing: Prism-Mask Based System



[Du'09]

Computed Tomography Imaging Spectrometer

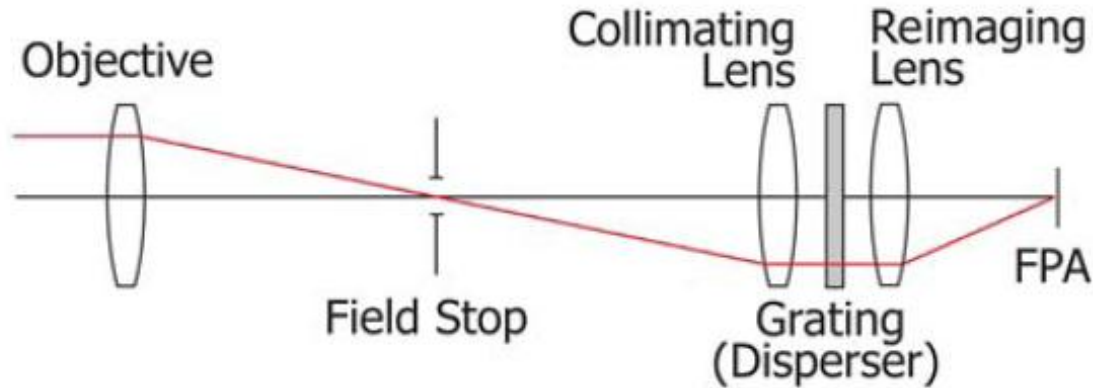
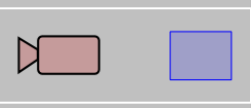
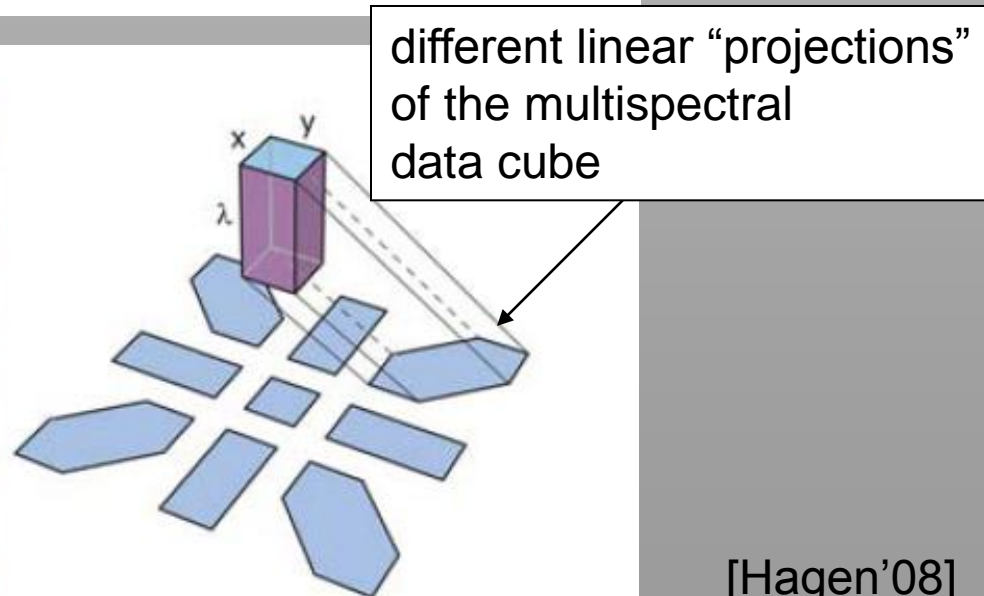
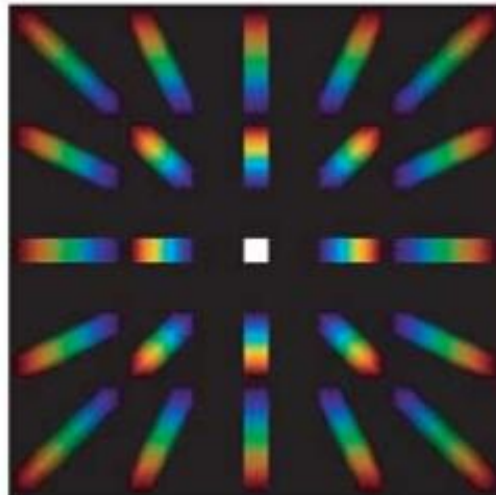
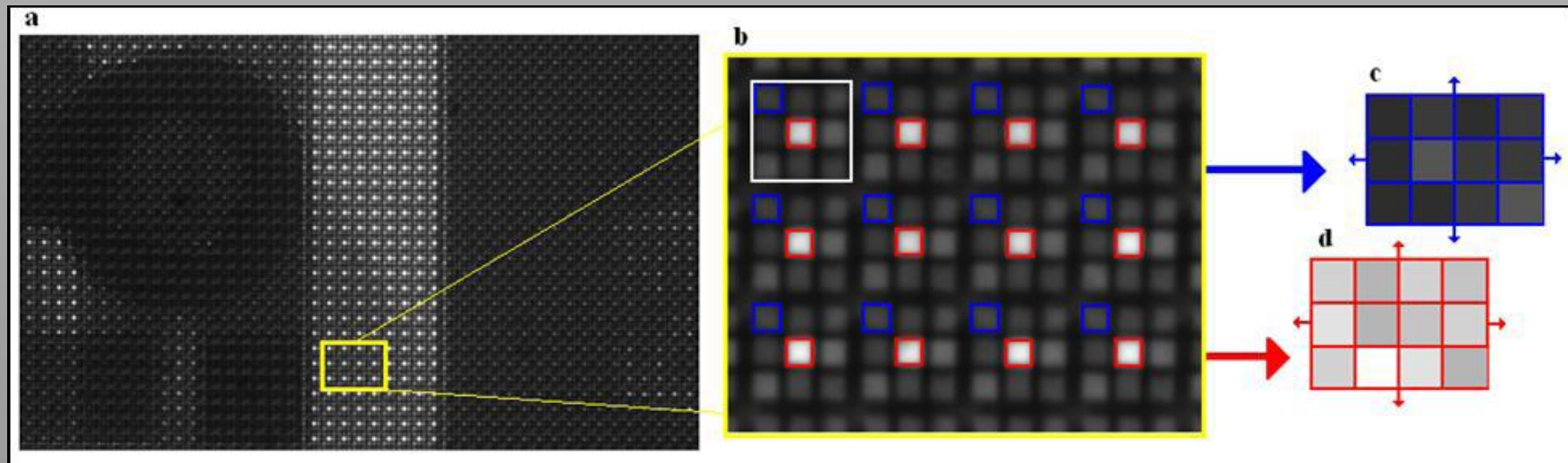
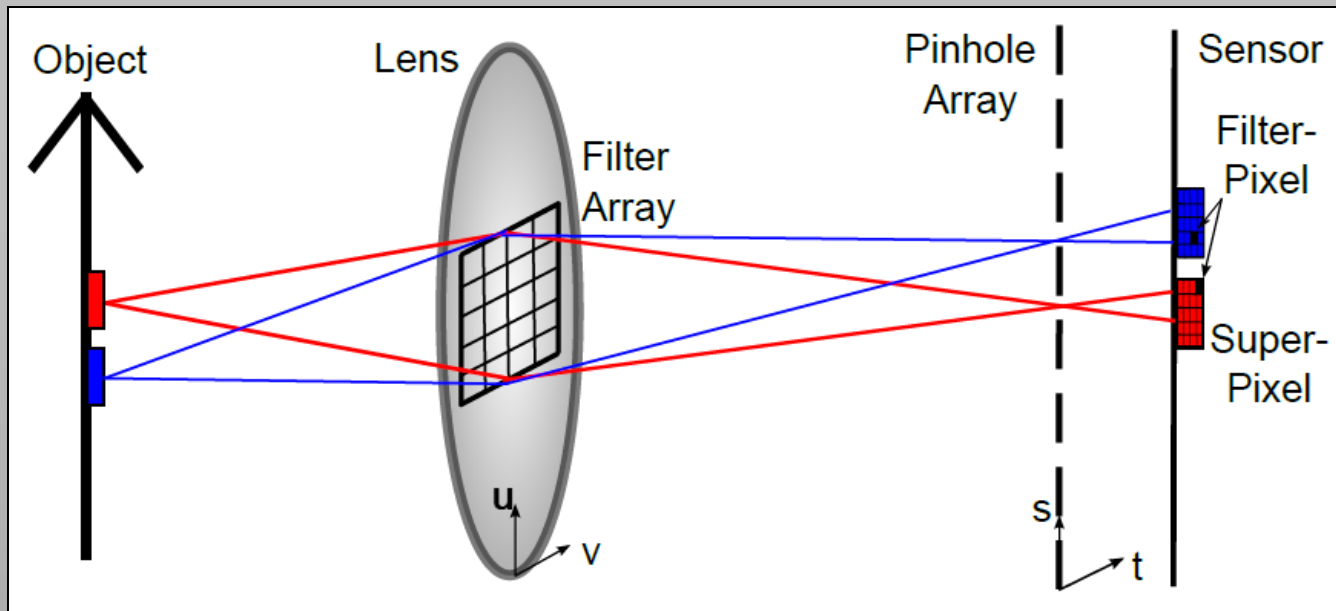


Fig. 1. (Color online) CTIS instrument views the target scene through a 2D grating. The field stop limits the field of view, such that the dispersed diffraction orders are spatially separated.



[Hagen'08]

Spectrally Filtered Light Fields



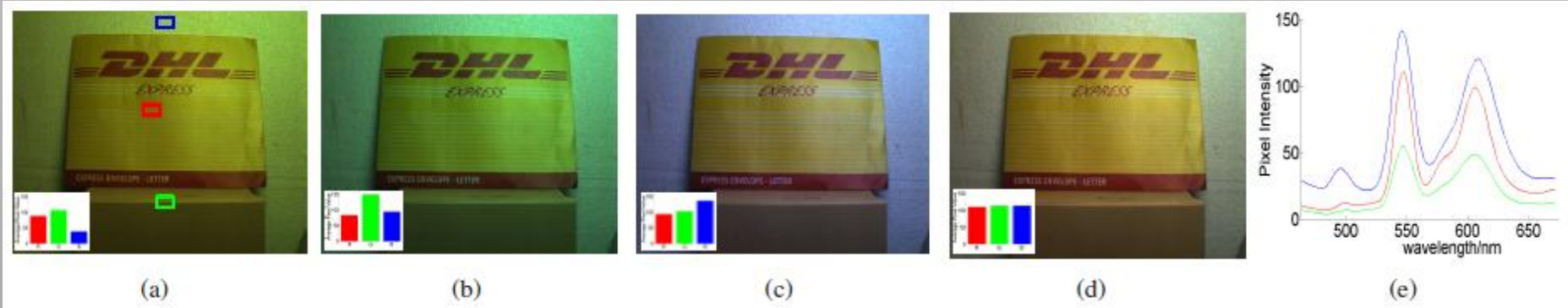
[Horstmeyer'09]

Applications

Applications

- automatic white balancing

Spatially uniform illumination



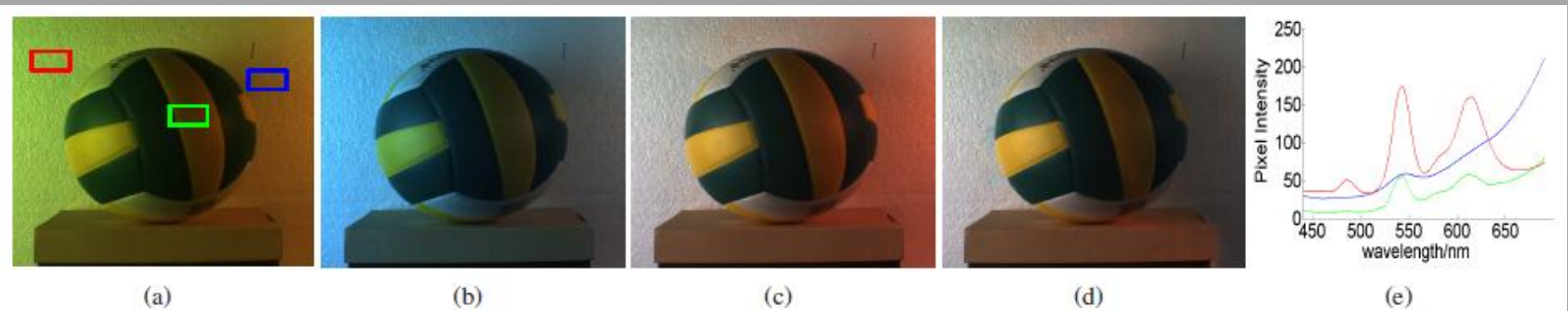
raw from RGB

tungsten WB

greyworld WB

spectral WB

spectra

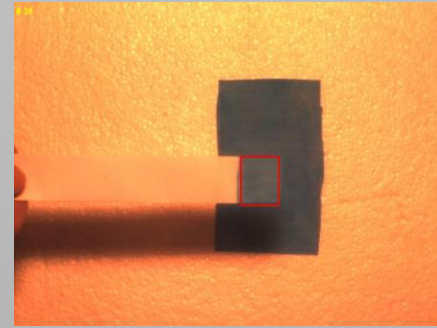
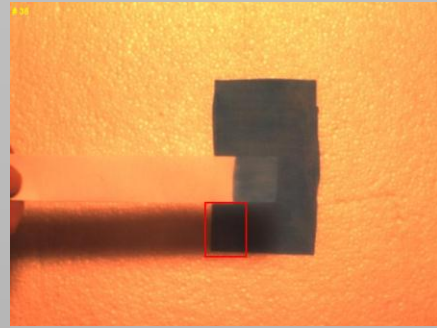
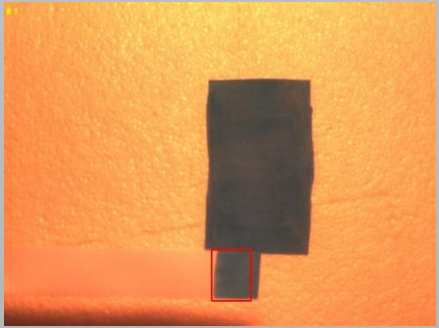


Spatially varying illumination

[Cao10]

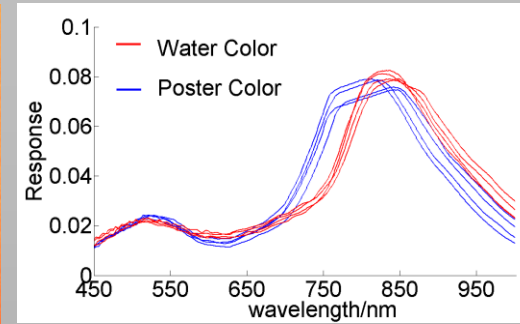
Applications

- improved tracking



RGB –
tracking lost

spectral –
tracking OK



- real and fake skin detection



[Cao10]

Applications

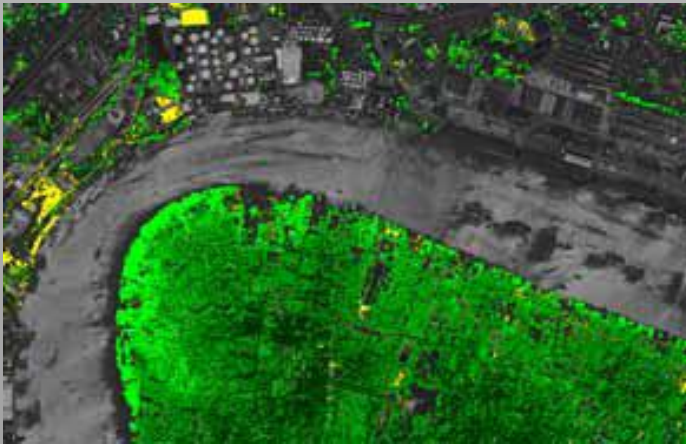
- analyze / restore paintings



[Calit]

Applications

- Satellite-Based Remote Sensing



vegetation mapping



urban land use



pollution monitoring

[DigitalGlobe'10]

Next: Light Field Acquisition

Computational Plenoptic Imaging

Gordon Wetzstein¹
Wolfgang Heidrich³

Ivo Ihrke²
Kurt Akeley⁴

Douglas Lanman¹
Ramesh Raskar¹

¹MIT Media Lab

²Saarland University

³University of British Columbia

⁴Lytro, Inc.

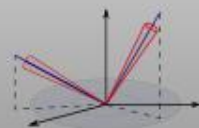
IV. Light Field Acquisition



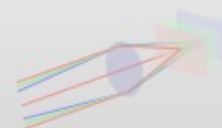
Dynamic Range



Color Spectrum



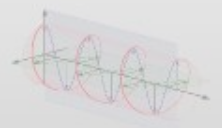
Directions | Light Fields



Space | Focal Surfaces



Time



Further Properties

IV.1 (Brief) Introduction to Light Fields

The 5D Plenoptic Function



$$P(\theta, \phi, \lambda, t)$$

Q: What is the set of all things that one can ever see?

A: The Plenoptic Function [Adelson and Bergen 1991]

(from *plenus*, complete or full, and *optic*)

The 5D Plenoptic Function



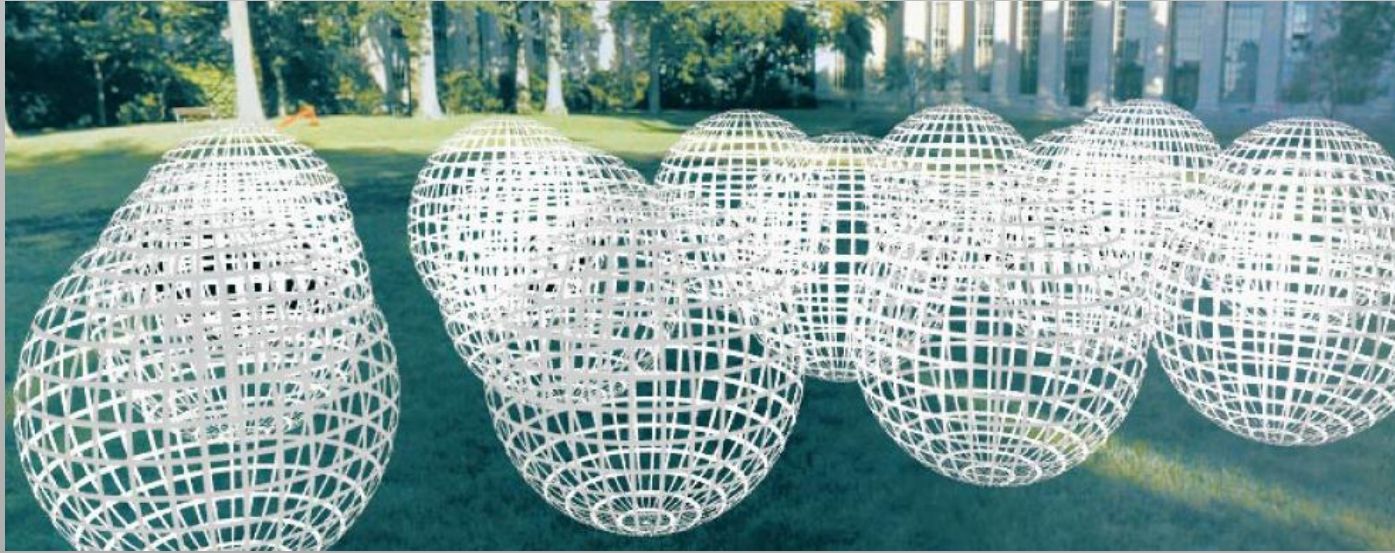
$$P(\theta, \phi, \lambda, t)$$

Q: What is the set of all things that one can ever see?

A: The Plenoptic Function [Adelson and Bergen 1991]

(from *plenus*, complete or full, and *optic*)

The 5D Plenoptic Function

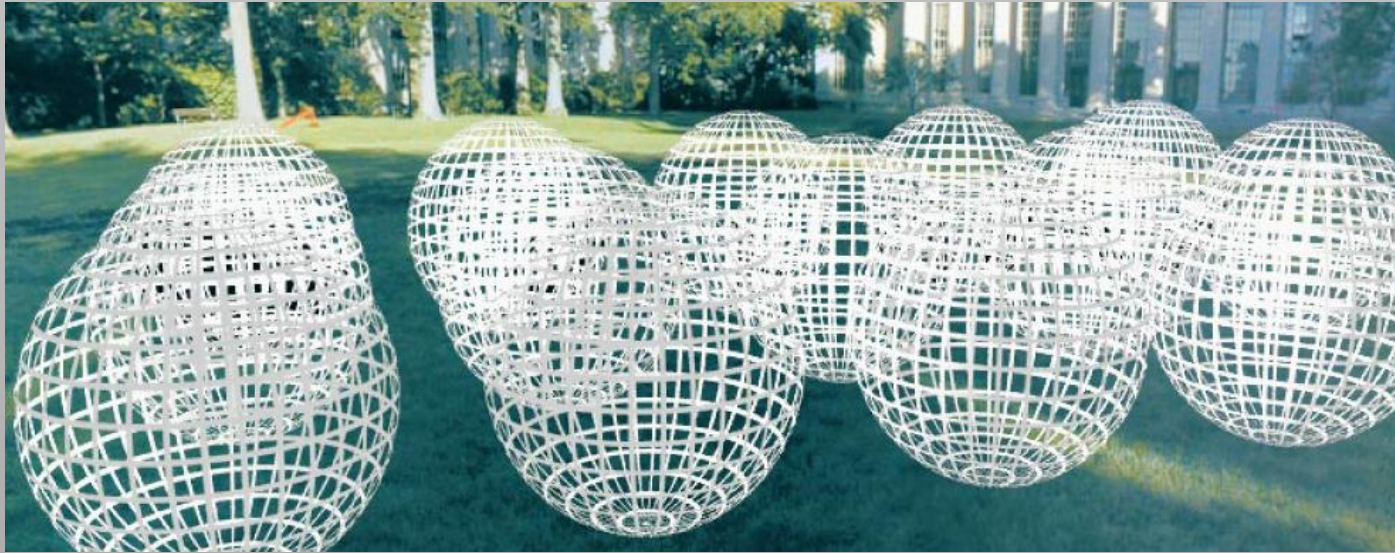


$$P(\theta, \phi, \lambda, t, p_x, p_y, p_z)$$

$P(\theta, \phi, \lambda, t, p_x, p_y, p_z)$ defines the intensity of light:

- as a function of viewpoint
- as a function of time
- as a function of wavelength

The 5D Plenoptic Function



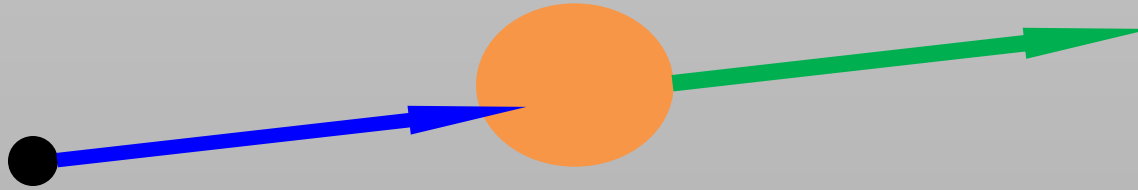
$$P(\theta, \phi, \lambda, t, p_x, p_y, p_z)$$

$P(\theta, \phi, \lambda, t, p_x, p_y, p_z)$ defines the intensity of light:

- as a function of viewpoint
- as a function of time
- as a function of wavelength

The 5D Plenoptic Function

Let's ignore color and time (i.e., these are attributes of rays)...



$$P(\theta, \phi, p_x, p_y, p_z)$$

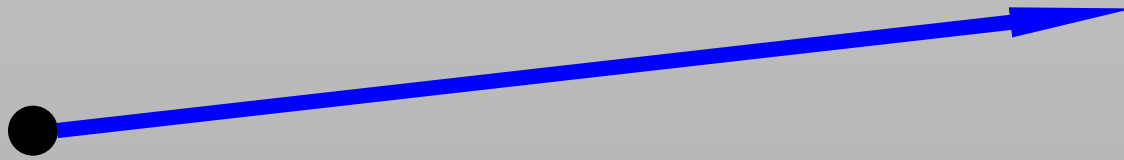
The plenoptic function is 5D:

- 3D position
- 2D direction

Require 5D to represent attributes across occlusions

The 4D Light Field

Consider a region free of occluding objects...



$$P(\theta, \phi, p_x, p_y, p_z)$$

The plenoptic function (light field) is 4D

- 2D position
- 2D direction

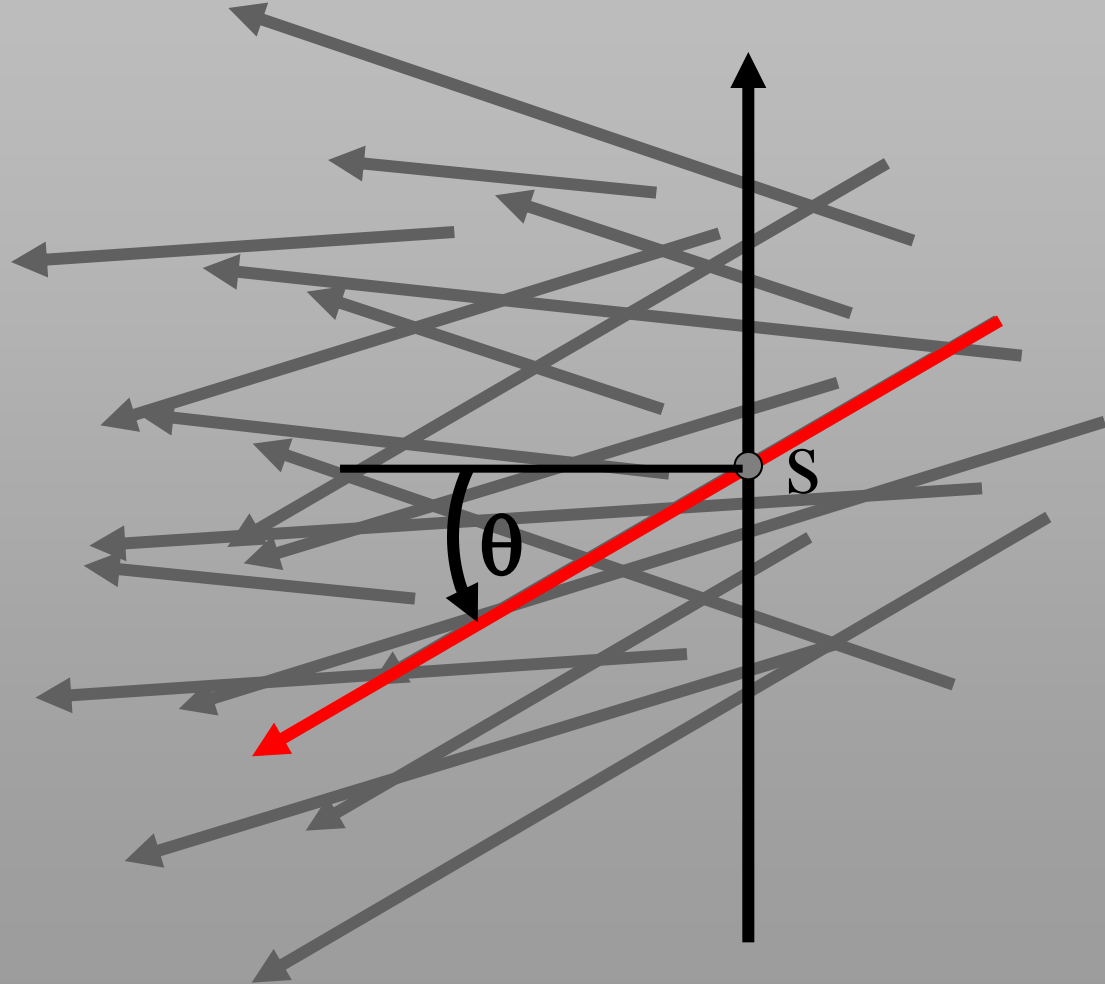
The space of all lines in a 3D space is 4D



[Levoy and Hanrahan 1996; Gortler et al. 1996]

Position-Angle Parameterization

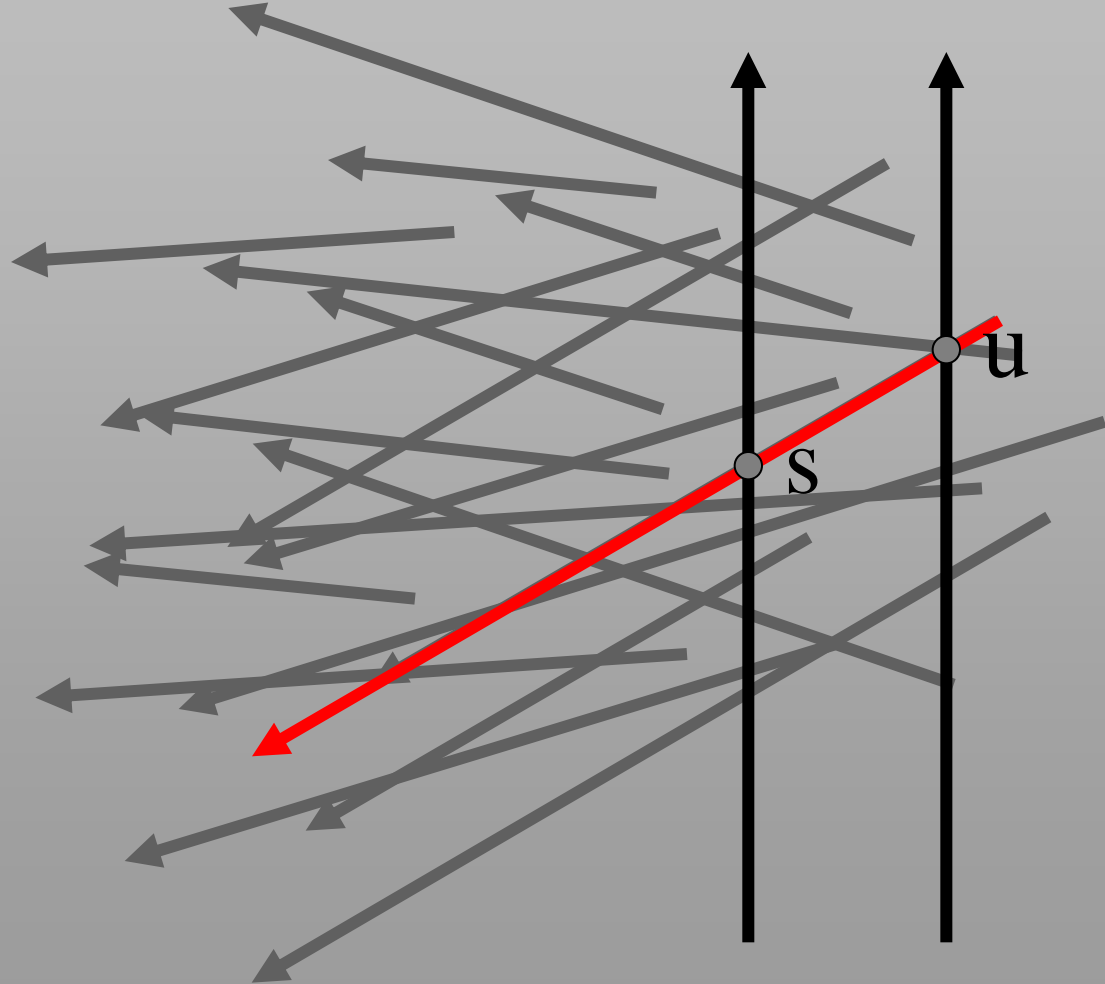
2D position
2D direction



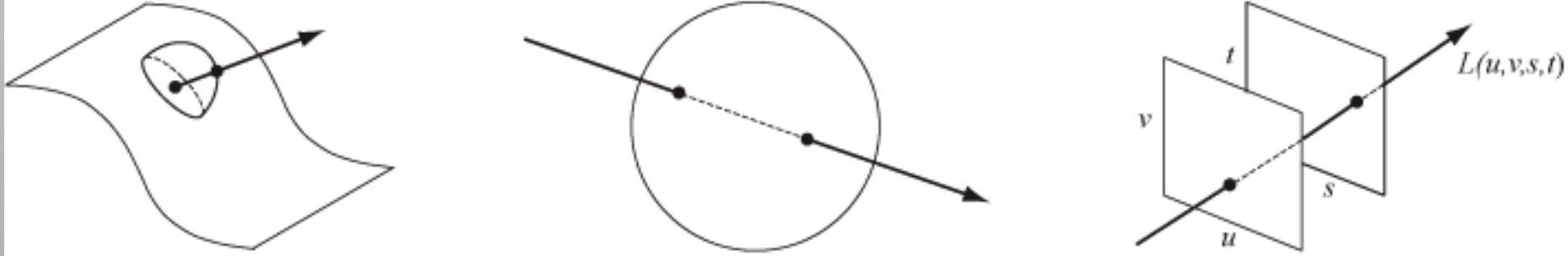
Two-Plane (Light Slab) Parameterization

2D position

2D position



Alternative Parameterizations

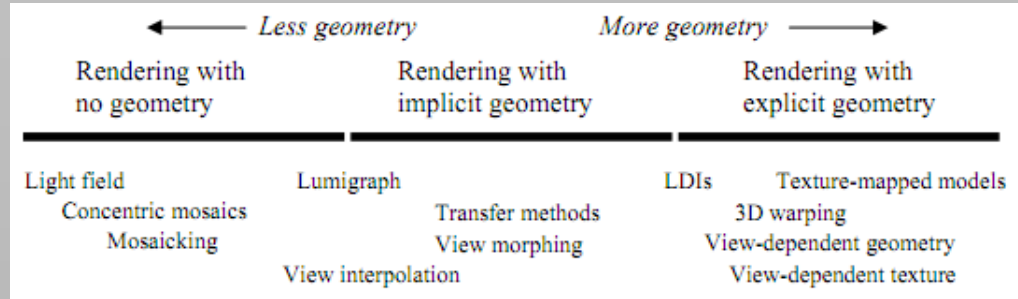


Left: Points on a plane or curved surface and directions leaving each point

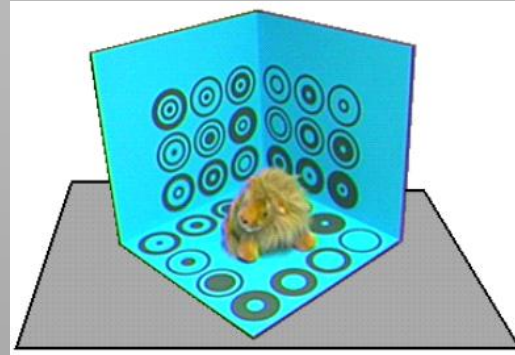
Center: Pairs of points on the surface of a sphere

Right: Pairs of points on two planes in general (meaning any) position

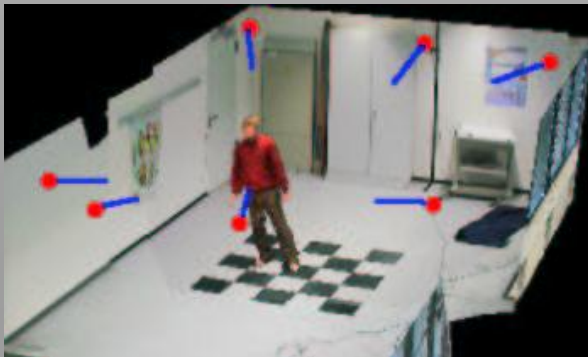
Image-Based Rendering



[Levoy and Hanrahan 1996]

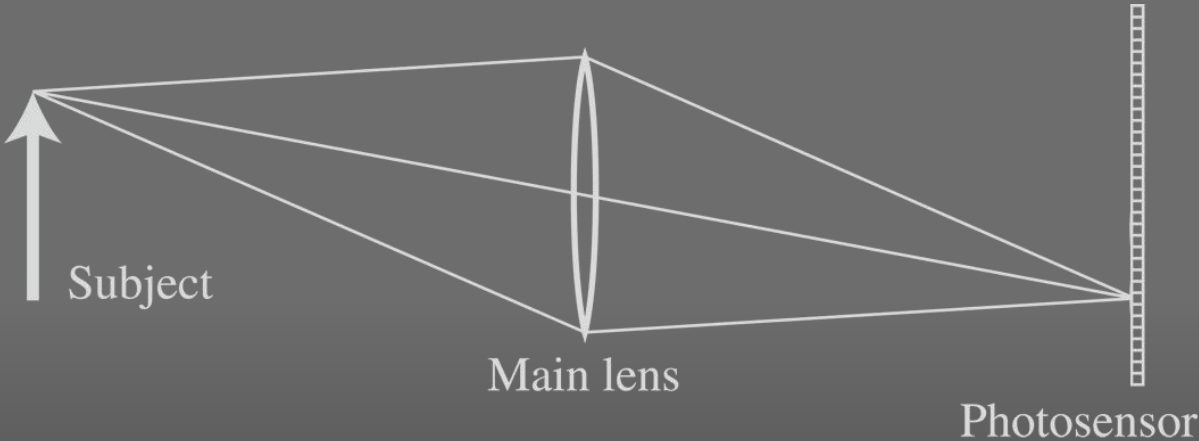


[Gortler et al. 1996]



[Carranza et al. 2003]

Conventional vs. Plenoptic Camera



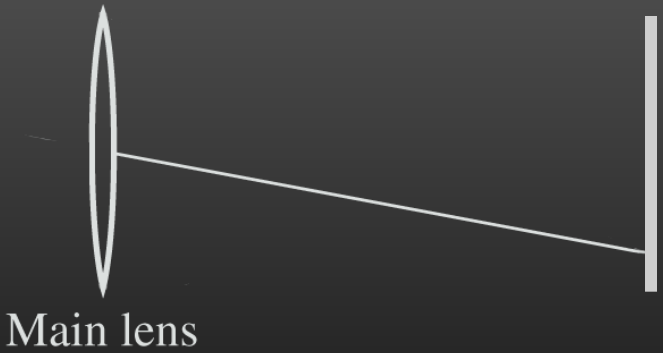
Scene Pixel = (s,t)

Virtual Camera = (u,v)

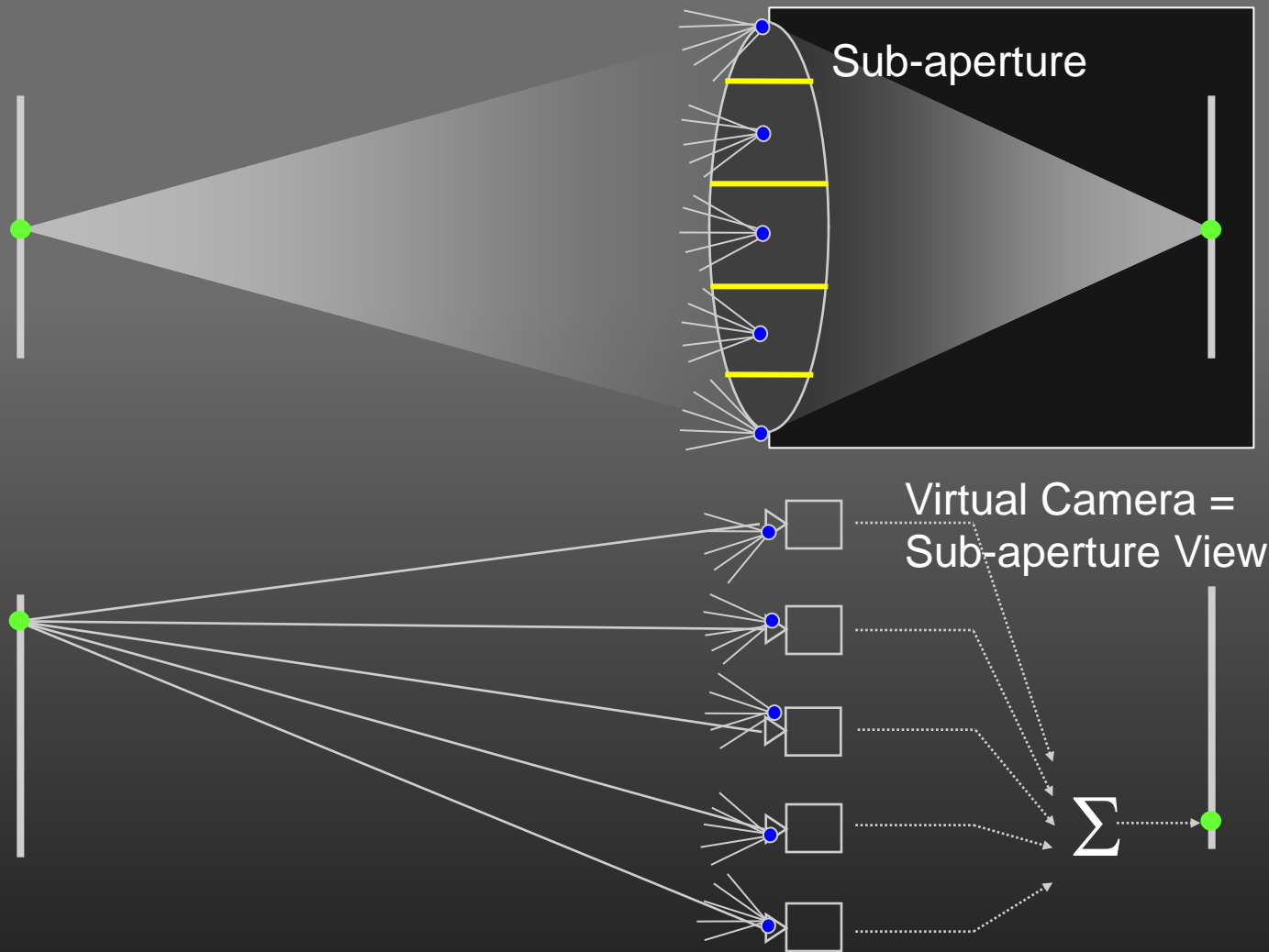
Pixel = (s,t)

uv-plane

st-plane



Light Field Photography = Array of (Virtual) Cameras

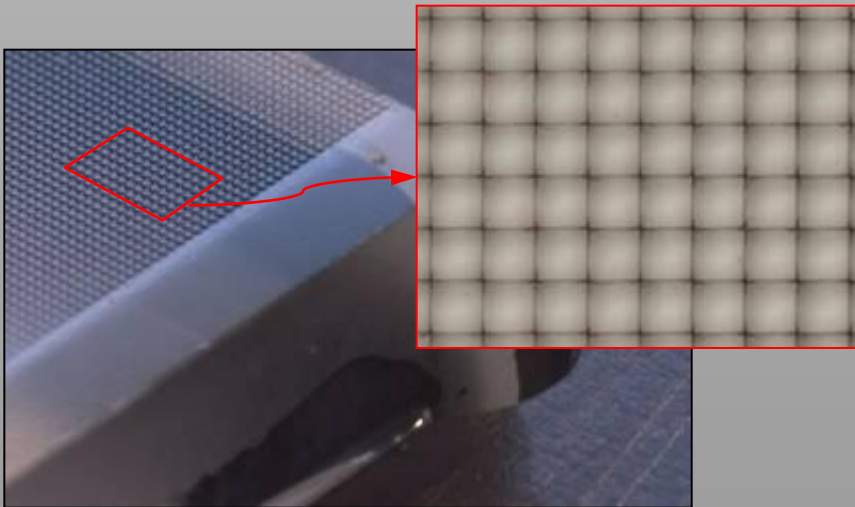




Digital Image Refocusing



Kodak 16-megapixel sensor



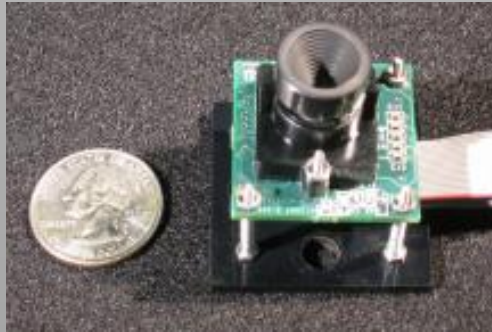
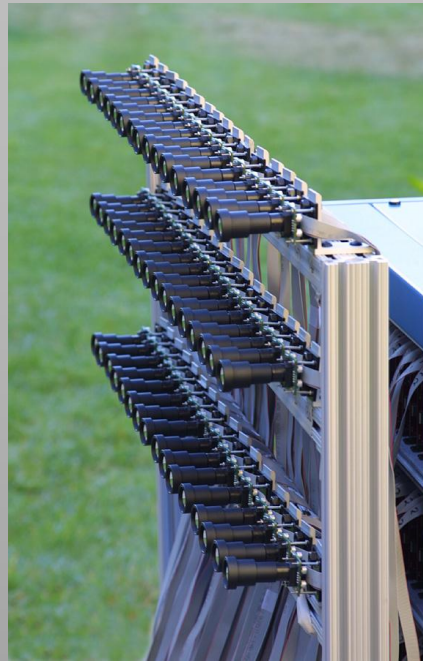
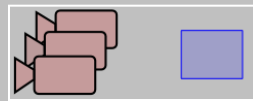
125 μ square-sided microlenses



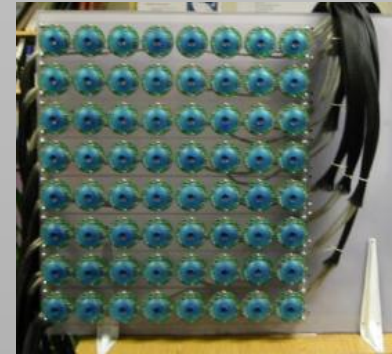
[Ng 2005]

IV.II Multiple Sensors

Static Camera Arrays

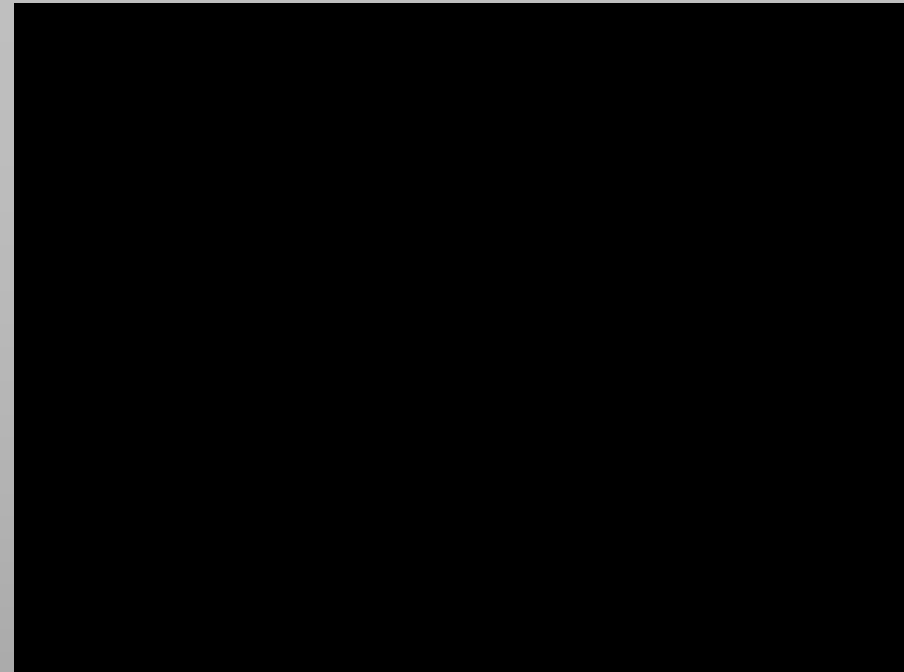
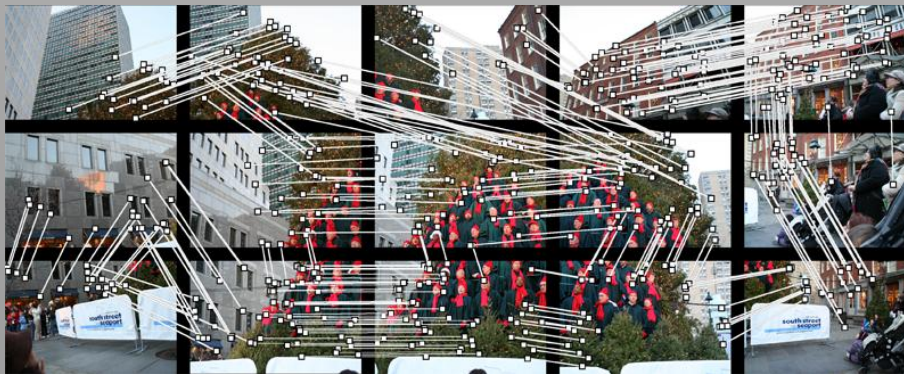
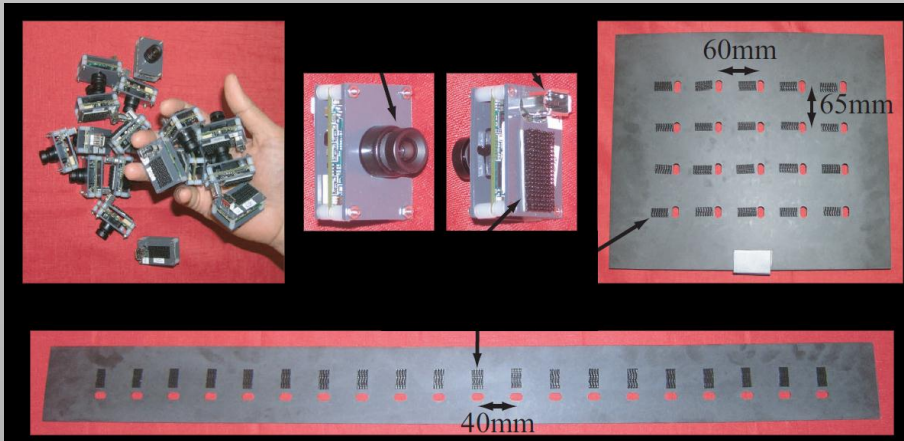
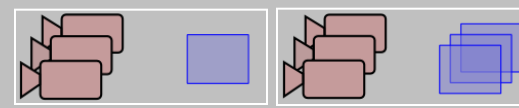


Stanford Multi-Camera Array
125 cameras using custom hardware
[Wilburn et al. 2002, Wilburn et al. 2005]



Distributed Light Field Camera
64 cameras with distributed rendering
[Yang et al. 2002]

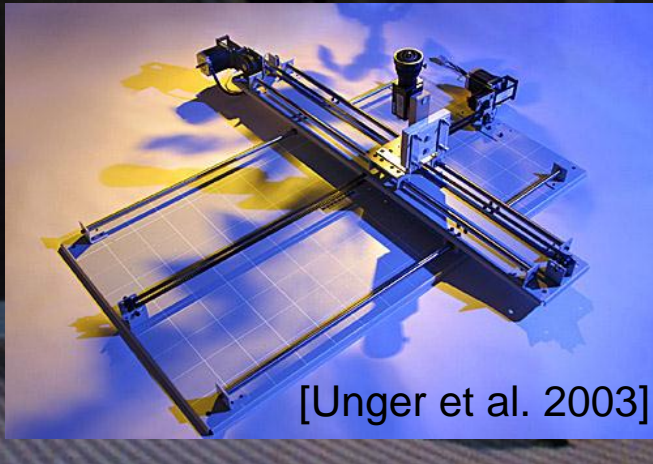
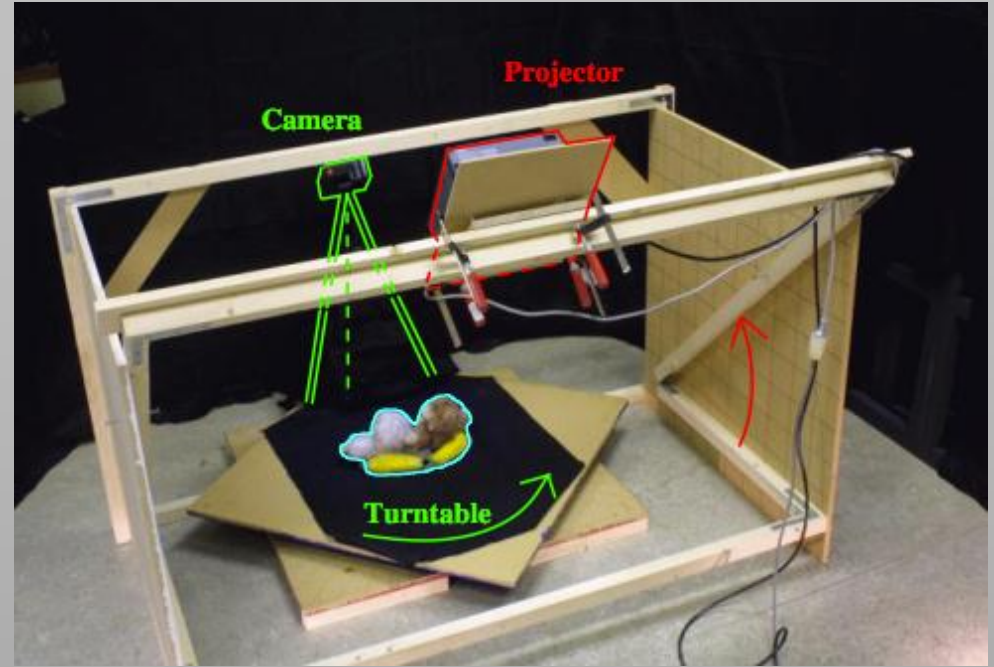
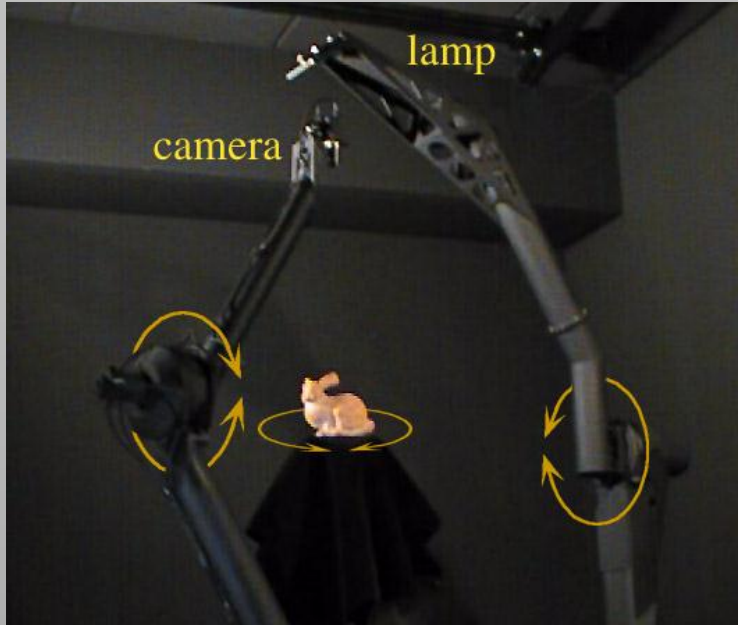
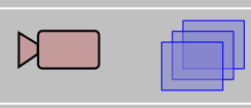
Flexible Camera Arrays



[Nomura et al. 2007]

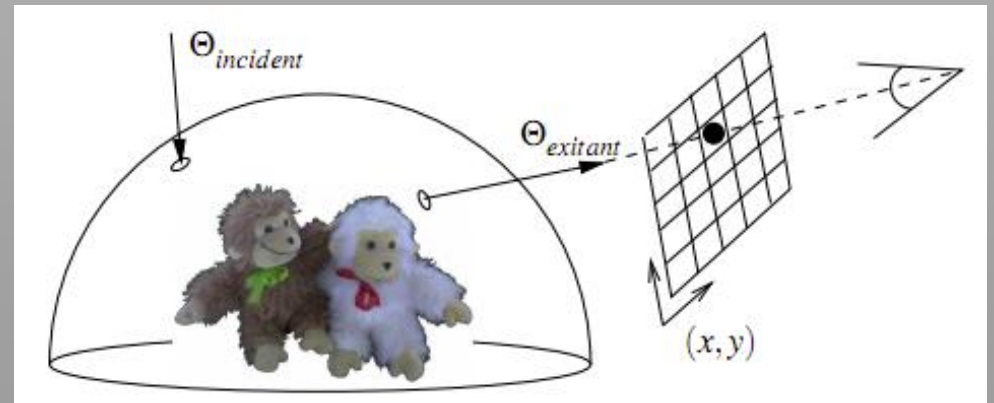
IV.III Temporal Multiplexing

Controlled Camera or Object Motion



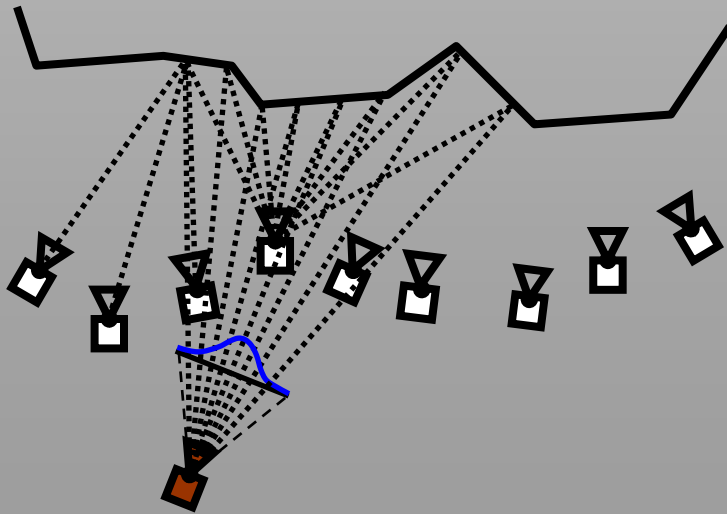
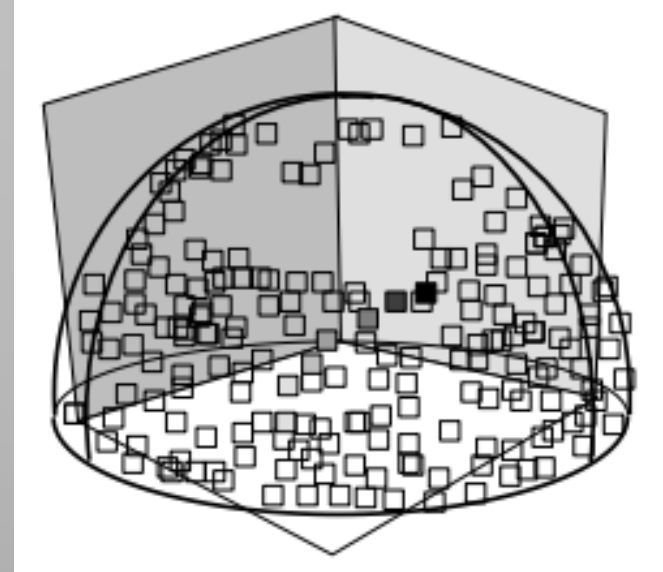
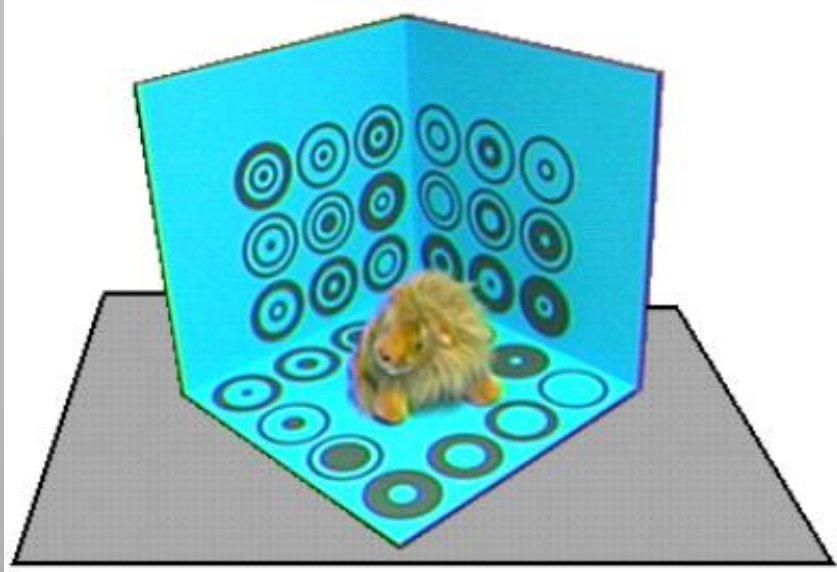
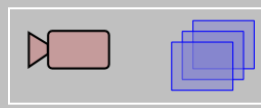
[Unger et al. 2003]

Stanford Spherical Gantry
[Levoy and Hanrahan 1996]



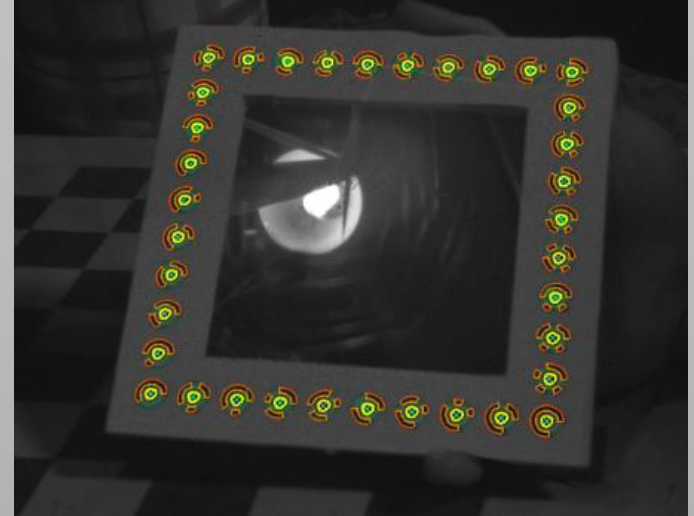
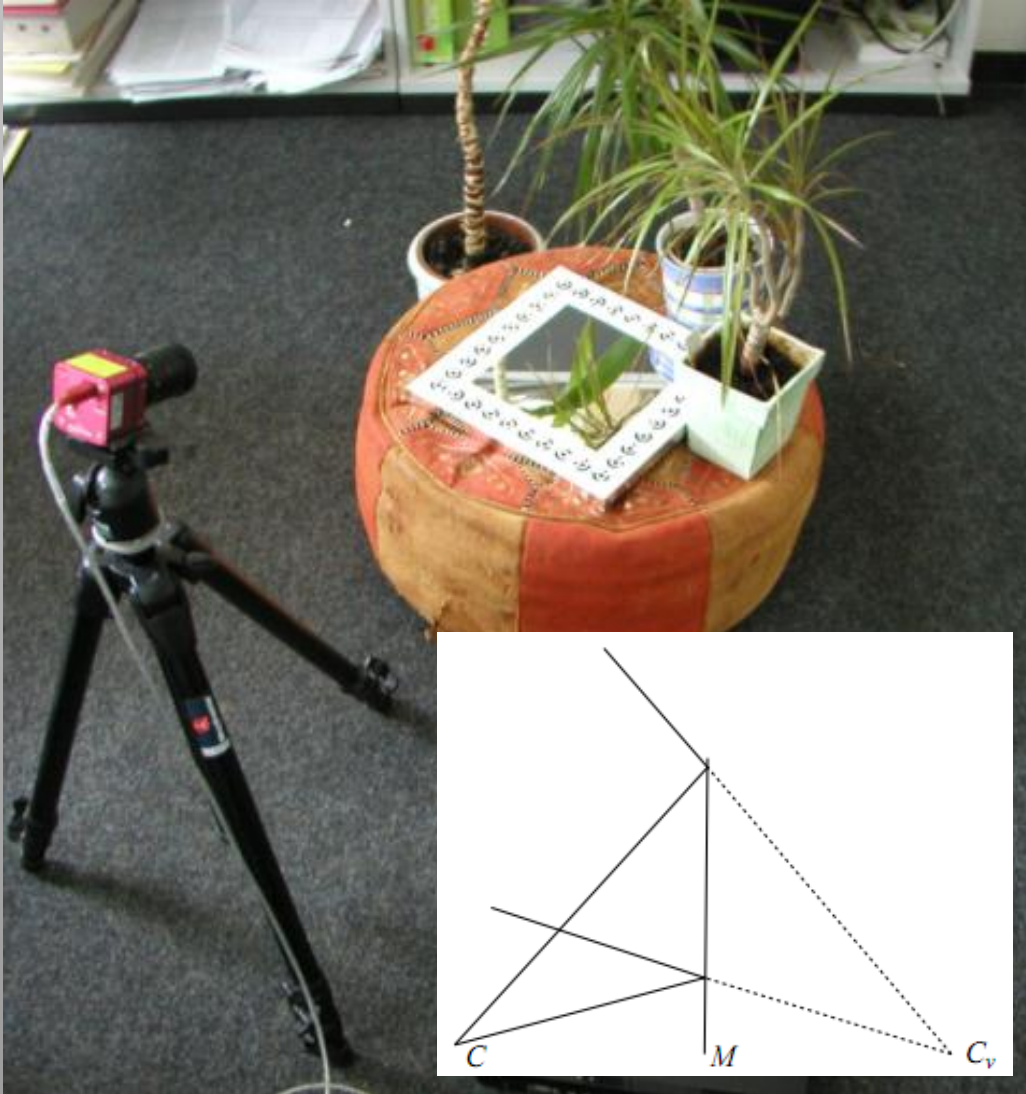
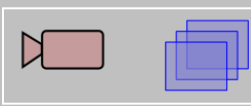
Relighting with 4D Incident Light Fields
[Masselus et al. 2003]

Uncontrolled Camera or Object Motion



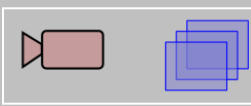
Unstructured Lumigraph Rendering
[Gortler et al. 1996; Buehler et al. 2001]

Virtual Cameras using a Steerable Mirror



Fast Incident Light Field Acquisition and Rendering
[Ihrke et al. 2008]

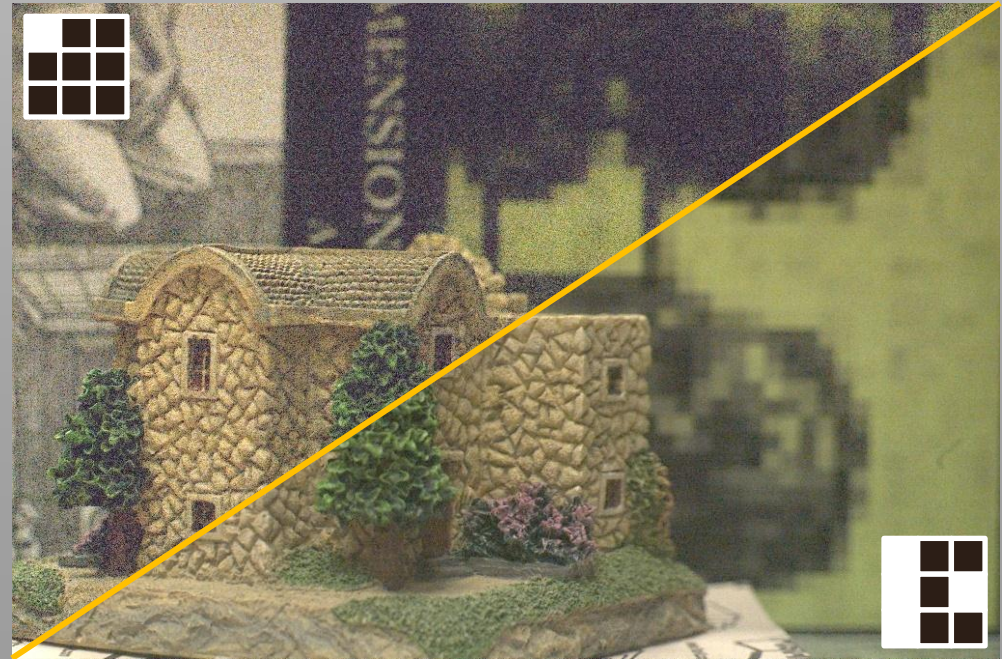
Programmable Aperture Photography



9 aperture patterns for 3x3 light field capture



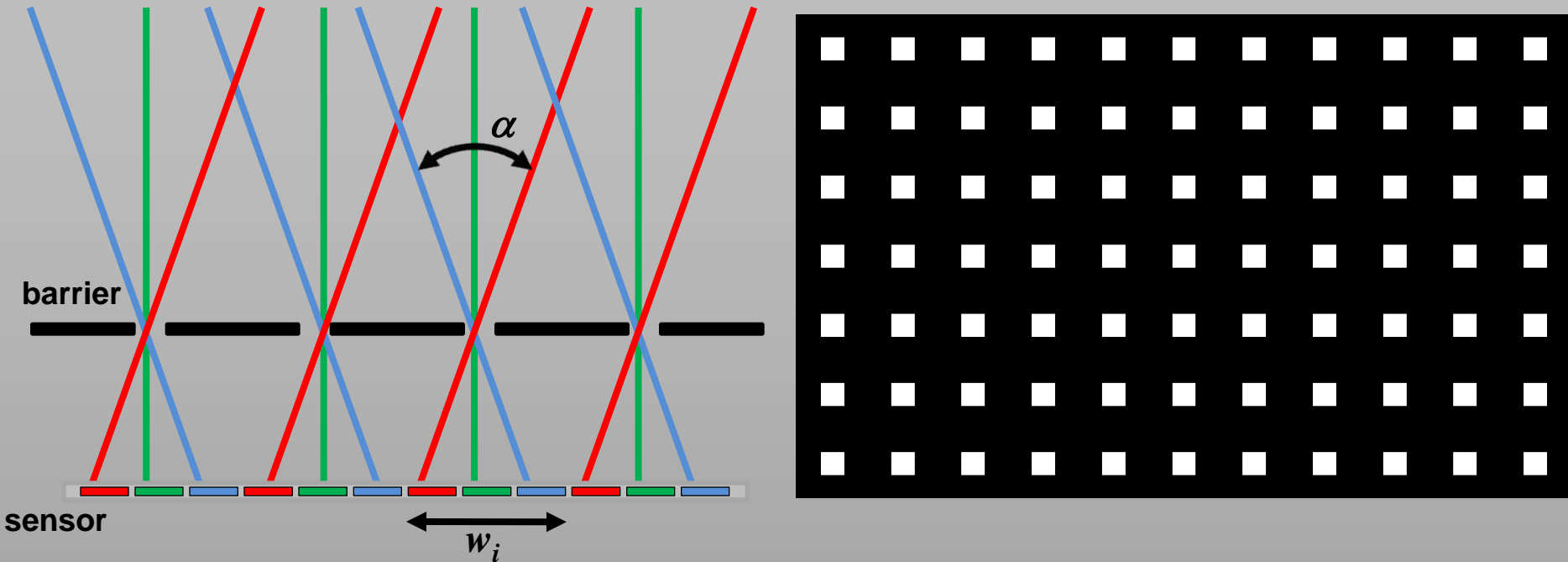
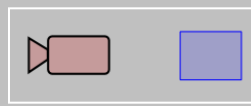
9 multiplexed aperture patterns for 3x3 light field capture



[Liang et al. 2008; Schechner and Nayar 2007]

IV.IV Spatial and Frequency Multiplexing

Parallax Barriers (“Slits” and “Pinhole Arrays”)

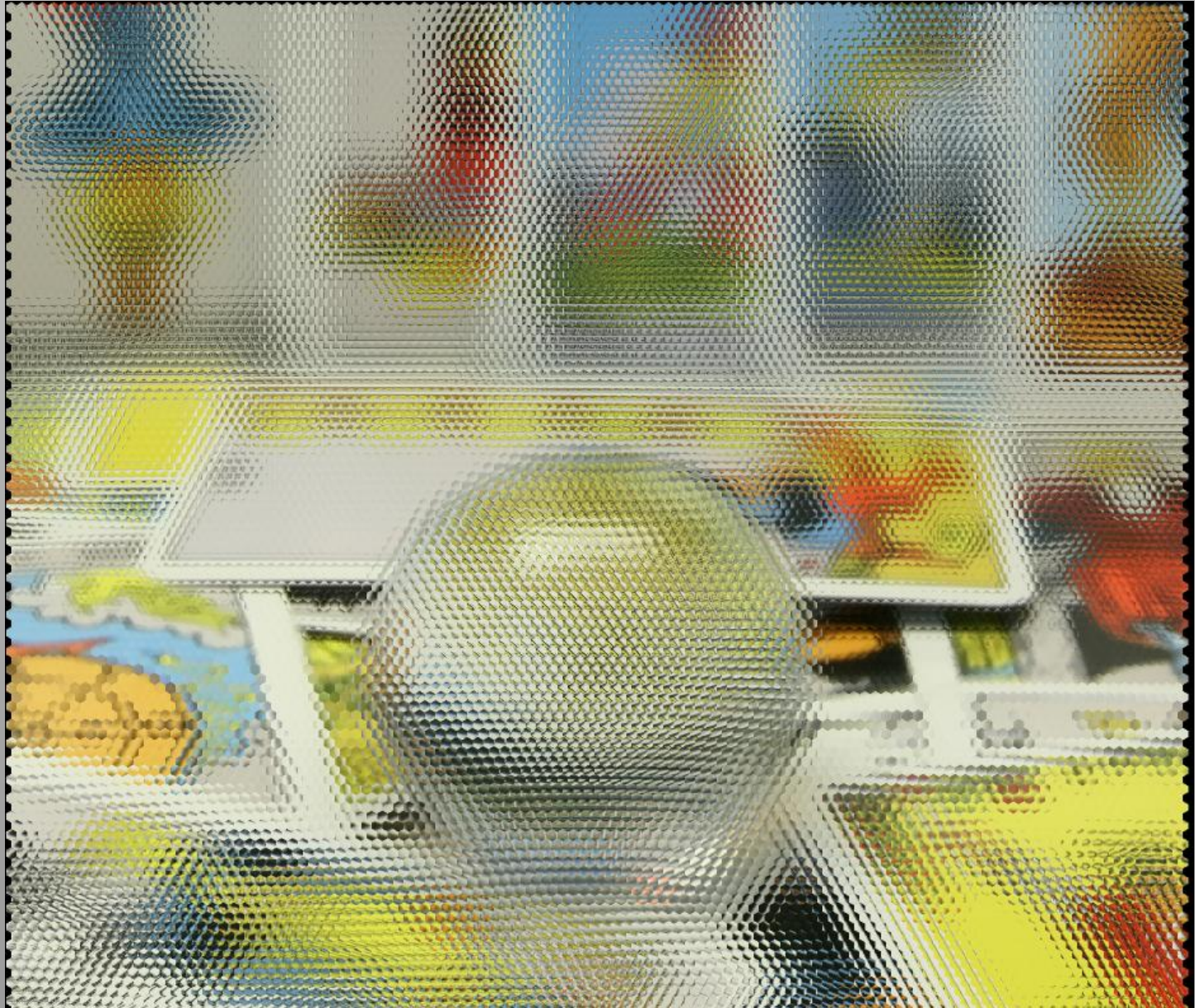
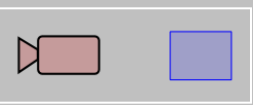


Spatially-multiplexed light field capture using masks (i.e., barriers):

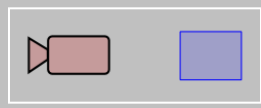
- Cause severe attenuation → long exposures or lower SNR
- Impose fixed trade-off between spatial and angular resolution (unless implemented with programmable masks, e.g. LCDs)

[Ives 1903]

Parallax Barriers (“Slits” and “Pinhole Arrays”)



Parallax Barriers (“Slits” and “Pinhole Arrays”) looking to the right →



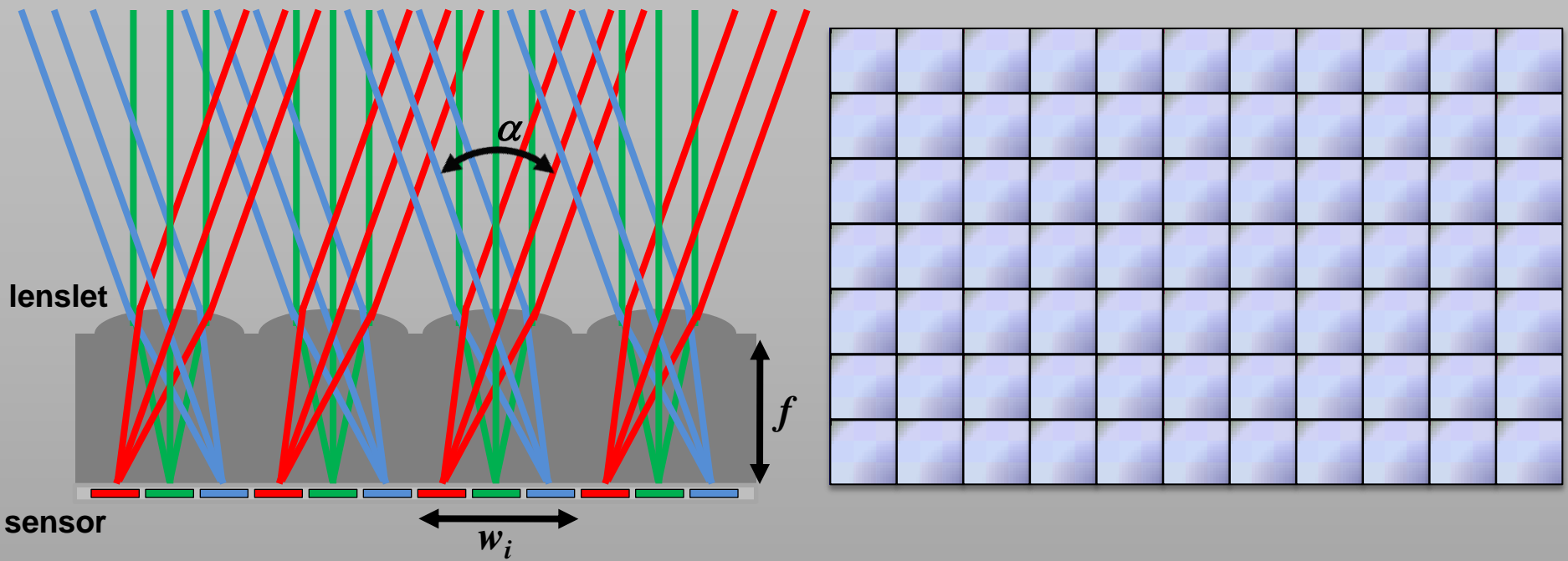
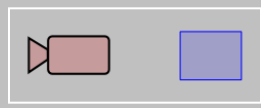
↑
looking up



Sample Image

[The (New) Stanford Light Field Archive]

Integral Imaging (“Lenticular” or “Fly’s Eye”)

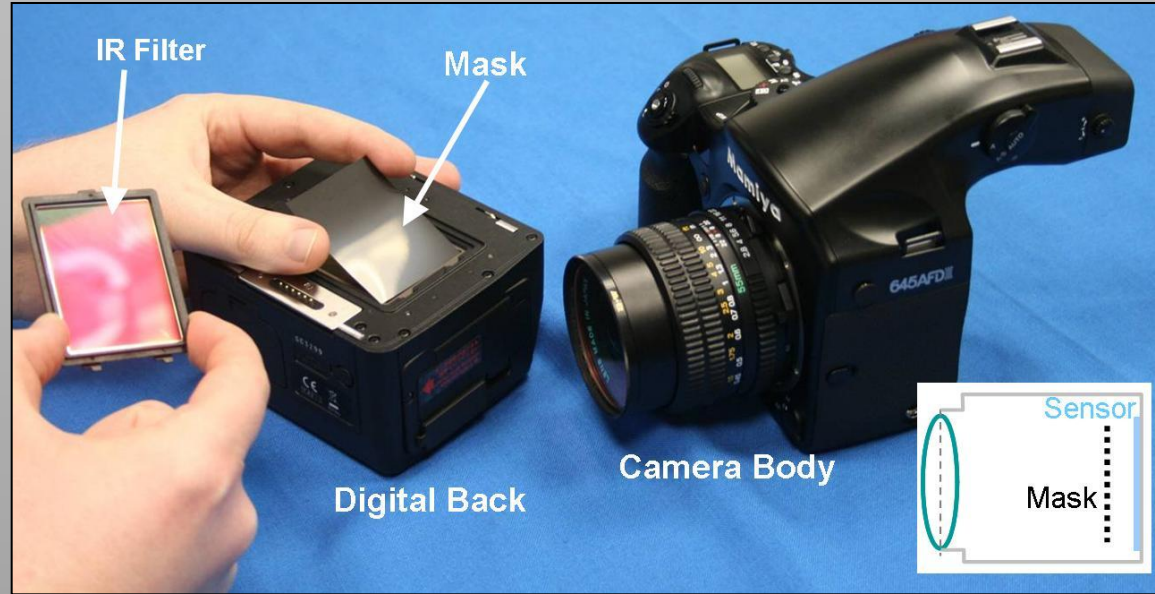
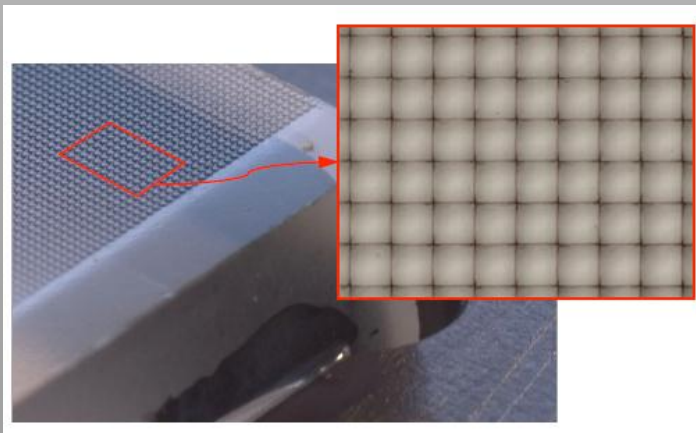


Spatially-multiplexed light field capture using lenslets:

- Impose fixed trade-off between spatial and angular resolution

[Lippmann 1908]

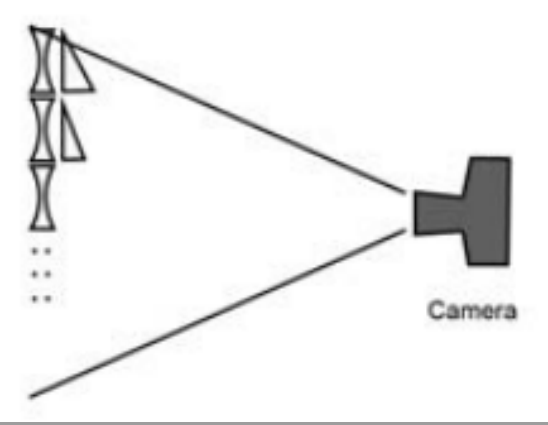
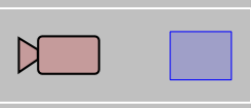
Modern, Digital Implementations



Digital Light Field Photography

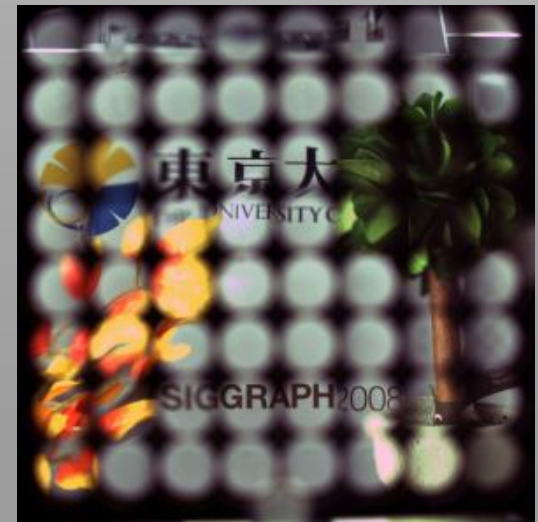
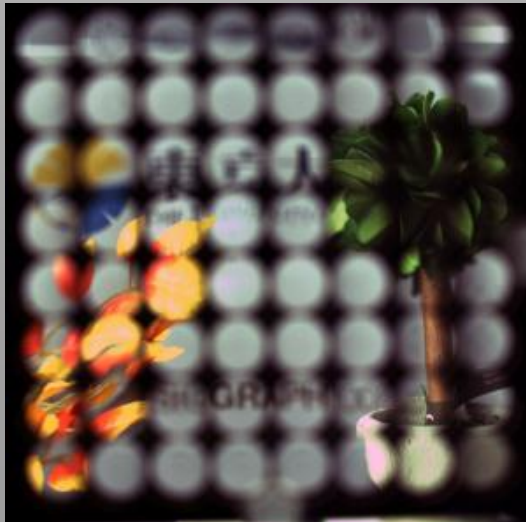
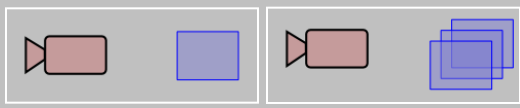
- Hand-held plenoptic camera [Ng et al. 2005]
- Heterodyne light field camera [Veeraraghavan et al. 2007]

External, Fixed Lens Attachments



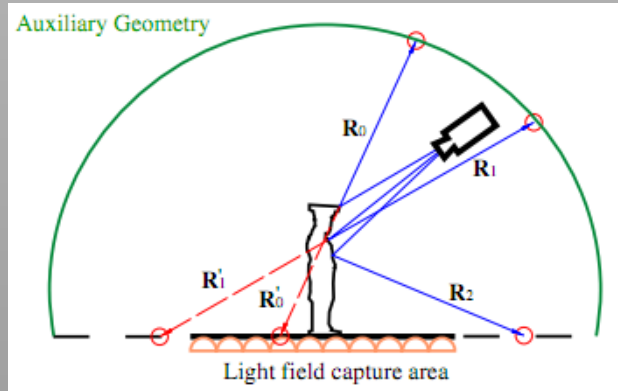
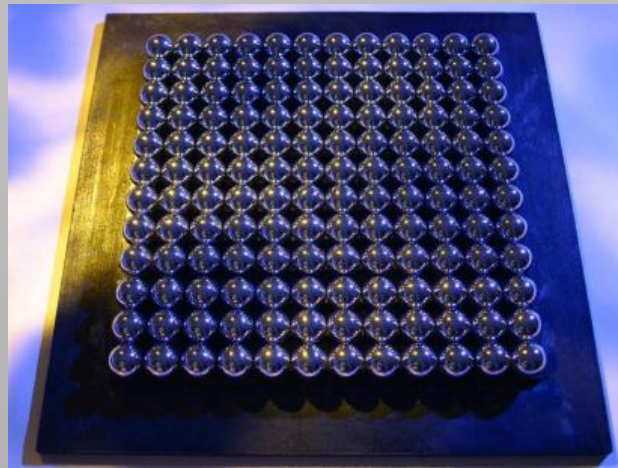
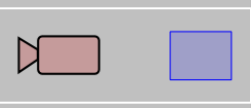
[Georgiev et al. 2006]

External, Adaptive Lens Attachments

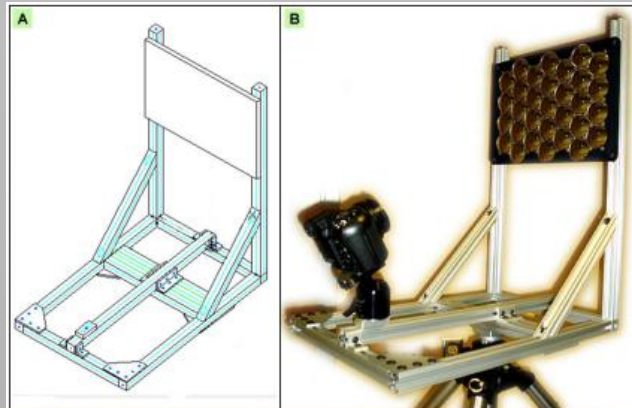


[Ueda et al. 2008]

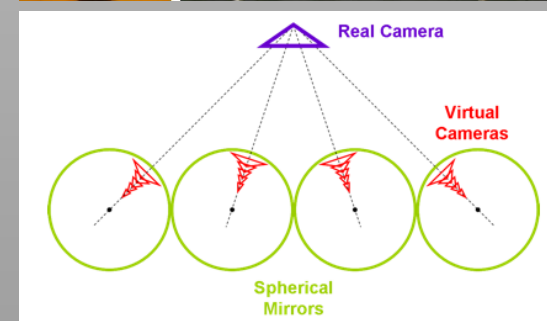
Virtual Cameras using Mirror Arrays



[Unger et al. 2003]

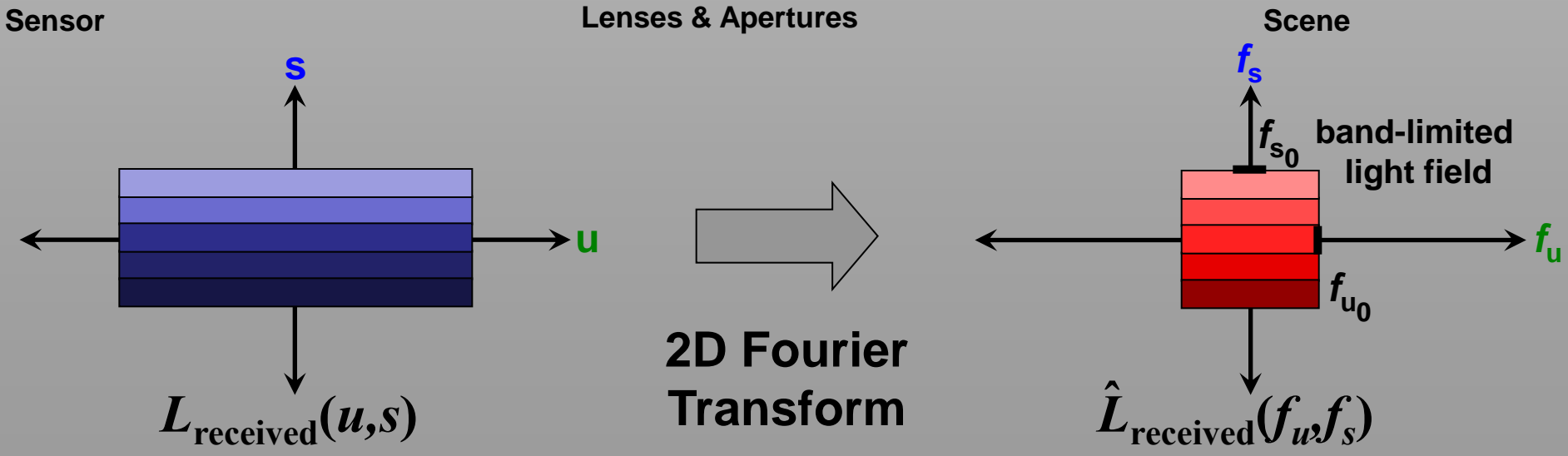
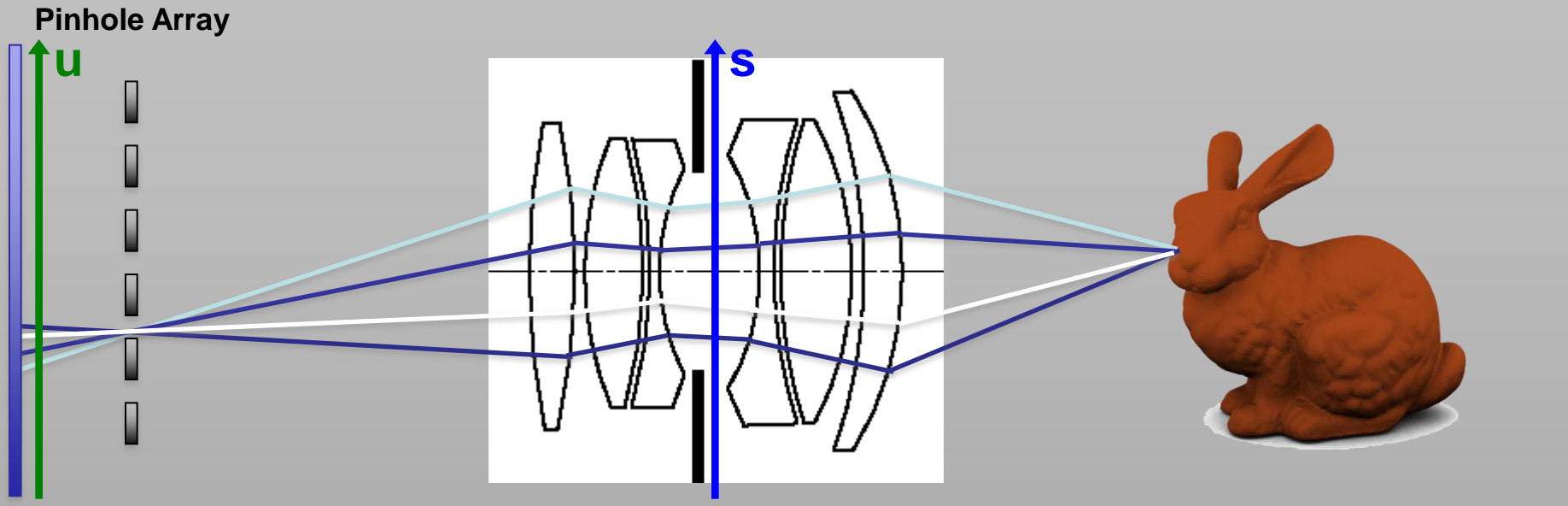
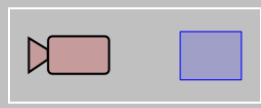


[Lanman et al. 2006]

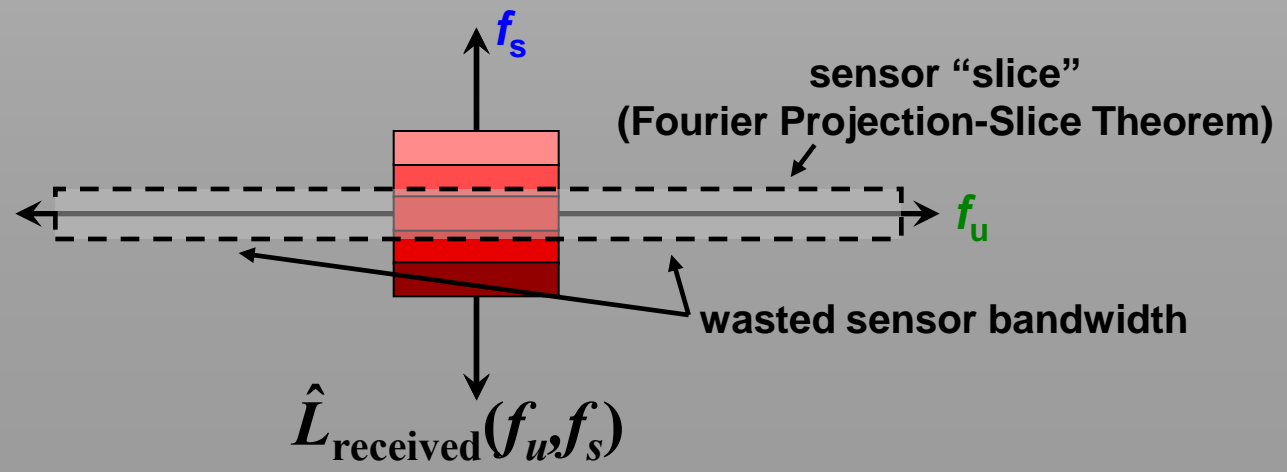
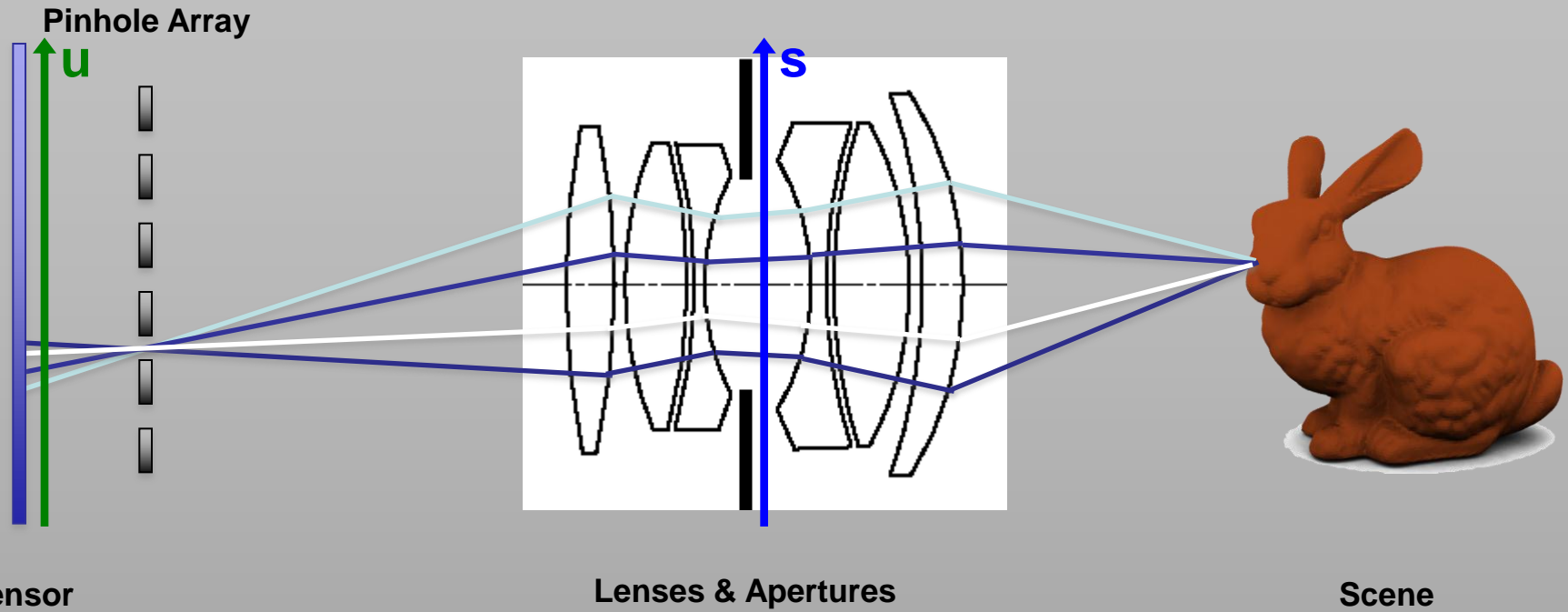
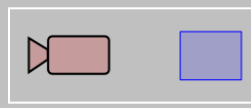


[Taguchi et al. 2010]

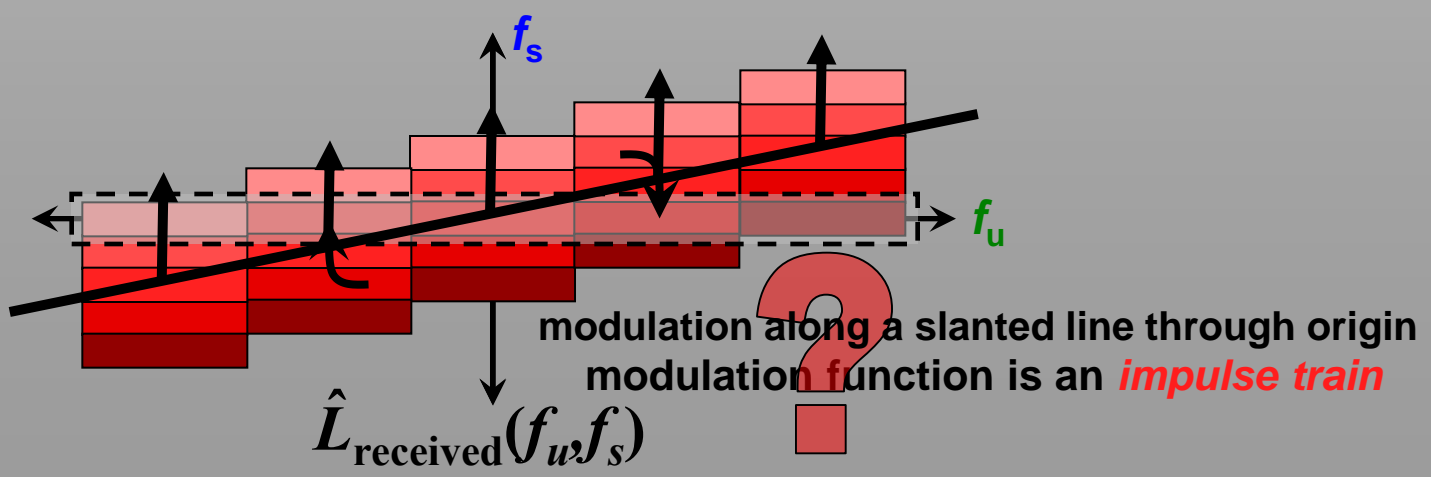
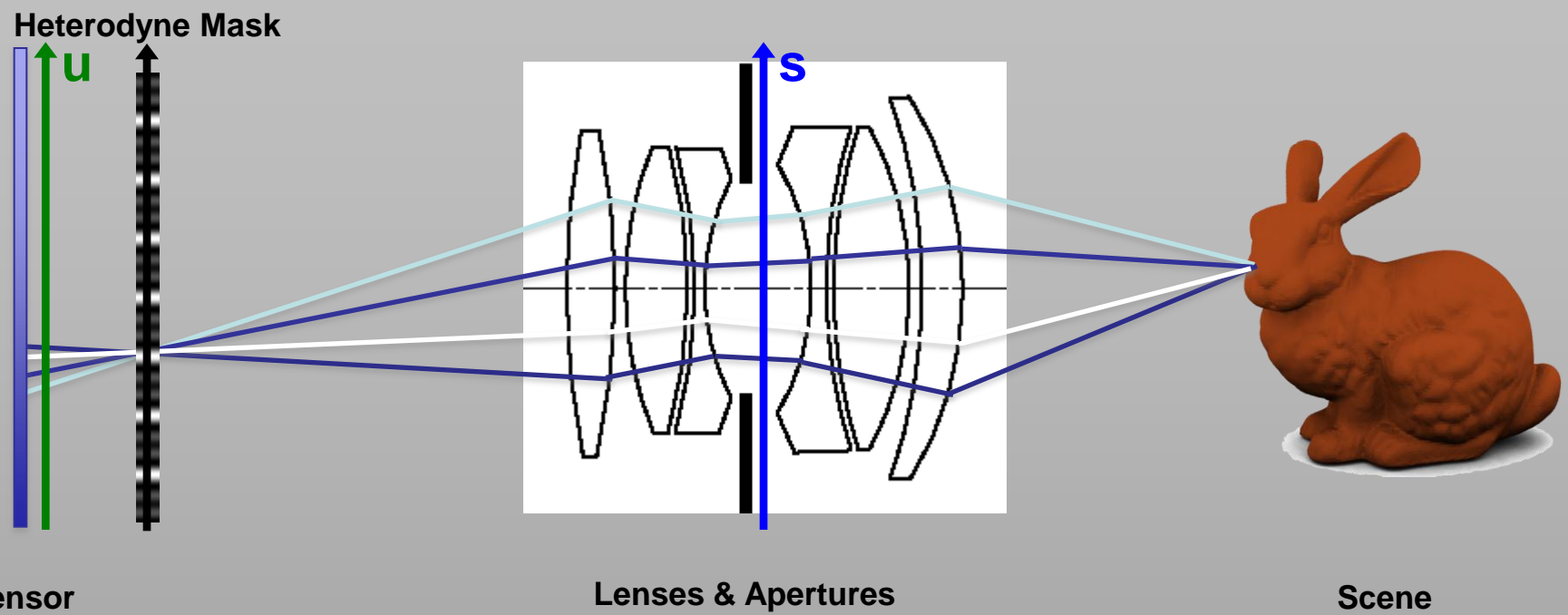
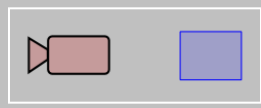
Light Field Analysis of Barrier Cameras



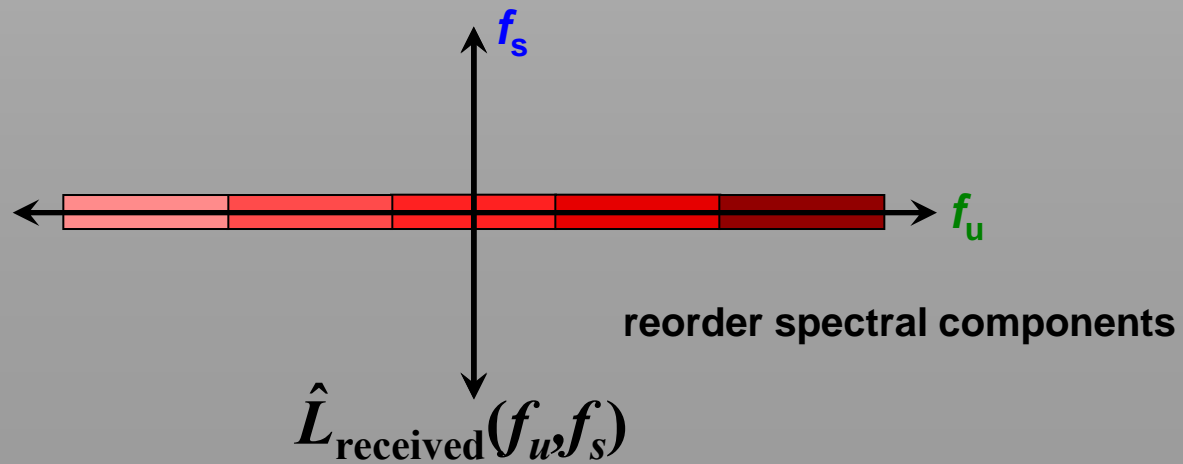
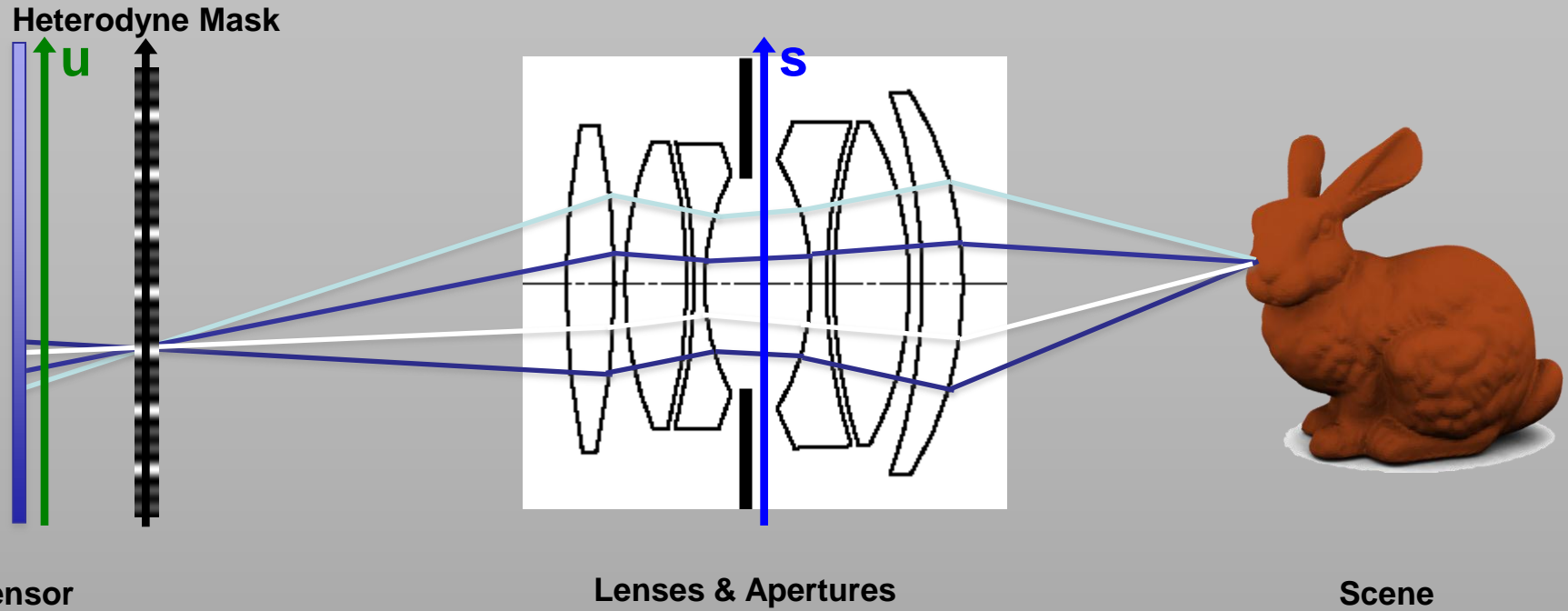
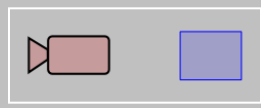
Modeling Sensors in the Frequency Domain



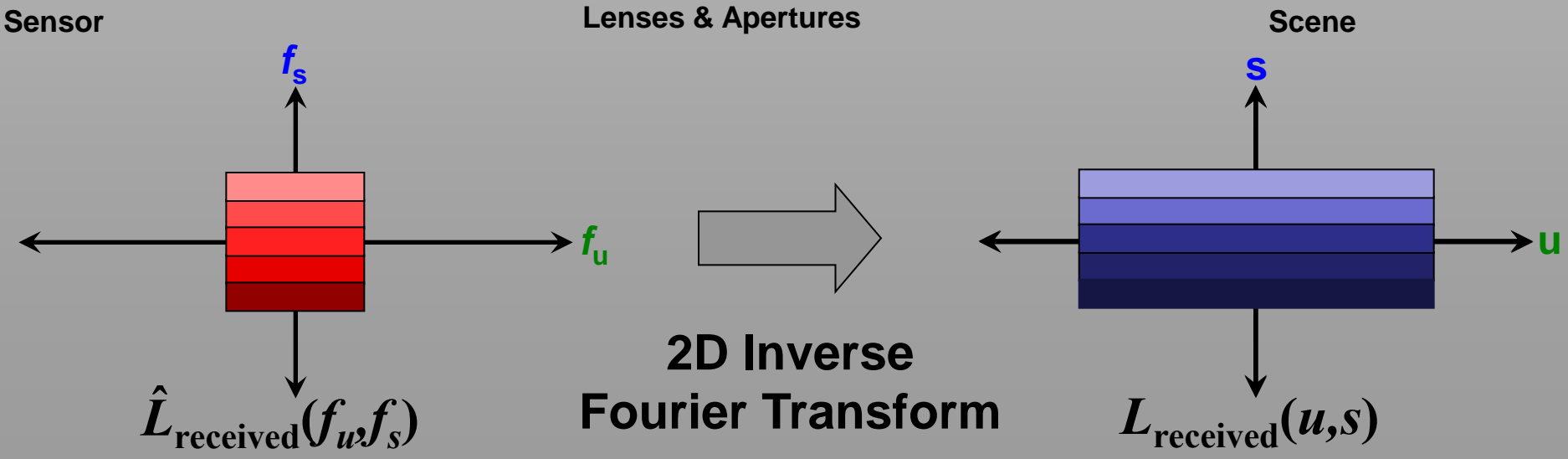
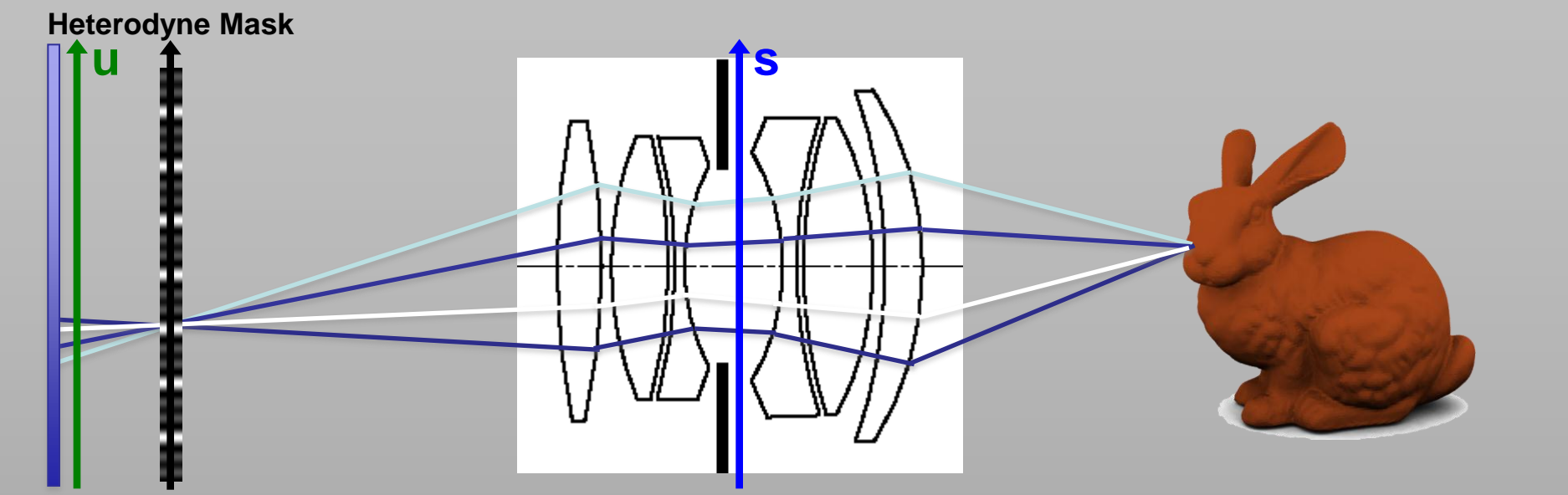
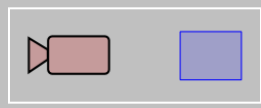
Heterodyne Light Field Cameras



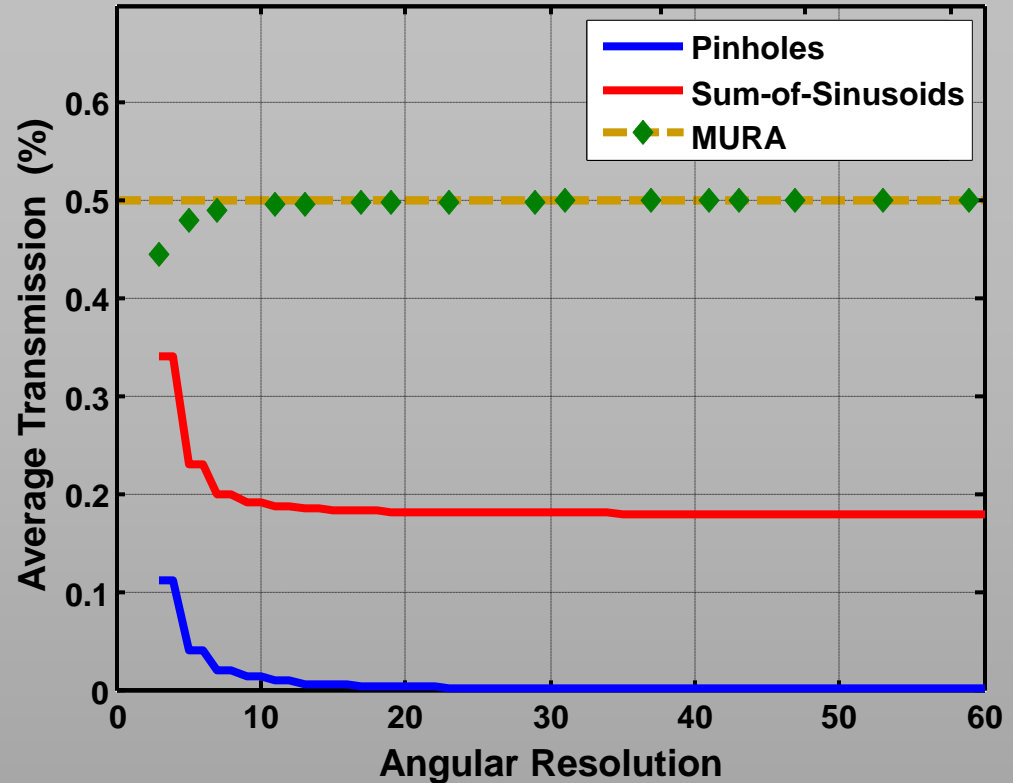
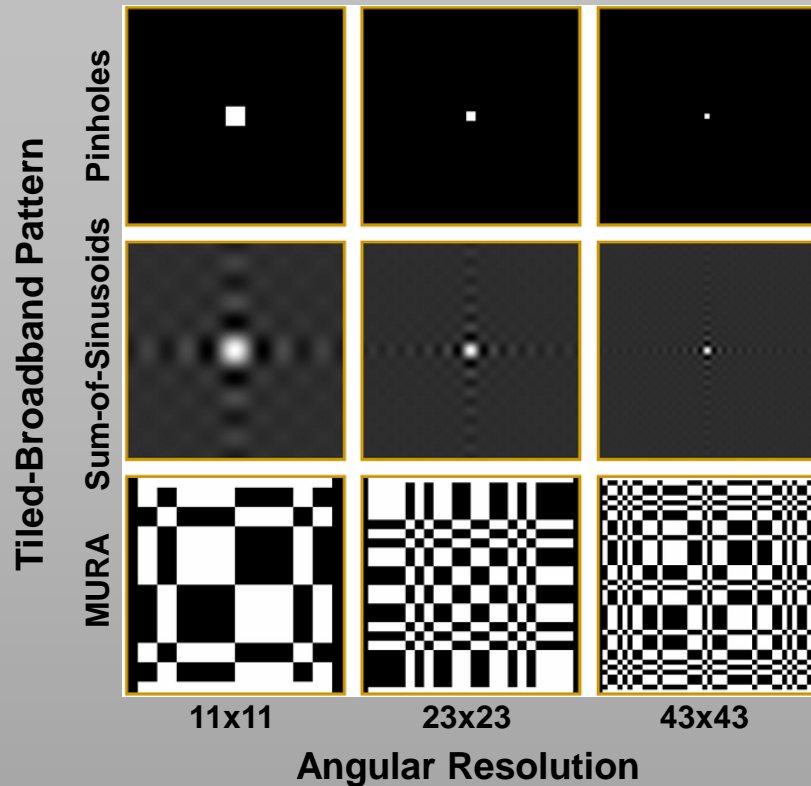
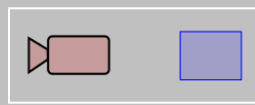
Heterodyne Light Field Cameras



Heterodyne Light Field Cameras



Optimal Masks for Optical Heterodyning

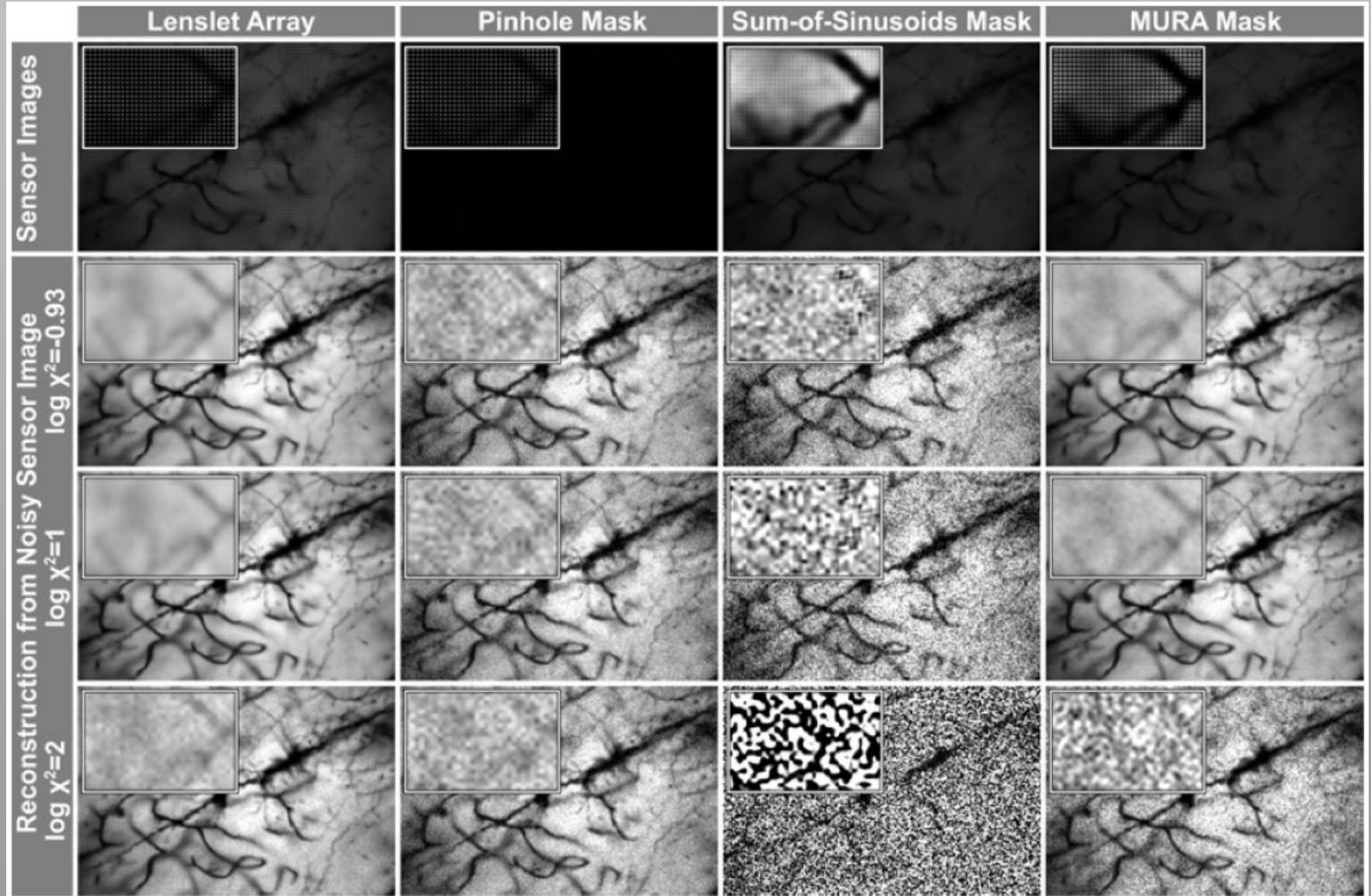
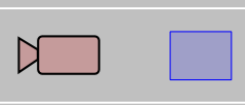


Benefits and Limitations

- Sum-of-Sinusoids converges to $\approx 18\%$ transmission
- Tiled-MURA near 50% (but only for prime-valued lengths)
- Binary vs. continuous-tone process (quantization)

[Lanman et al. 2008]

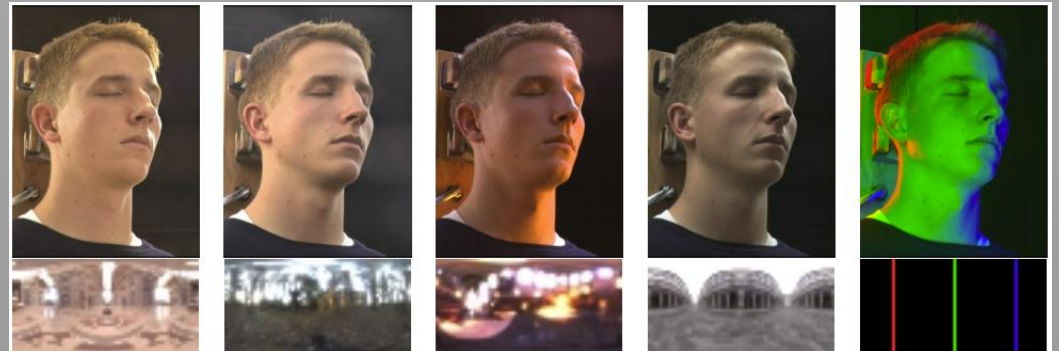
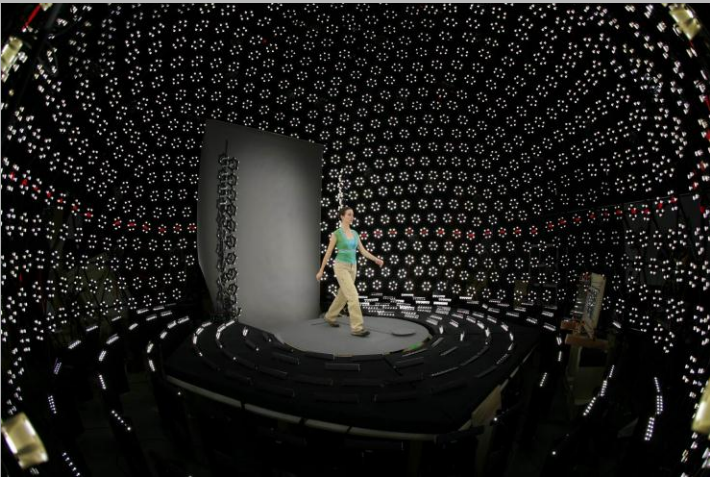
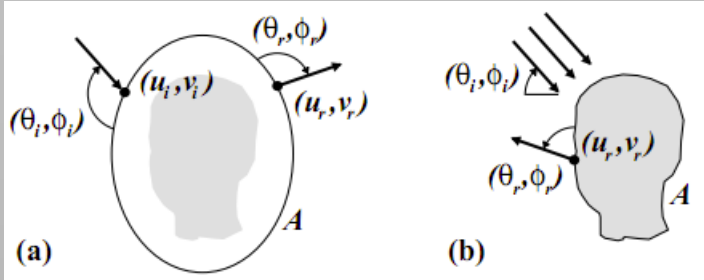
Optimal Masks for Optical Heterodyning



[Ihrke et al. 2010]

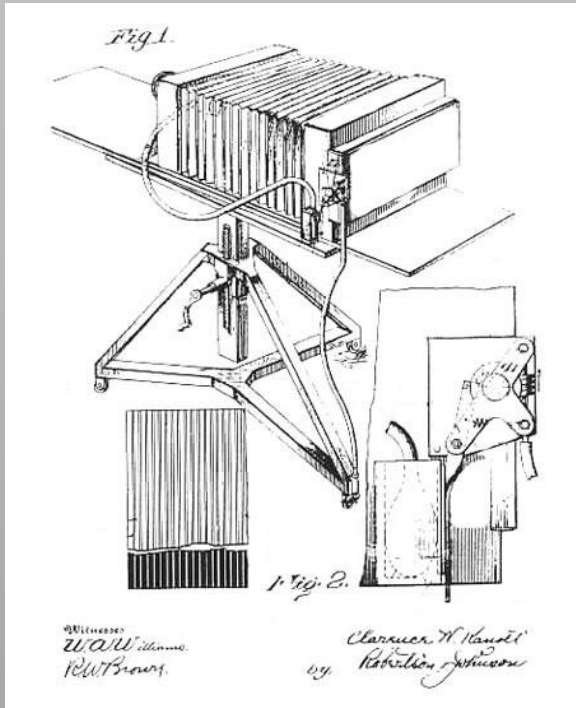
IV.V Applications

Scene Relighting with 8D Reflectance Fields

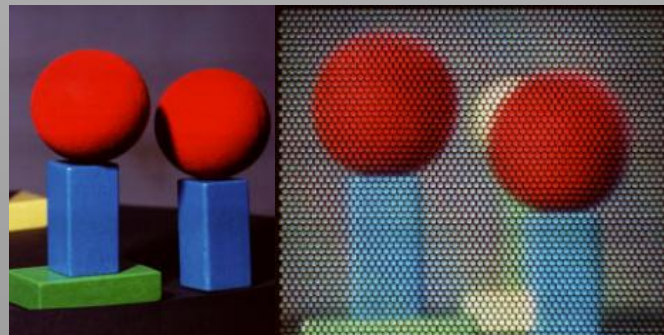
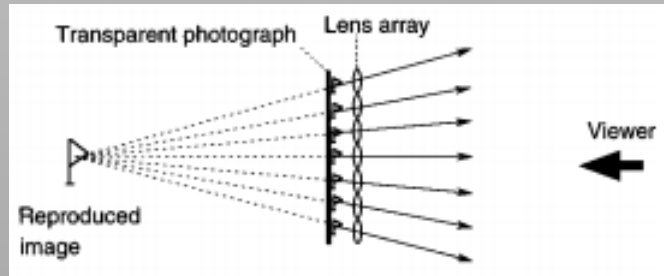
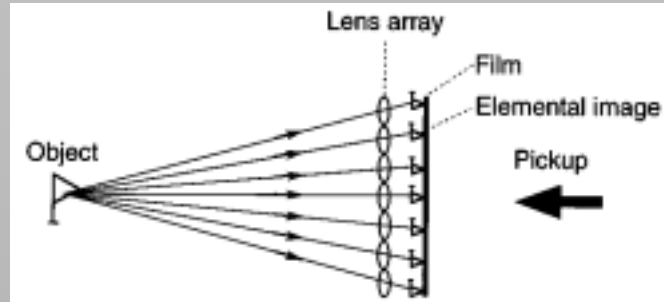


[Debevec et al. 2000]

3D Displays



Parallax Panoramagram
[Kanolt 1918]

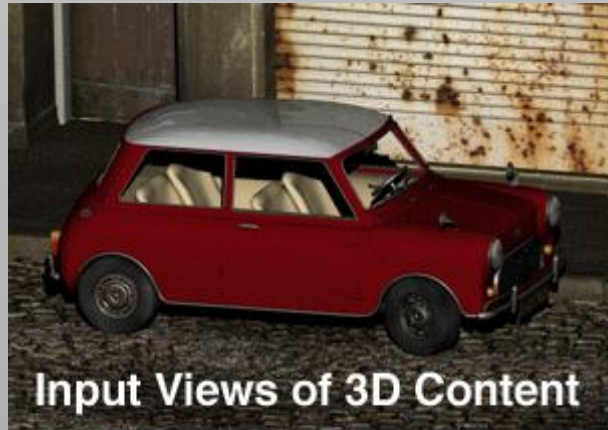


3DTV with Integral Imaging
[Okano et al. 1999]



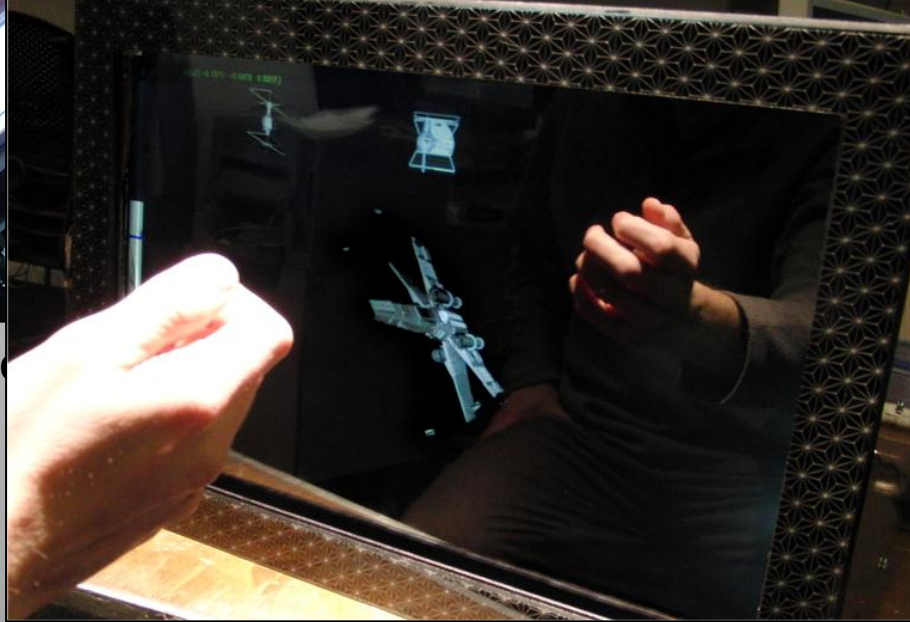
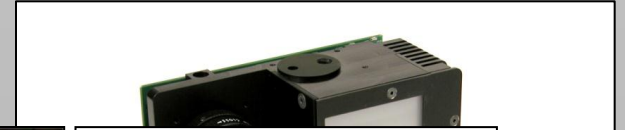
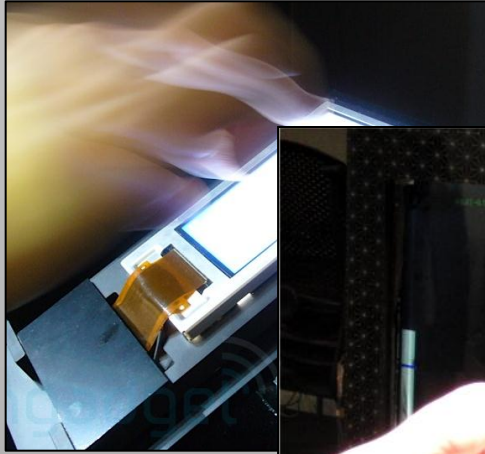
MERL 3DTV
[Matusik and Pfister 2004]

3D Displays



High-Rank 3D (HR3D)
[Lanman et al. 2010]

(BiDi)irectional Screens



Embed

-of-flight)

Sensing multi-touch, gestures, and ambient light using LCDs

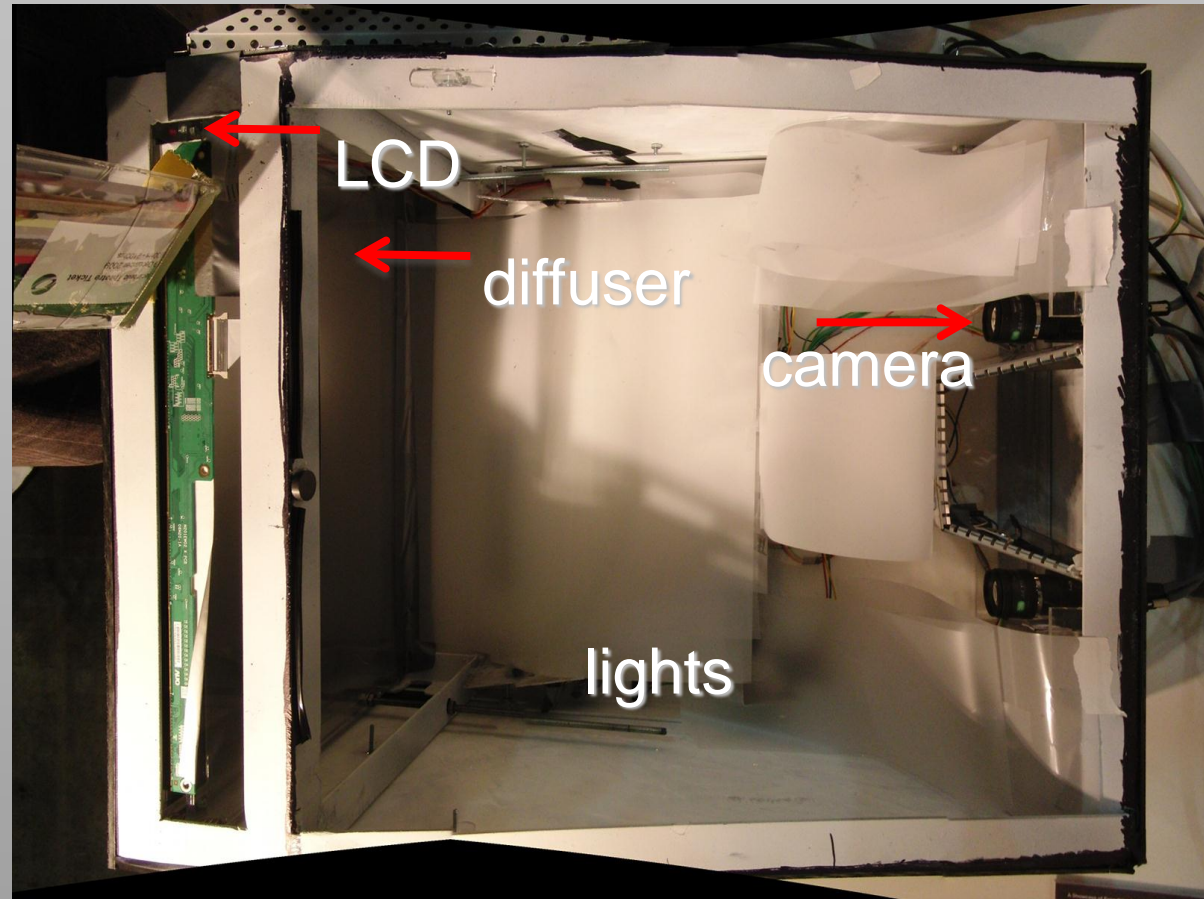
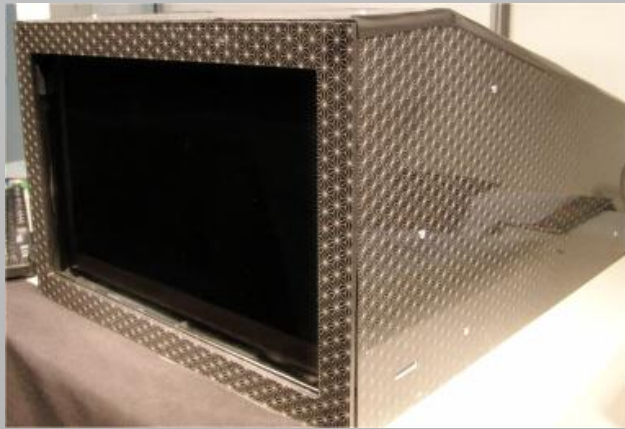
Design Goals

- Capture light field reflected or emitted by objects
- Prevent image capture from interfering with image display
- Support unencumbered interaction with thin form factor

Ambient light sensors

[Hirsch et al. 2009]

(BiDi)irectional Screens

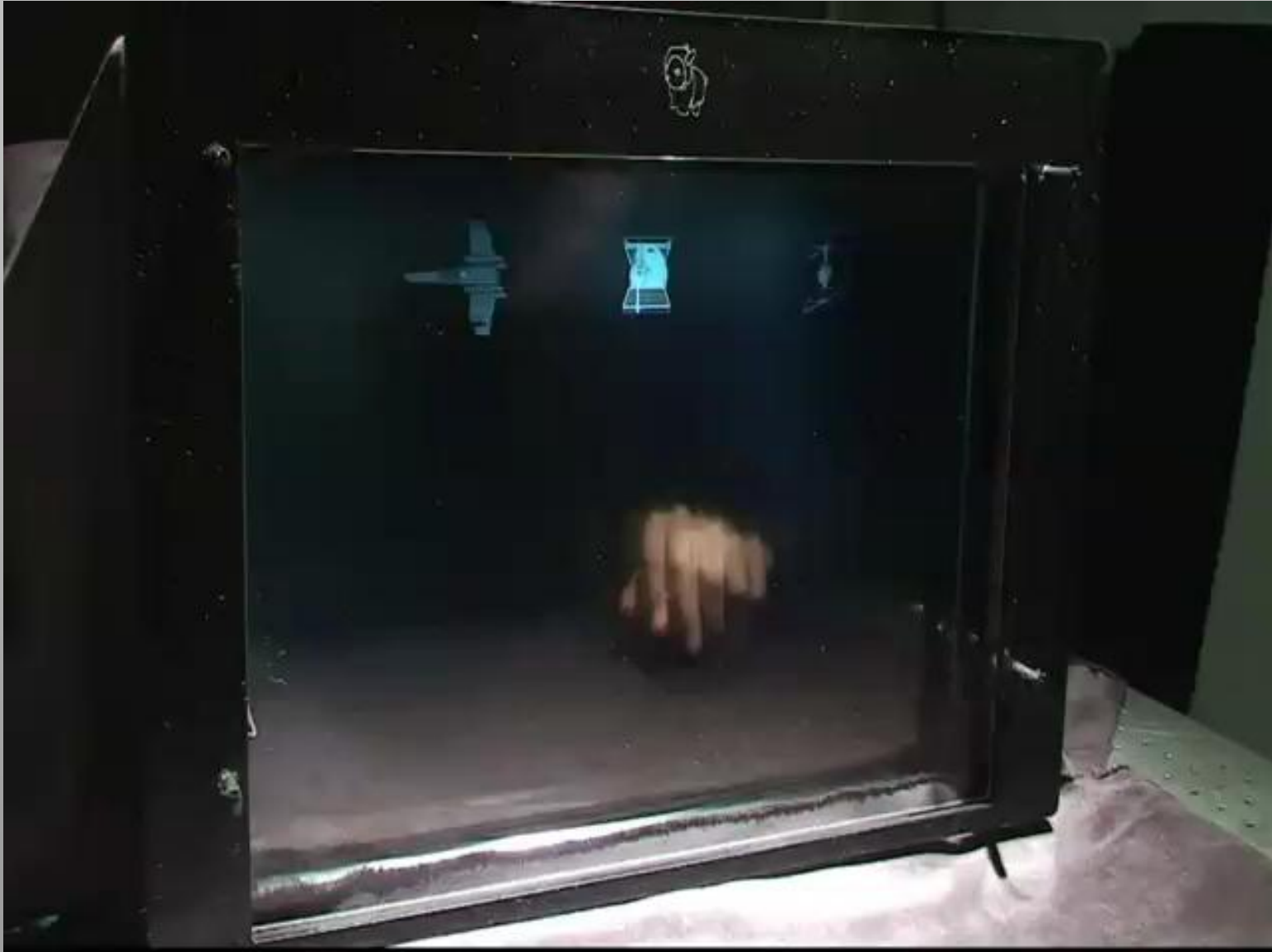


Components

- 20.1 inch Sceptre X20WG-Nagall LCD [1680 × 1050 @ 60 fps]
- Point Grey Flea2 video cameras [1280 × 960 @ 7 fps]

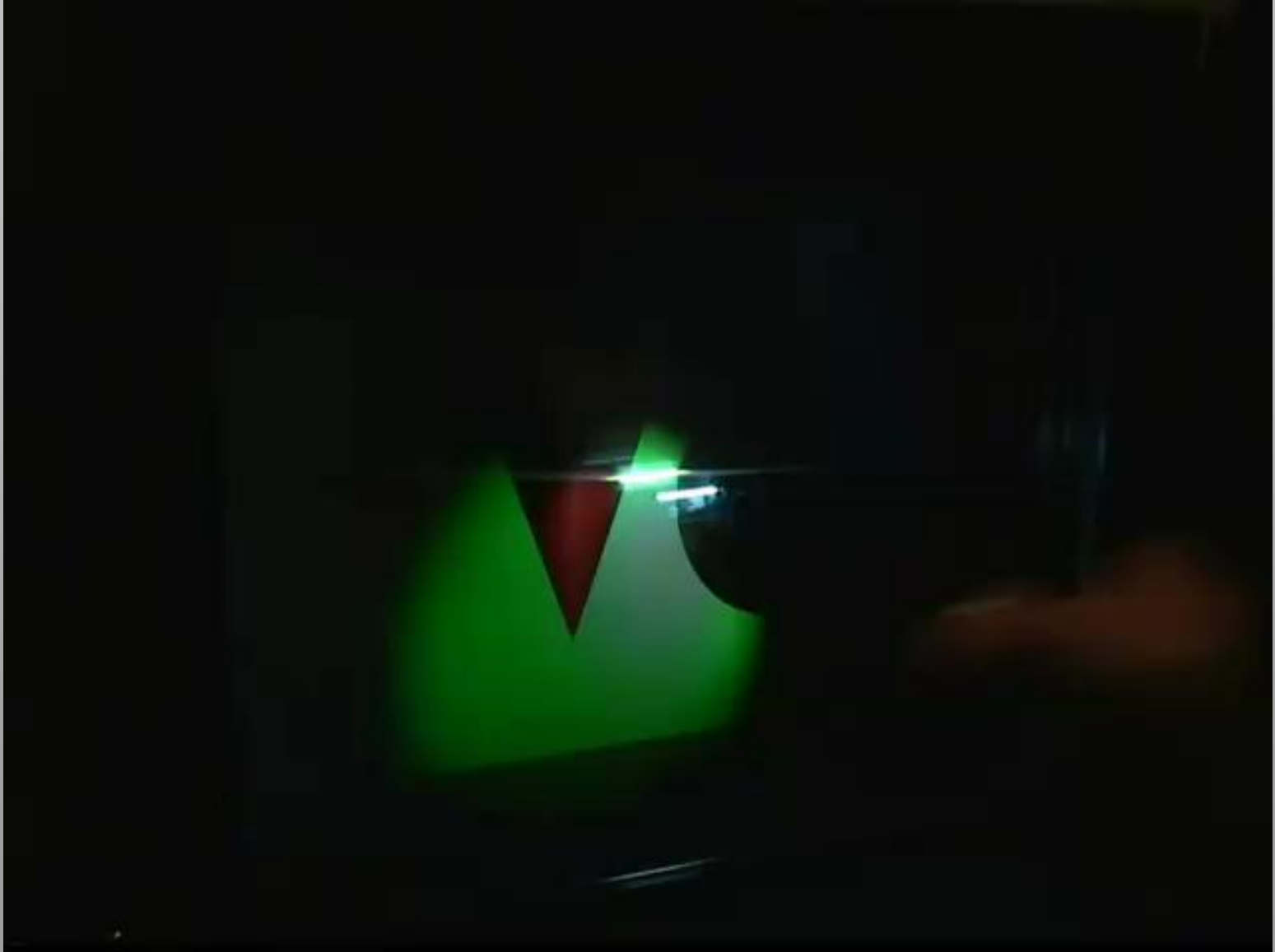
[Hirsch et al. 2009]

(BiDi)rectional Screens



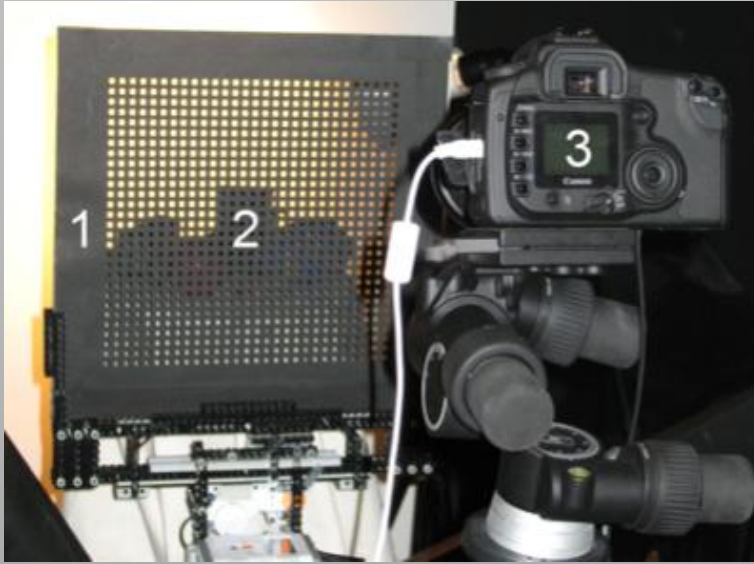
[Hirsch et al. 2009]

(BiDi)irectional Screens

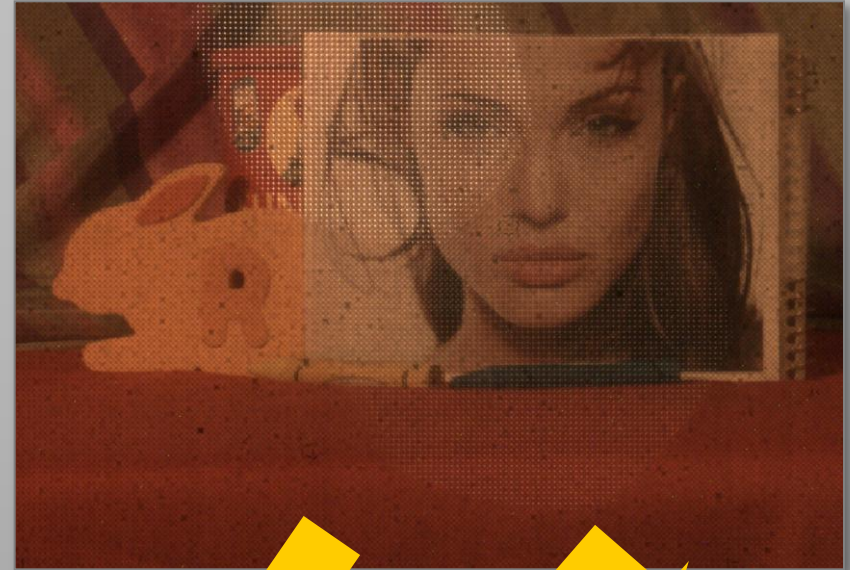


[Hirsch et al. 2009]

Mitigating Glare

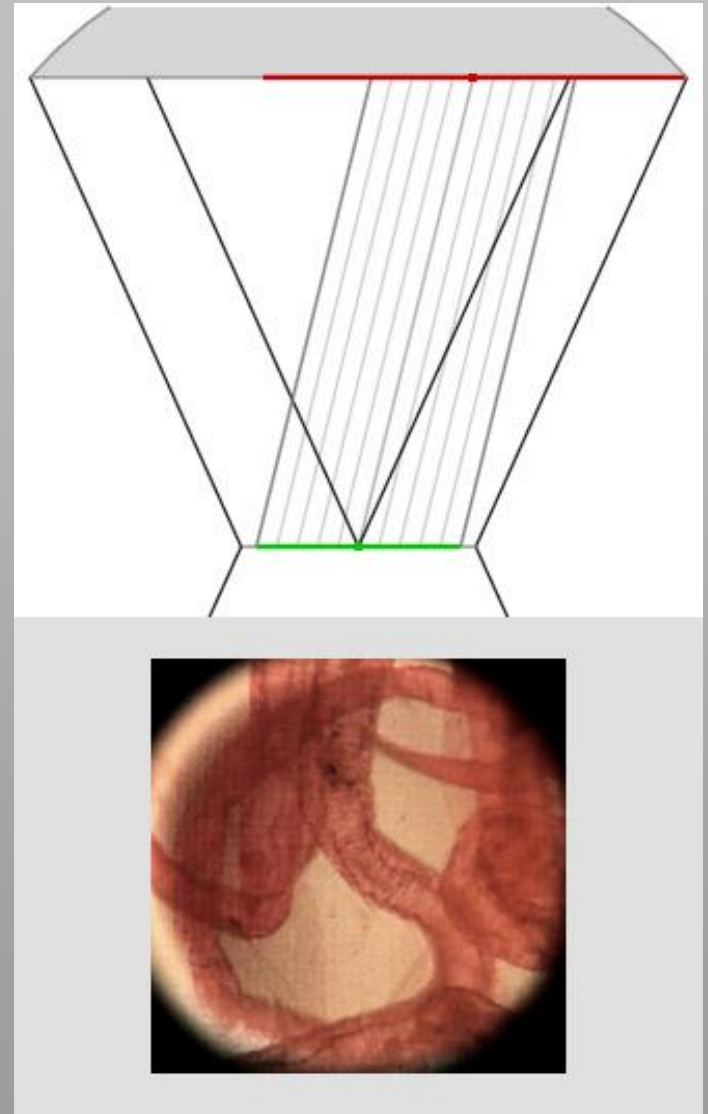
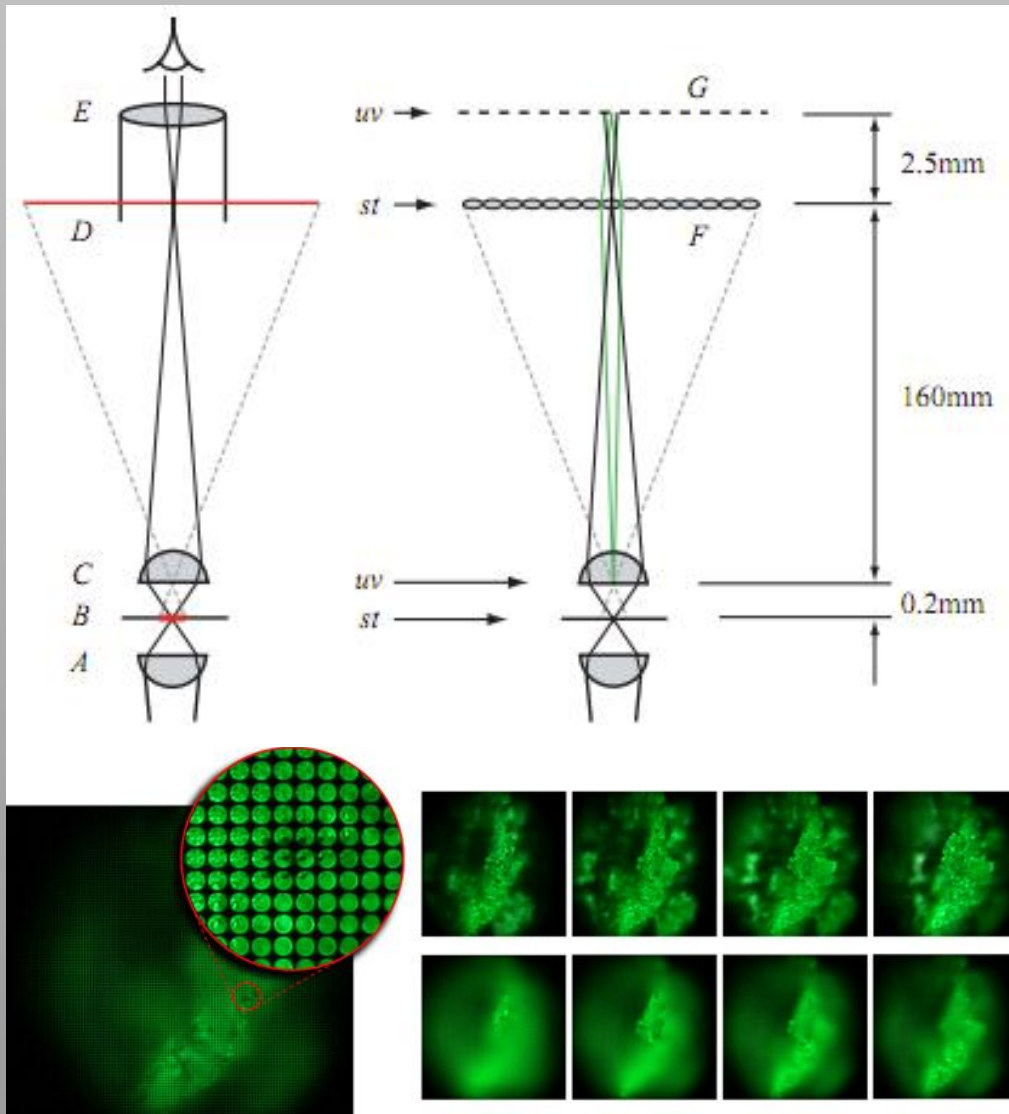


Veiling Glare in HDR Imaging
[Talvala et al. 2007]



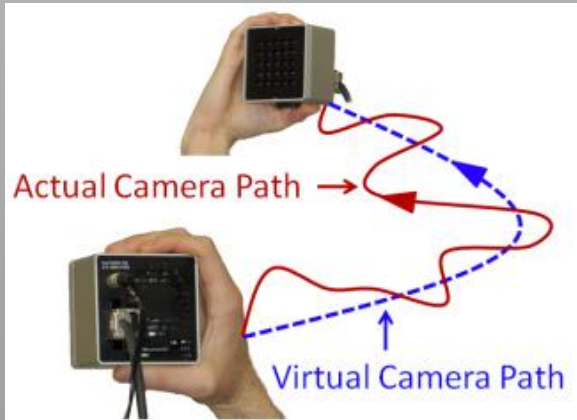
Glare-Aware Photography
[Raskar et al. 2008]

Light Field Microscopy

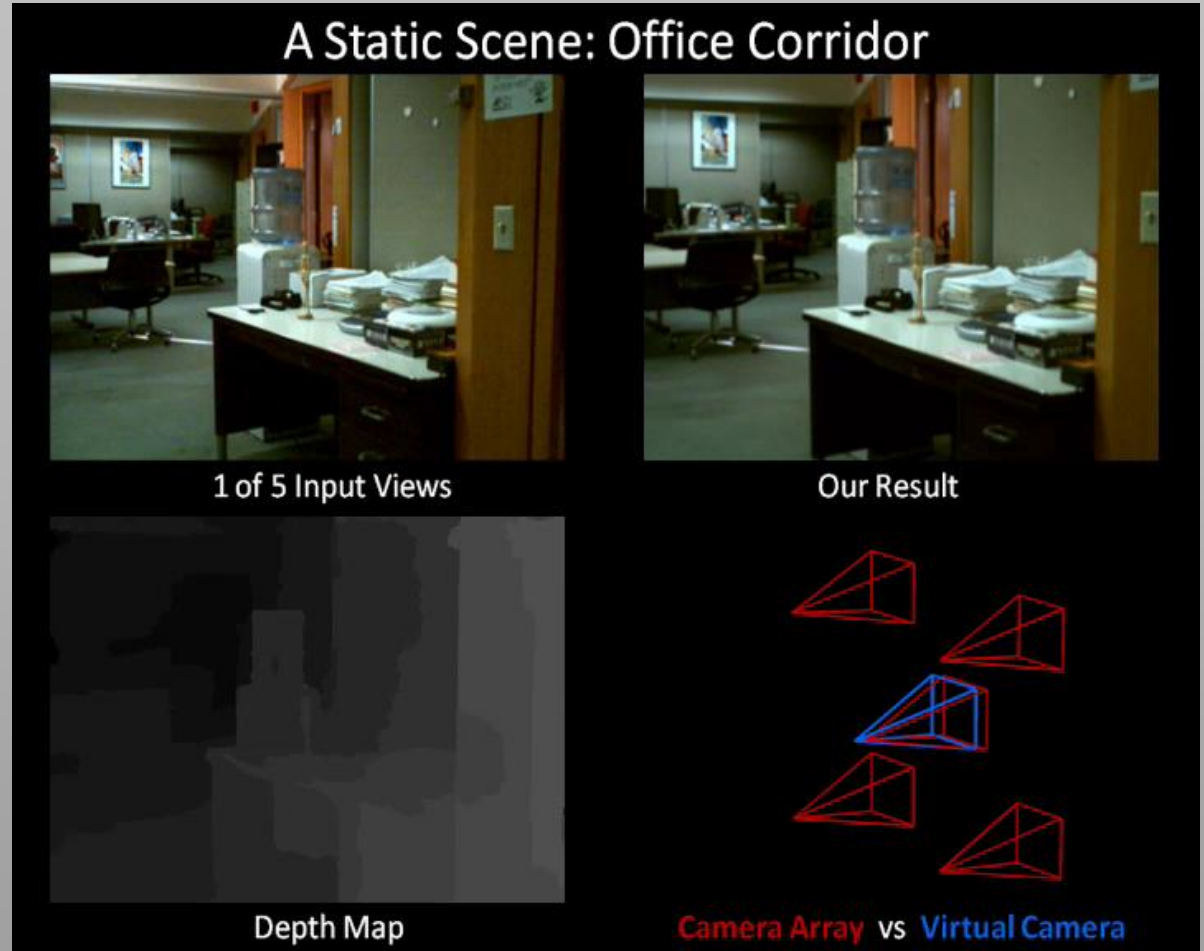


[Levoy et al. 2004; Levoy et al. 2006]

Image Stabilization



A Static Scene: Office Corridor



1 of 5 Input Views

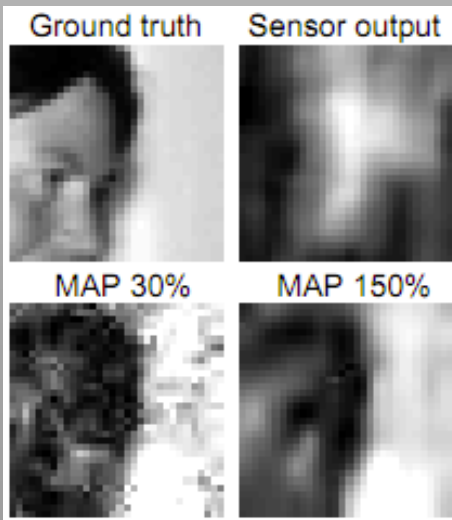
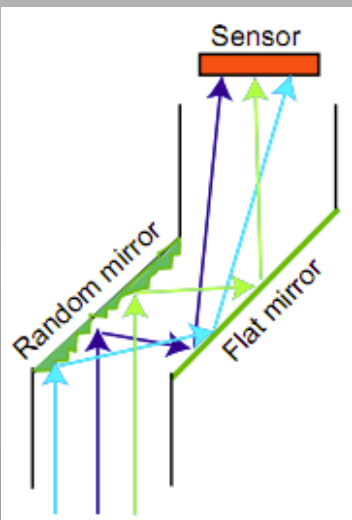
Our Result

Depth Map

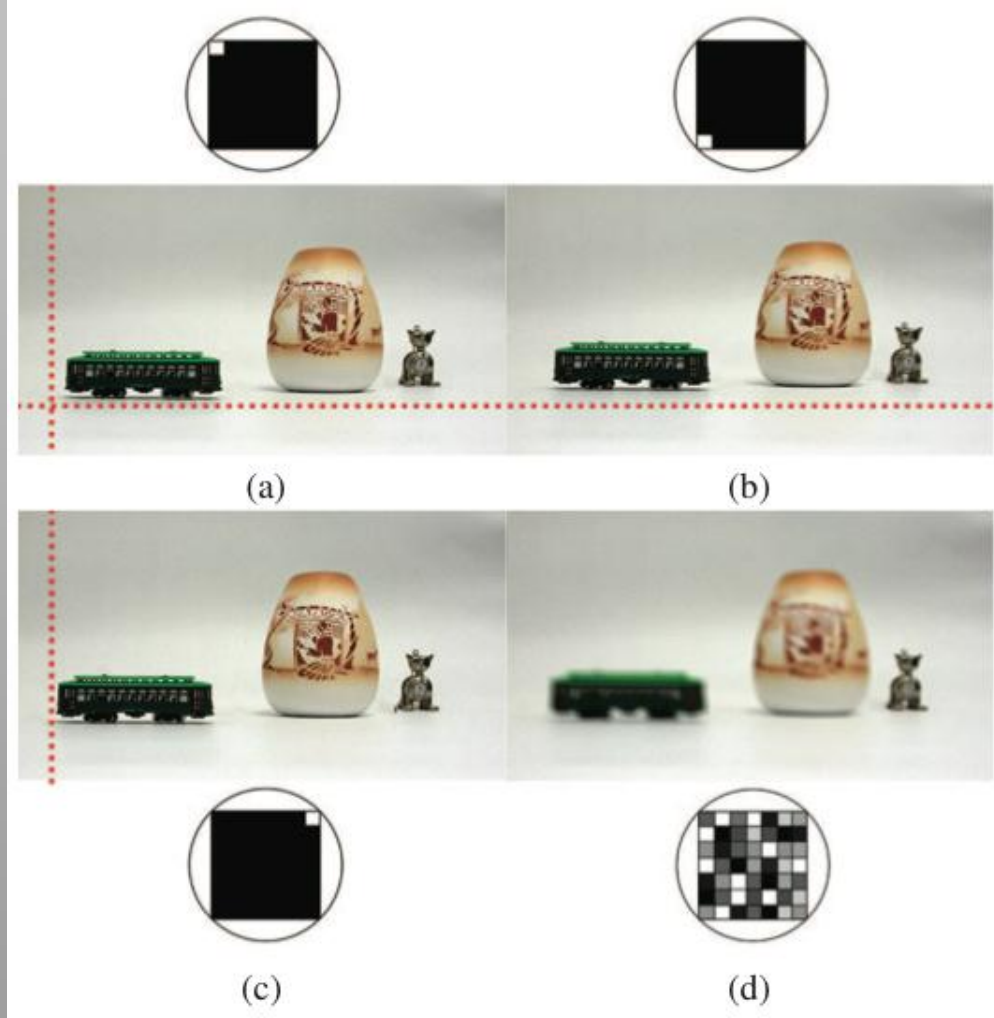
Camera Array vs Virtual Camera

[Smith et al. 2009]

Compressive Light Field Capture



Random Lens Imaging
[Fergus et al. 2006]



Compressive Light Field Imaging
[Babacan et al. 2009; Ashok and Neifeld 2010]

Next: Multiplexing Space & Focal Surfaces

Computational Plenoptic Imaging

Gordon Wetzstein¹
Wolfgang Heidrich³

Ivo Ihrke²
Kurt Akeley⁴

Douglas Lanman¹
Ramesh Raskar¹

¹MIT Media Lab

²Saarland University

³University of British Columbia

⁴Lytro, Inc.

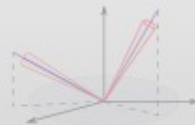
V. Multiplexing Space & Focal Surfaces



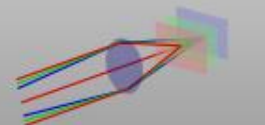
Dynamic Range



Color Spectrum



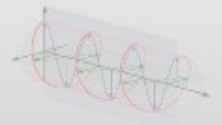
Directions | Light Fields



Space | Focal Surfaces



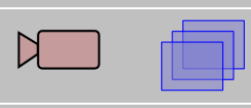
Time



Further Properties

V.I Gigapixel Imaging & Superresolution

Gigapixel Panorama



Rotate camera!

[Kopf et al. 07]



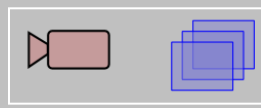
Auto-exposure for each image



Processed HDR image

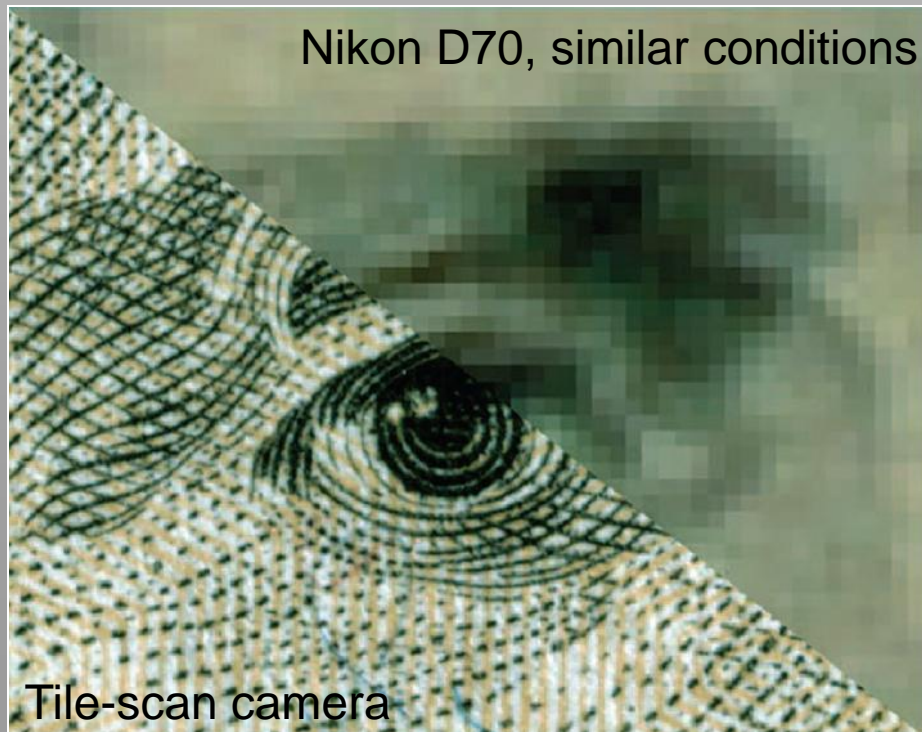
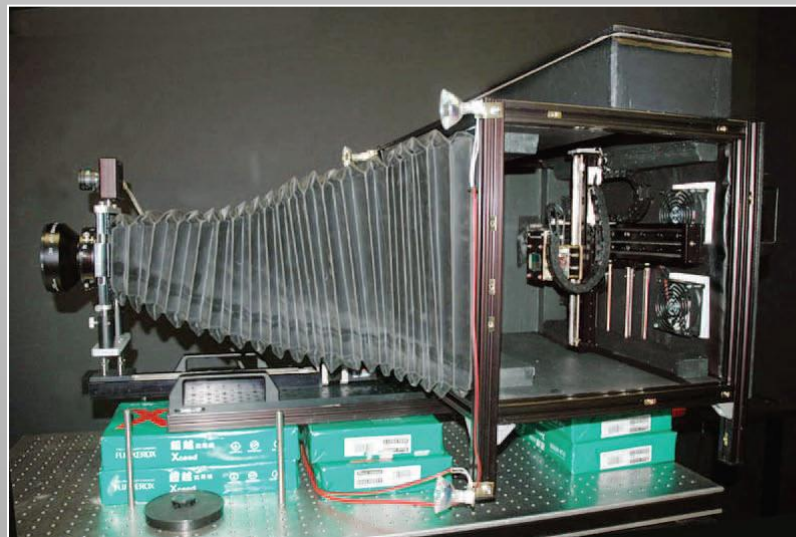


Gigapixel Large-Format Tile-Scan Camera



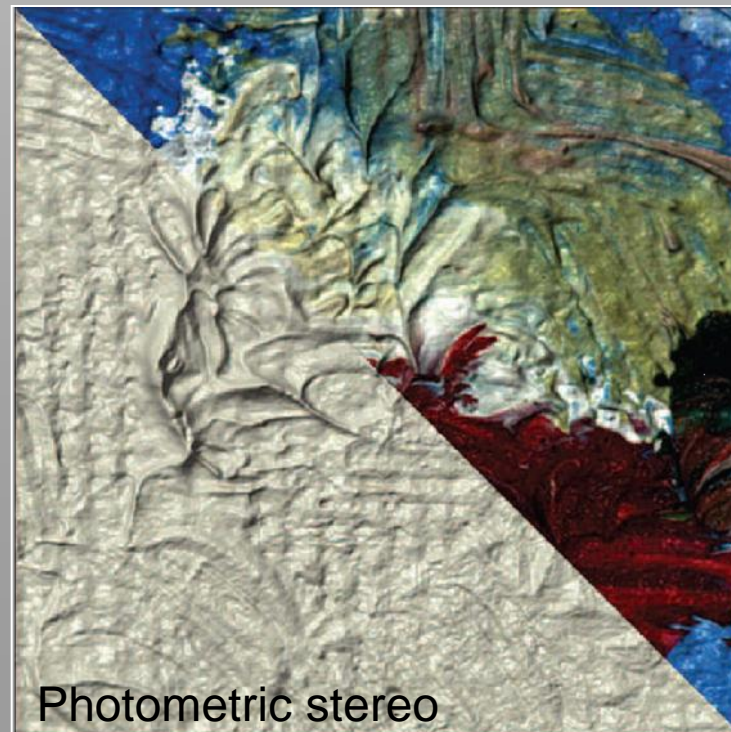
Translate sensor!

[Ben-Ezra 11]



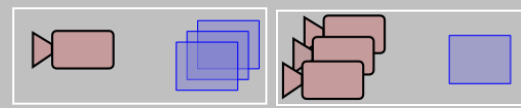
Nikon D70, similar conditions

Tile-scan camera

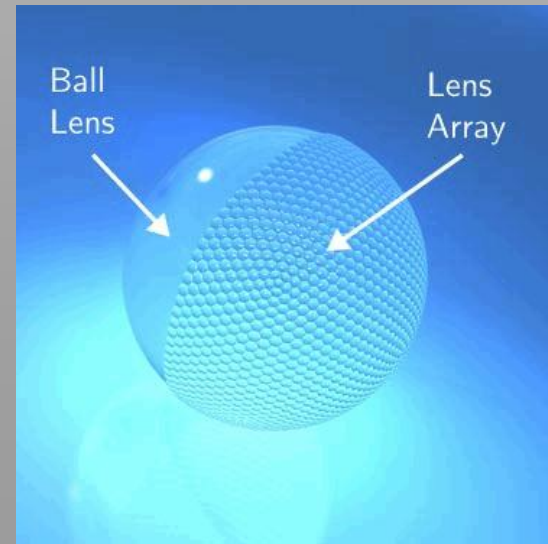
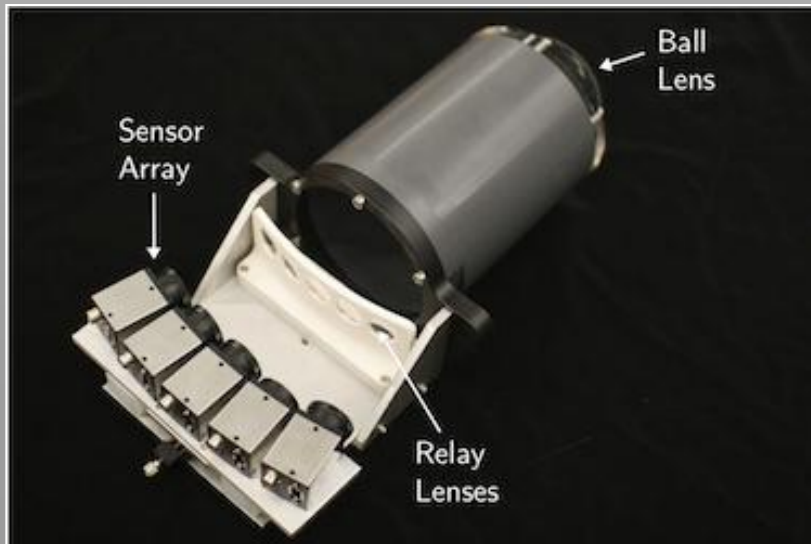
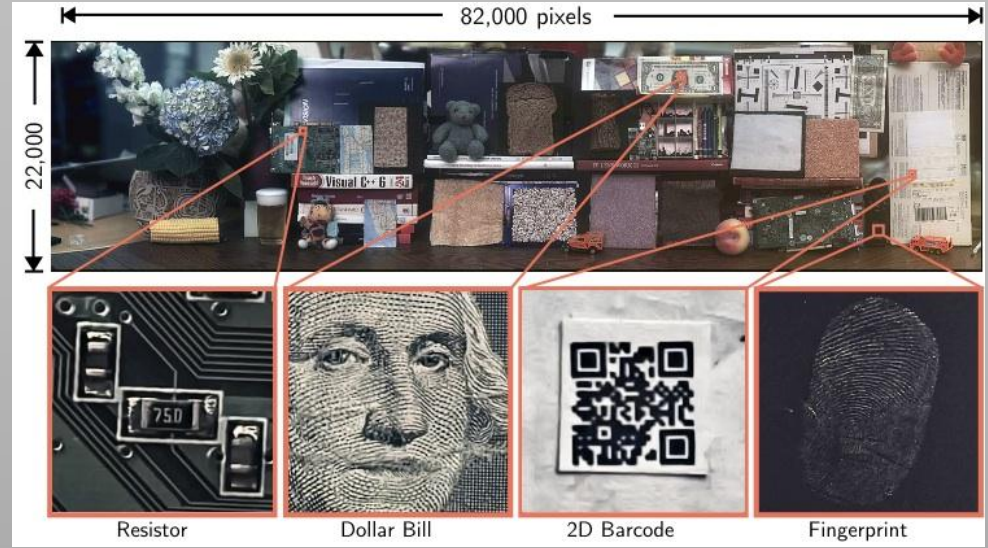
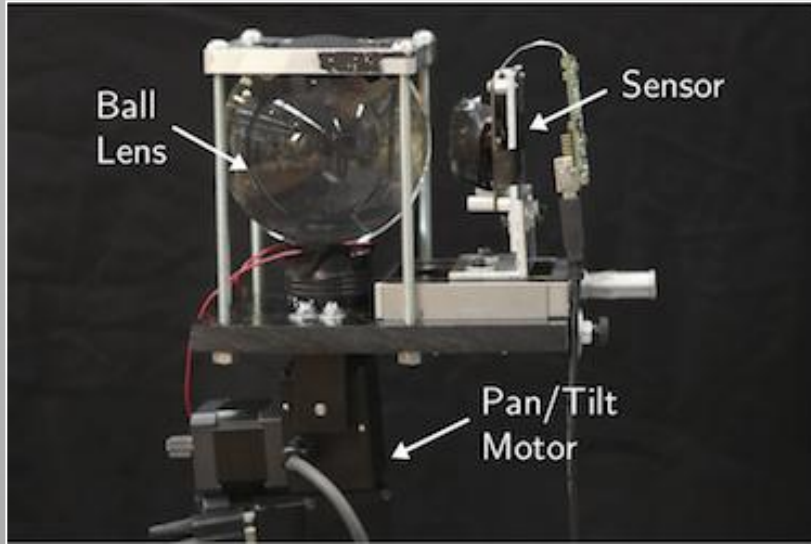


Photometric stereo

Gigapixel Computational Imaging



Beat lens scaling laws with computational processing!



[Cossairt et al. 11]

Superresolution Overview

- Input

One of the input images

- Camera

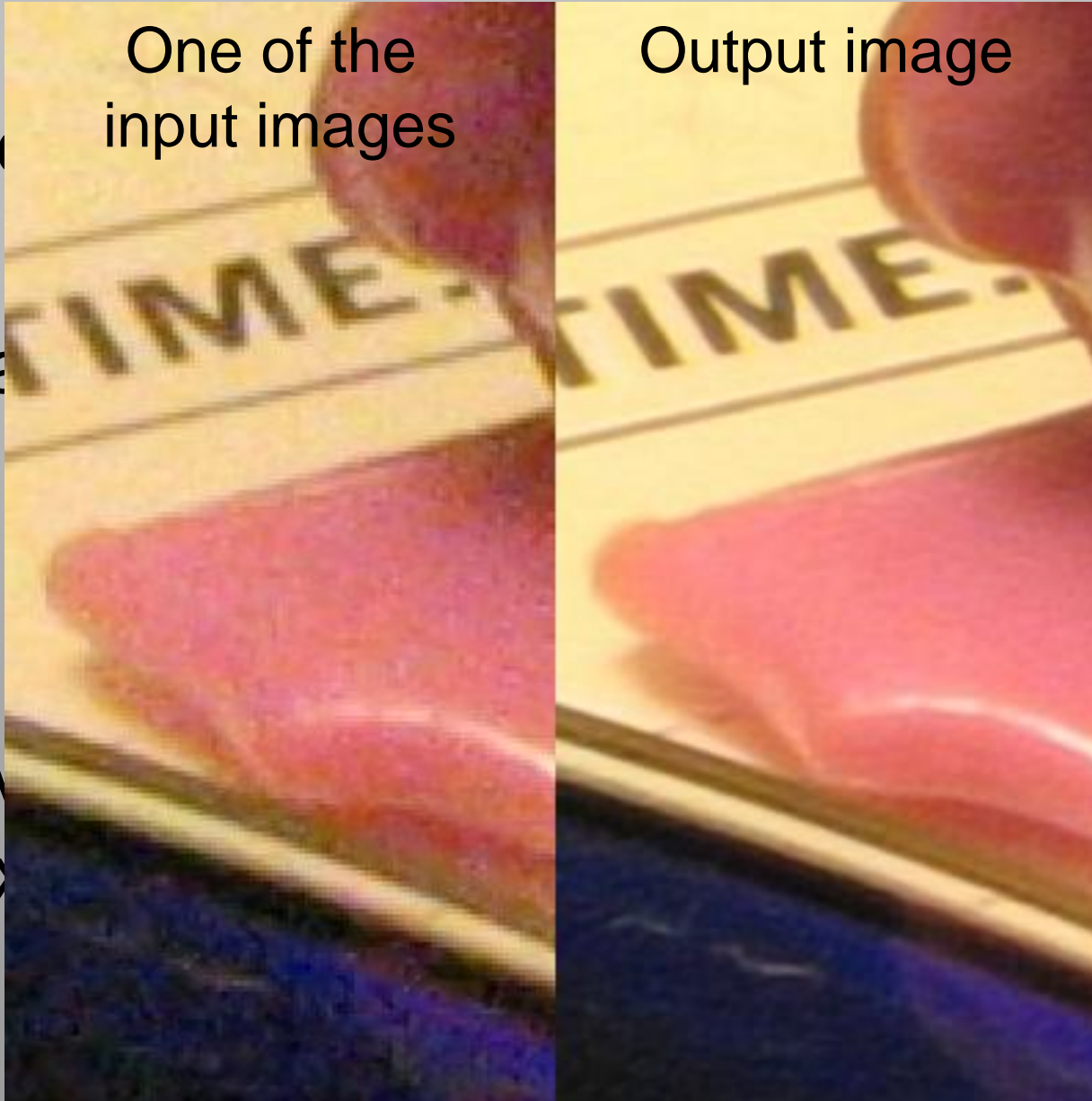
1.

2.

- Output
pixels

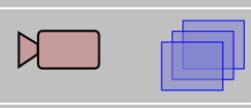
Output image





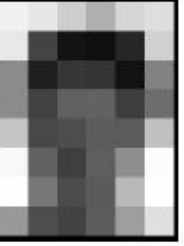





area



<http://en.wikipedia.org/wiki/Super-resolution>

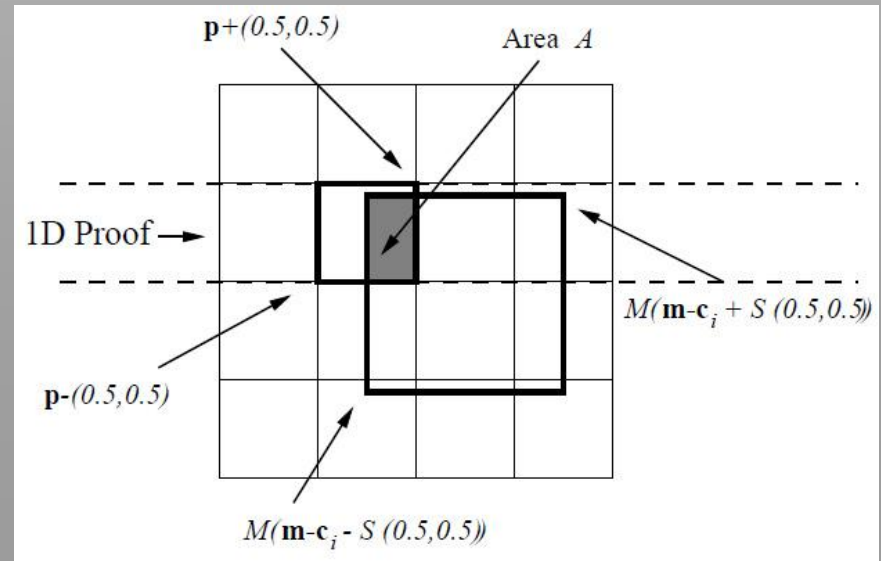
Superresolution Overview



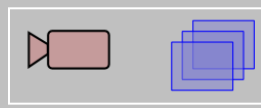
No. of Input Images	 1 (Original)	 4	 16	 64	 256
Linear Magnification	 $\times 1$	 $\times 2$	 $\times 4$	 $\times 8$	 $\times 16$

[Baker & Kanade 02]

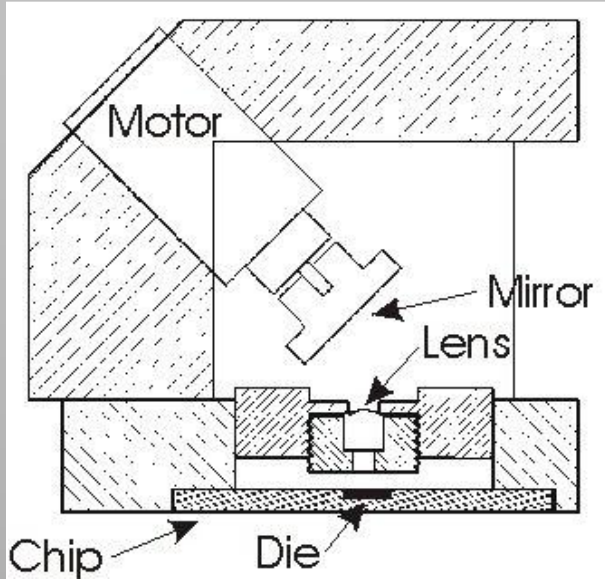
- Solve big linear system for high-res image



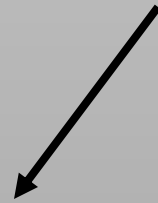
Superresolution with Mechanical Vibrations



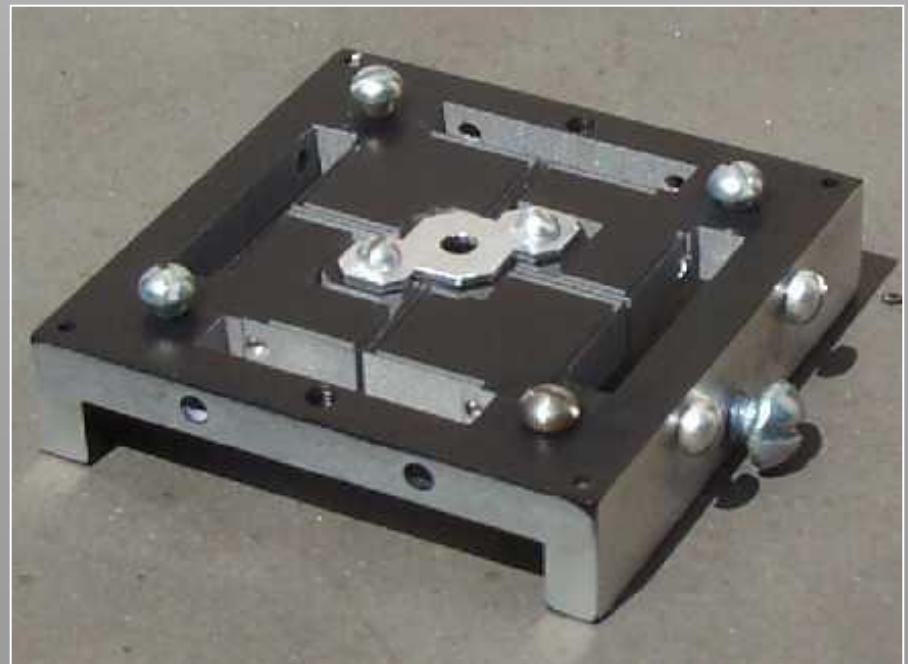
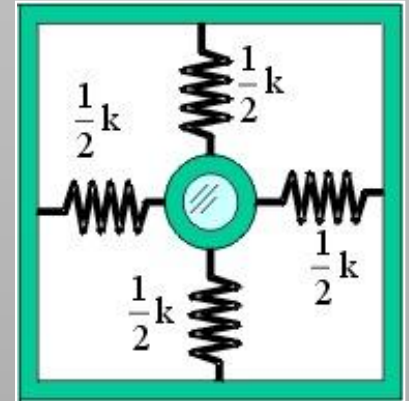
- multiple, successive images



vibrating mirror

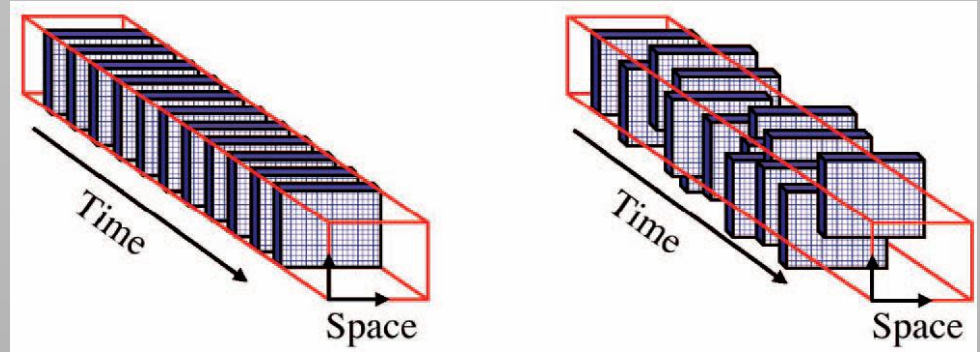
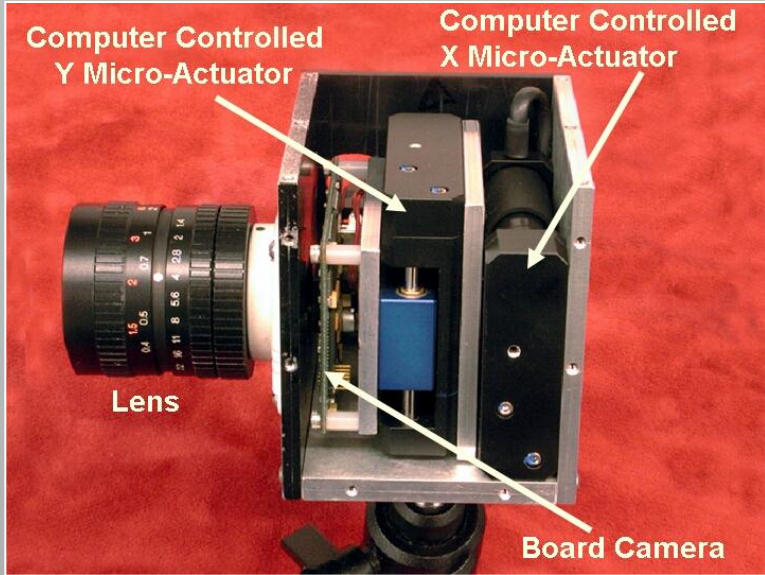
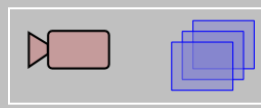


vibrating sensor



[Landolt et al. 01]

Jitter Camera



RAW Jitter Camera output

Super-resolved image



83279 50
5028841971
1 6939937510
10 5820974944 592
5923078164 062862085



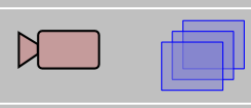
83279 50
5028841971
1 6939937510
10 5820974944 592
5923078164 062862085

1415926535 8979323846
1415926535 8979323846 264338
1415926535 8979323846 2643383279 50
1415926535 8979323846 2643383279 5028841971
1415926535 8979323846 2643383279 5028841971 6939937510
1415926535 8979323846 2643383279 5028841971 6939937510 5820974944 592
1415926535 8979323846 2643383279 5028841971 6939937510 5820974944 592 5923078164 062862085

1415926535 8979323846
1415926535 8979323846 264338
1415926535 8979323846 2643383279 50
1415926535 8979323846 2643383279 5028841971
1415926535 8979323846 2643383279 5028841971 6939937510
1415926535 8979323846 2643383279 5028841971 6939937510 5820974944 592
1415926535 8979323846 2643383279 5028841971 6939937510 5820974944 592 5923078164 062862085

[Ben-Ezra et al. 05]

Superresolution with Aperture Masks



Aperture mask L_0 and corresponding photo

Aperture mask L_1 and corresponding photo

Aperture mask L_2 and corresponding photo

Aperture mask L_3 and corresponding photo

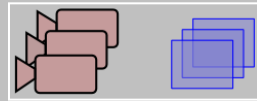
Aperture mask L_2 and corresponding photo

Aperture mask L_3 and corresponding photo

Result: Combination of the photos captured with different aperture masks to get a higher resolution image with reduced aliasing.

[Mohan et al. 08]

Space-Time Superresolution



- Increase resolution in space & time from multiple low-resolution videos

3 of 18 low-temporal input sequences



temporal super-resolution



close-ups



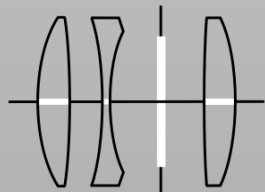
[Shechtman et al. 05]

V.II Optical Field Correction

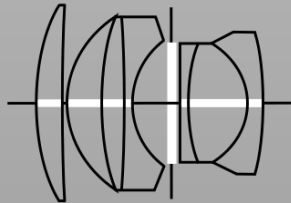
What is Field Curvature?

- Focal surface of most lenses is not planar
- Lots of engineering effort!

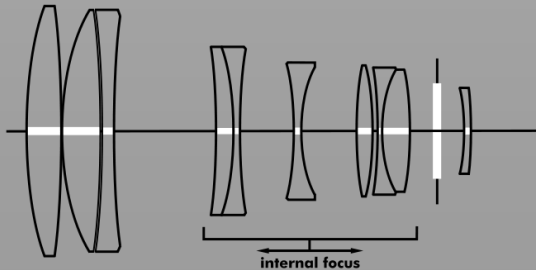
http://en.wikipedia.org/wiki/Camera_lens



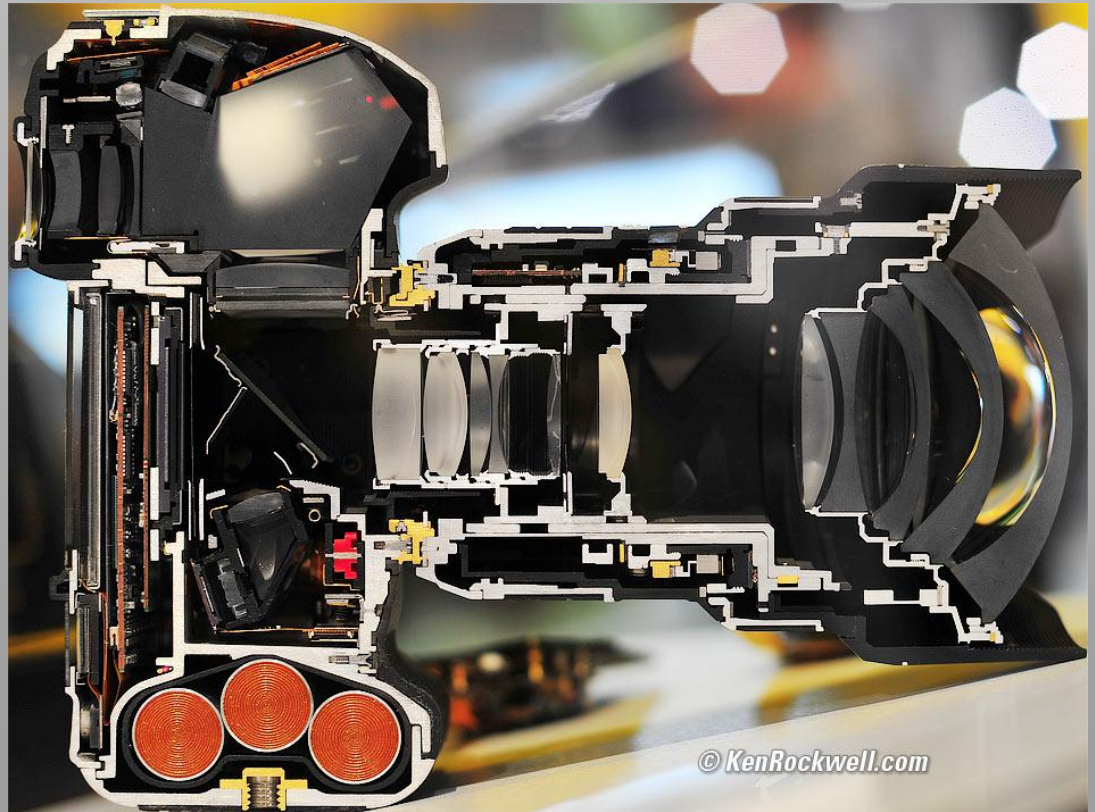
Taylor, Taylor & Hobson
Cooke Triplet
1893



Zeiss Sonnar 50mm f/1.5, 1932



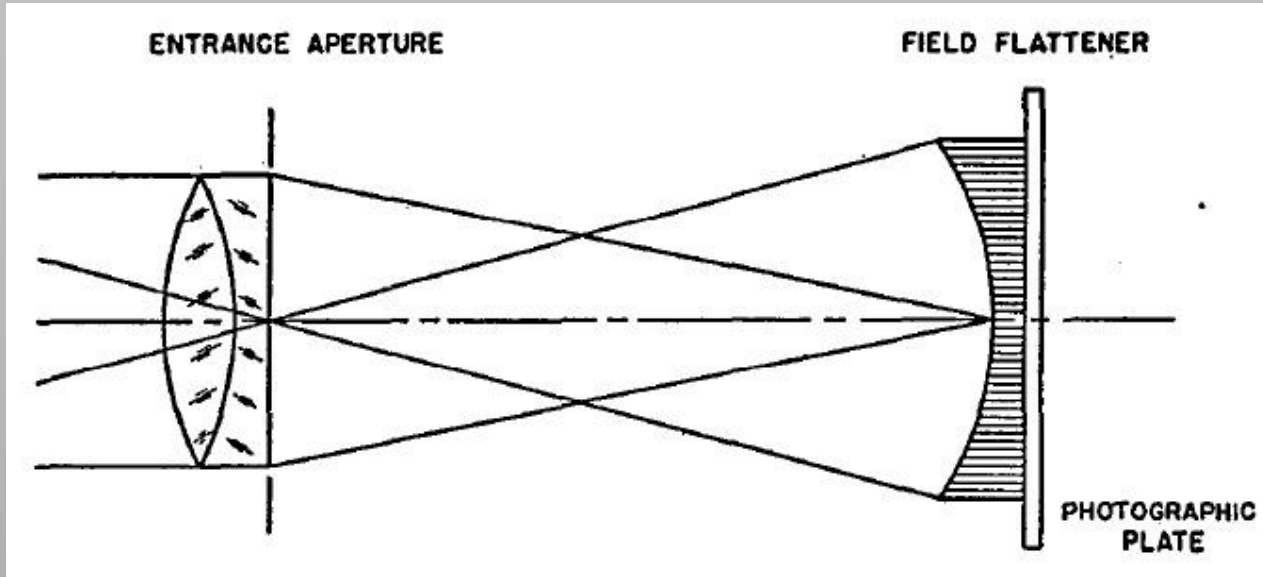
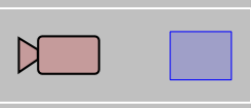
Nippon Kogaku Nikkor 200mm f/2 ED IF, 1977



© KenRockwell.com

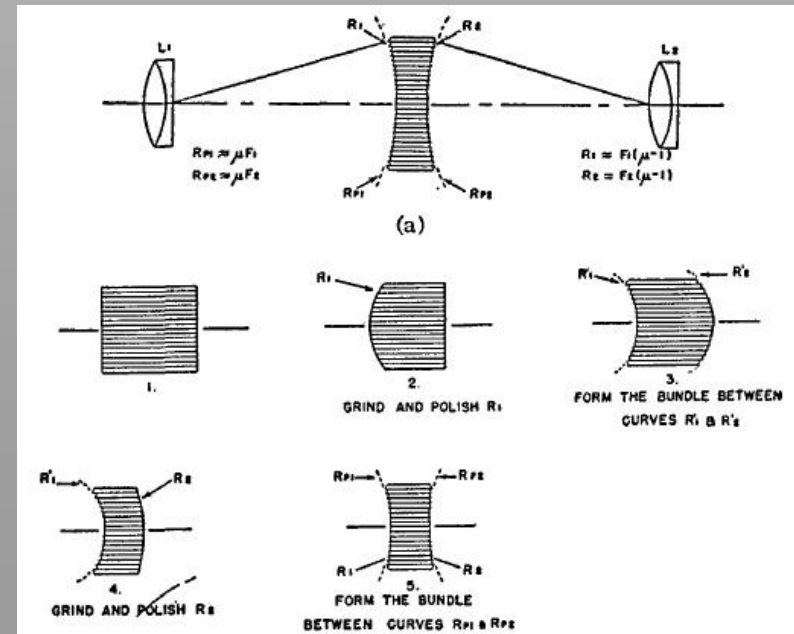
www.kenrockwell.com

Fiber Optics



[Kapany & Hopkins 57]

- Aligned fiber bundle for reshaping curvature
- Limited resolution...

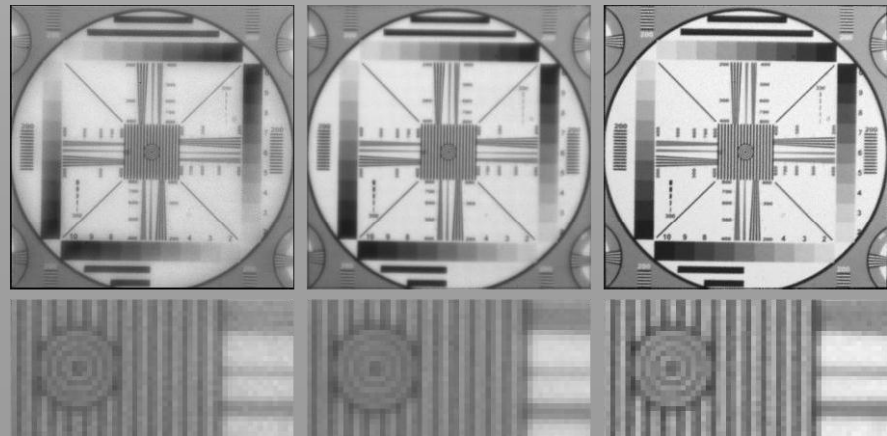
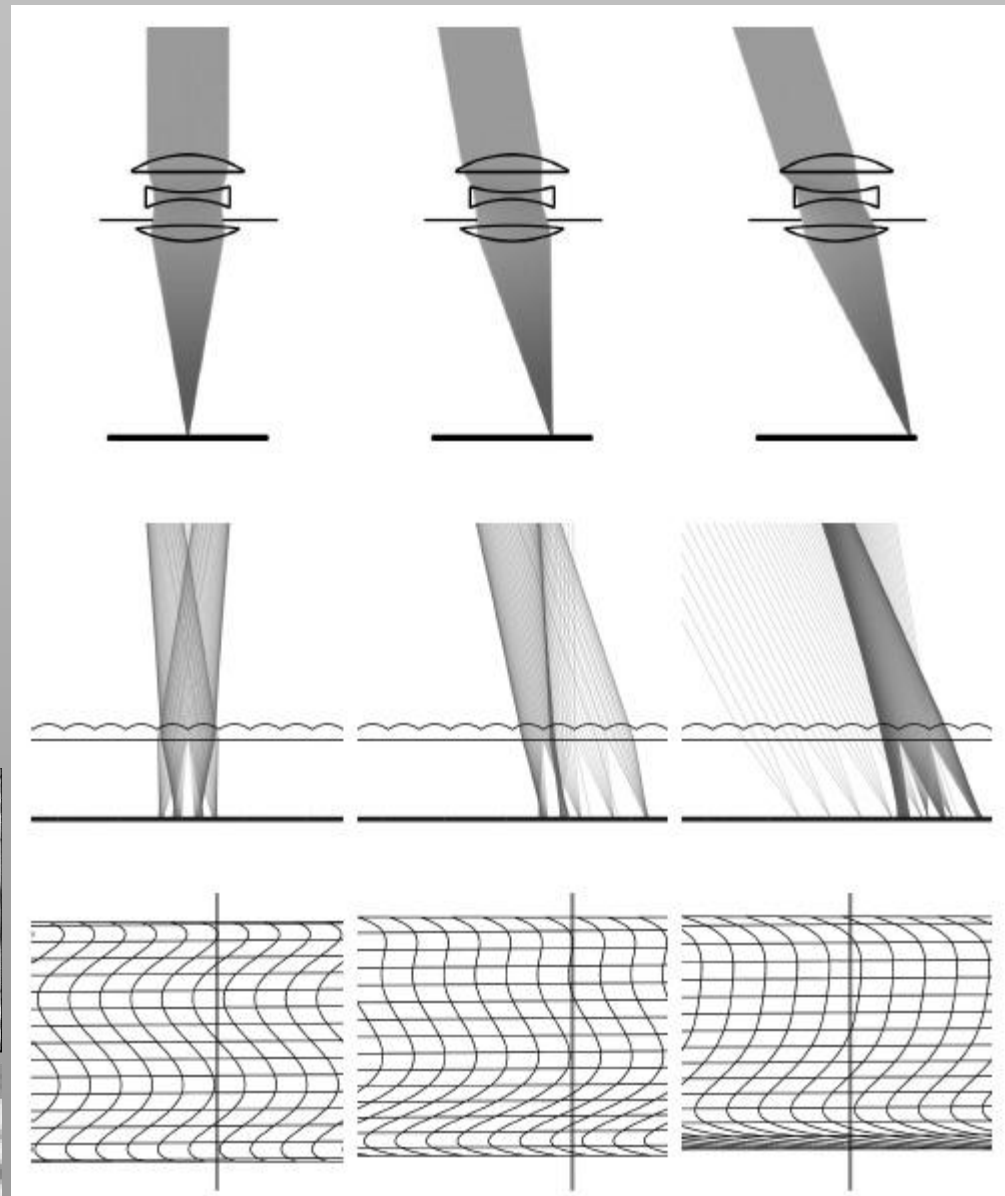


Aberration Correction with Light Field Camera

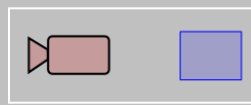


Digital curvature correction
by resorting rays of
4D light field

[Ng & Hanrahan 06]

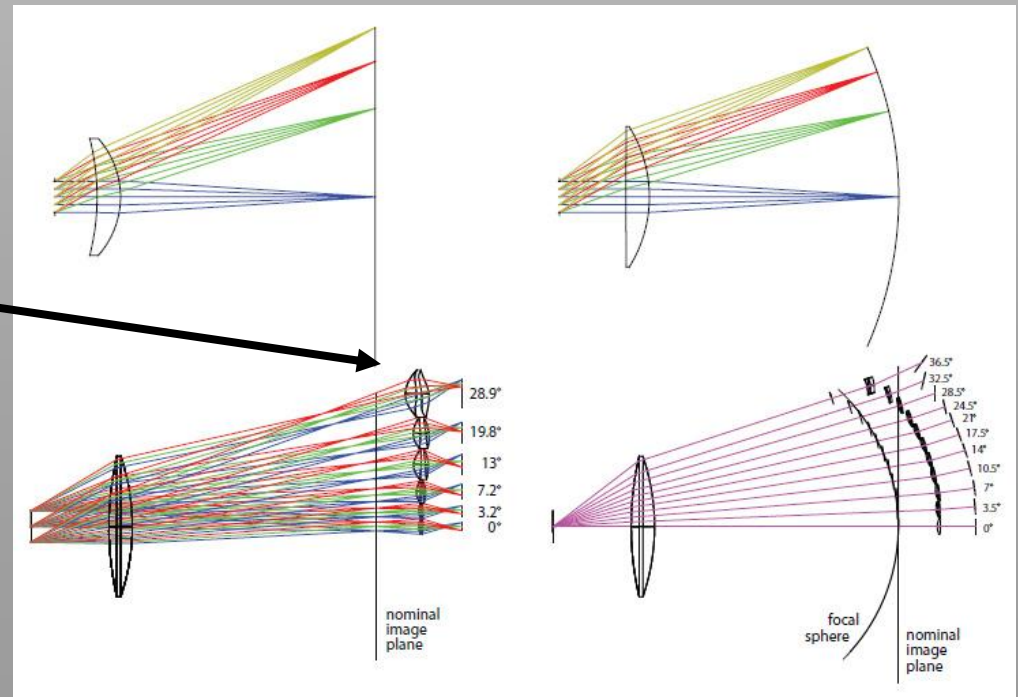


Multiscale Lens Design



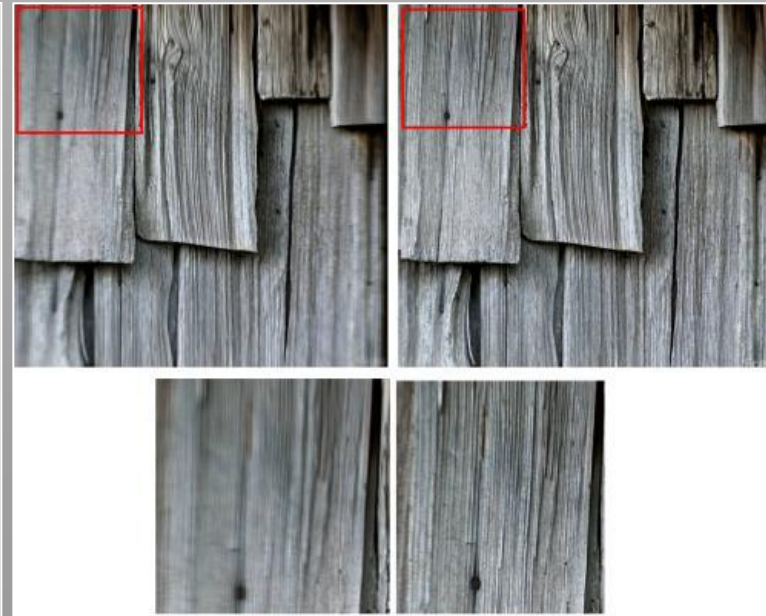
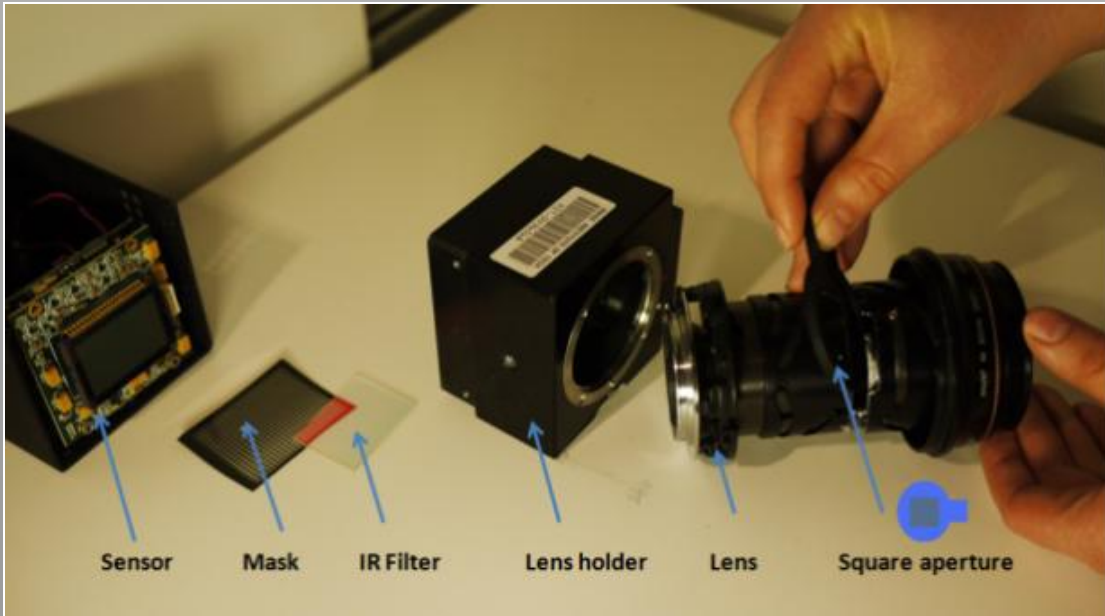
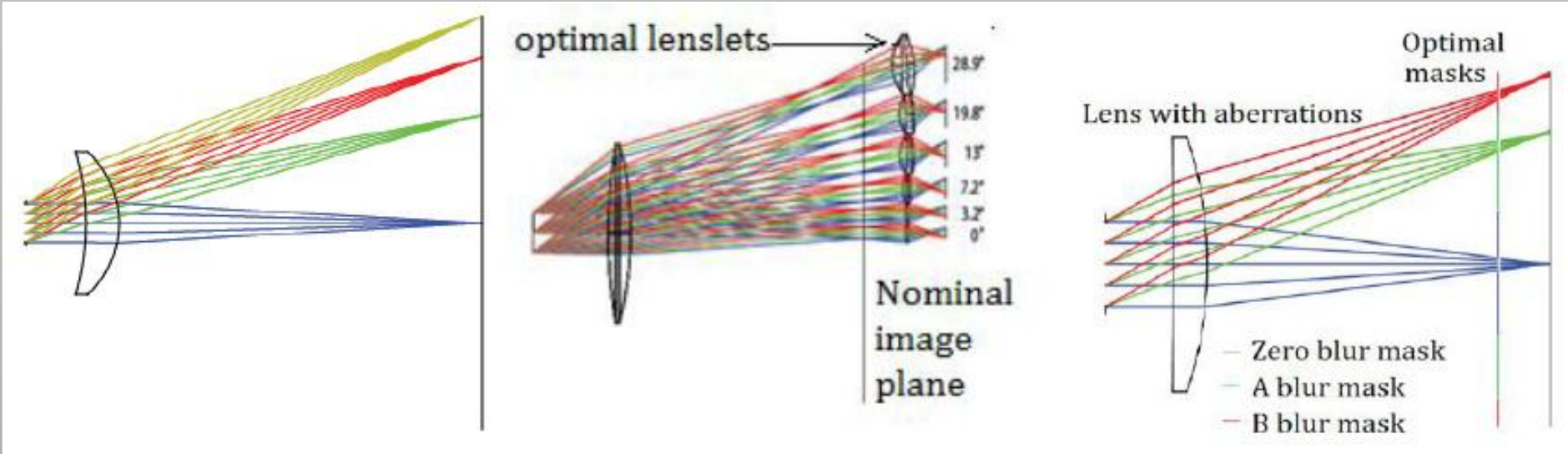
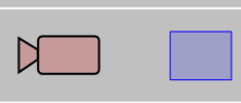
- Optimized multi-layered, multi-scale optics
- Joint optical encoding and computational processing
- Extended FOV & DOF, super-resolution, smaller form factor, lower cost

Locally optimal masks



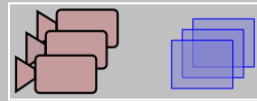
[Brady & Hagen 09]

Coded Attenuation Masks

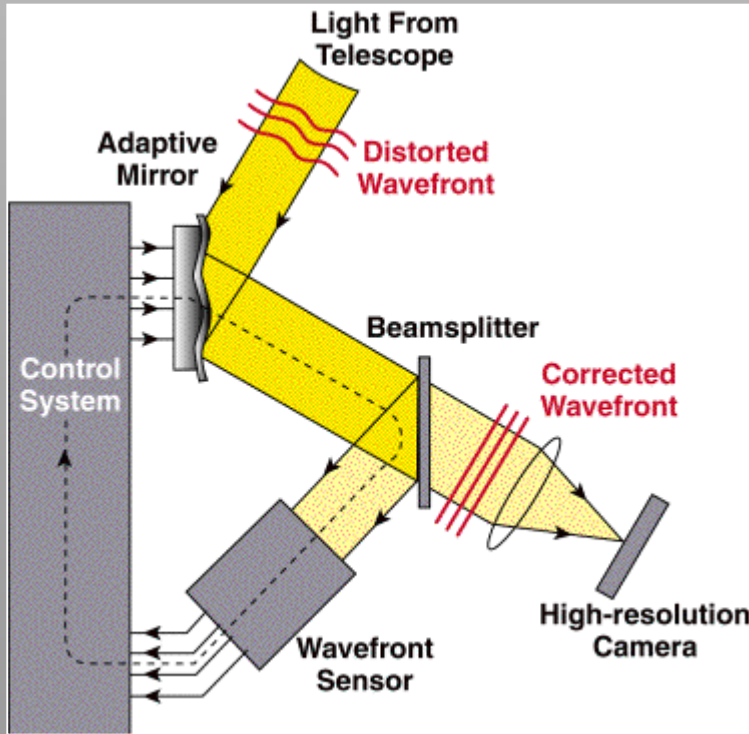


[Prandharkar et al. 10]

Adaptive Optics

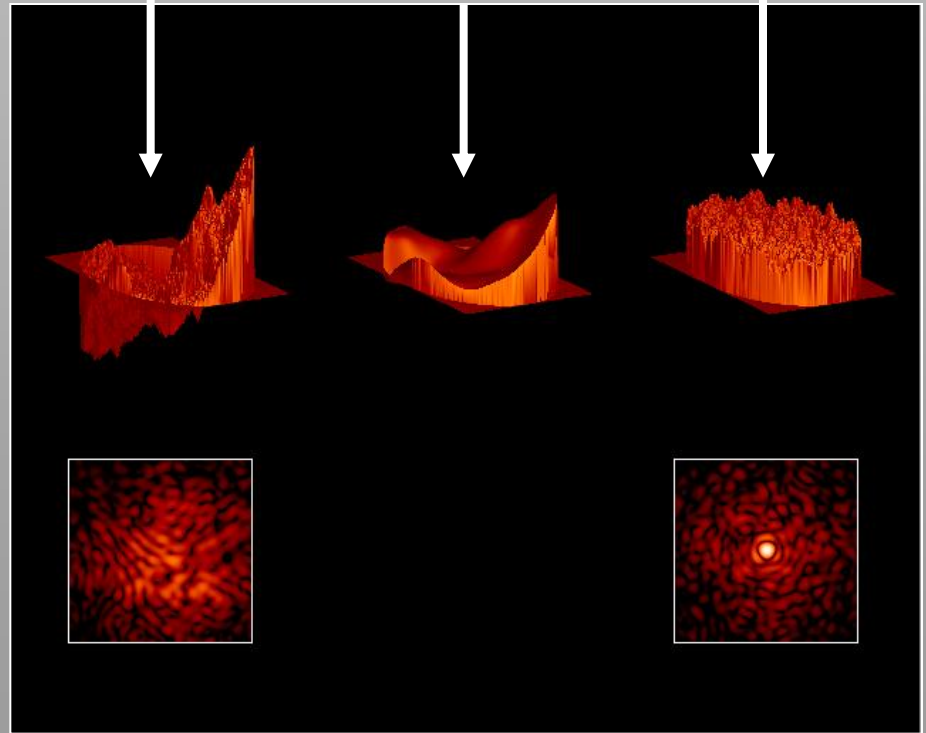


- Deformable mirrors to correct wavefront
- Used in astronomy and retinal imaging



C. Max, Center for Adaptive Optics

uncorrected wavefront corrected wavefront
deformable mirror



www.lyot.org

V.III Extended Depth of Field Imaging

Depth of Field

- Defocus blur is depth-dependent
- Can be described as convolution, where
 - Kernel size is related to depth
 - Kernel shape is that of aperture

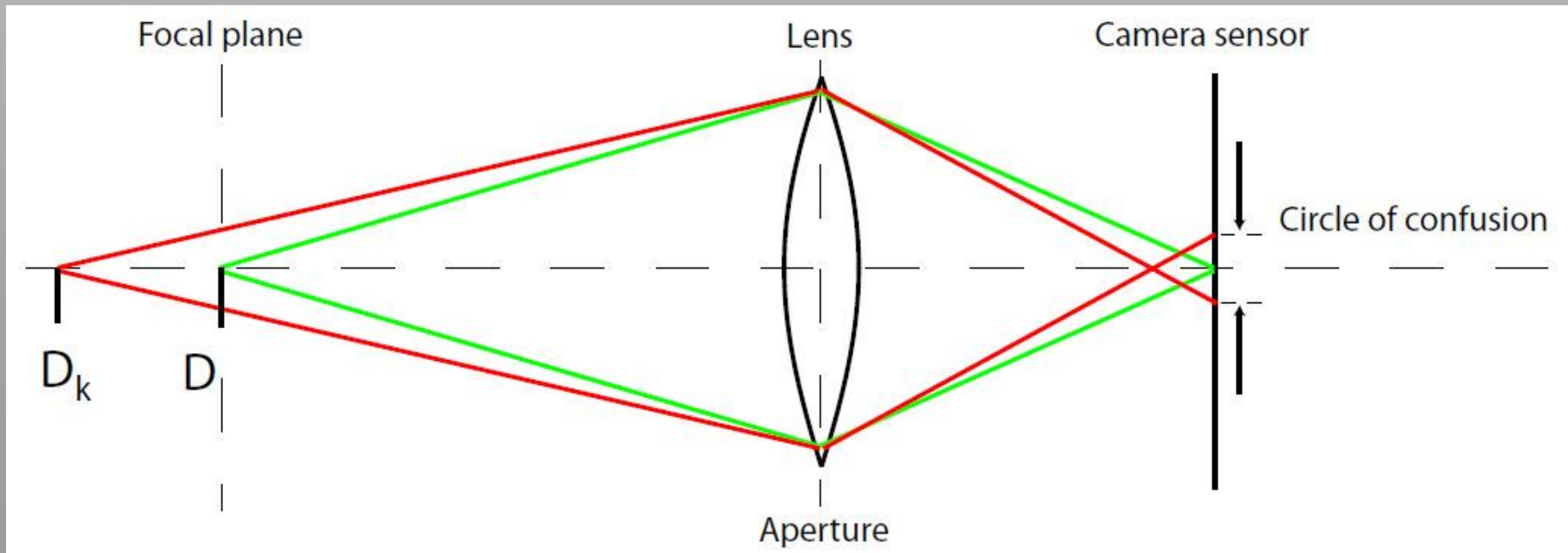


Image from [Levin et al. 07]

Deconvolution is Hard

- Low



(a) Captured image



(b) Richardson-Lucy



(c) Gaussian prior



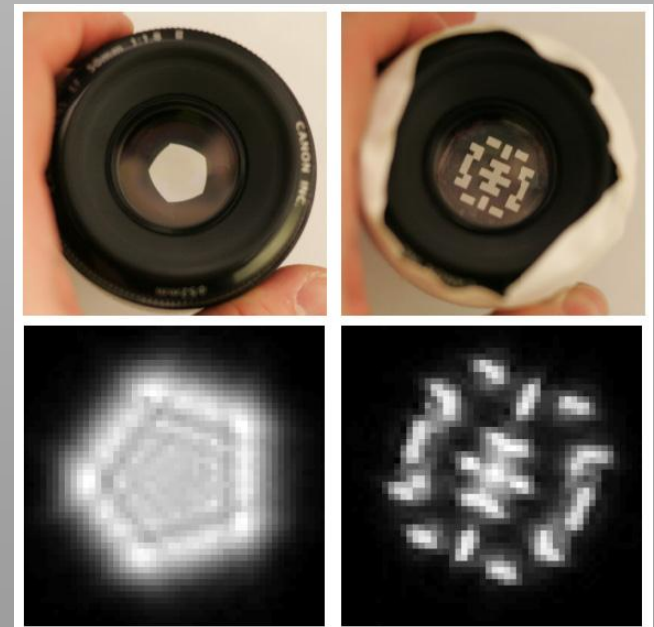
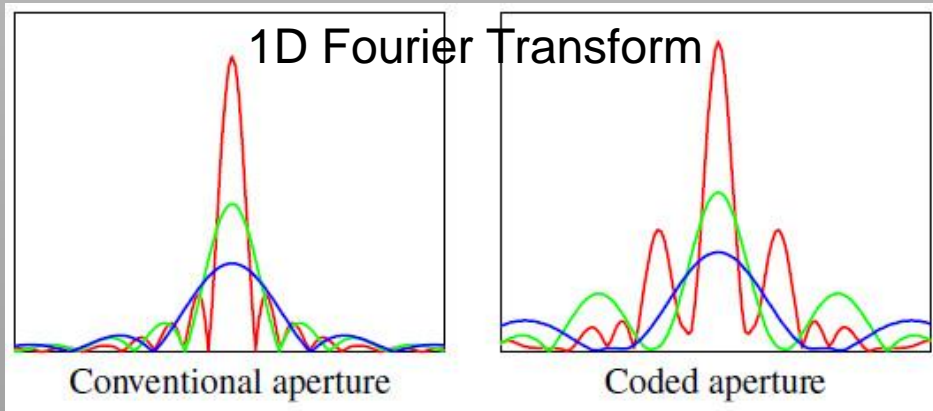
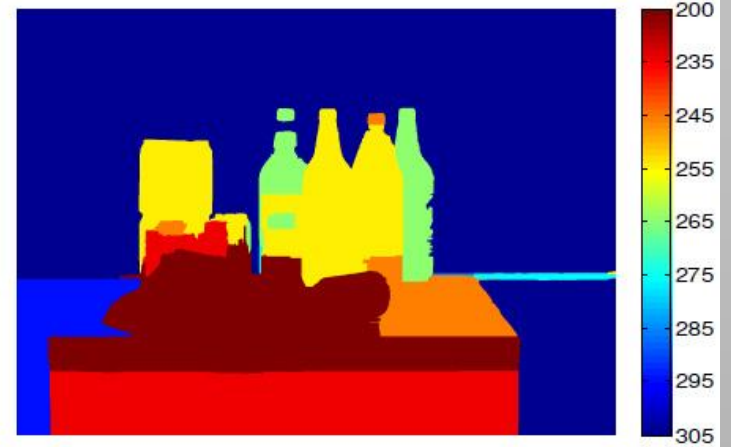
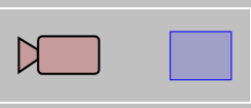
(d) Sparsity prior

e
je

Approaches to Extend DOF

- Make PSF invertible \rightarrow deconvolution becomes well-posed
- Make PSF depth-invariant \rightarrow shift-invariant deconvolution
- Capture multiple images with different focus settings (no deconvolution)

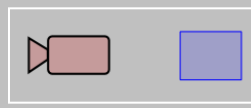
Coded Apertures – Attenuation Masks



- Depth from coded aperture
- Spatially varying deconvolution

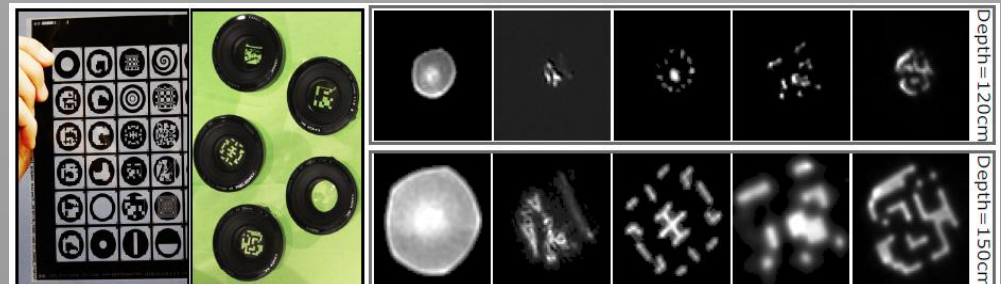
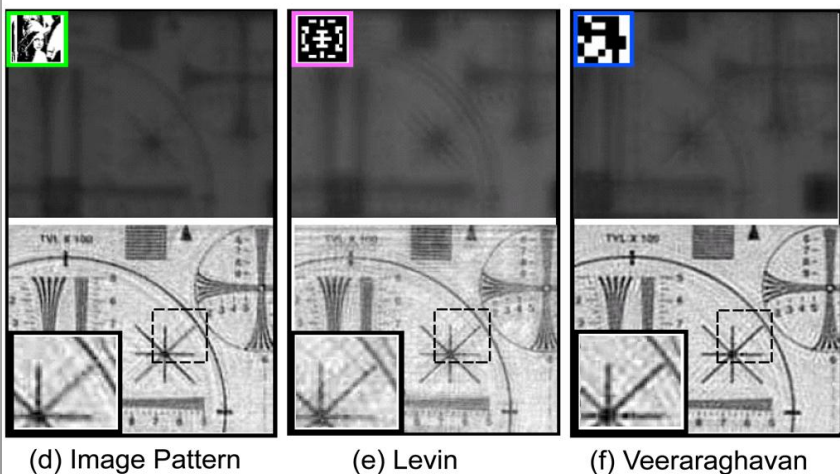
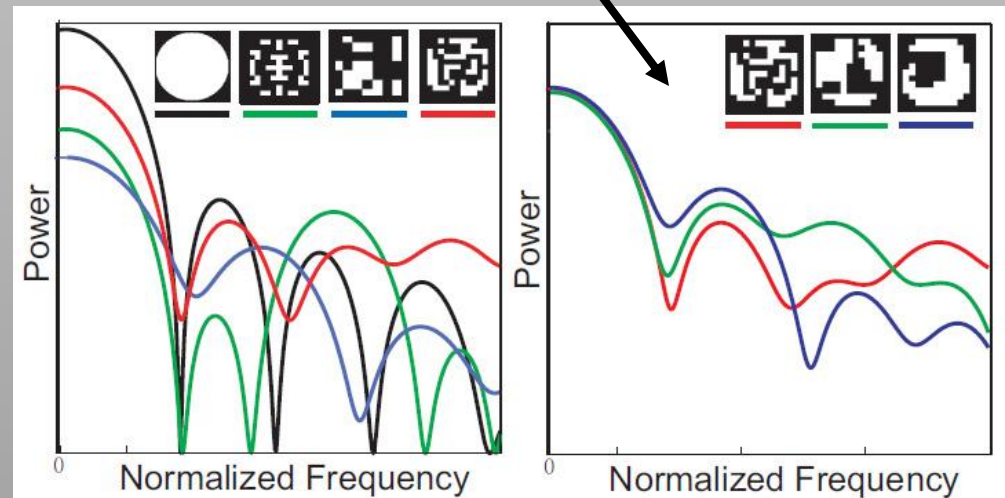
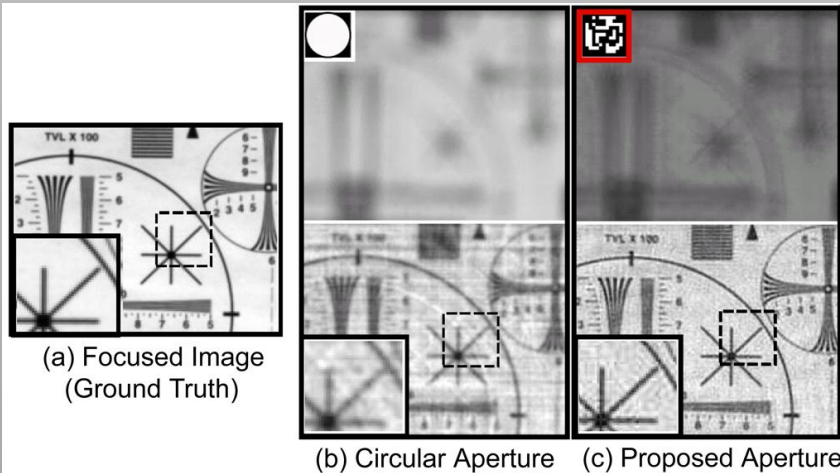
[Levin et al. 07]

Coded Apertures – Analysis

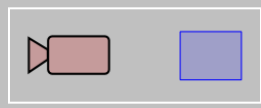


- Analysis of aperture patterns with respect to camera noise

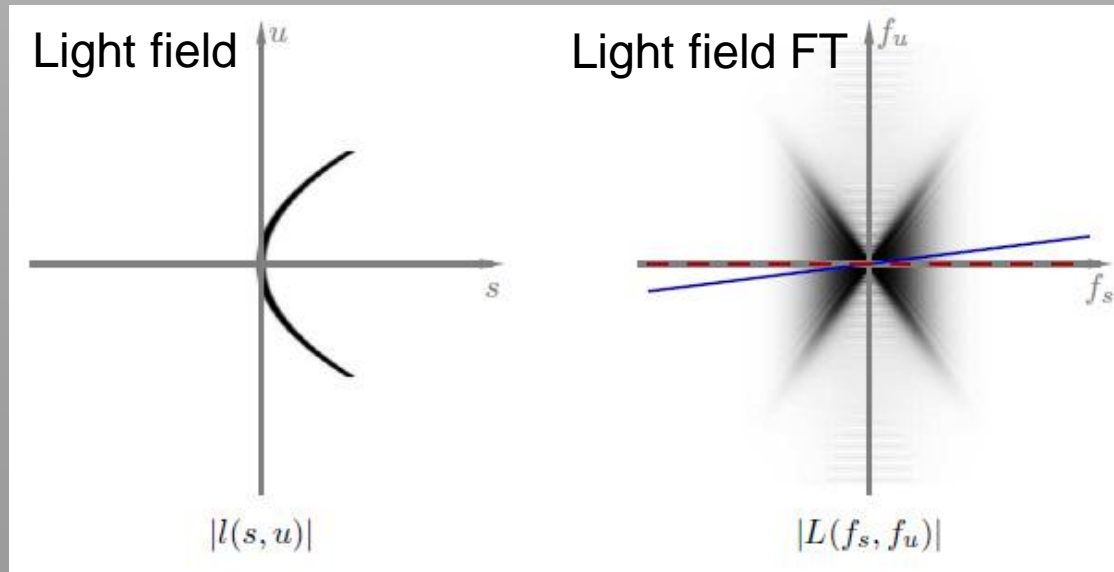
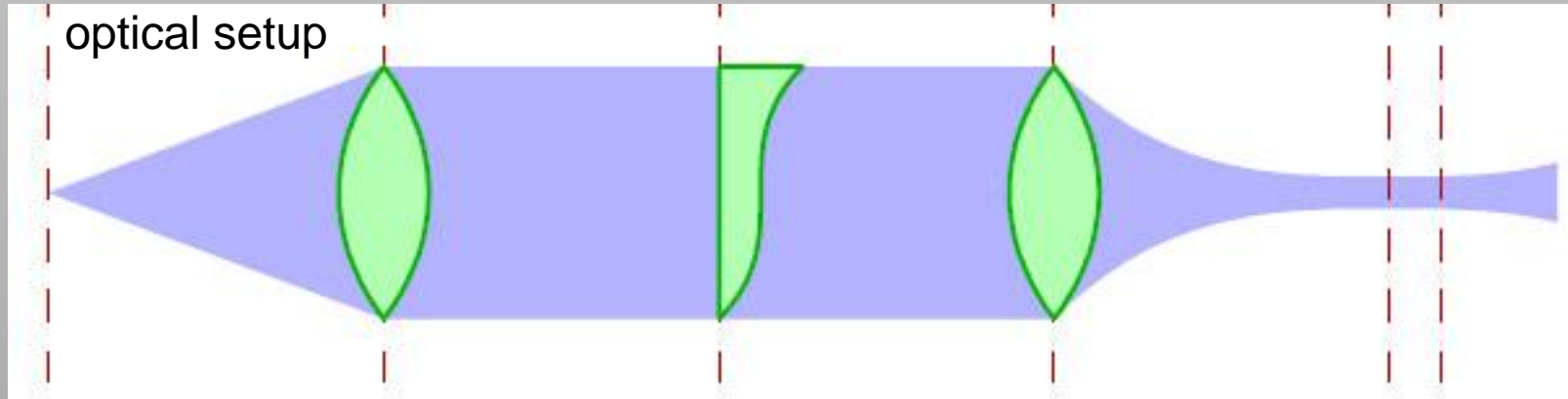
Optimized patterns for increasing noise



Coded Apertures – Cubic Phase Plate



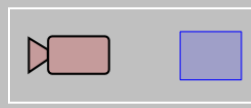
- (almost) depth-independent PSF



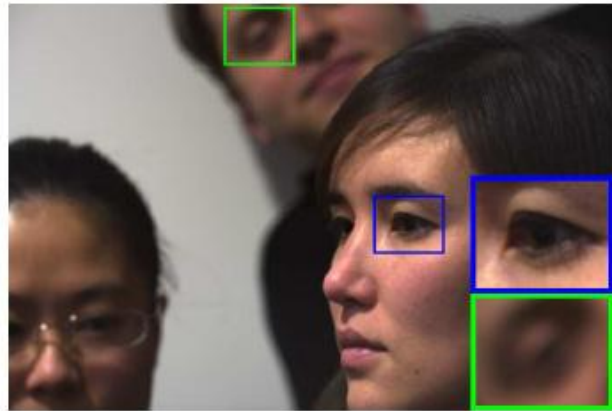
[Dowski & Cathey 97]

Images from [Zhang & Levoy 09]

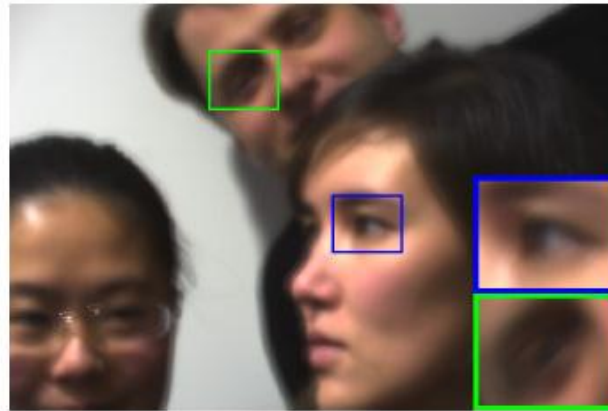
Coded Apertures – Lattice Focal Lens



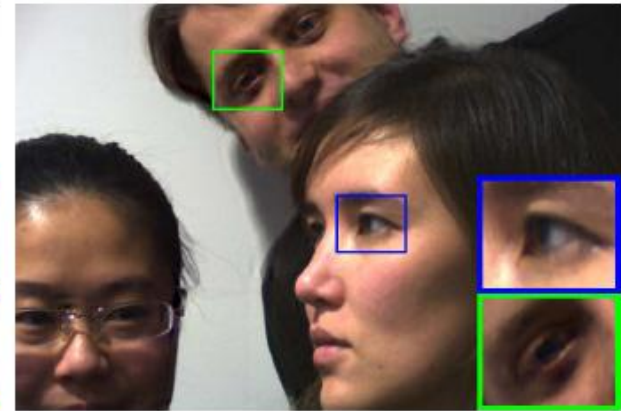
- 4D is different than 2D – dimensionality gap



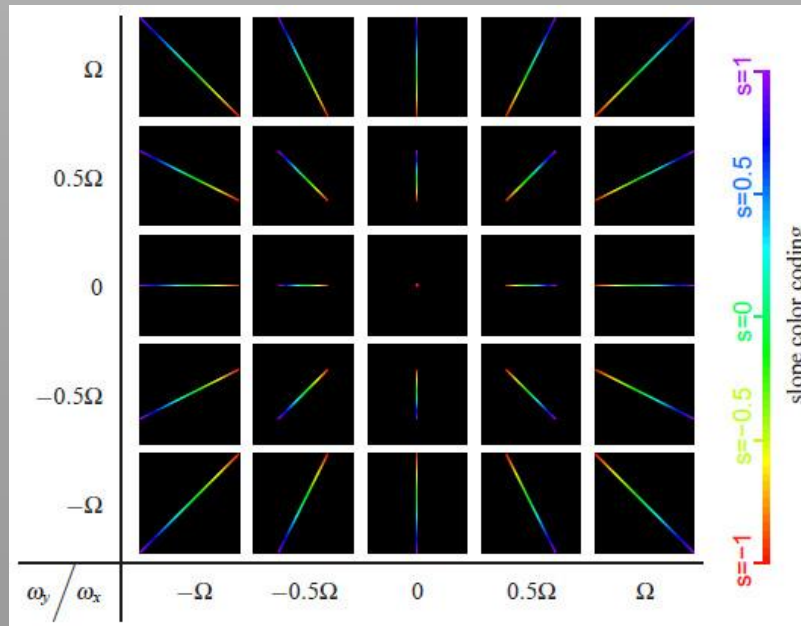
Standard lens image



Our lattice-focal lens: input

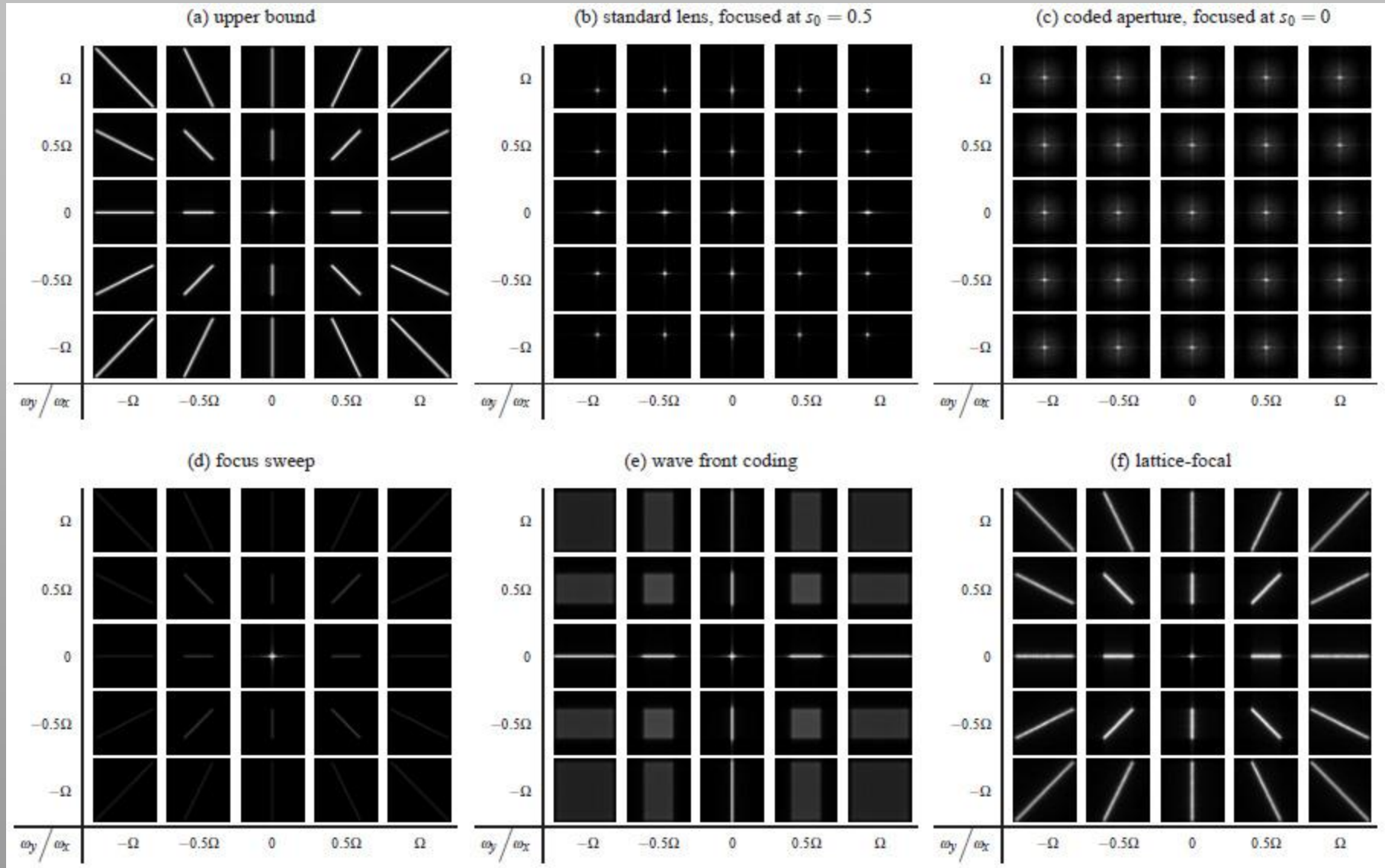
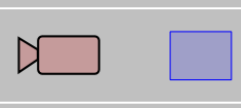


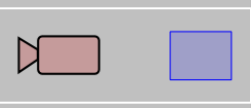
Lattice-focal lens: all-focused output



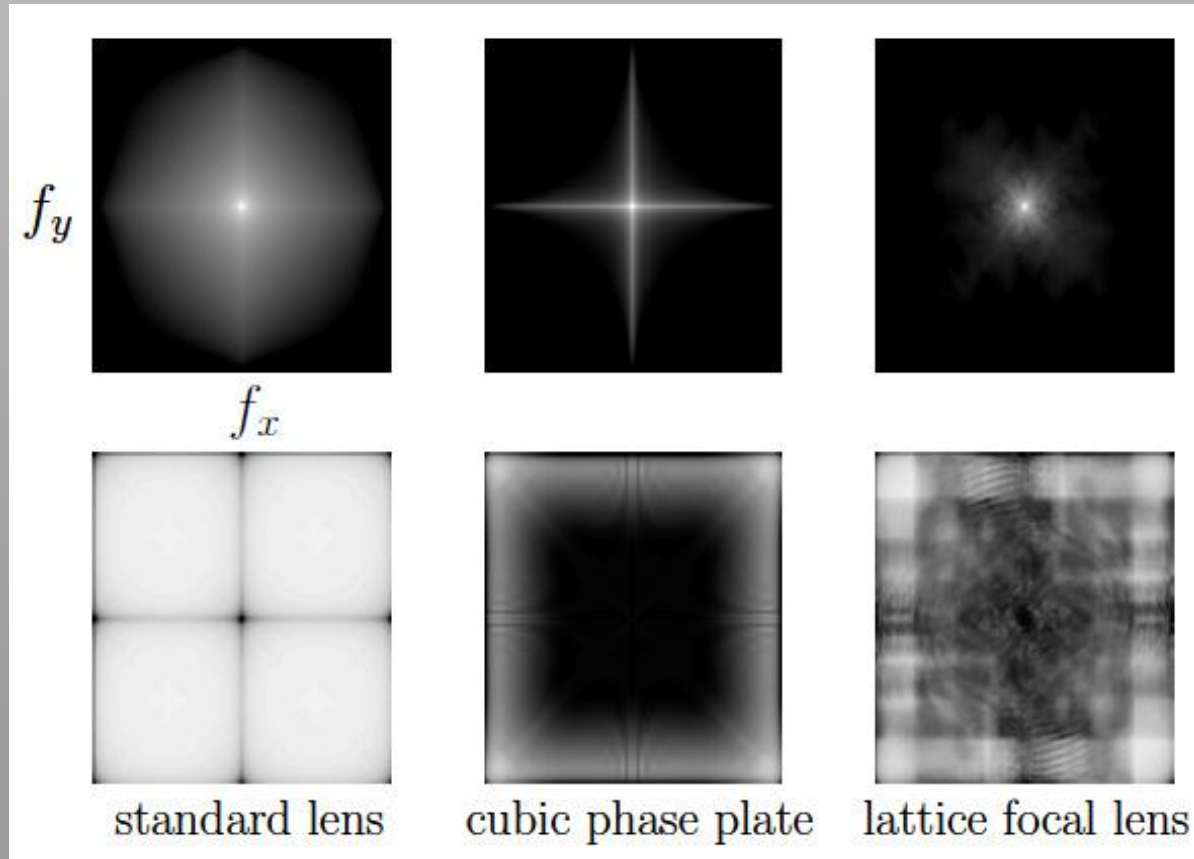
[Levin et al. 09]

Coded Apertures – Frequency Analysis

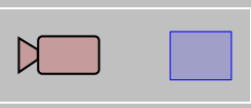




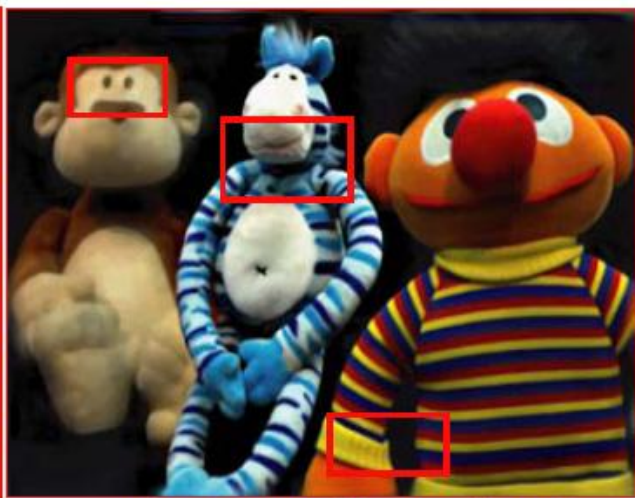
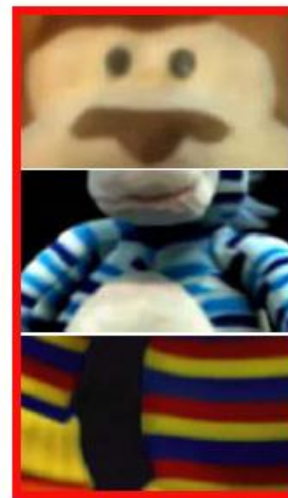
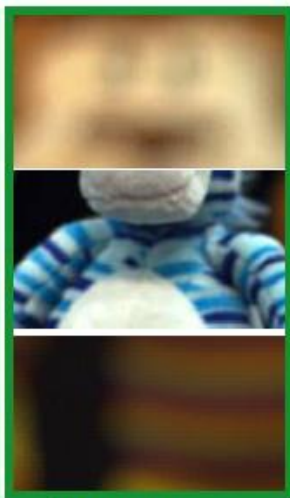
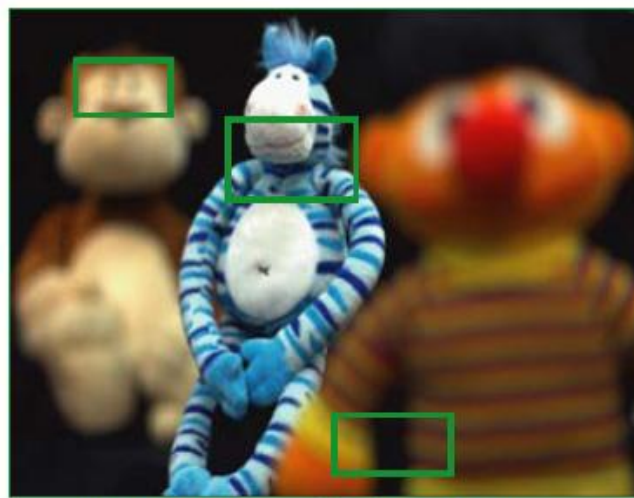
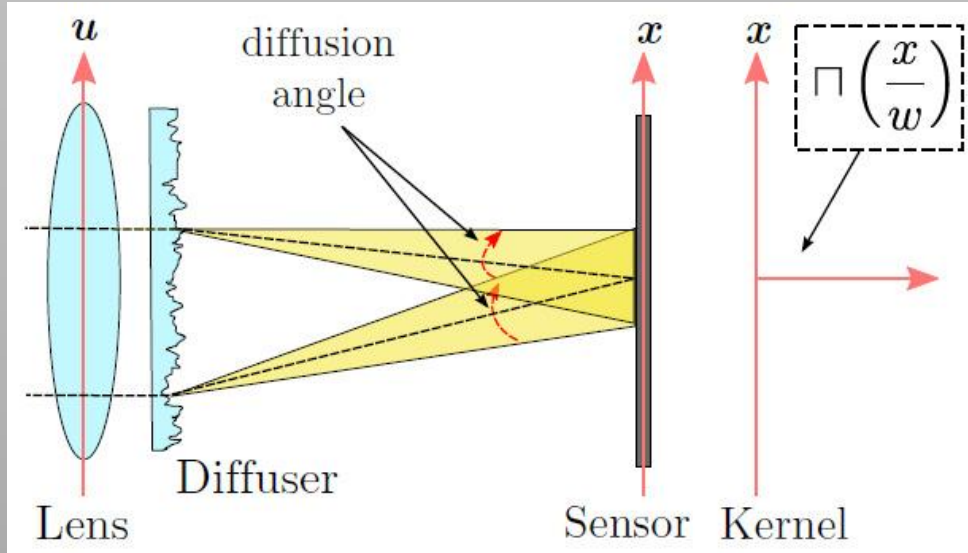
- Analysis of computational cameras with respect to **depth-invariance** and **transfer efficiency**



Coded Apertures – Diffusion Masks



- depth-invariance and invertibility



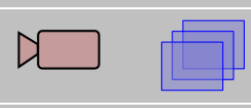
(a) An image captured with a F/1.8 lens

(b) Detail from (a)

(c) Detail from (d)

(d) A diffusion coded image (deblurred)

Focal Stacks



- Capture multiple images with different focus settings
- Merge into single extended DOF image



One of 28 Image of a Focal Stack
(Focused on Ball 5)



One of 28 Image of a Focal Stack
(Focused on Ball 7)



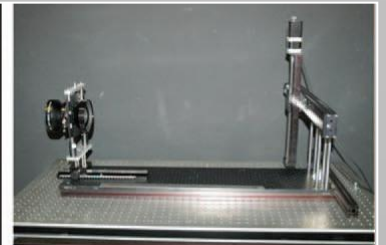
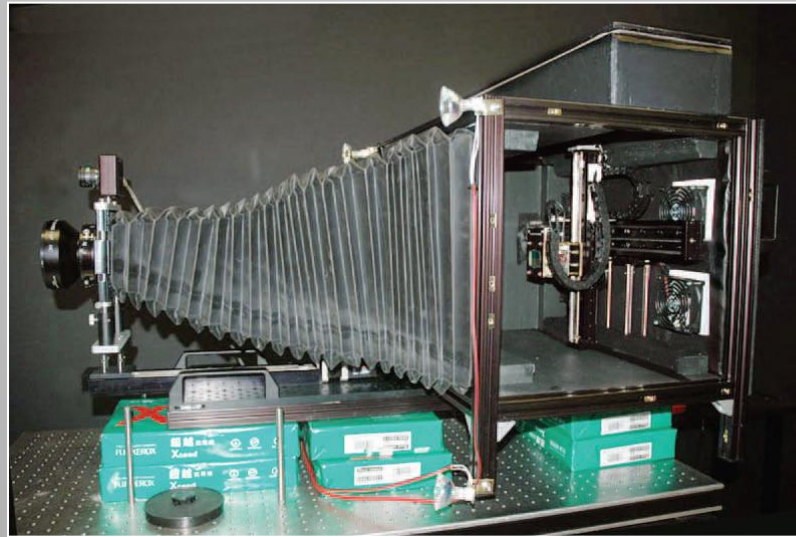
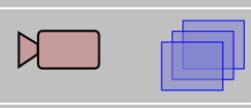
One of 28 Image of a Focal Stack
(Focused on the Potted Plants in the Back)



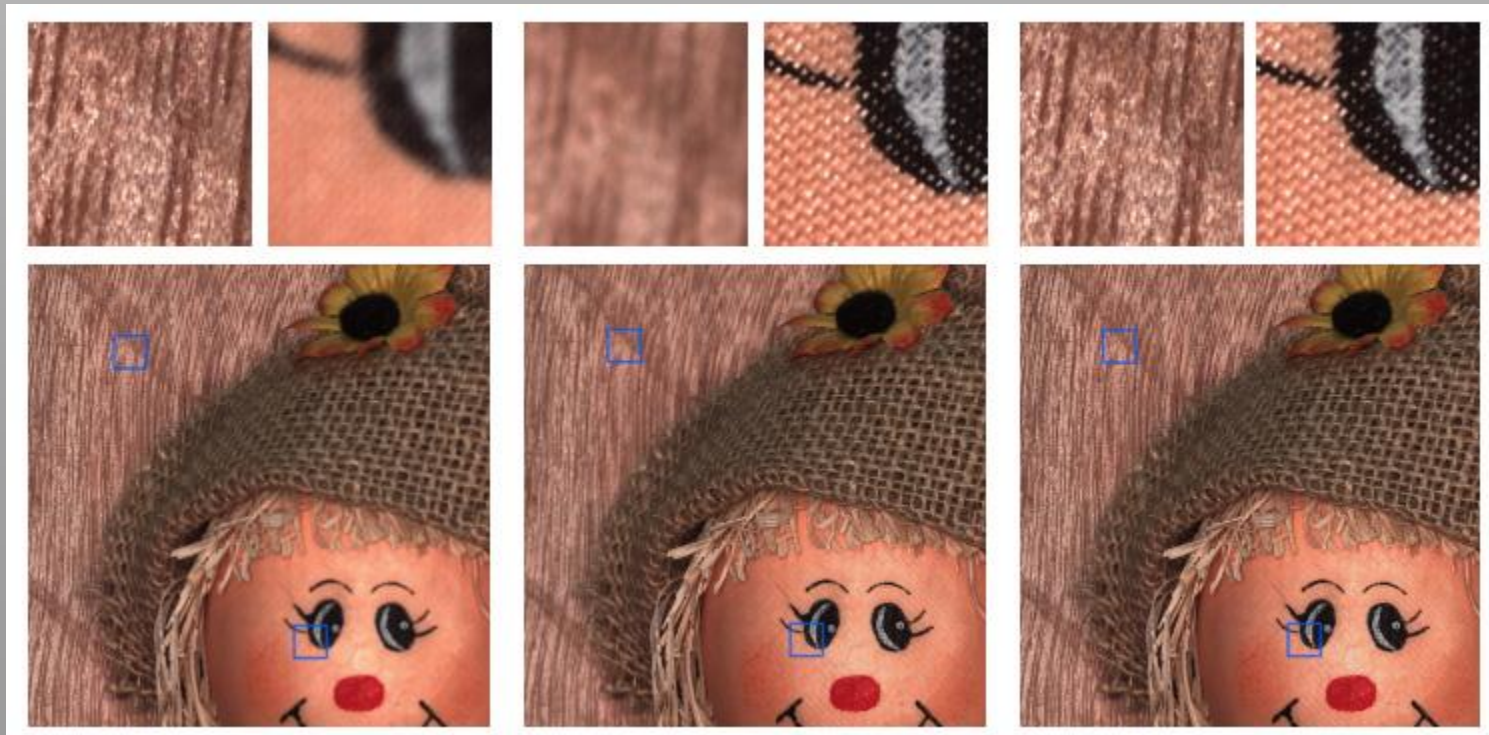
All-Focused Image Computed from the Focal Stack

[Nagahara et al. 08]

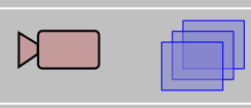
Gigapixel Focal Stack



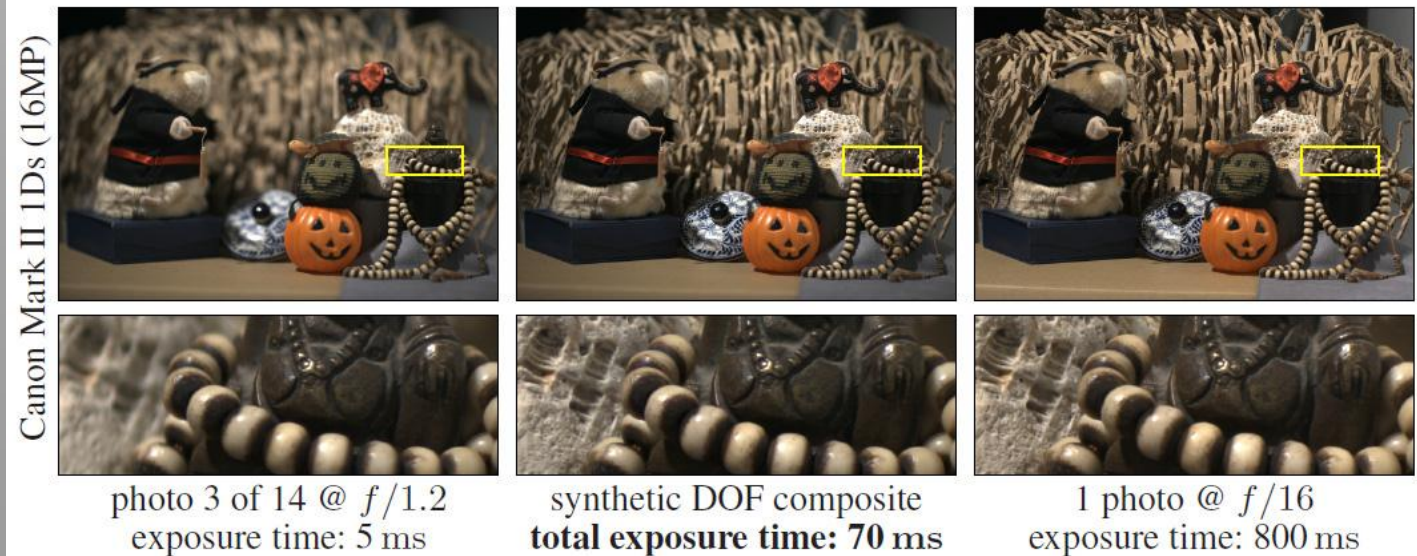
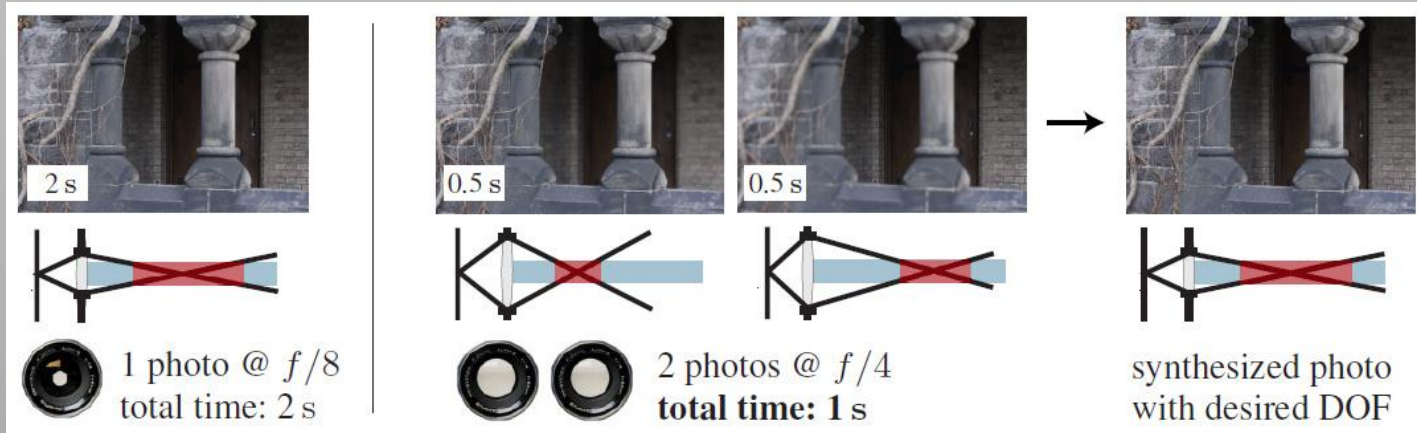
[Ben-Ezra 11]



Light Efficient Photography

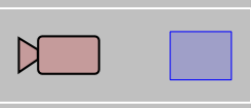


- Efficient selection of focus and aperture settings for desired DOF – less noise

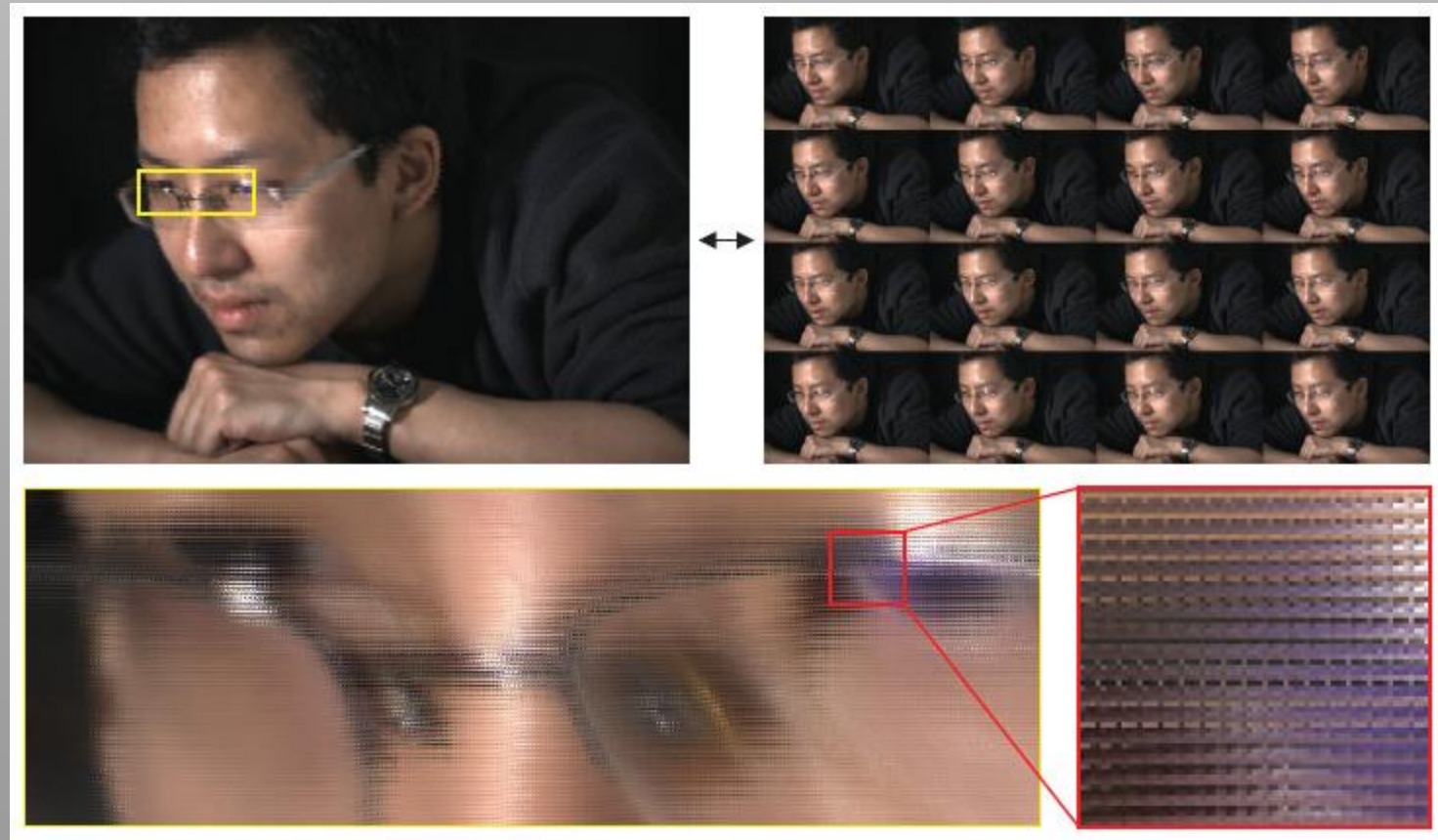


[Hasinoff & Kutulakos 08]

Focal Stack Multiplexing – Single Image

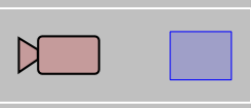


- Multiplex focal stack like Bayer pattern
- Interesting concept – no implementation

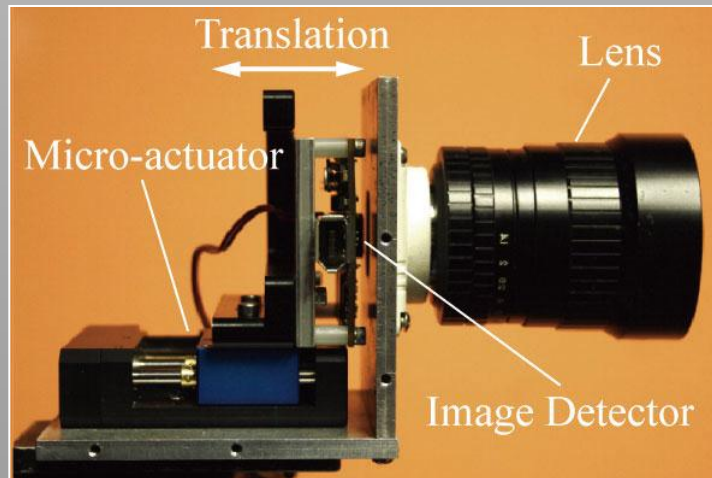


[Kutulakos & Hasinoff 09]

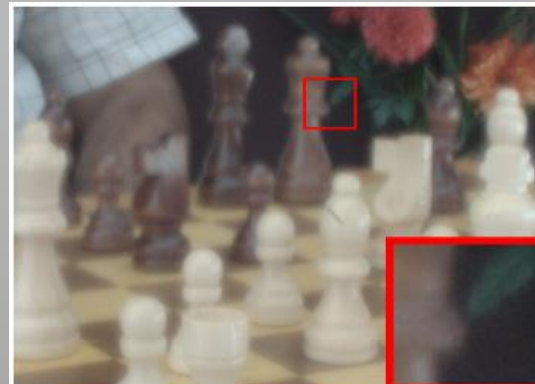
Focal Sweep – Moving Sensor



- Move sensor along optical axis over exposure time
- Movement makes PSF depth-independent



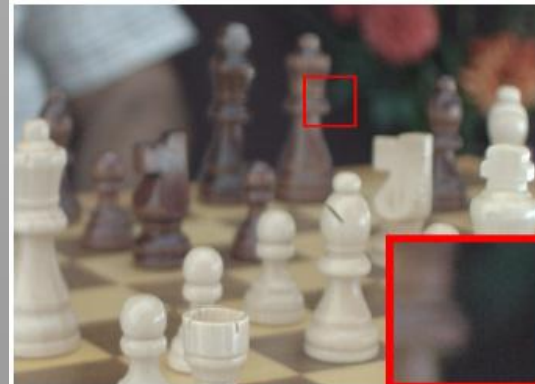
[Nagahara et al. 08]



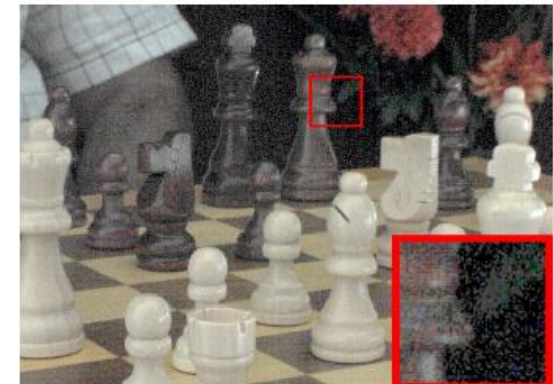
(a) Captured Image ($f/1.4$, $T=0.36\text{sec}$)



(b) Computed EDOF Image

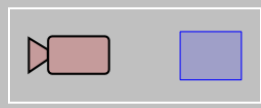


(c) Image from Normal Camera ($f/1.4$, $T=0.36\text{sec}$, Near Focus)

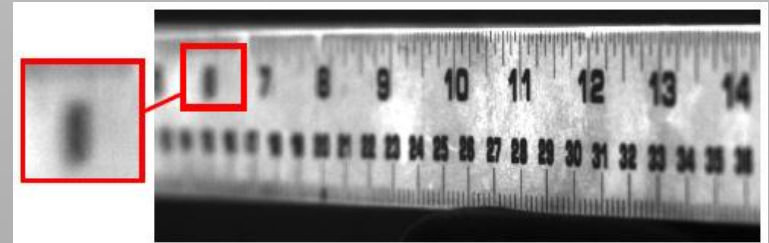
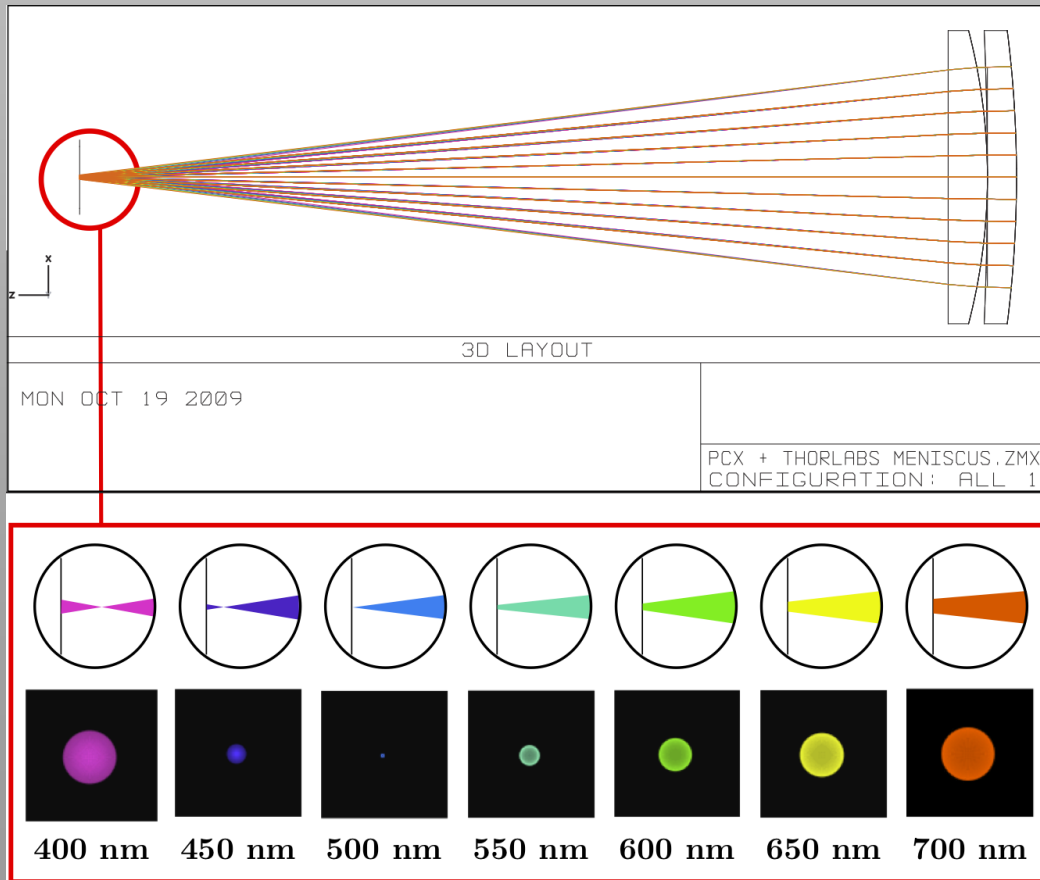


(d) Image from Normal Camera ($f/8$, $T=0.36\text{sec}$, Near Focus) with Scaling

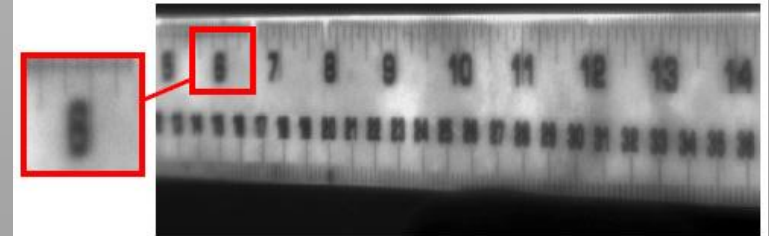
Spectral Focal Sweep



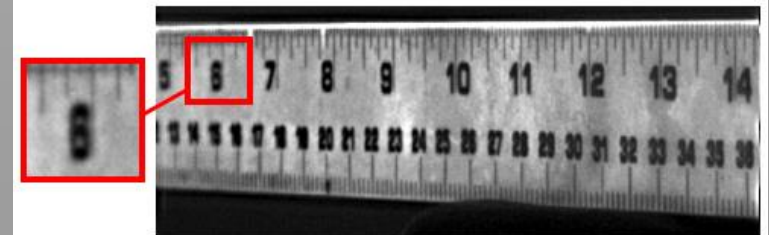
- Use chromatic aberrations
- Multiple focal planes on sensor plane



(a) An image captured with a corrected lens (8ms exposure)

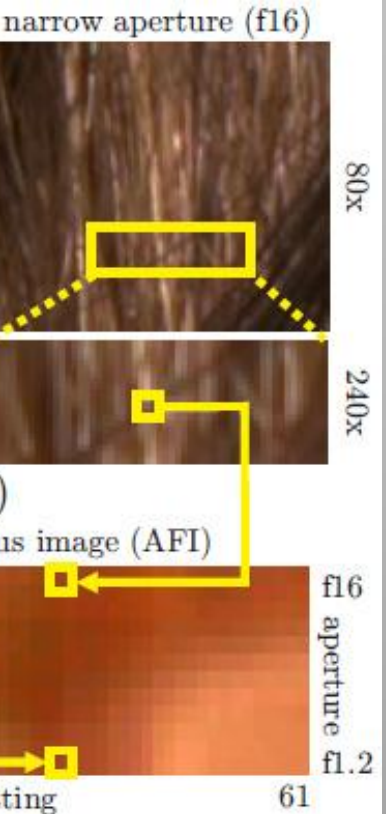
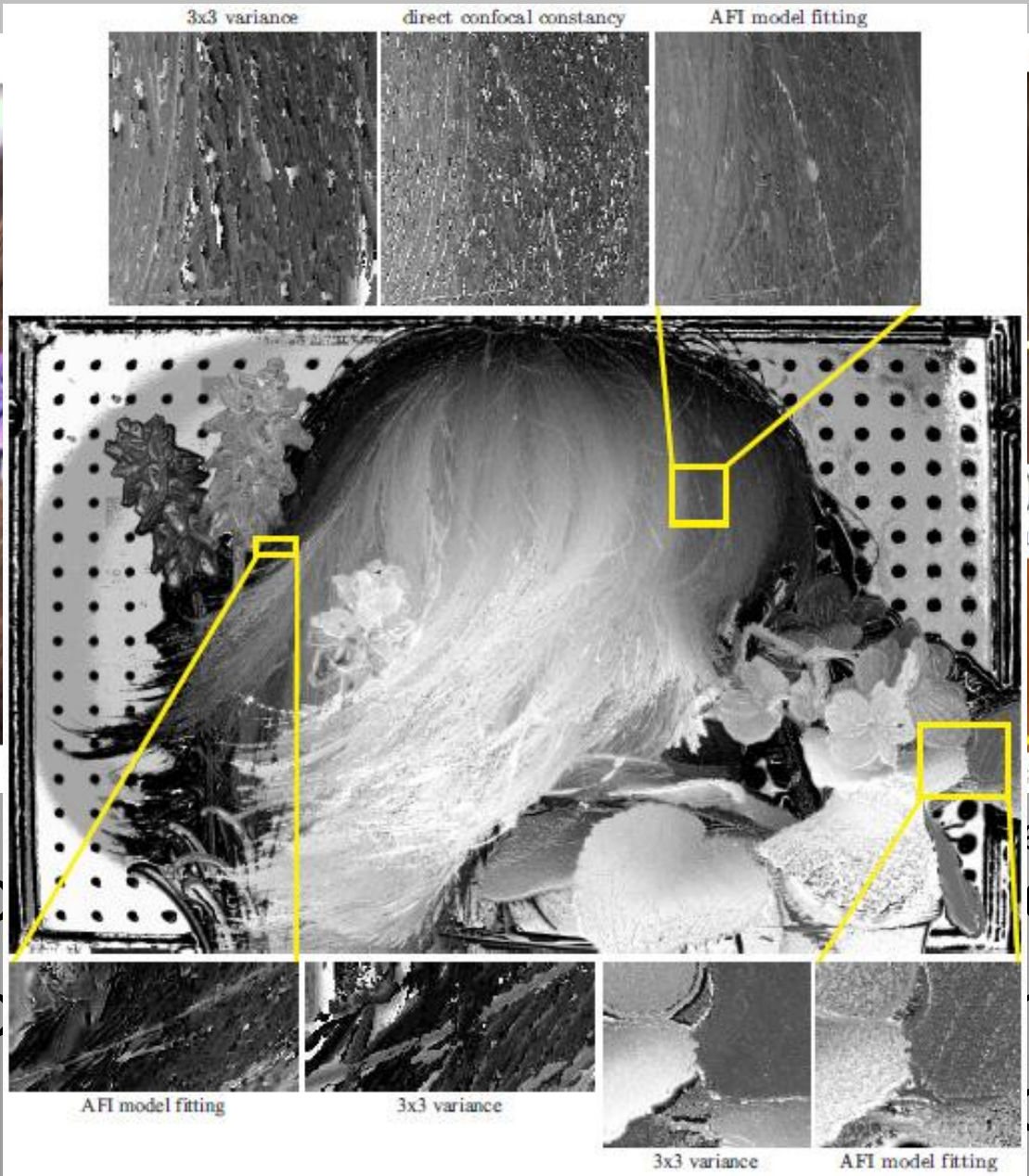
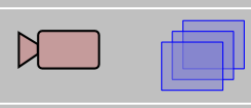


(b) An image captured with a SFS camera (8ms exposure)



(c) The image from Figure 1(b) after deblurring

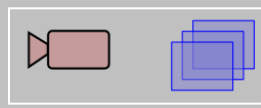
Applications – Confocal Stereo



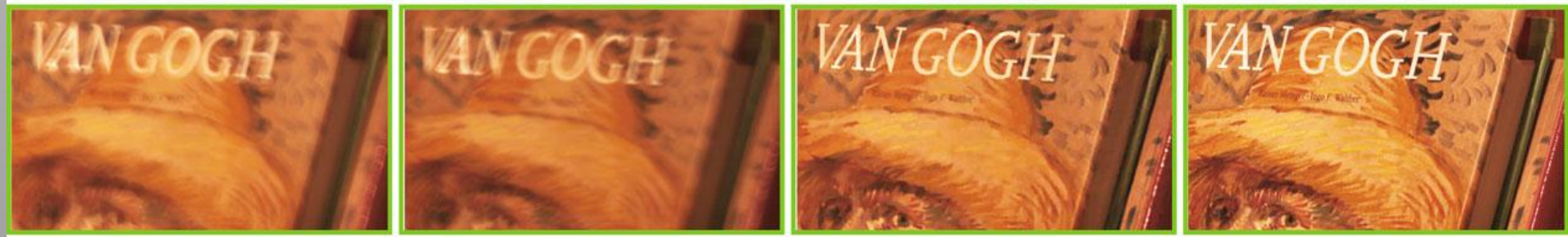
- Multiplexed
- Confocal
- Variational

and Kutulakos 09]
 focus
 intensity
 focus

Applications – Depth from Coded Apertures



(d) Estimated Depth Map



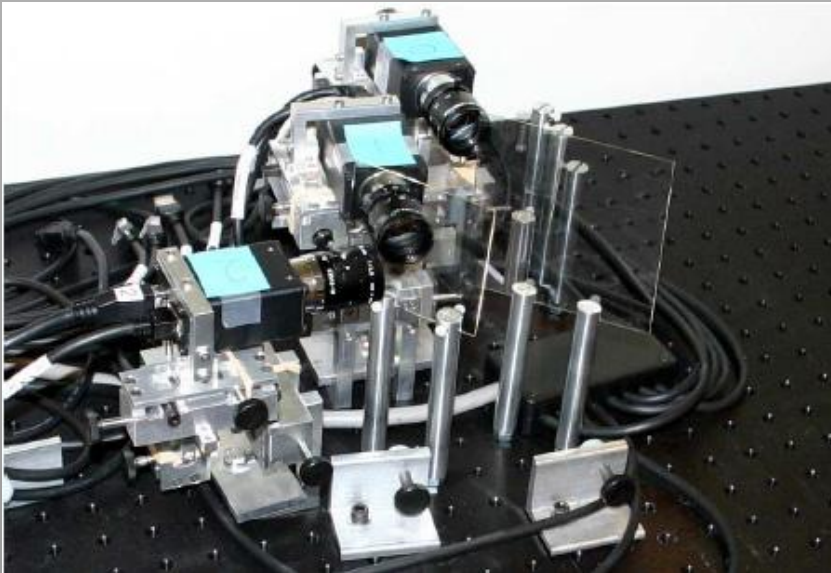
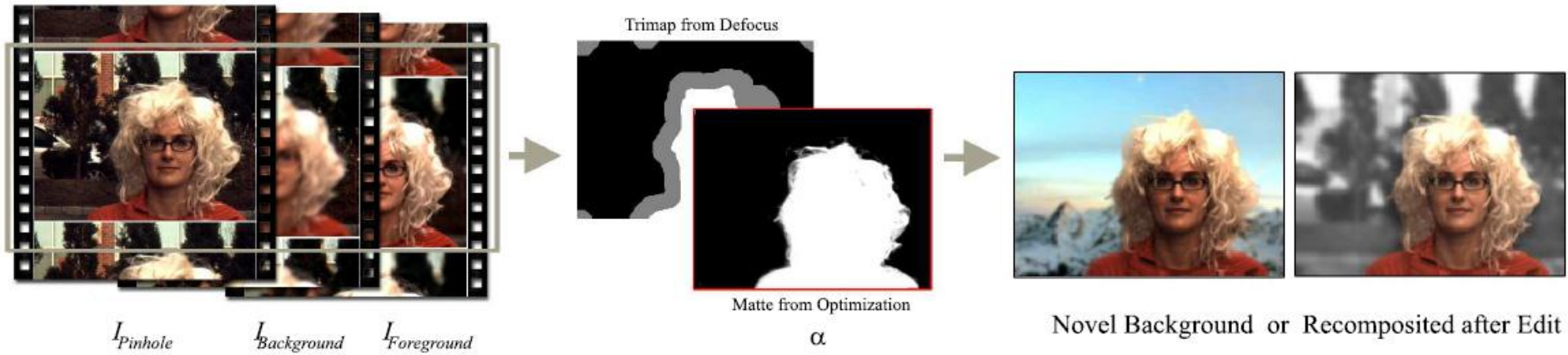
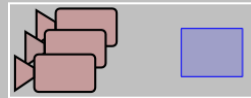
(a) Captured Image 1

(b) Captured Image 2

(c) Recovered All-focused Image

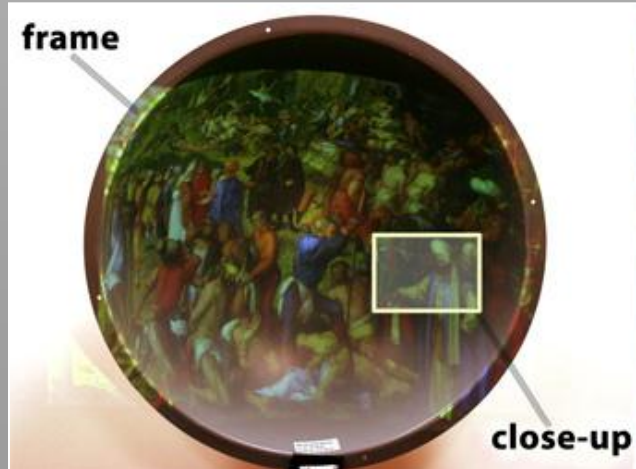
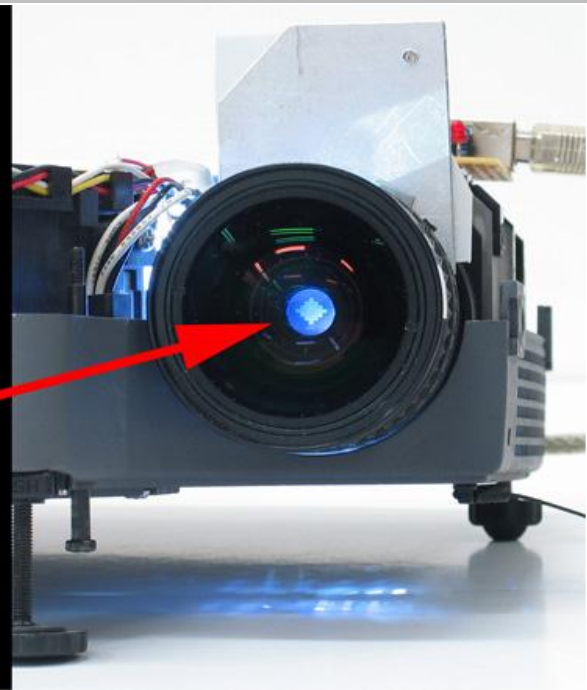
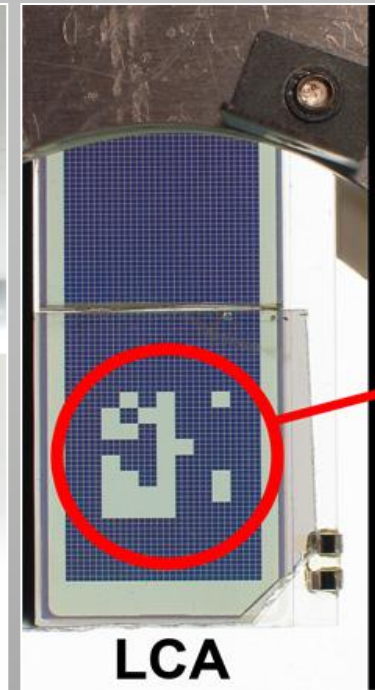
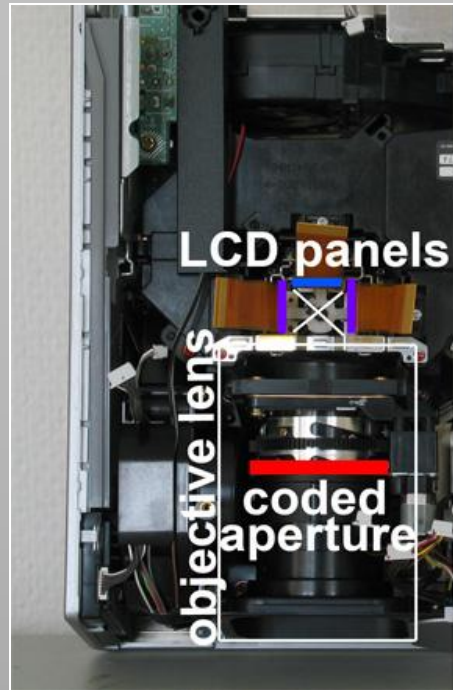
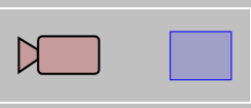
(e) Ground Truth

Applications – Defocus Video Matting



[McGuire et al. 05]

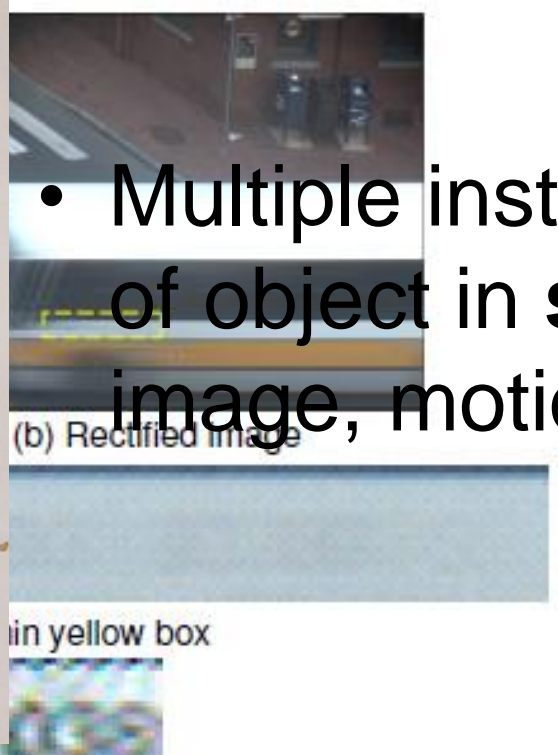
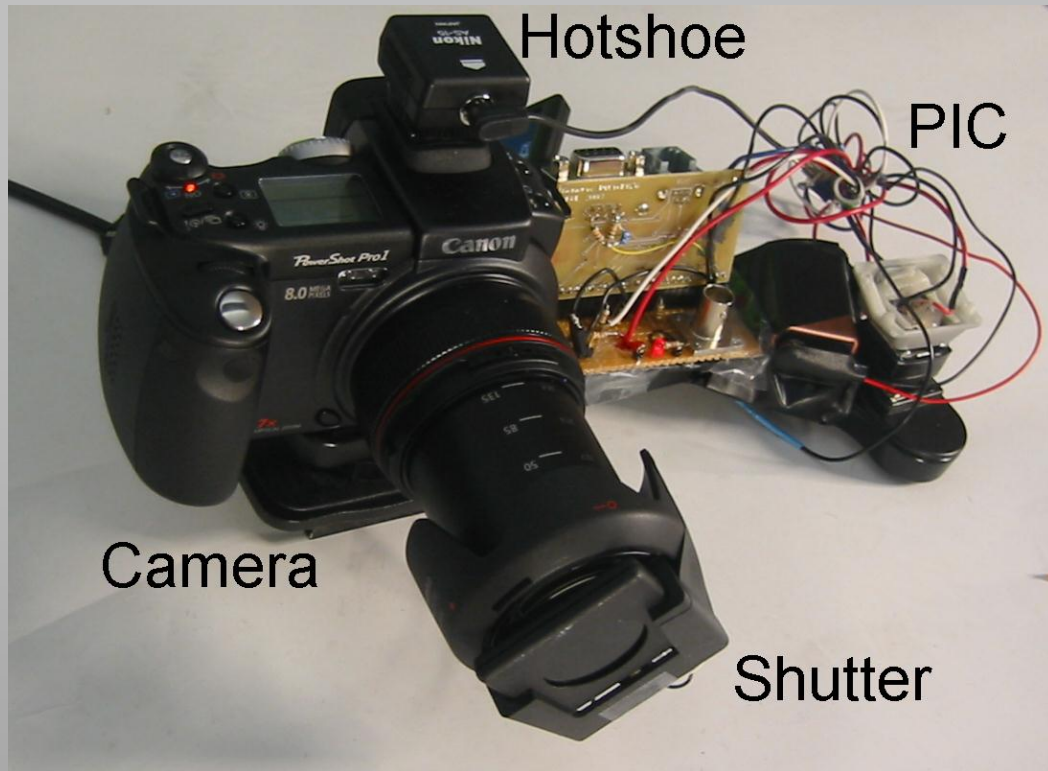
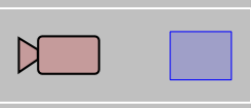
Applications – Extended DOF Projection



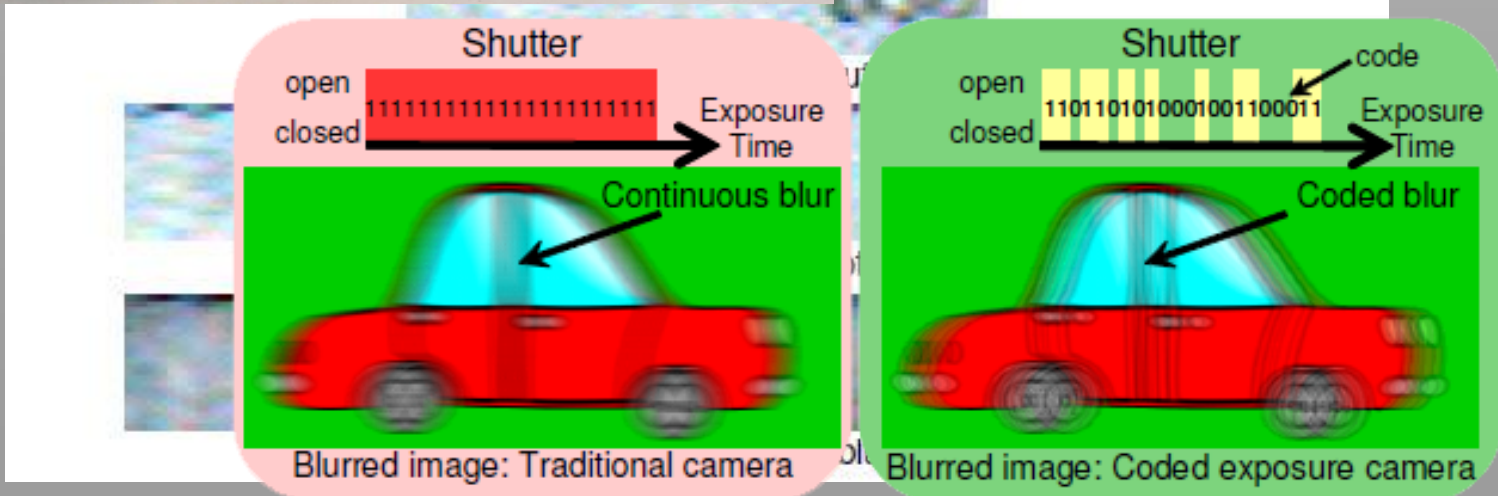
[Grosse et al. 10]

Next: High Speed Imaging

Object Motion and Coded Exposure Camera

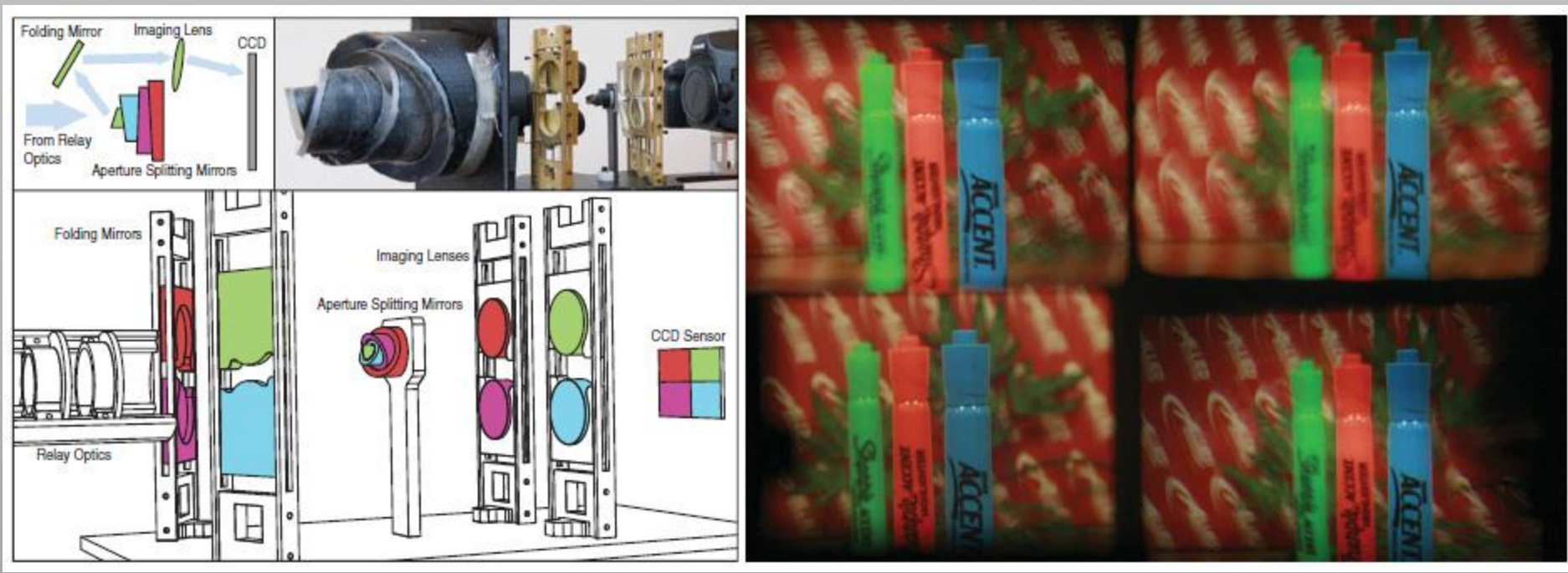
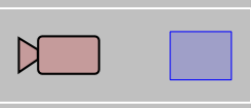


- Multiple instances of object in **single image**, motion blur



[Agrawal & Raskar 07]

Multi-Aperture Photography



[Green et al. 07]

- Multiple images, single sensor
- Synthesize different apertures in post-processing

Computational Plenoptic Imaging

Gordon Wetzstein¹
Wolfgang Heidrich³

Ivo Ihrke²
Kurt Akeley⁴

Douglas Lanman¹
Ramesh Raskar¹

¹MIT Media Lab

²Saarland University

³University of British Columbia

⁴Lytro, Inc.

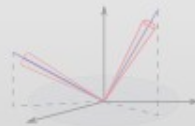
VI. Multiplexing Time



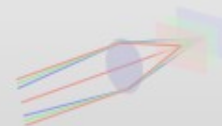
Dynamic Range



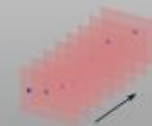
Color Spectrum



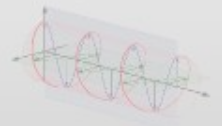
Directions | Light Fields



Space | Focal Surfaces

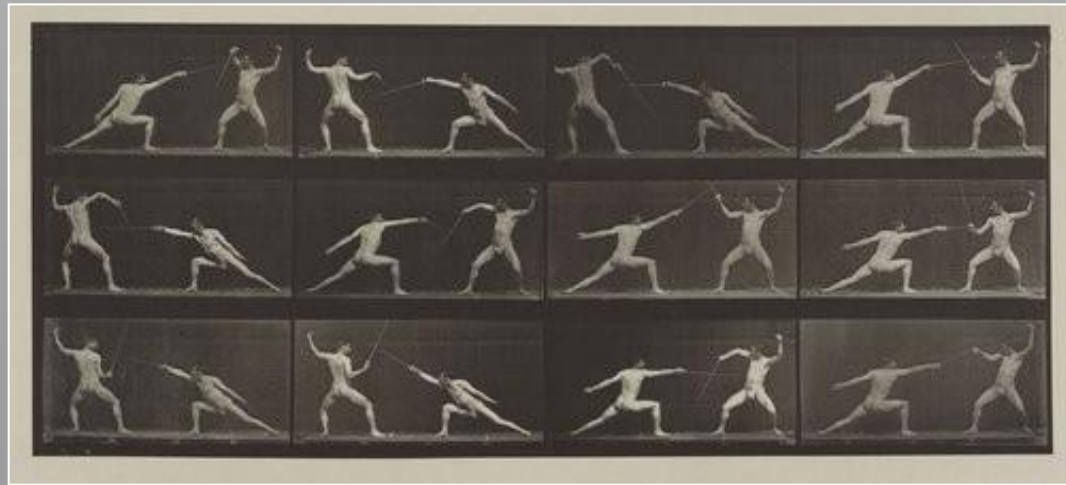
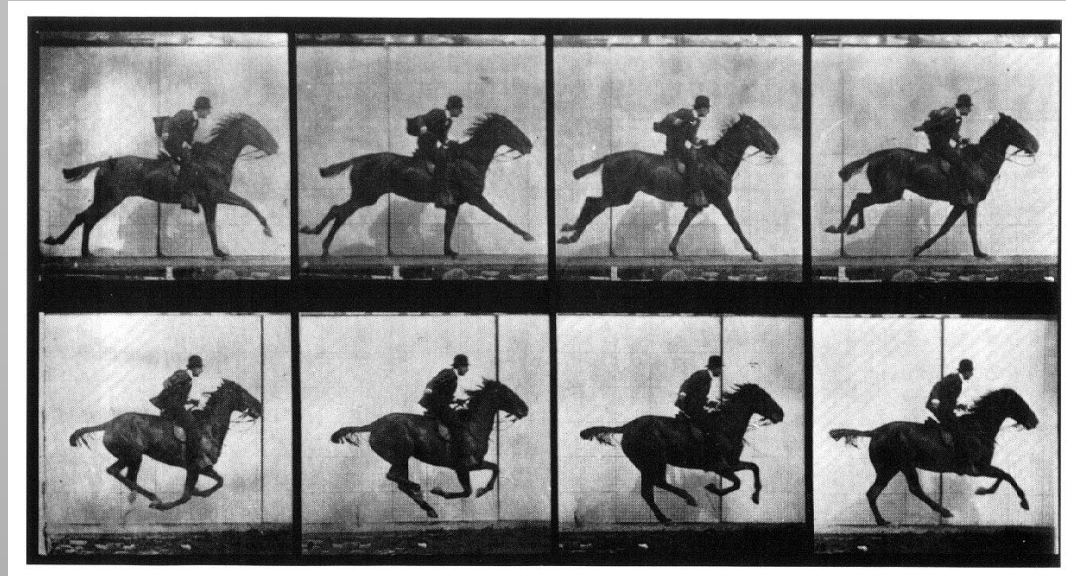
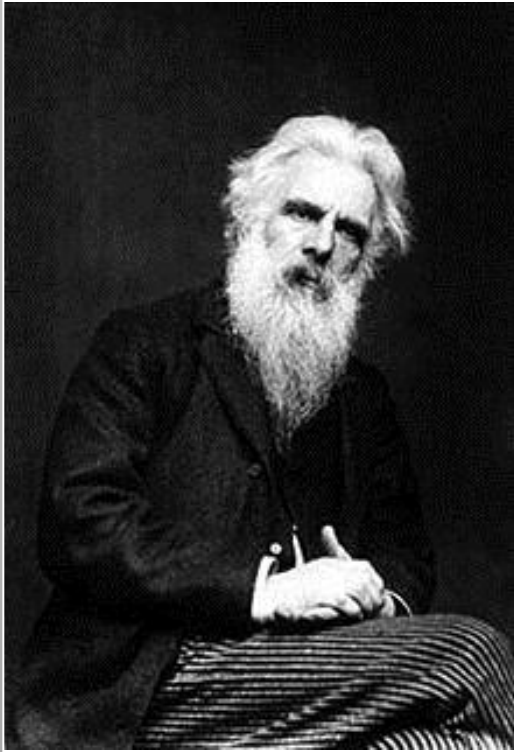


Time



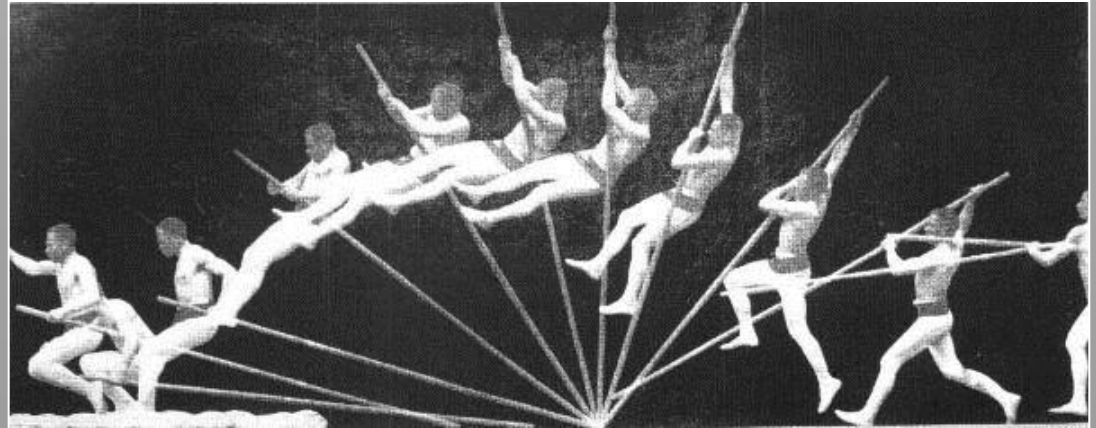
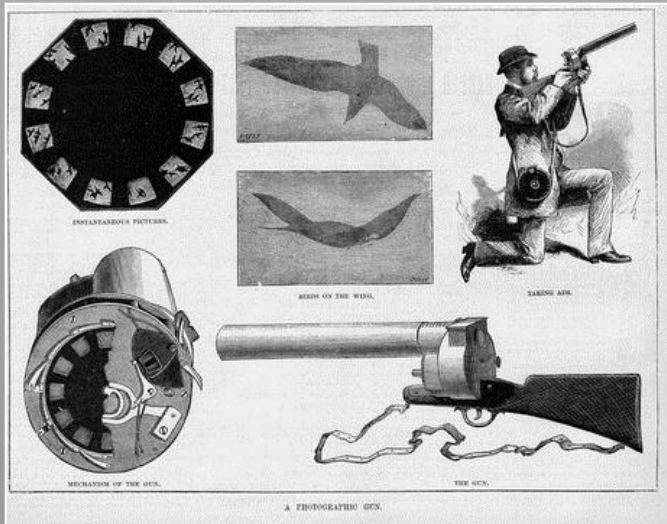
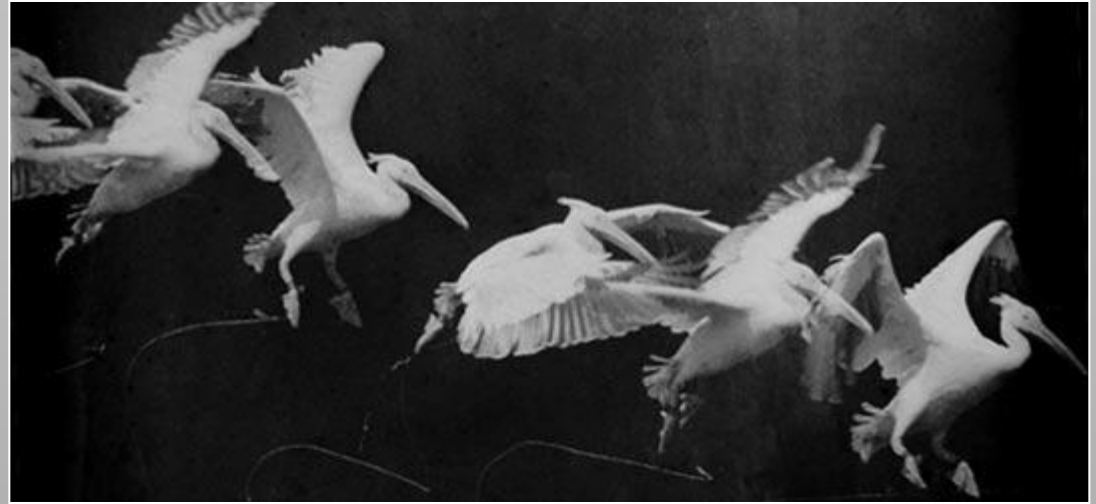
Further Properties

History – Eadweard Muybridge 1830-1904



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

History – Étienne-Jules Marey 1830-1904



www.wikipedia.org

VI.I Time Lapse Photography

BBC Time Lapse – Look It!

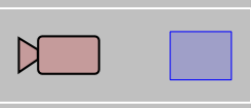
- Long exposures to avoid temporal aliasing



ND 3.0 filter, f22, 1 minute exposure

VI.II High-Speed Imaging

High-Speed Cameras



Vision Research Phantom Flex (CMOS)
2570 fps at HD resolution



Shimadzu HyperVision HPV-2 (CCD)
one million fps at 312x260 pixels

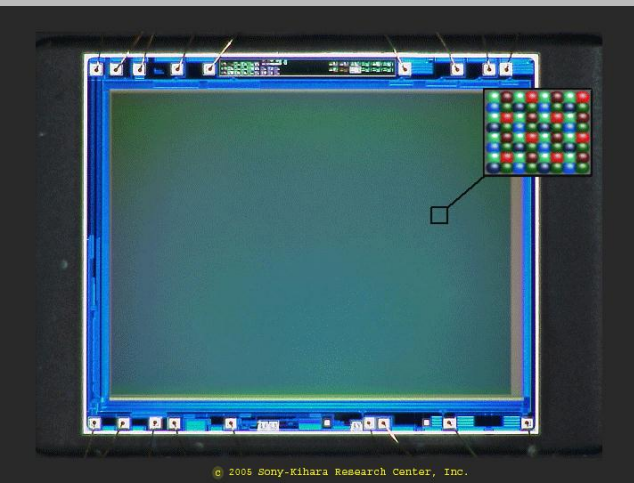
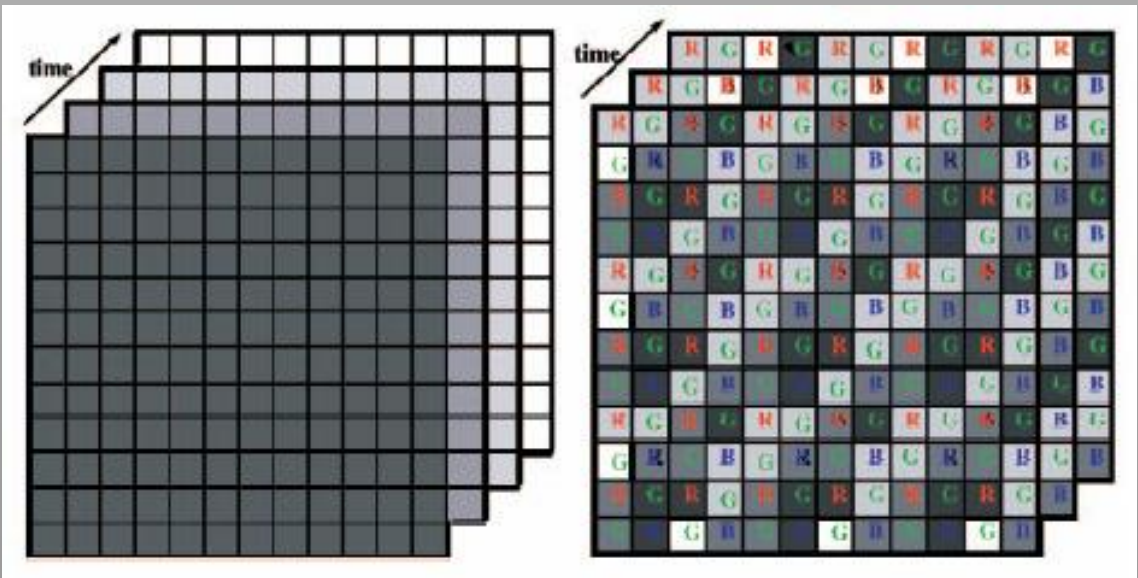
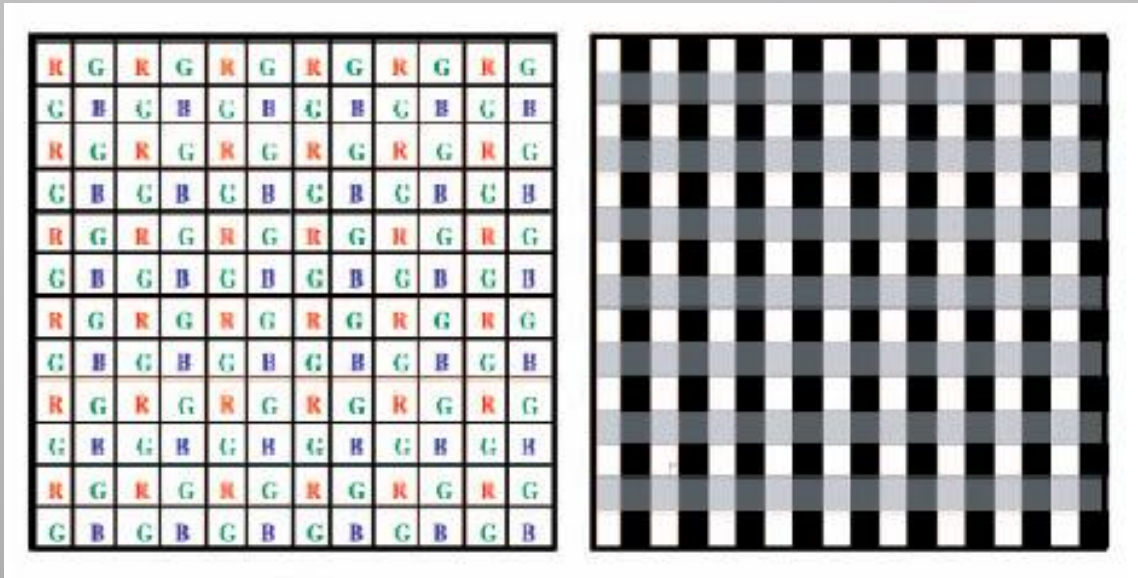
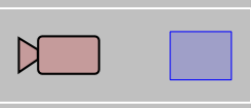


Photron FASTCAM SA5 (CMOS)
7500 fps at megapixel resolution
one million fps at 64x64 pixels



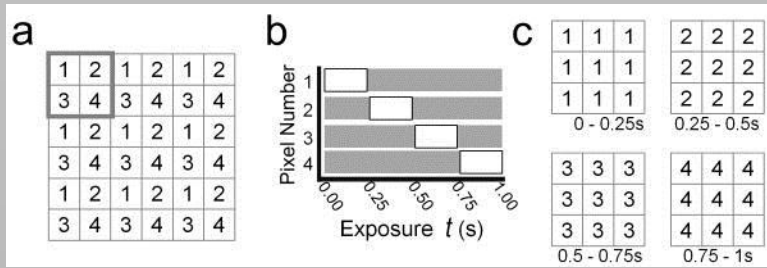
Casio Exilim Series (consumer cam)
1000 fps at reduced resolution

Assorted Pixels

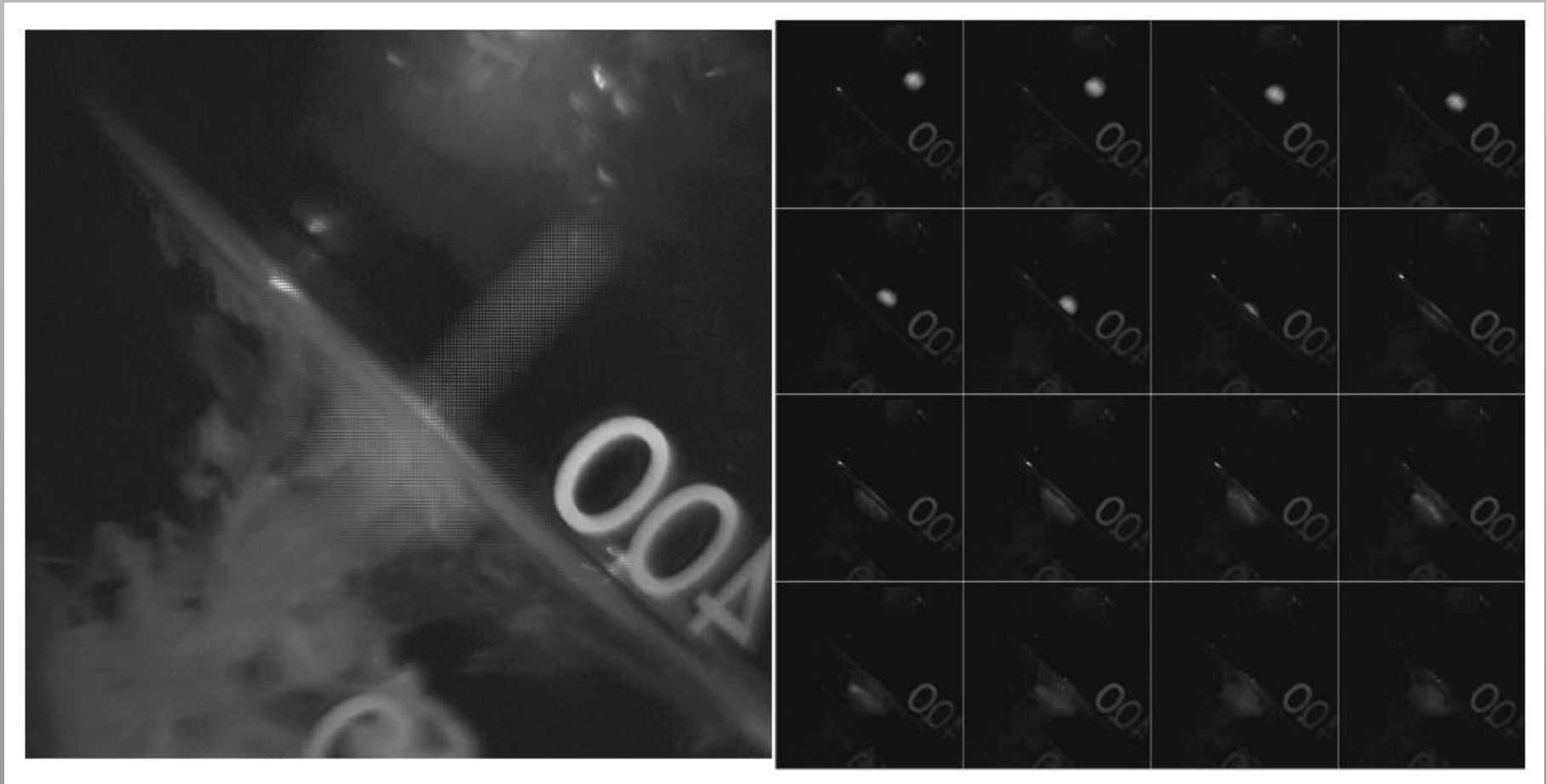


[Narasimhan & Nayar 05]

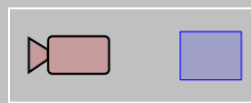
Temporal Mosaic with DMD



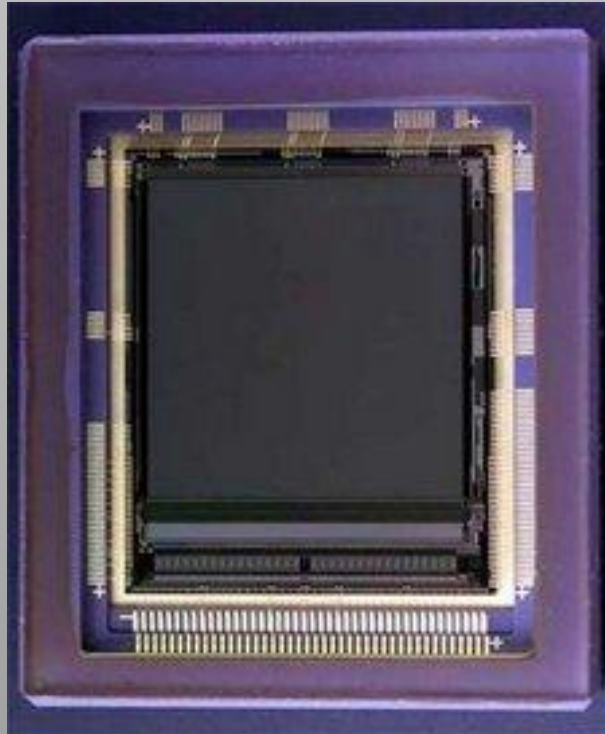
DMD aligned with CCD in microscope



[Bub et al. 10]



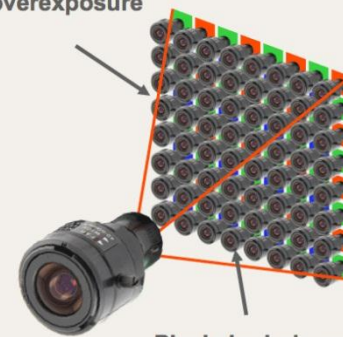
www.pixim.com



Pixim's Digital Pixel System® Technology



Pixels in bright areas automatically adjust to eliminate overexposure

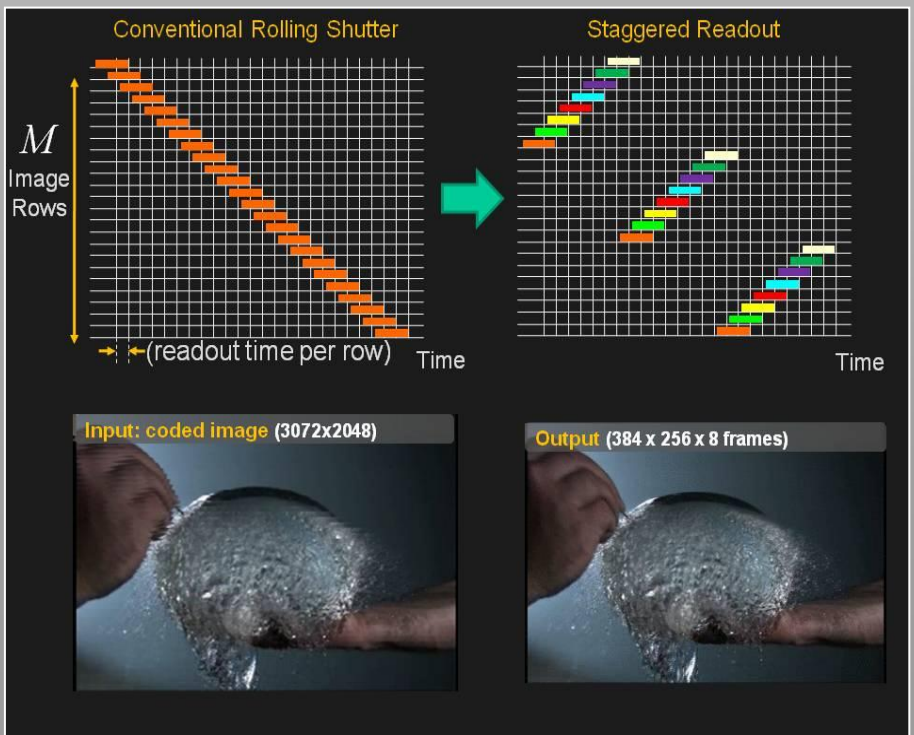
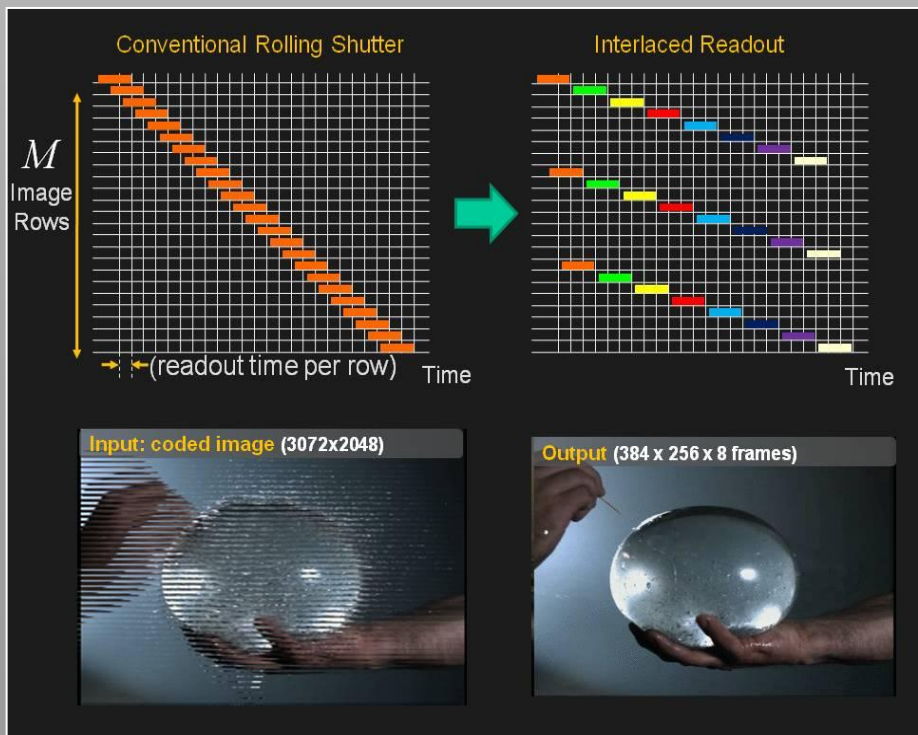
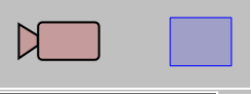


Pixels in dark areas automatically adjust to eliminate underexposure

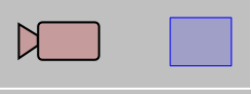
- Only all-digital solution
- Every pixel automatically adjusts to produce an optimal exposure
- Its like having over 400,000 self-adjusting cameras inside – one for every pixel
- *Every Pixel Tells a Story*

Cypress Semiconductor LUPA 3000
3 megapixels, 485 fps

Coded Rolling Shutter



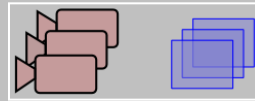
Reinterpretable Imager



- Moving pinhole over time in aperture
- Capture with light field camera

The diagram illustrates the process of recovering facial expressions from a captured photo. On the left, a large, blurry image of a man's face is labeled "Captured Photo". A green arrow points from this image to a 3x3 grid of nine smaller images on the right, labeled "Recovered Facial Expressions". These nine images show the same man with various facial expressions: neutral, hands covering his eyes, smiling, frowning, looking to the side, and others. A blue circular arrow icon is positioned between the two main image areas, indicating a transformation or process. Below the "Captured Photo" and "Recovered Facial Expressions" labels, there are three white boxes containing the text "SIGGRAPH 2007", "SIGGRAPH 2007", and "This Paper" respectively.

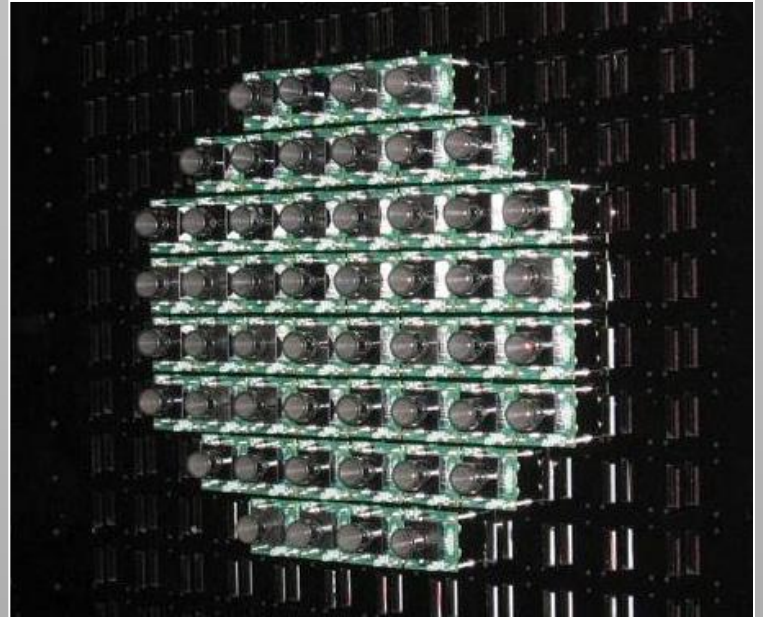
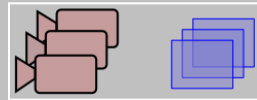
Bullet Time Effect



from 'The Matrix'

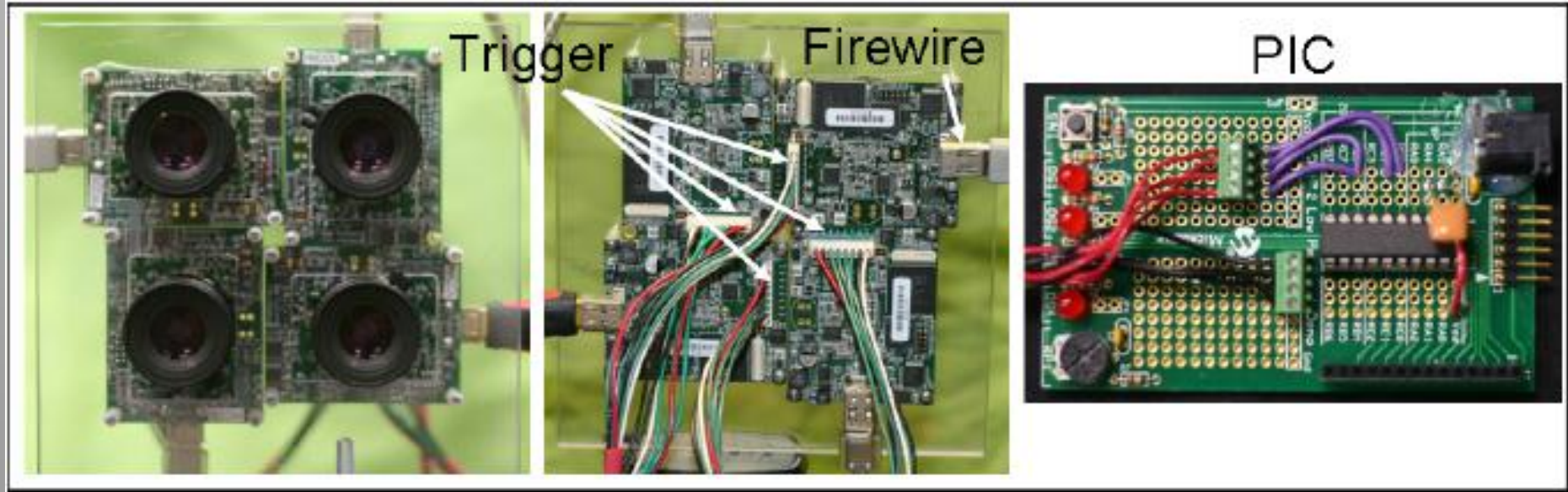
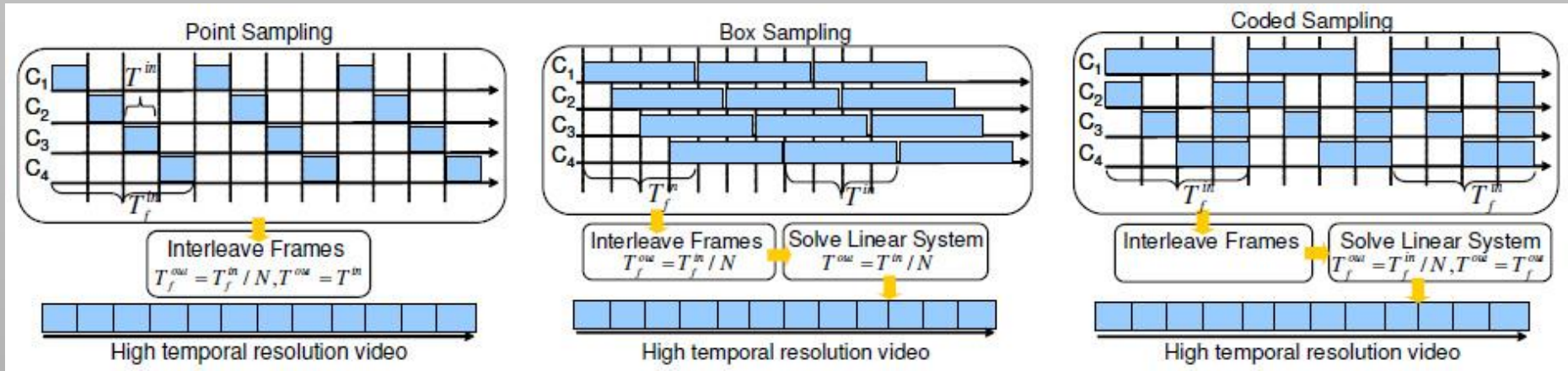
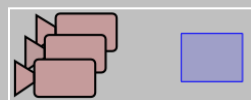


Stanford Multi-Camera Array

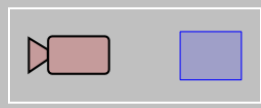


[Wilburn et al. 04]

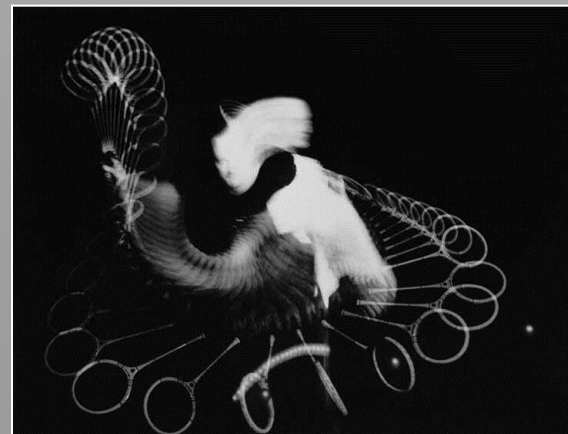
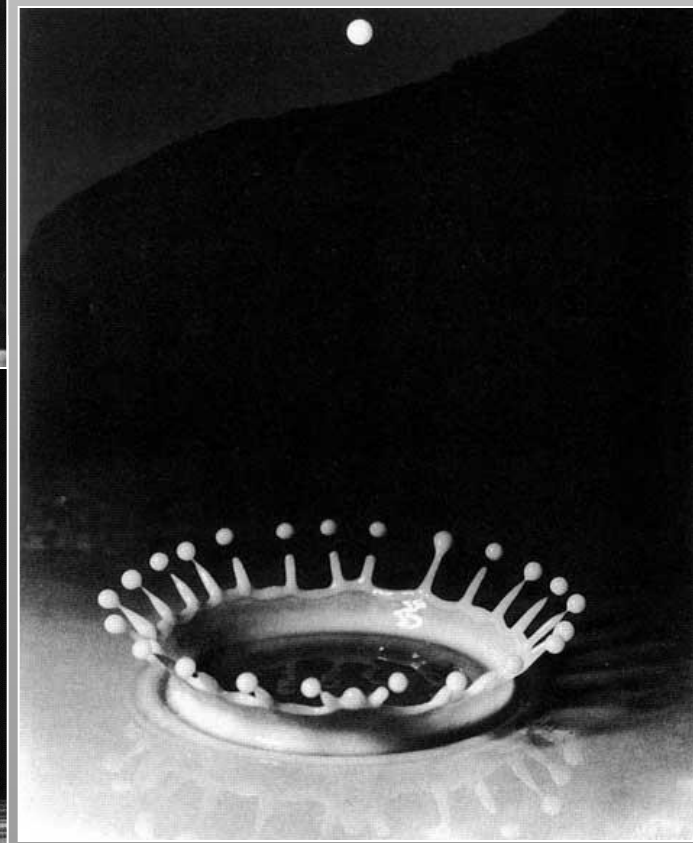
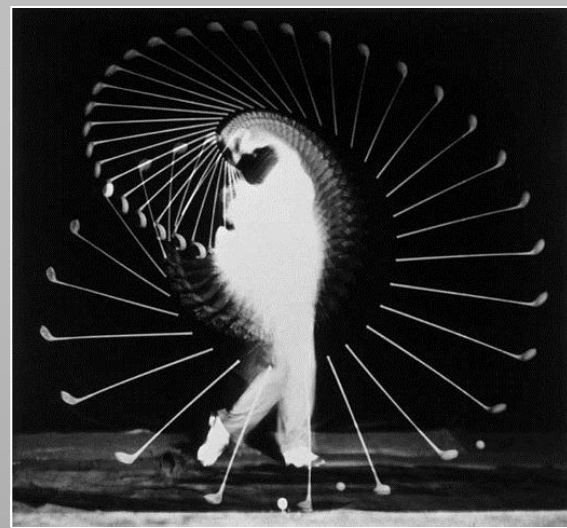
Coded Temporal Sampling



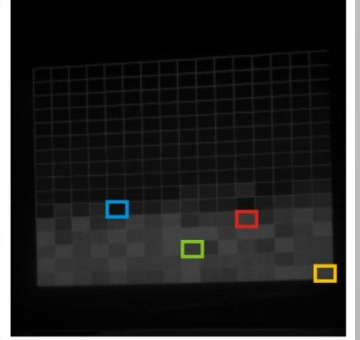
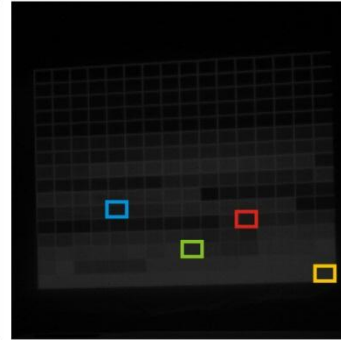
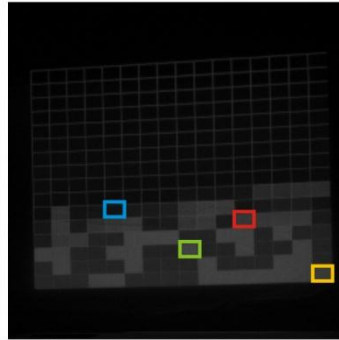
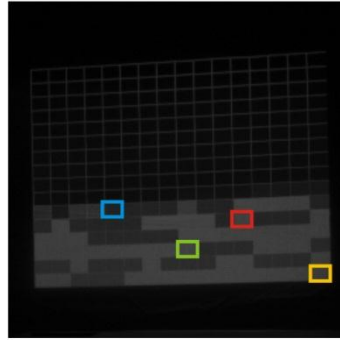
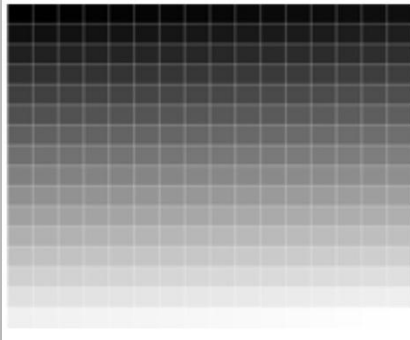
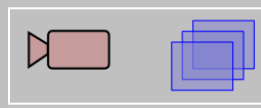
High-Speed Illumination – Electronic Strobes



Harold 'Doc' Edgerton
1903-1990



Temporal Dithering with DLP Illumination



Calibration image input to projector

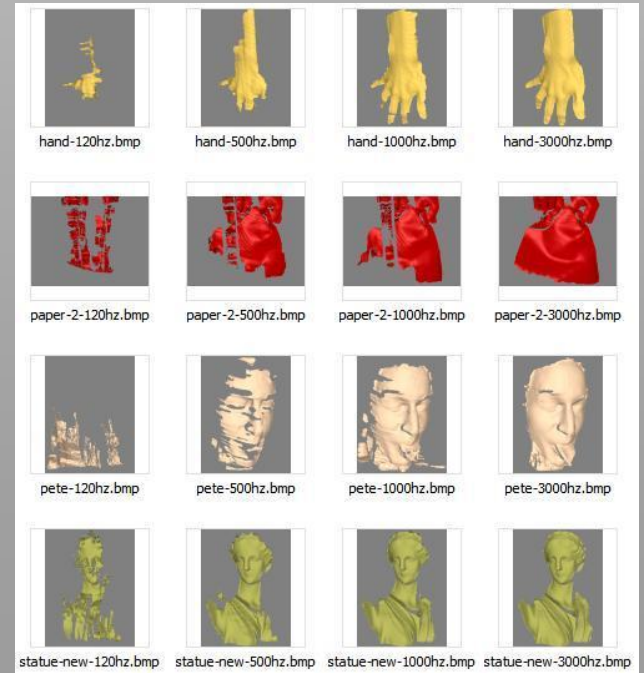
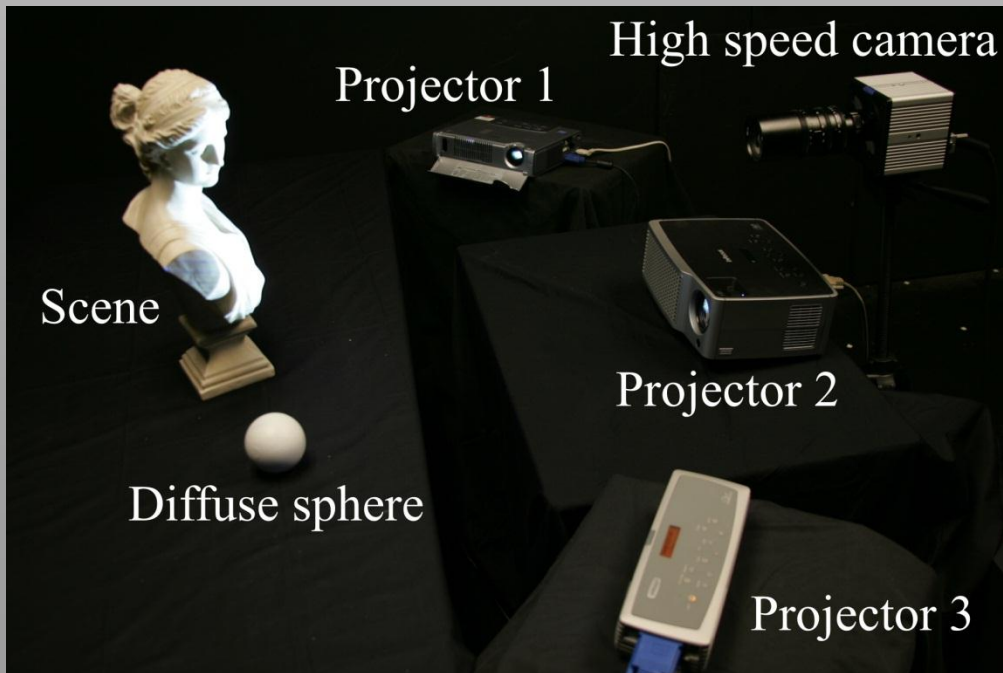
Frame 5

Frame 27

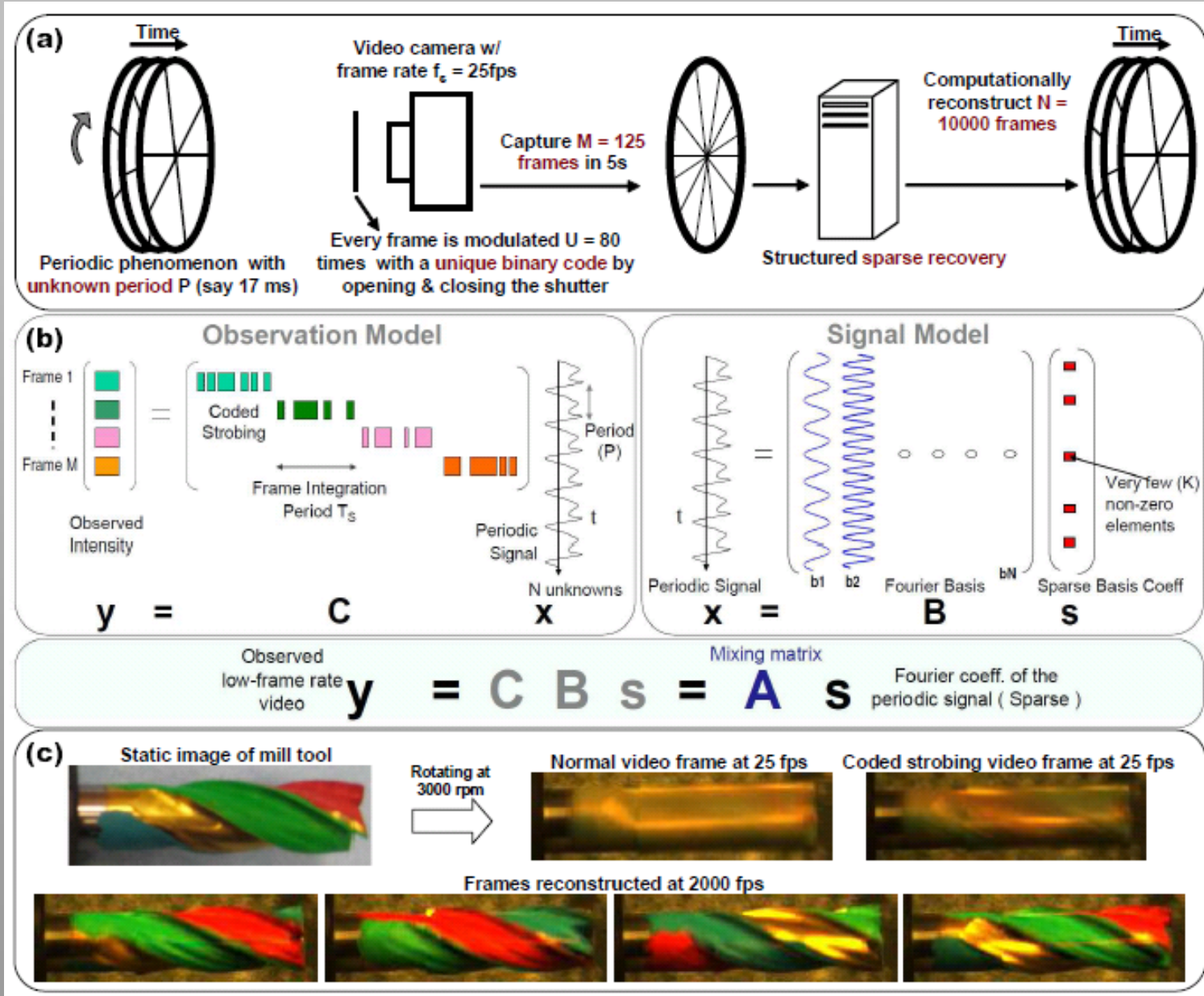
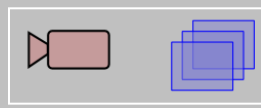
Frame 85

Frame 91

Measured images at 10000fps



Coded Strobing Photography



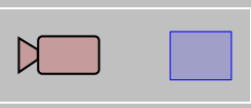
$$y = C B s = A s$$

Observed low-frame rate video

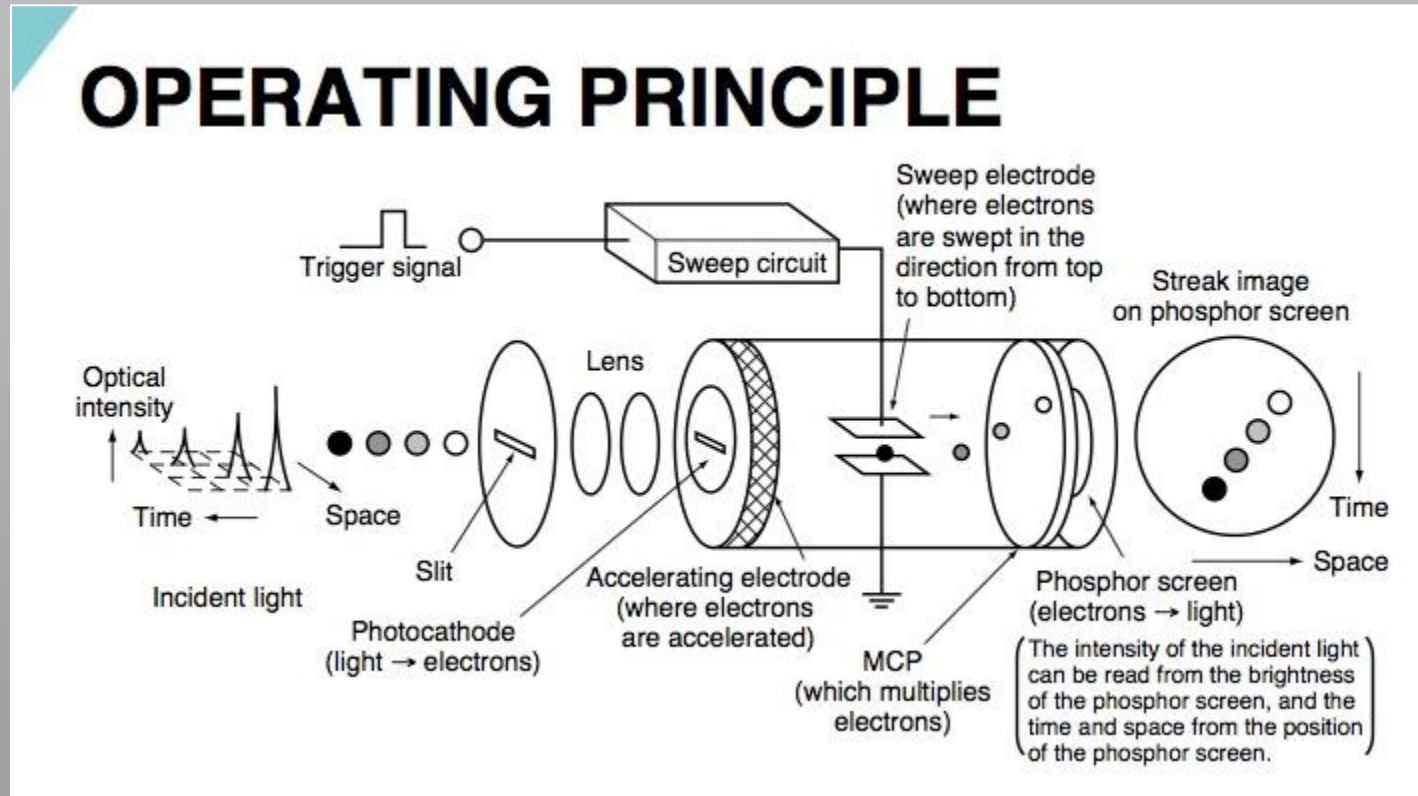
Mixing matrix

Fourier coeff. of the periodic signal (Sparse)

Streak Cameras



C5680
\$200K



www.hamamatsu.com

VI.1 Motion Deblurring

Motion Deblurring Overview

- Motion blur is velocity-dependent
- Can be described as convolution, where
 - Kernel shape is motion trajectory
 - Trajectory is modulated by exposure function



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

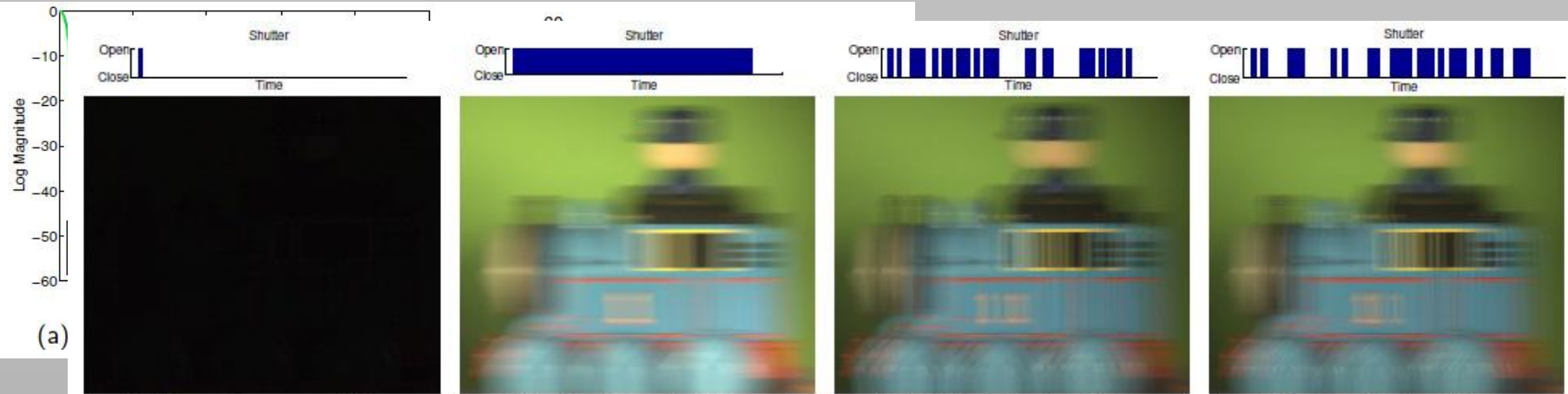
Deconvolution is Still Hard

- Again – problems:
 - Camera noise
 - Spatially varying kernel (velocity-dependent)
 - Unknown motion trajectory
 - Ill-posed problem, kernel of box integration function is not invertible (optical cancellation of image frequencies)

Approaches to Improve Motion Deblurring

- Make PSF invertible → coded exposure
- Make PSF velocity-invariant → shift-invariant deconvolution
- Automate PSF estimation

Flutter Shutter



(a)

(a) Short Exposure Photo

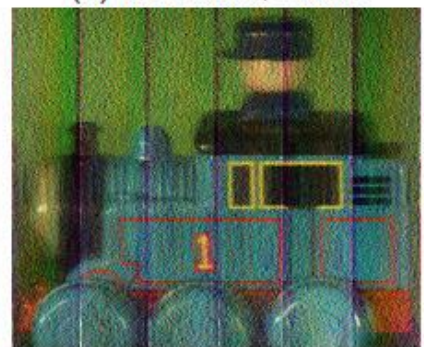
(b) Traditional, 200ms

(c) MURA Code, 200ms

(d) Our Code, 200ms



(e) Log intensity of (a)



(f) Deblurred Image



(g) Deblurred Image



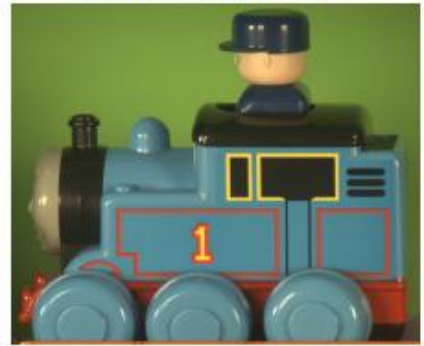
(h) Deblurred Image



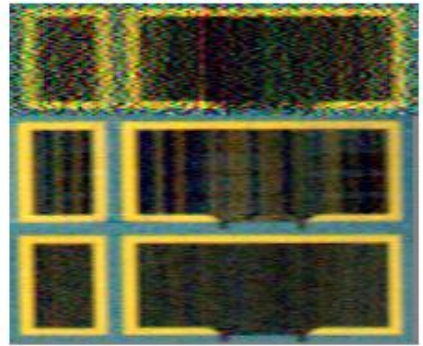
(i) Toy Train Setup



(j) RL deblur result of (b)

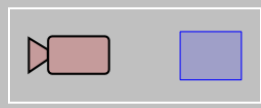


(k) Photo of Static Toy Train



(l) Deblurring details

Optimal Motion PSFs

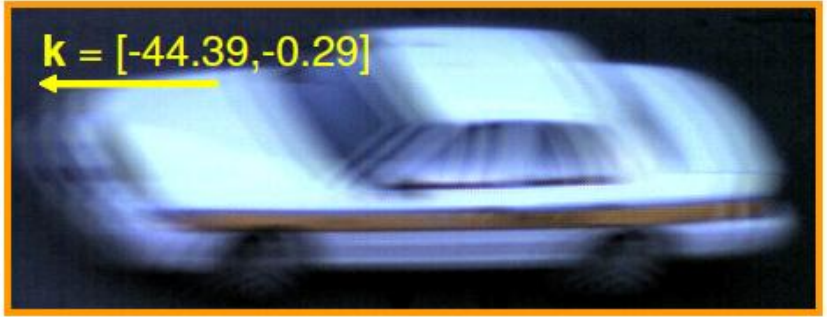


- Optimality criteria PSF invertibility & estimation

	Traditional Camera	Coded Exposure Raskar et al.	Coded Exposure Ours
PSF Estimation	✓	✗	✓
PSF Invertibility	✗	✓	✓
Traditional		111111111111111111111111111111111111	
Coded (Raskar et al.)		1010101011100111101110101111011	
Coded (Ours)		1111111111111000010011101000111	



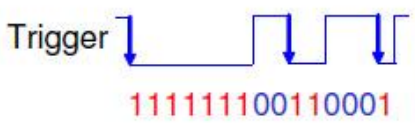
Captured frame



Cropped input frame



Coded exposure camera



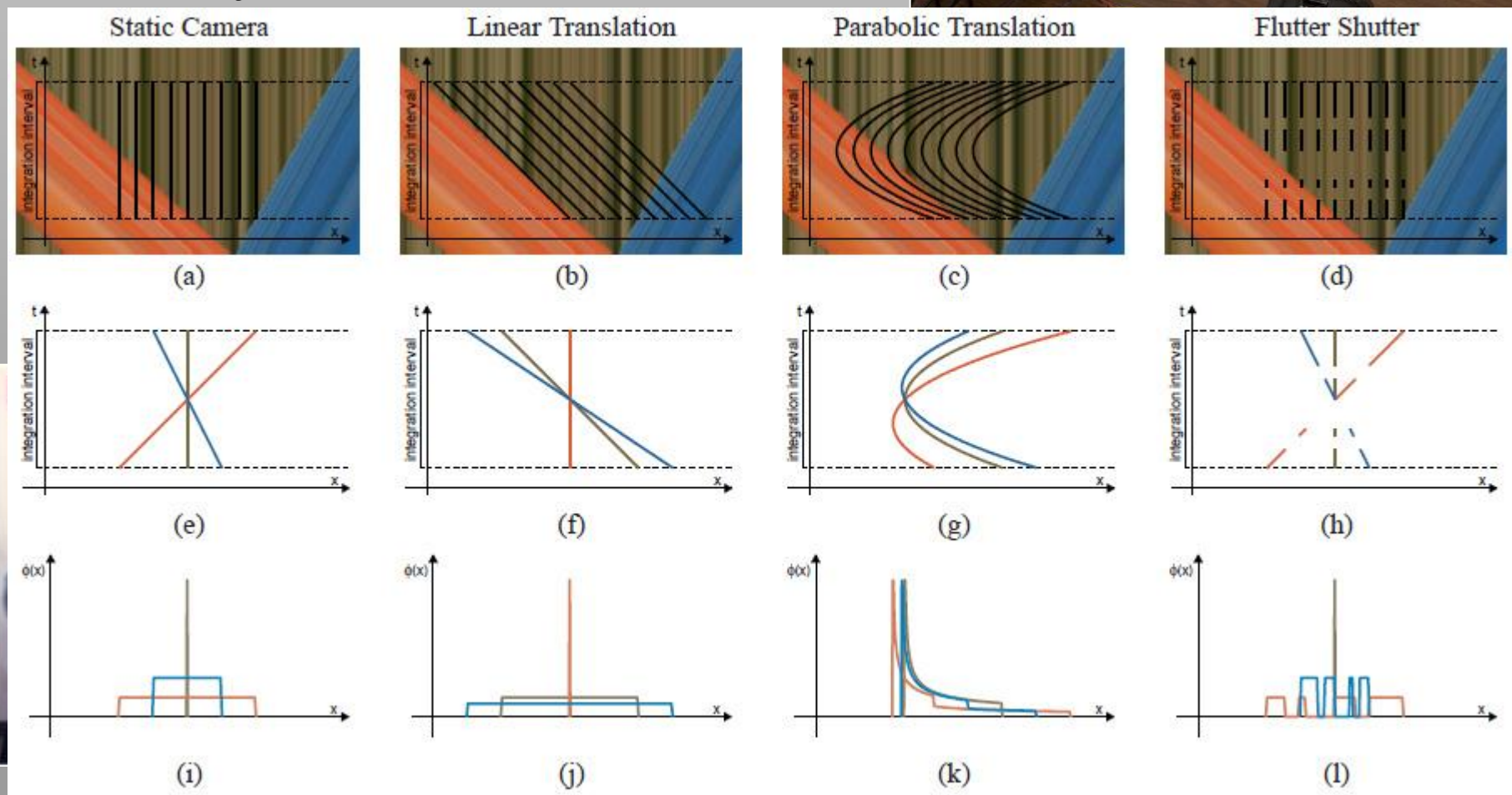
Deblurring result

[Agrawal & Xu 07]

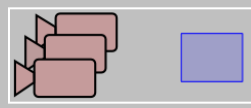
Motion Invariant Photography



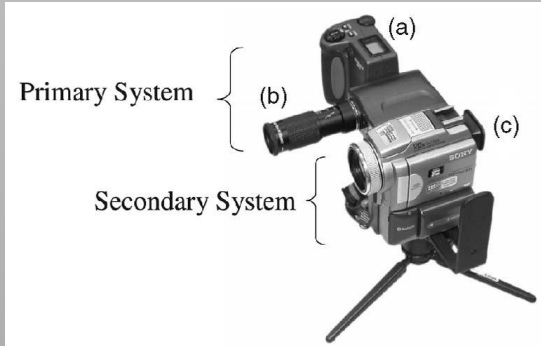
- Engineer PSF to be motion invariant
- Only for 1D motion



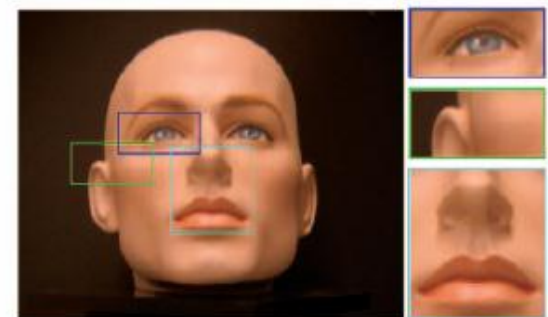
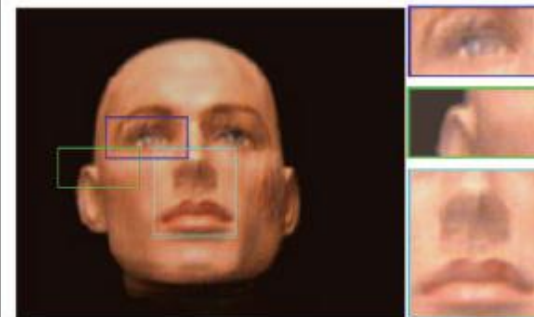
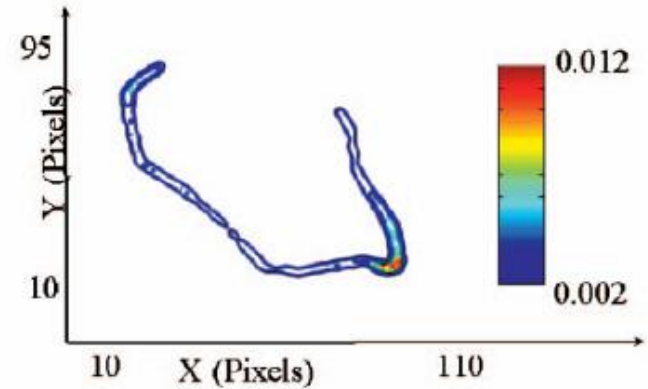
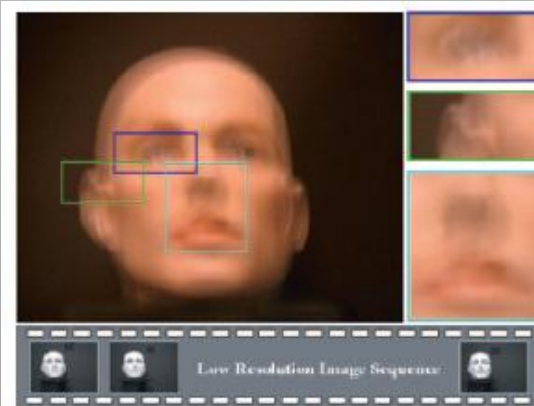
Hybrid Cameras



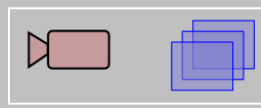
- Combined high-speed low-quality & low-speed high-quality camera



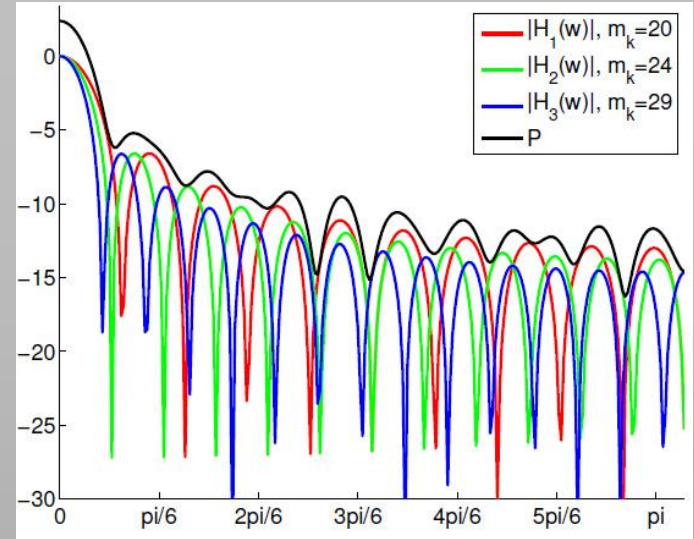
[Ben-Ezra & Nayar 04]



Motion Blur in Video



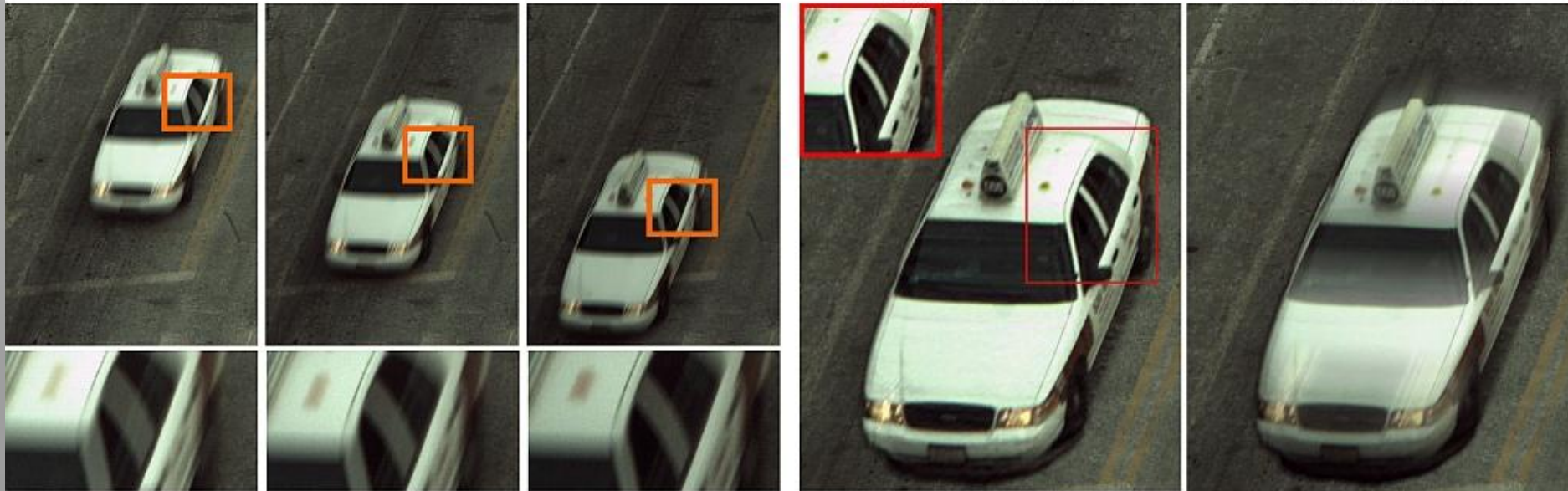
- Coded exposure & super-resolution in successive video frames



Captured Photos

Deblurred Result

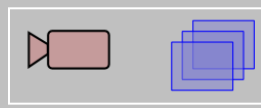
Motion Streaks



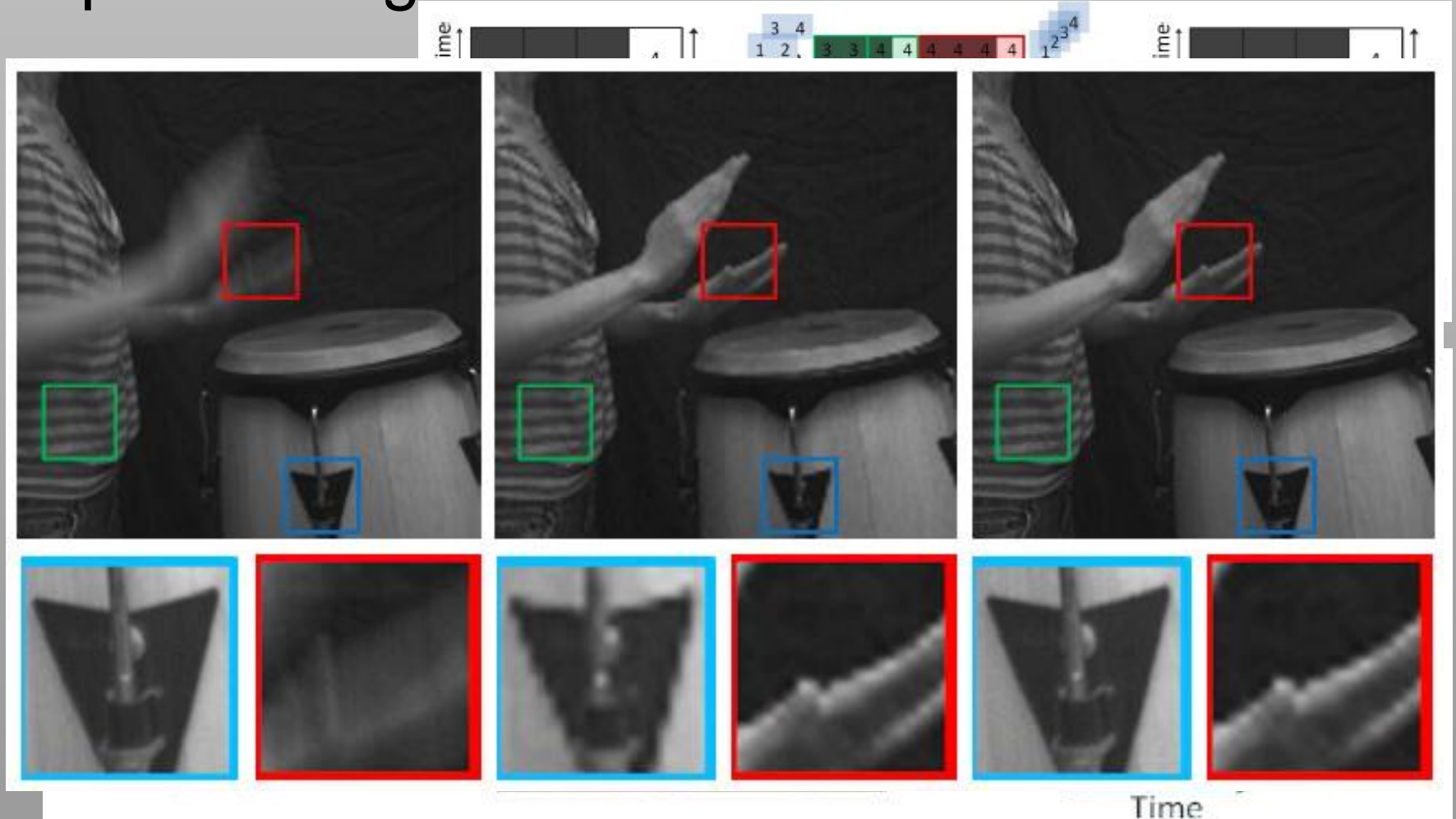
[Agrawal et al. 09]

Next: Further Light Properties

Flexible Voxels

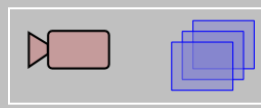


- Flexible space-time resolution as post-processing

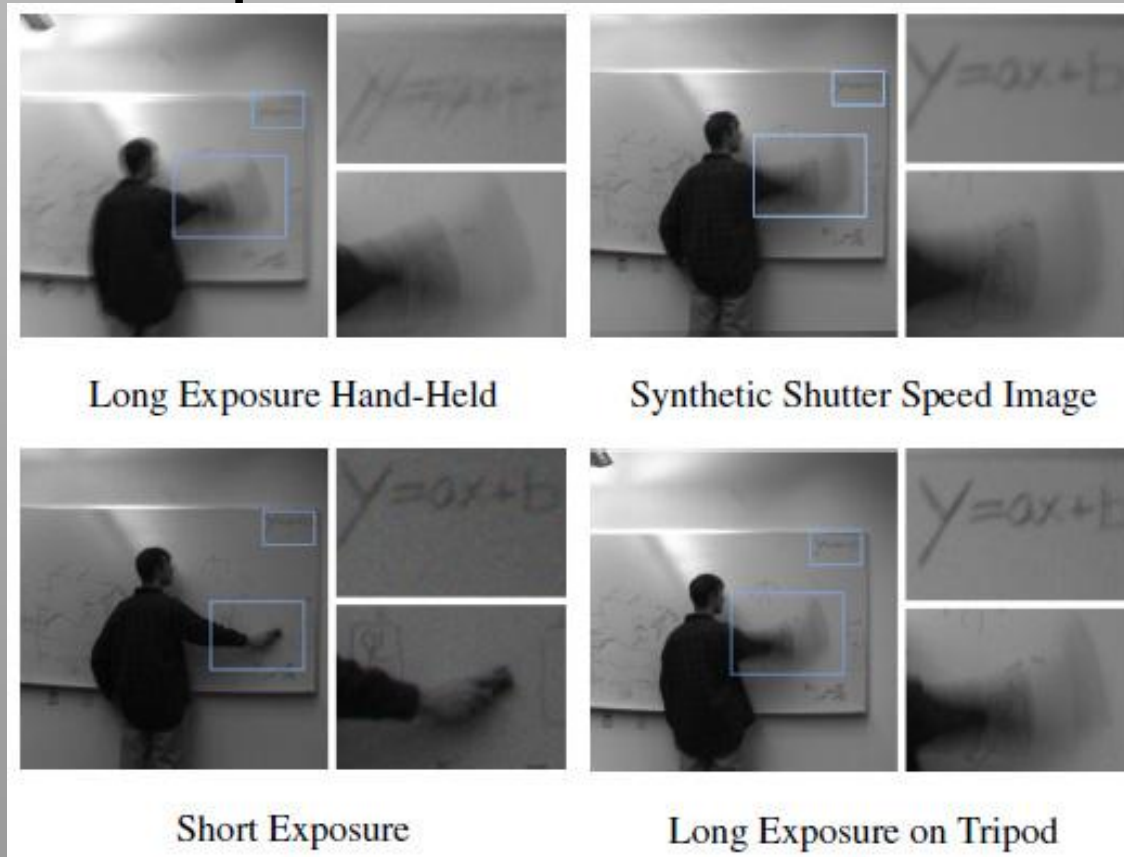


[Gupta et al. 10]

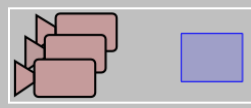
Synthetic Shutter Speed Imaging



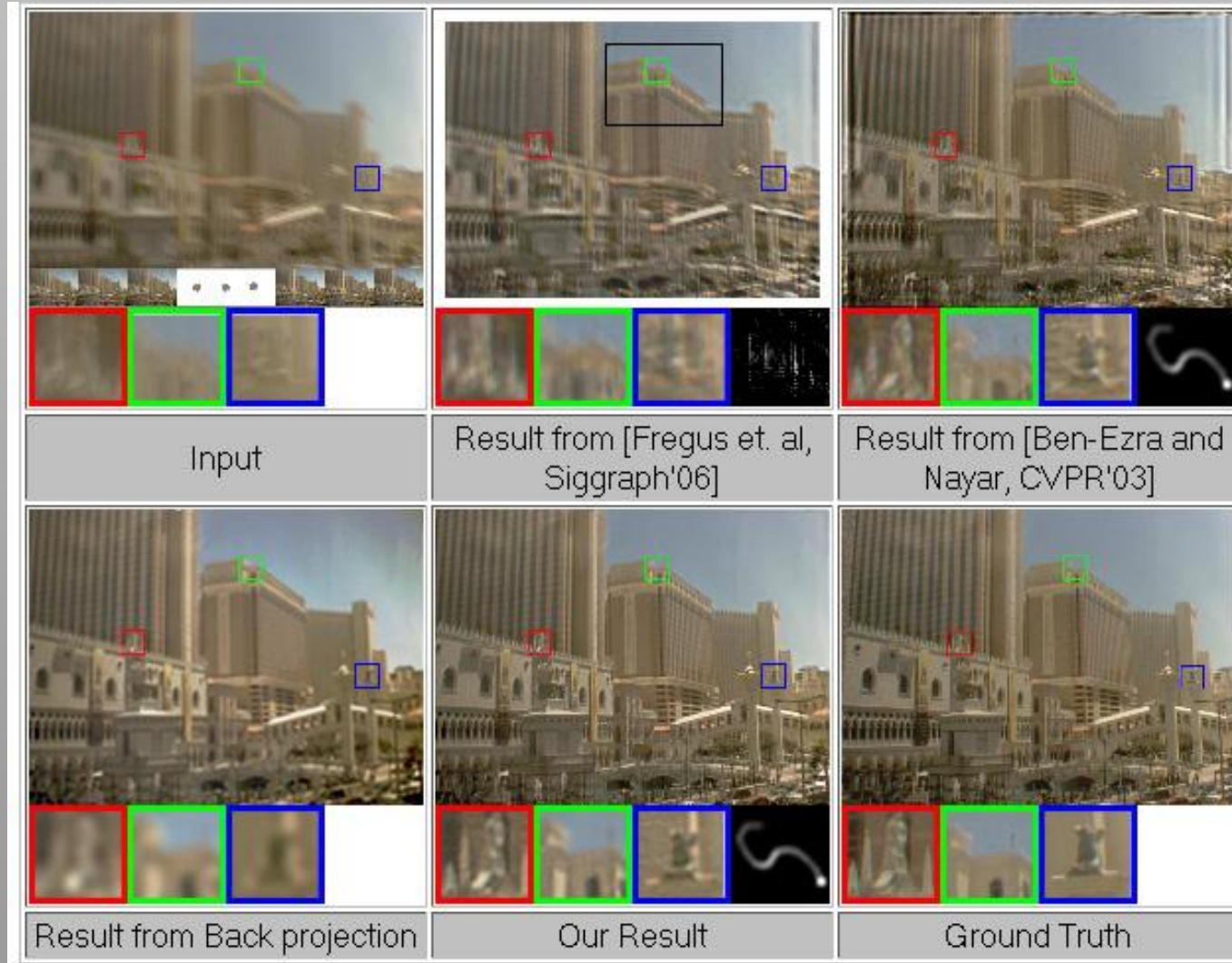
- Combine multiple short exposures to reduce noise
- Align with optical flow



[Telleen 07]

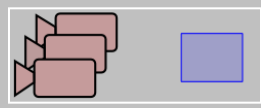


- Motion deblurring & super-resolution

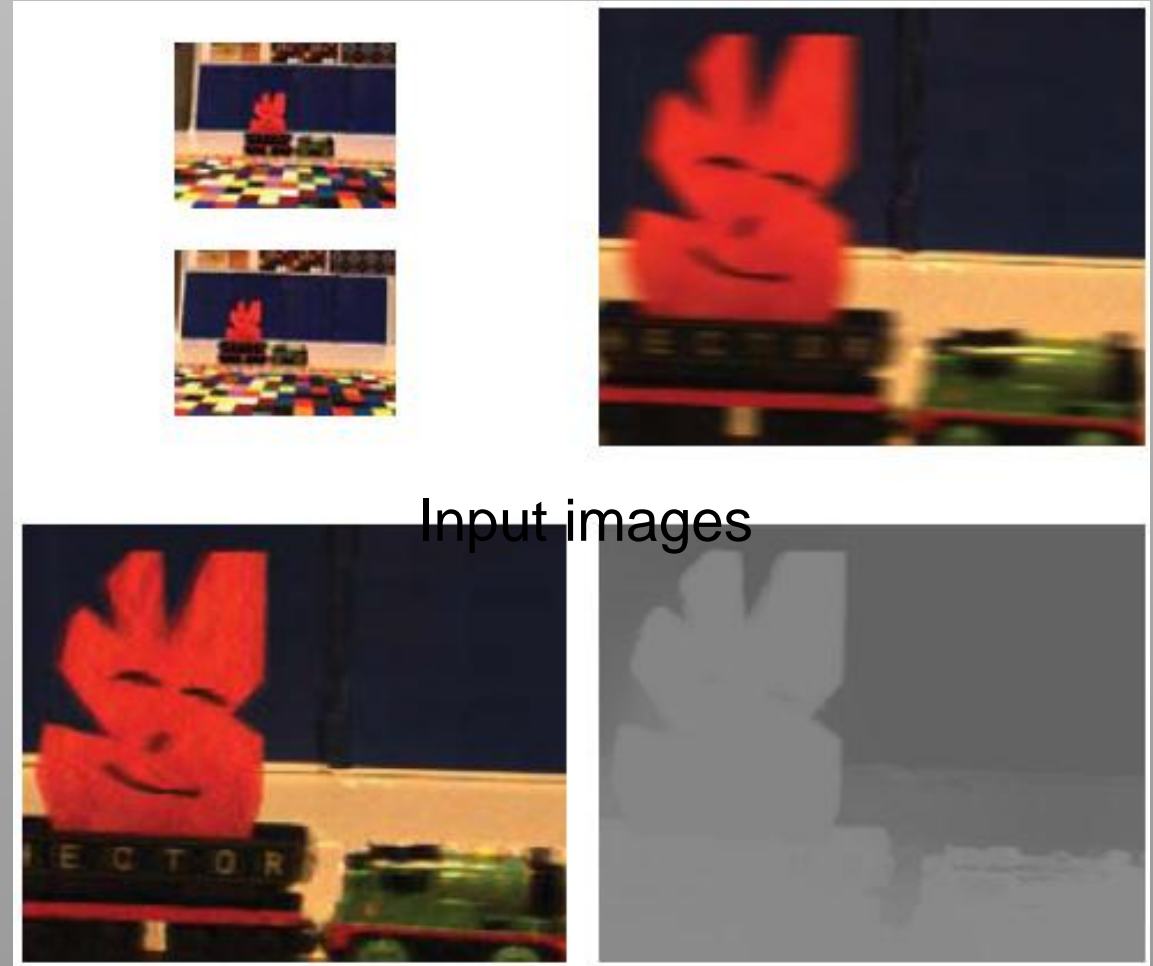
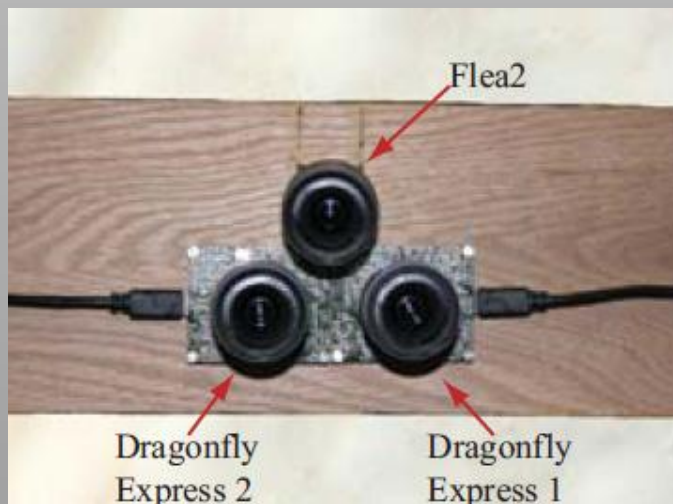


[Tai et al. 08]

Hybrid Cameras



- Motion deblurring & depth from two low-resolution high-speed cameras



Input images

Deblurred result

Recovered Depth

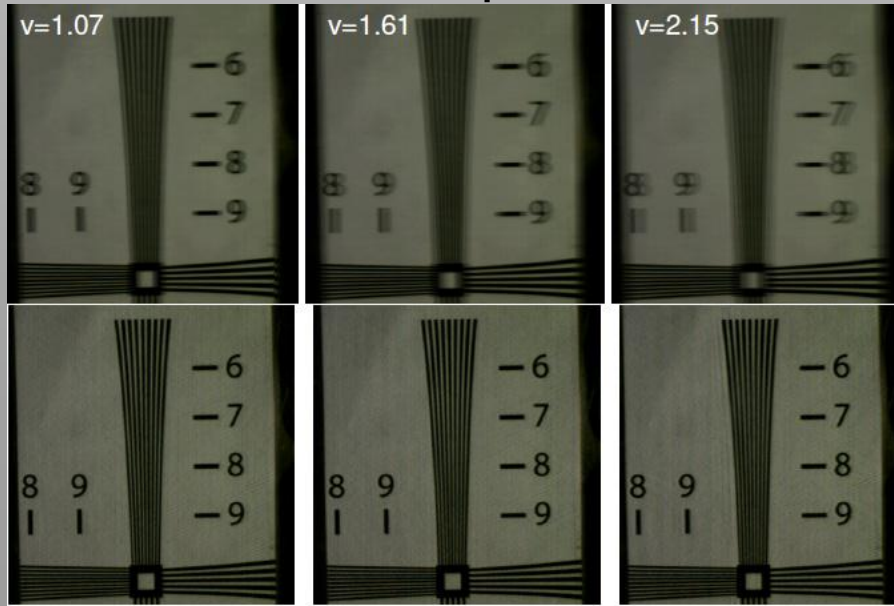
[Li et al. 08]

Analysis

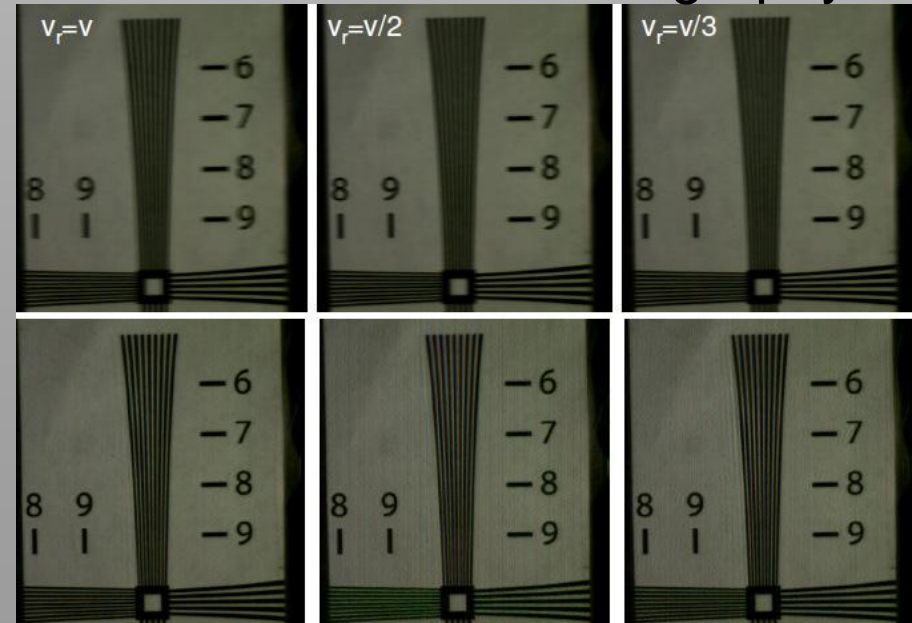


- Analysis of optimal coded, single image deblurring
- MIP becomes worse when velocities exceed expectations

Coded Exposure



Motion Invariant Photography



[Agrawal & Raskar 09]

Computational Plenoptic Imaging

Gordon Wetzstein¹
Wolfgang Heidrich³

Ivo Ihrke²
Kurt Akeley⁴

Douglas Lanman¹
Ramesh Raskar¹

¹MIT Media Lab

²Saarland University

³University of British Columbia

⁴Lytro, Inc.

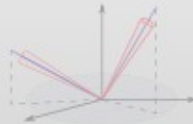
VII. Further Light Properties



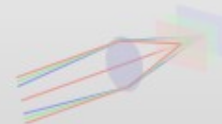
Dynamic Range



Color Spectrum



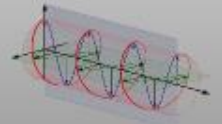
Directions | Light Fields



Space | Focal Surfaces



Time



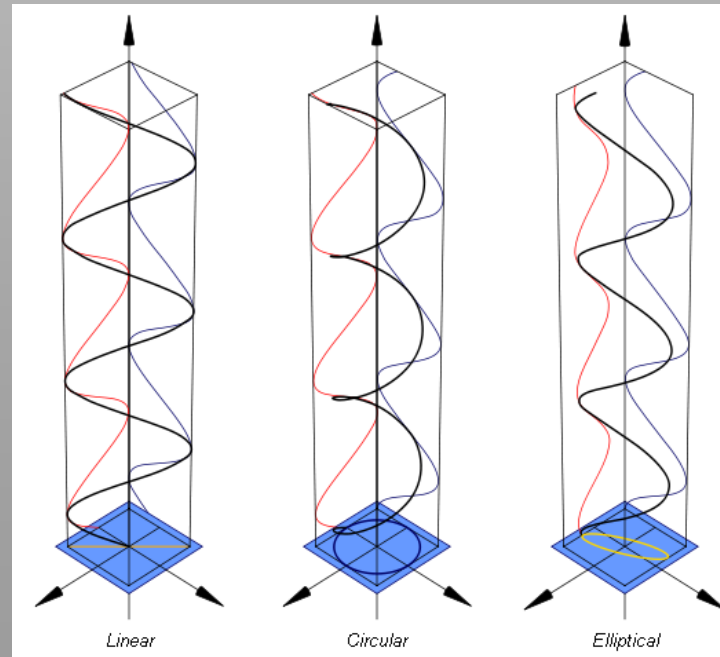
Further Properties

VII.1 Polarization

What is Polarization?

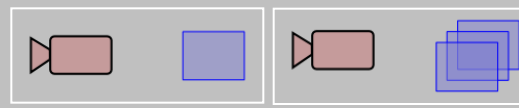
- Wave property of light
- Oscillation perpendicular to propagation
- *Field Guide to Polarization*, Collett 05

Polarization State

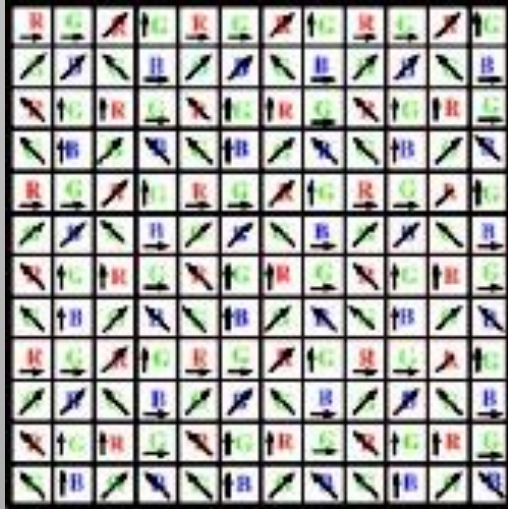


[http://en.wikipedia.org/wiki/Polarization_\(waves\)](http://en.wikipedia.org/wiki/Polarization_(waves))

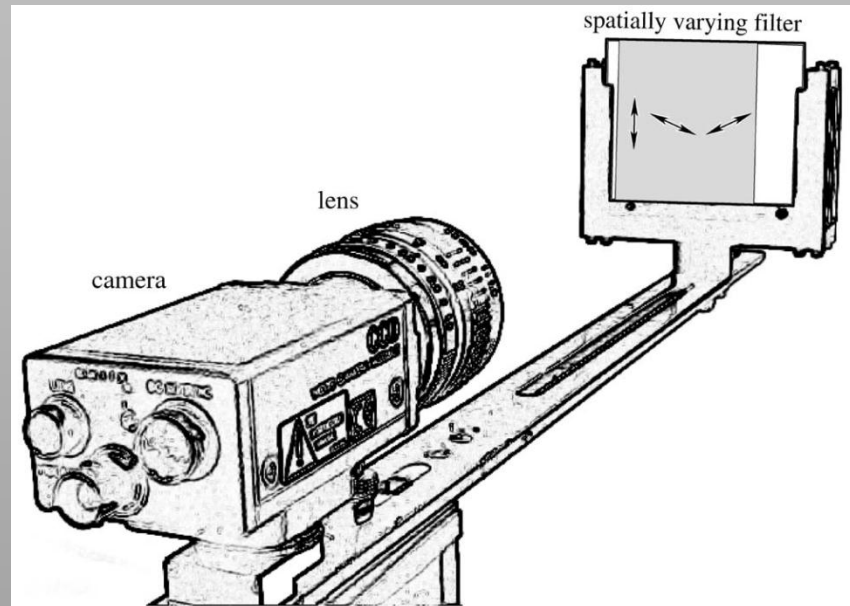
Polarization Acquisition



Assorted Pixels



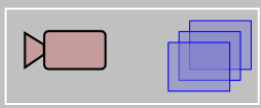
Generalized Mosaics Polarization Panoramagrams



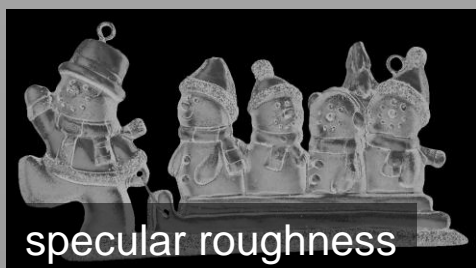
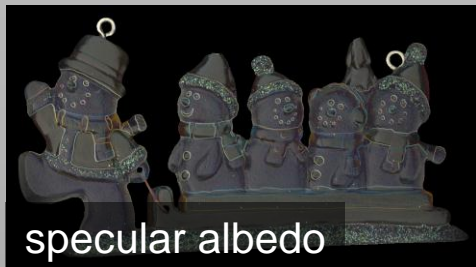
[Narasimhan & Nayar 05]

[Schechner & Nayar 05]

Polarization Acquisition - Applications



Measure Surface Properties



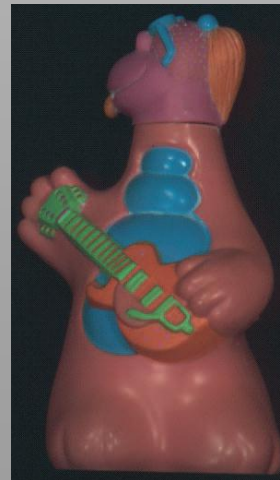
[Ghosh et al. 10]

Dehazing



[Namer & Schechner 05]

Shape Estimation



[Miyazaki et al. 03]

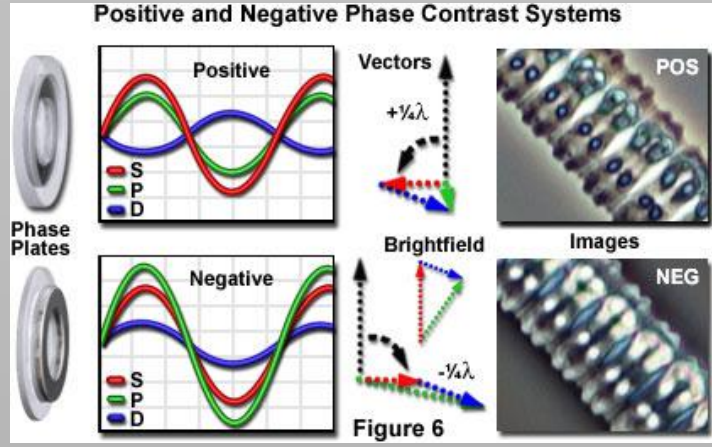
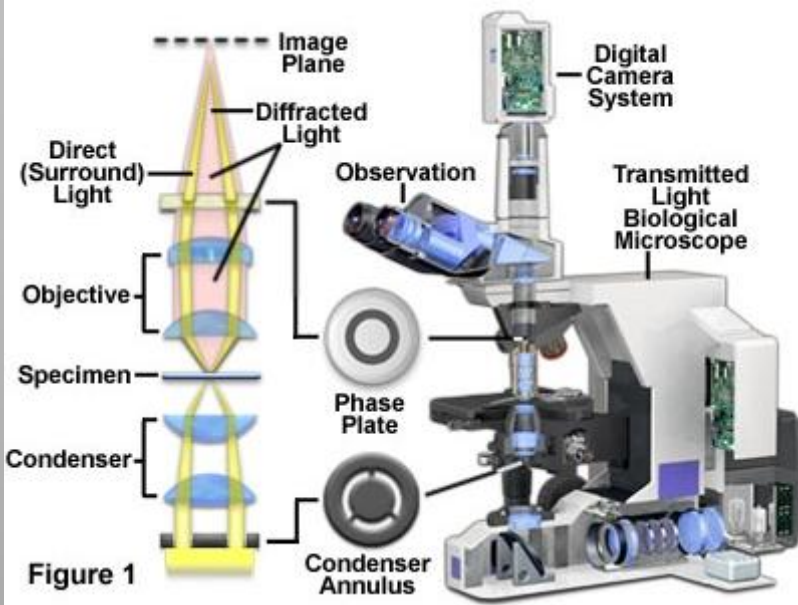
Also: optical communication, underwater vision, BRDF & IOR acquisition, ...

VII.II Phase Imaging

Phase Contrast Imaging

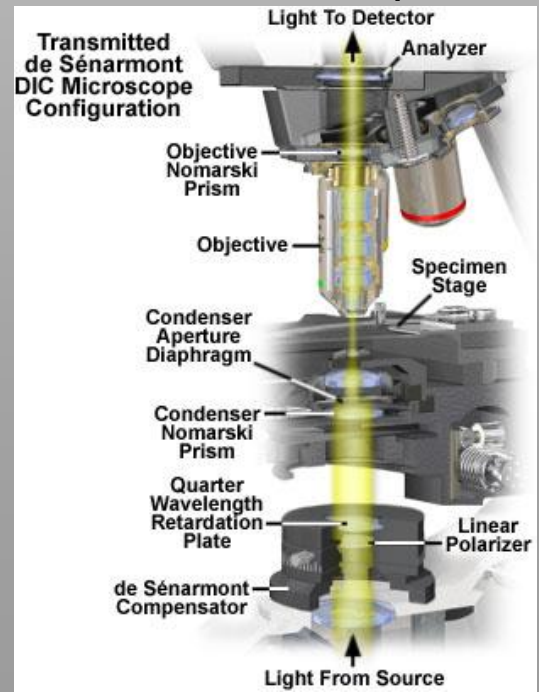
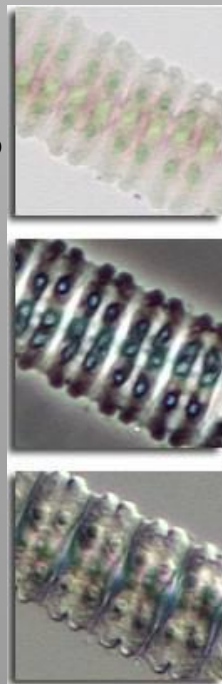


Phase Contrast Microscope



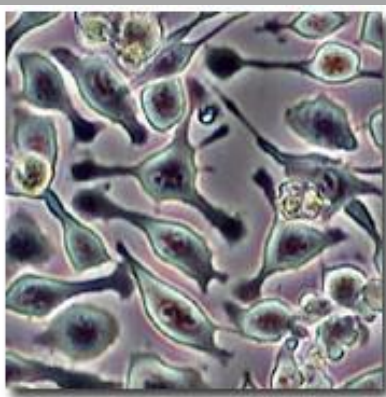
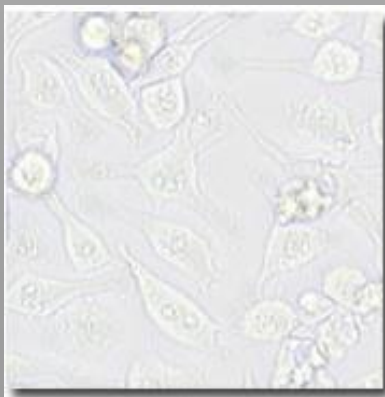
DIC Microscope

Phase Contrast Brightfield
DIC



Brightfield

Phase Contrast



<http://www.microscopyu.com/articles/phasecontrast/phasemicroscopy.html>

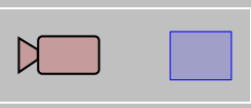
Schlieren Imaging



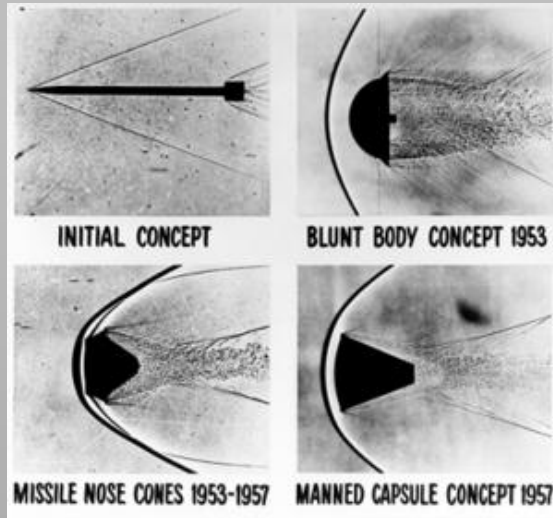
images courtesy Gary Settles



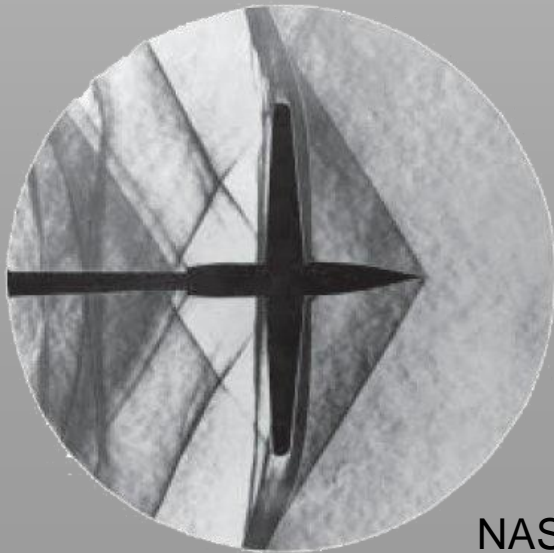
Schlieren Imaging – Applications



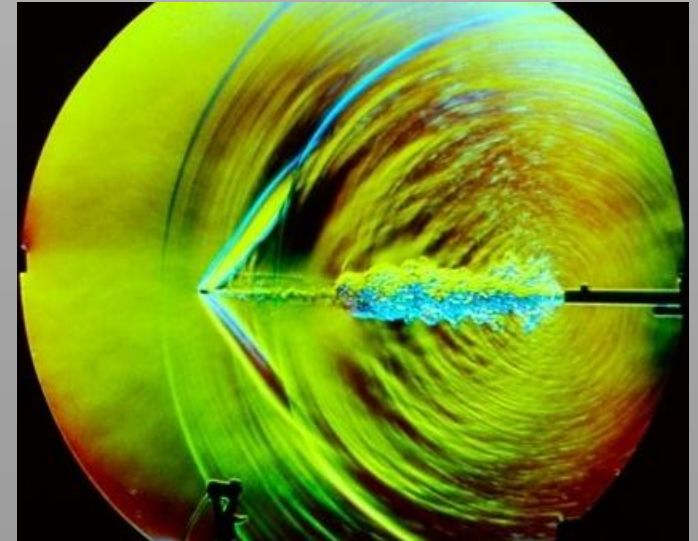
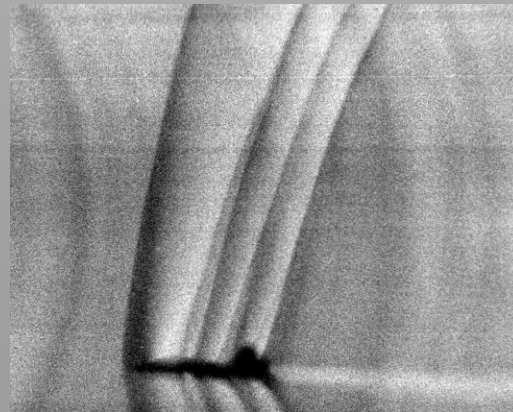
Aeronautical engineering



Ballistics

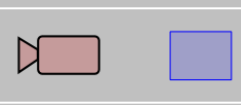


NASA



Gary Settles

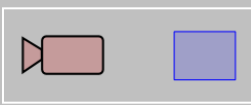
Schlieren Imaging – Optical Setups



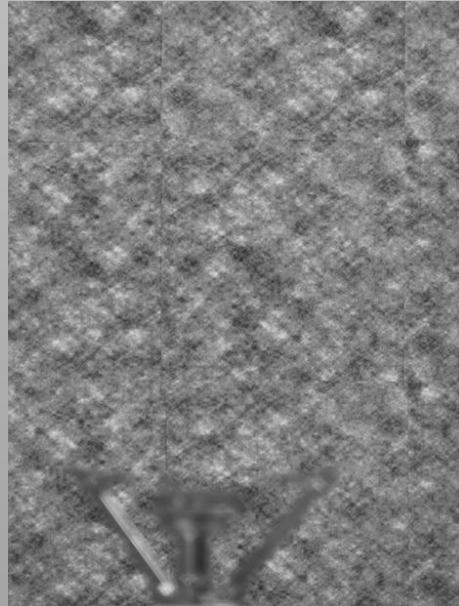
Imperial College London NASA



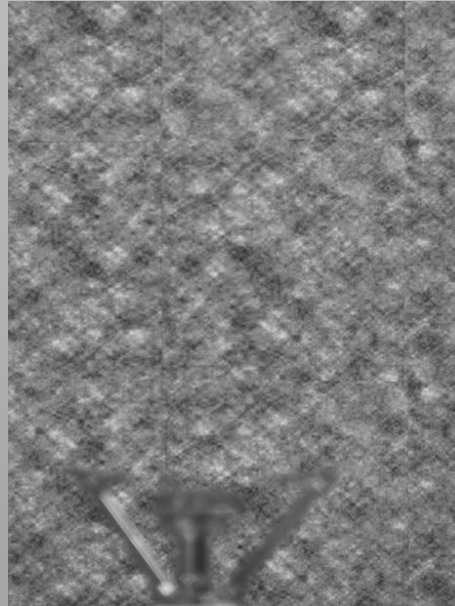
Background Oriented Schlieren Imaging



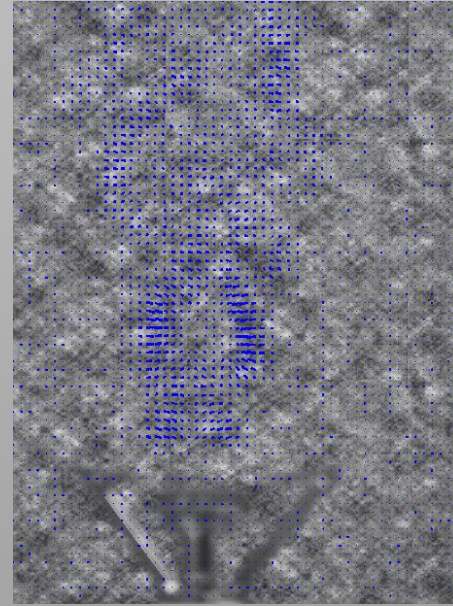
Optical Encoding \rightarrow Computational Processing



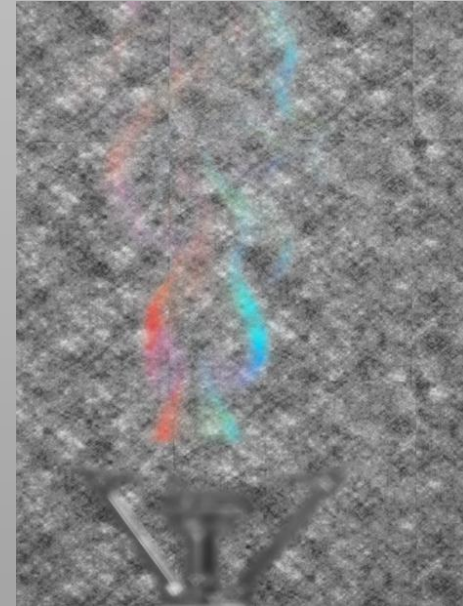
Undistorted
Reference Image



Distorted Image



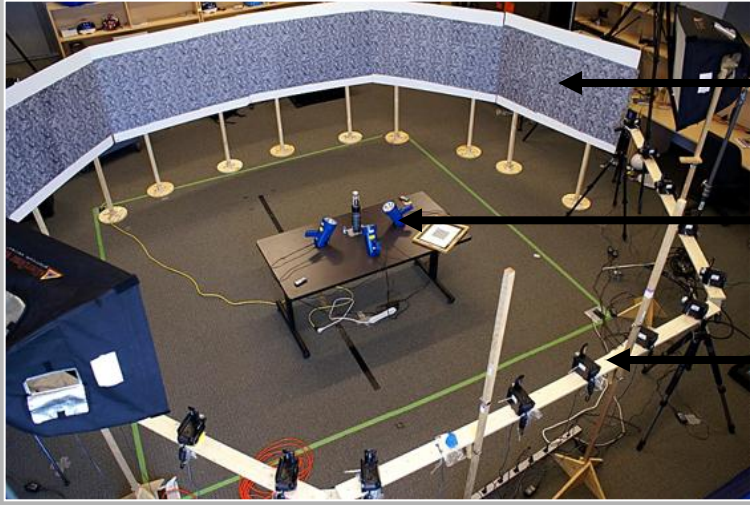
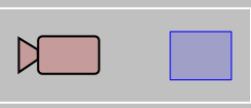
Optical Flow Vectors



Color Coded
Optical Flow Vectors

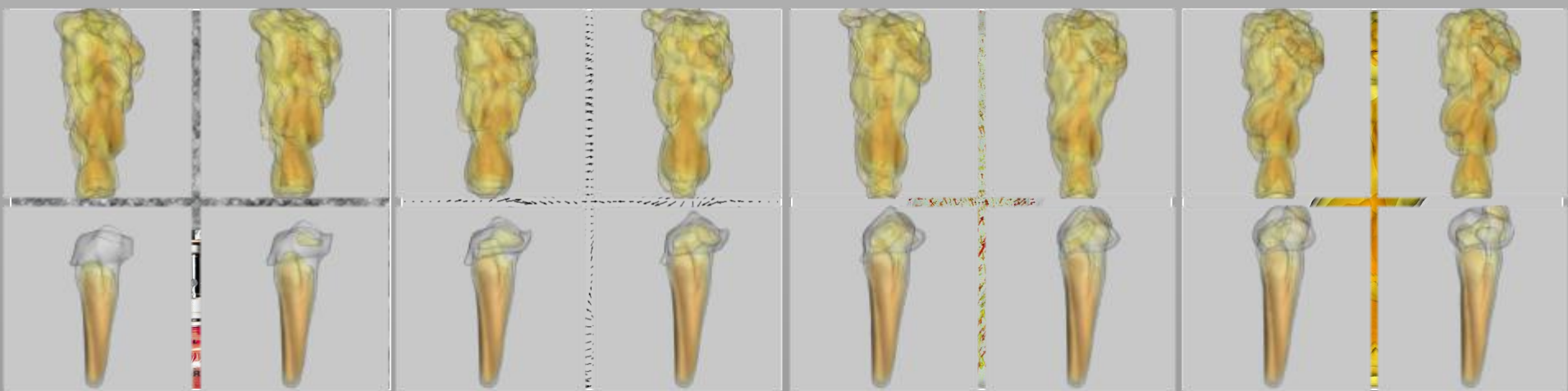
[Atcheson et al. 2008]

BOS – Tomographic Reconstruction



- High Frequency Background
- Transparent, Refractive Object
- Camera Array

Optical Setup



Camera Image

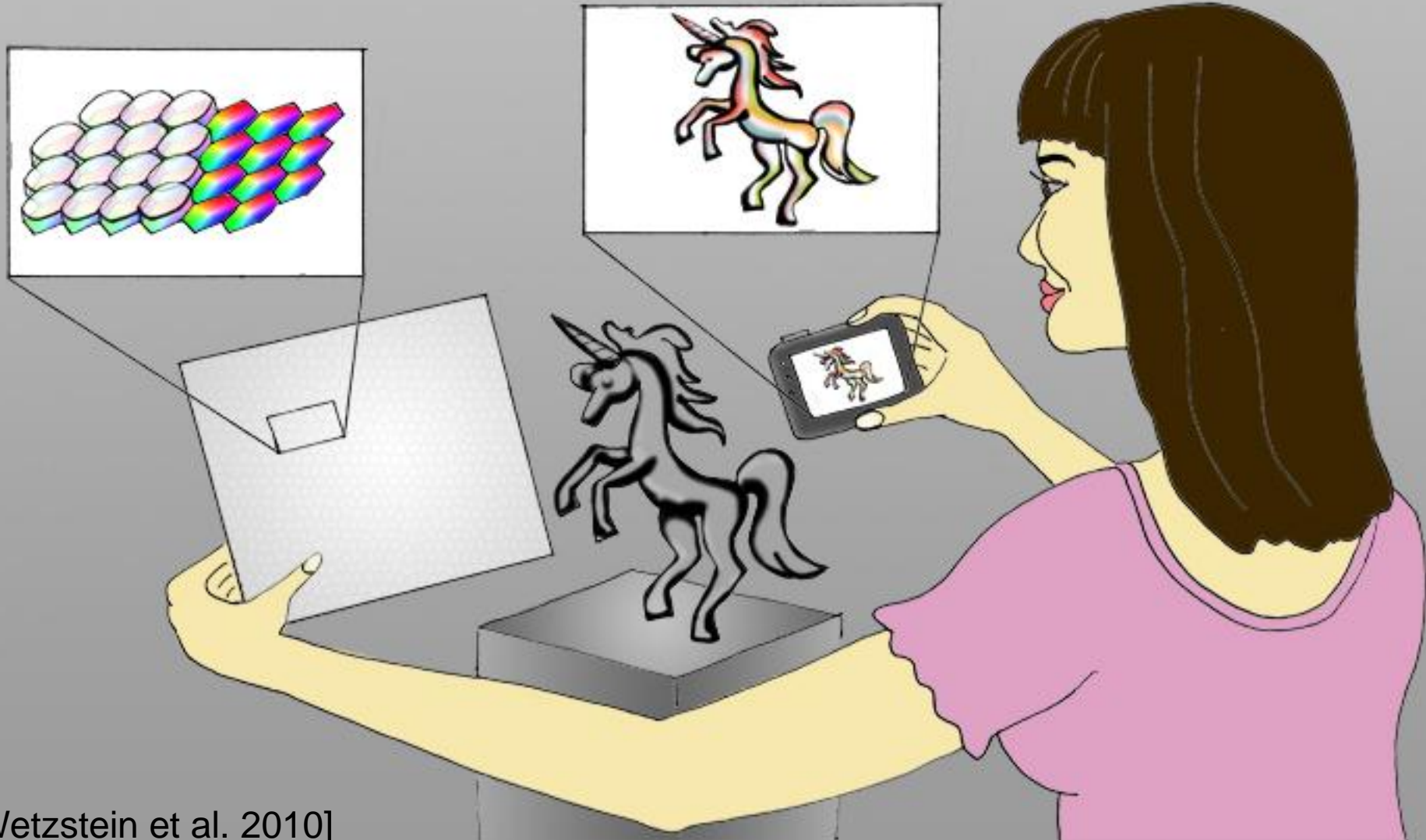
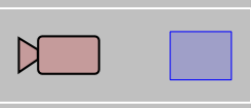
2D Optical Flow

3D Gradient Field

3D Refractive Index

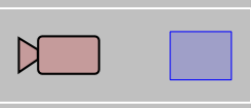
[Atcheson et al. 2008]

Light Field Background Oriented Schlieren



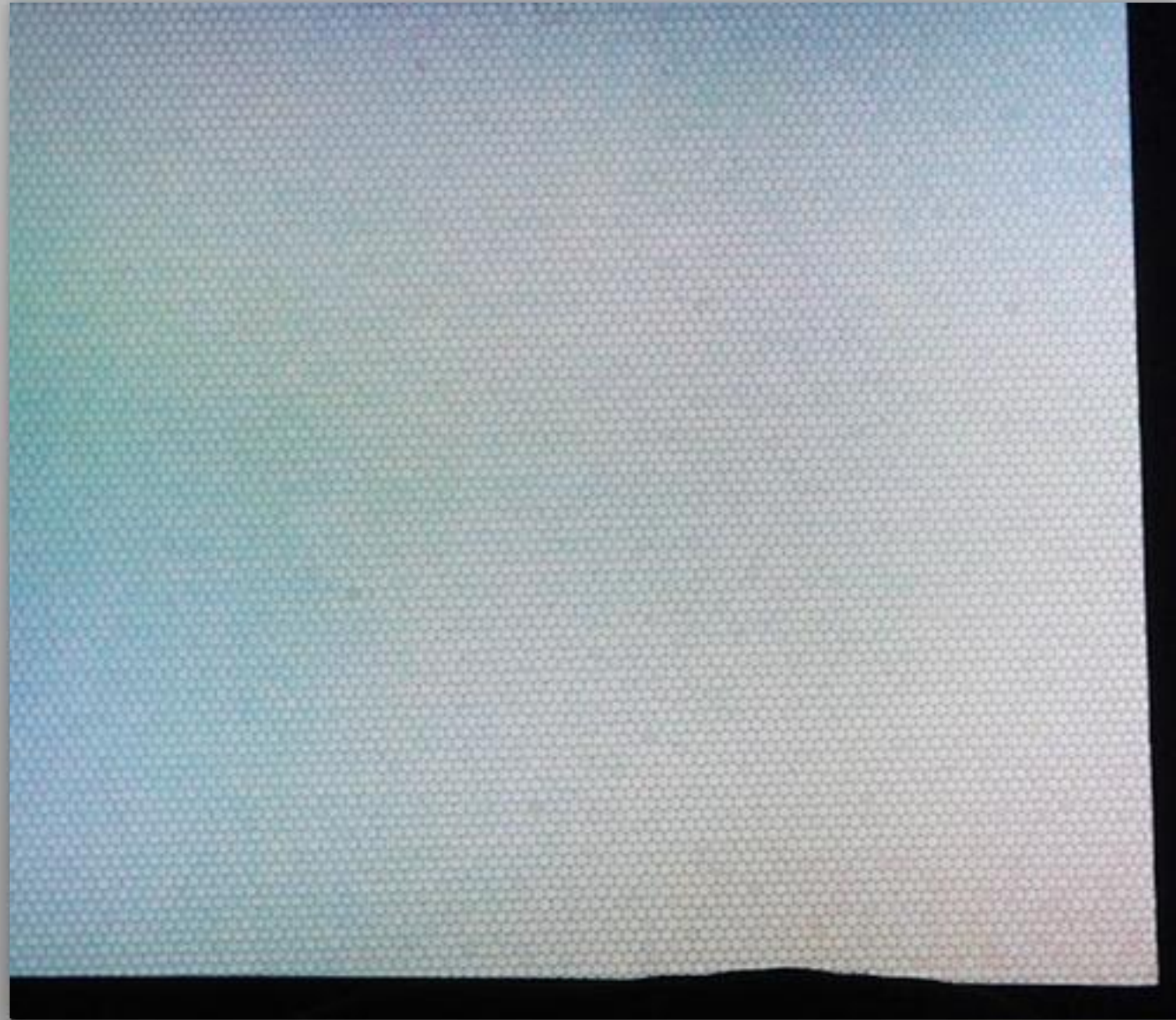
[Wetzstein et al. 2010]

Light Field Background Oriented Schlieren



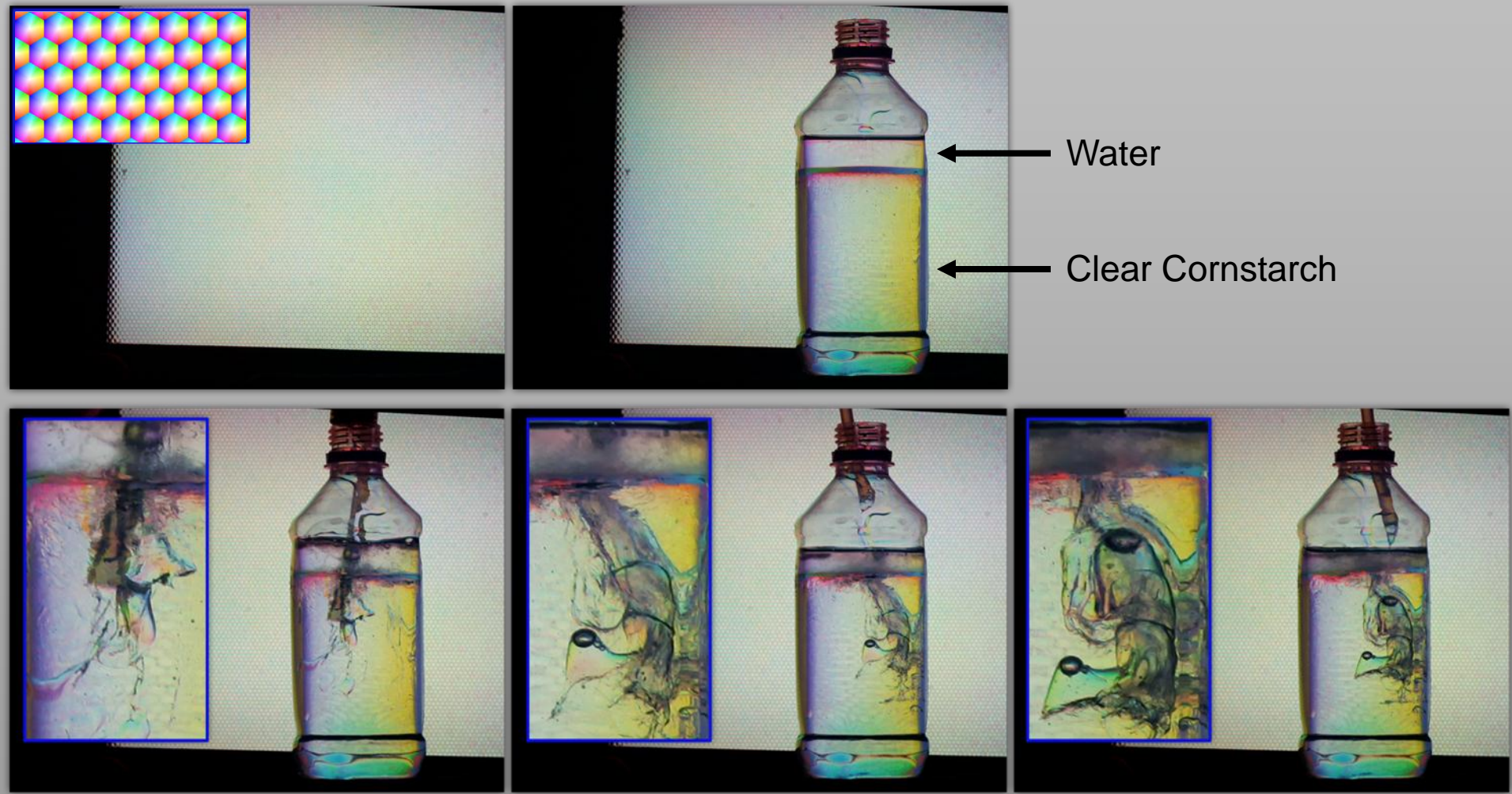
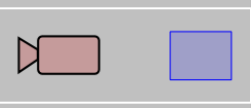
Uniform Illumination

Light Field Probe



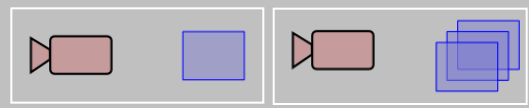
[Wetzstein et al. 2010]

Light Field Background Oriented Schlieren

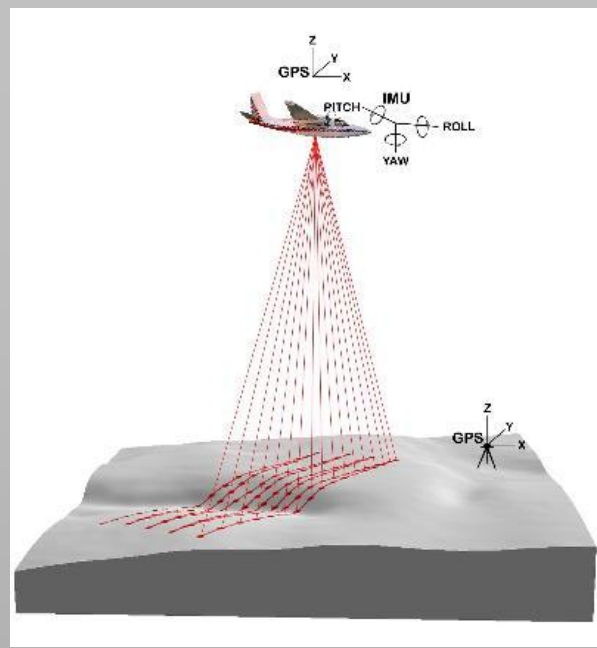


[Wetzstein et al. 2010]

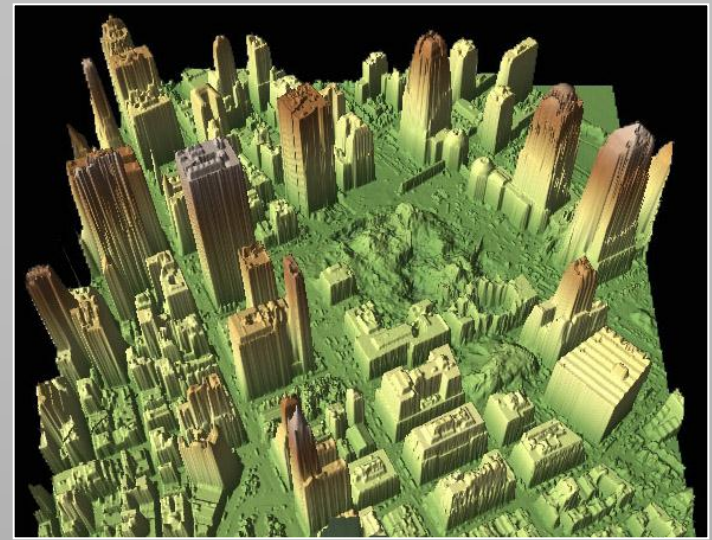
VII.III LIDAR and Time-of-Flight



LIDAR



University of Washington



NASA

Time-of-Flight Cameras



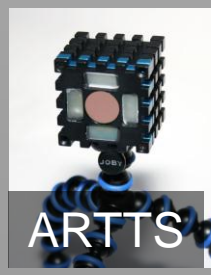
PMD



Panasonic



Fotonics



ARTTS



MESA

See [Kolb et al. 10], EG STAR for more details

Next: Discussion

Computational Plenoptic Imaging

Gordon Wetzstein¹
Wolfgang Heidrich³

Ivo Ihrke²
Kurt Akeley⁴

Douglas Lanman¹
Ramesh Raskar¹

¹MIT Media Lab

²Saarland University

³University of British Columbia

⁴Lytro, Inc.

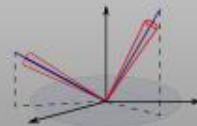
VIII. Discussion



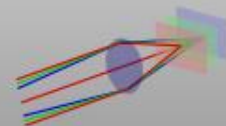
Dynamic Range



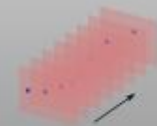
Color Spectrum



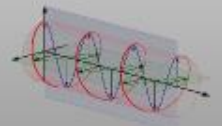
Directions | Light Fields



Space | Focal Surfaces



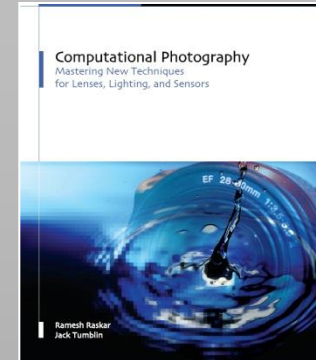
Time



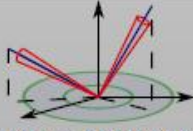
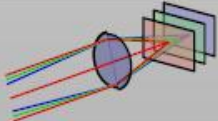
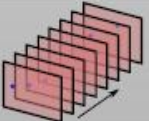
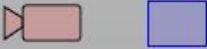

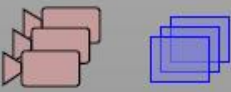


Further Properties

Summary

- Survey of plenoptic image acquisition
- Classification based on plenoptic dimension & hardware setup
- Also see computational photography

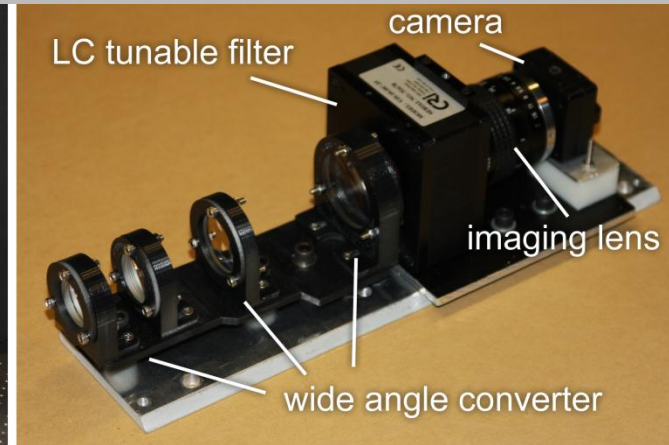
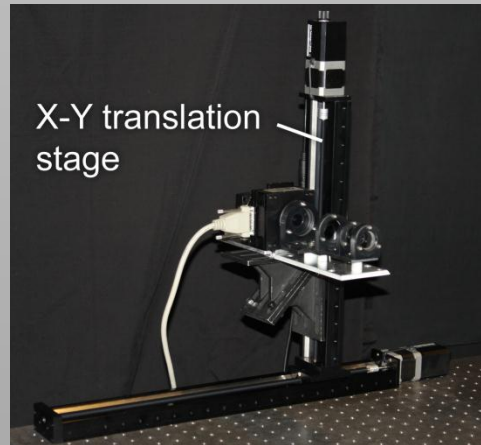
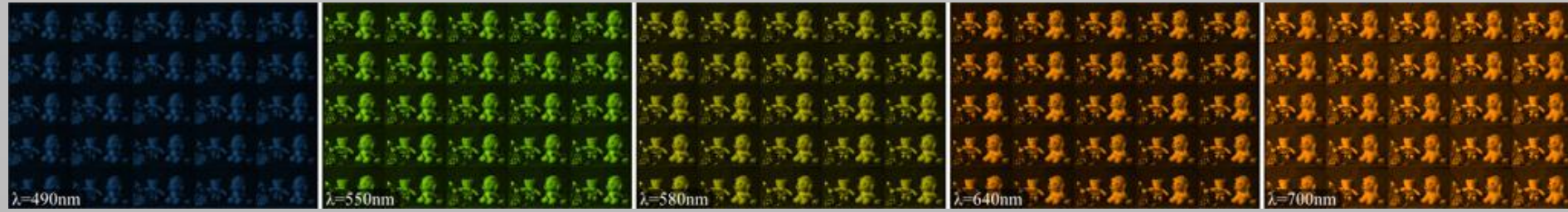


Plenoptic Dimension					
Acquisition Approach	Dynamic Range	Color Spectrum	Directions Light Fields	Space Focal Surfaces	Time
 <p>Single Shot Acquisition</p>	<p>Assorted Pixels Gradient Camera Adaptive DR Imaging</p>	<p>Color Filter Arrays Assorted Pixels Dispersive Optics</p>	<p>Plenoptic Cameras w/ Lenses, Masks, or Mirrors Compound Eye Cameras</p>	<p>Coded Apertures Focal Sweep Field Correction</p>	<p>Assorted Pixels Flutter Shutter Reinterpretable Imager Sensor Motion</p>
 <p>Sequential Image Capture</p>	<p>Exposure Brackets Generalized Mosaics HDR Video</p>	<p>Narrow Band Filters Generalized Mosaicing Agile Spectrum Imaging</p>	<p>Programmable Aperture Camera & Gantry</p>	<p>Focal Stack Jitter Camera Super-Resolution</p>	<p>High-Speed Imaging Temporal Dithering</p>
 <p>Multi-Device Setup</p>	<p>Split Aperture Imaging Optical Splitting Trees</p>	<p>Multi-Camera Arrays Optical Splitting Trees</p>	<p>Multi-Camera Arrays</p>	<p>Multi-Camera Arrays</p>	<p>Multi-Camera Arrays Hybrid Cameras</p>

Observations

- Most approaches use **fixed plenoptic resolution tradeoffs**
- Strong **correlations** between plenoptic dimensions
- Need for **sophisticated reconstruction** techniques (e.g. compressive sensing)

Future Directions – Exploit Plenoptic Redundancy



- Plenoptic datasets
- Simulate acquisition & reconstruction
- Explore redundancies

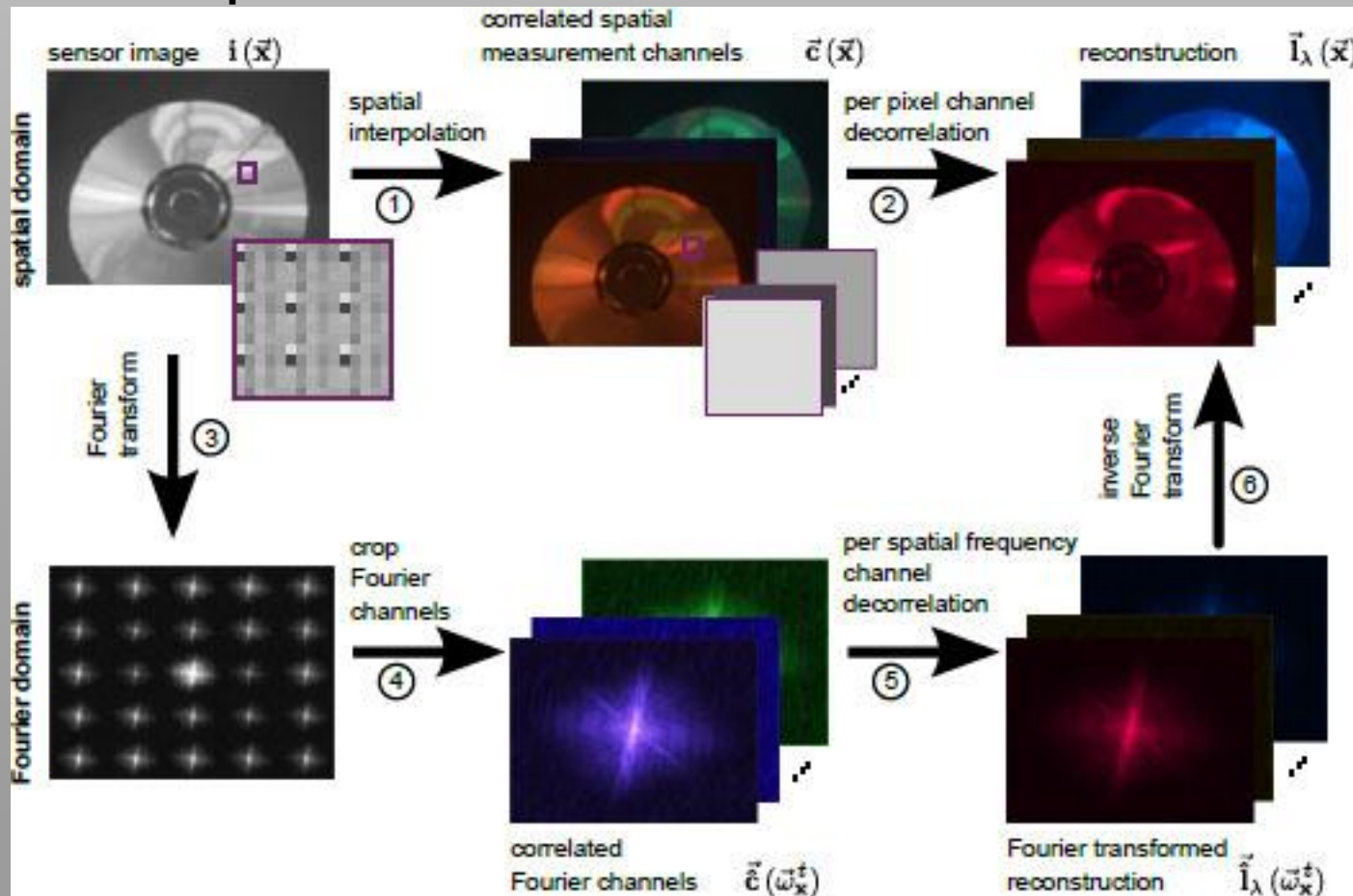
The End

Future Directions – Exploit Plenoptic Redundancy

- Explore plenoptic priors – mathematical formulations for correlations between and within dimensions
- Common practice in
 - Color demosaicking
 - Extended DOF and light field acquisition (dimensionality gap prior) [Levin et al. 09,10]
- Extend to time, polarization, plenoptic manifolds

Future Directions – Unified Plenoptic Reconstruction

- Unified reconstruction in terms of
 - Domain (image space vs. Fourier)
 - Plenoptic dimension



[Ihrke et al. 10]