Computational Photography

SIGGRAPH 2007 Course 1 Notes

Organizers

Ramesh Raskar
Senior Research Scientist
MERL - Mitsubishi Electric Research Labs
201 Broadway
Cambridge, MA 02139
Email: raskar@merl.com
http://www.merl.com/people/raskar/

Jack Tumblin
Assistant Professor
Computer Science Dept., Northwestern University
1890 Maple Avenue, Evanston IL 60201,
Email: jet@cs.northwestern.edu
http://www.cs.northwestern.edu/~jet

Course Web Page

http://www.merl.com/people/raskar/photo/
Acknowledgements

We would like to thank Amit Agrawal for his help on several aspects of this course and for providing the pseudo-code included in these notes.

We would also like to thank the following for providing us their slides and videos which have been extensively used throughout this course.

Stanford University: Bennett Wilburn, Vaibhav Vaish, Ren Ng

Washington University: Aseem Agrawala

MERL/Brown University: Morgan McGuire

University of Pennsylvania: Jianbo Shi
Course Description

Computational photography combines plentiful computing, digital sensors, actuators, and lights to escape the limitations of traditional film cameras. Unbounded dynamic range, variable focus, resolution, and depth of field, hints about shape, reflectance, and lighting, and new interactive forms of photos that are partly snapshots and partly videos are just some of the new applications found in computational photography.

Many computational photography ideas are new to digital artists and programmers, even those very familiar with film and digital photography techniques. This emerging multi-disciplinary field combines new and old ideas from both image capture and computational methods for images that may present a steep learning curve. For example, few photographers may know recent high dynamic range imaging methods, and image processing researchers face rapidly changing capture, alignment and noise issues in arrays of digital cameras. These topics, however, can be easily learned without extensive background. The goal of this course is to present both aspects in a compact form.

The course briefly reviews fundamental topics in digital imaging and then provides a practical guide to underlying techniques beyond image processing such as gradient domain operations, graph cuts, bilateral filters and optimizations. We also cover topics in applied optics such as ray-transfer matrix and Fourier optics. Computational capture methods include sophisticated sensors, light sources, and on-board processing. Examples include adaptation to sensed scene depth and illumination, interactive pictures made by varying camera parameters or actively modifying the flash illumination parameters. Computational reconstruction methods include 'photomontage' that optimally fuses information from multiple images, improves signal to noise ratio and extracts scene features such as depth edges.

The participants learn about topics in image capture and manipulation methods for generating compelling pictures for computer graphics and for extracting scene properties for computer vision, with several examples. We hope to provide enough fundamentals to satisfy the technical specialist without intimidating the curious graphics researchers interested in photography.

Prerequisites

A basic understanding of camera operation and image processing is required. Familiarity with concepts of linear systems, convolution, and machine vision will be useful.

Photographers, digital artists, image processing programmers and vision researchers using or building applications for digital cameras or images will learn about camera fundamentals and powerful computational tools, along with many real world examples.
Course Schedule

A.1 Introduction (Raskar, 10 minutes)
A.2 Concepts in Computational Photography (Tumblin, 30 minutes)
A.3 Understanding Film-like Photography (Raskar, 30 minutes)
A.4 Image Processing Tools (Tumblin, 30 mins)
   Q&A (5 minutes)

B.1 Improving Film-like Photography (Tumblin, 20 minutes)
B.2 Computational Camera (Nayar, 60 minutes)
B.3 Image reconstruction techniques (Tumblin, 20 minutes)
   Q&A (5 minutes)

C.1 Multi-perspective Photography (Davidhazy, 35 minutes)
C.2 Lightfield photography and microscopy (Levoy, 30 minutes)
C.3 Fourier Analysis of Light Fields (Georgiev, 35 minutes)
   Q&A (5 minutes)

D.1 Computational Illumination (Raskar, 40 minutes)
D.2 Computational Imaging in the Sciences (Levoy, 30 minutes)
D.3 Future Cameras (Raskar, 20 minutes)
   Summary and Discussion (15 minutes)

Topics not covered: film cameras, advanced optics, traditional image processing, image based rendering (IBR) and novel view synthesis, lighting and illumination hardware technologies, camera assisted projector display systems, geometric calibration and photometric calibration techniques, compression, storage, photo-editing software packages, file formats and standards.

Course Web Page  http://www.merl.com/people/raskar/photo/
Lecturer Biographies

Marc Levoy

Marc Levoy is an Associate Professor of Computer Science and Electrical Engineering at Stanford University. He received his PhD in Computer Science from the University of North Carolina in 1989. In the 1970's Levoy worked on computer animation, developing an early computer-assisted cartoon animation system. In the 1980's Levoy worked on volume rendering, a family of techniques for displaying sampled three-dimensional functions, such as CT and MR data. In the 1990's he worked on technology and algorithms for 3D scanning. This led to the Digital Michelangelo Project, in which he and a team of researchers spent a year in Italy digitizing the statues of Michelangelo using laser rangefinders. His current interests include light field sensing and display, computational imaging, and digital photography. Levoy received the NSF Presidential Young Investigator Award in 1991 and the SIGGRAPH Computer Graphics Achievement Award in 1996 for his work in volume rendering.

http://graphics.stanford.edu/~levoy/

Shree Nayar

Shree K. Nayar received his PhD degree in Electrical and Computer Engineering from the Robotics Institute at Carnegie Mellon University in 1990. He heads the Columbia Automated Vision Environment (CAVE), which is dedicated to the development of advanced computer vision systems. His research is focused on three areas; the creation of novel vision sensors, the design of physics based models for vision, and the development of algorithms for scene understanding. His work is motivated by applications in the fields of digital imaging, computer graphics, and robotics. Professor Nayar has received best paper awards at ICCV 1990, ICPR 1994, CVPR 1994, ICCV 1995, CVPR 2000 and CVPR 2004. He is the recipient of the David and Lucile Packard Fellowship (1992), the National Young Investigator Award (1993), the NTT Distinguished Scientific Achievement Award (1994), and the Keck Foundation Award for Excellence in Teaching (1995). He has published over 100 scientific papers and has several patents on inventions related to vision and robotics.

http://www.cs.columbia.edu/~nayar/
Andrew Davidhazy

Professor Andrew Davidhazy is a professor in the Imaging and Photographic Technology Department of the School of Photographic Arts and Sciences at the Rochester Institute of Technology (RIT) in Rochester, NY. He is a teacher with over 30 years of experience and while specializing in scientific and technical aspects of photography he is almost equally active in the application of technical imaging concepts to aesthetic purposes.

He was a NASA/ASEE Research Fellow in 1994 at NASA Langley Research Center, VA. and the inaugural Kodak Visiting Professor to RMIT in Australia. He was the recipient of the Eisenhart Award for Outstanding Teaching at RIT and the 1990 Professor Raymond C. Bowman Award from the Society for Imaging Science and Technology and was awarded the Society Fellowship in 2001. He is a Fellow of SPIE, the International Society for Optical Engineering and an Associate member of the Royal Photographic Society. He has published and lectured widely on the general topic of "Simplified approaches to Strip and Streak Photography and Scanning Photographic Systems", as well as many other topics related to photographic instrumentation, as invited speaker to conferences, workshops and seminars worldwide. He collaborated with Drs. Leslie Stroebel and Ronald Francis on an investigation for the House Select Committee on the assassination of President Kennedy and has been a consultant in photographic instrumentation for many industrial and governmental agencies.

His writings and photographs have been published in numerous books, articles and journals including Popular Photography, American Photo, Industrial Photography, Camera, etc. His specialty is in the area of scanning photography, especially panoramic and peripheral portraiture. He has developed an unusual rotating film panoramic and peripheral camera and built several prototype scanning cameras and enlargers. With the latter, he has made enlargements that exceed 100 continuous feet in length.

http://www.rit.edu/~andpph

Todor Georgiev

Todor Georgiev is a Senior Research Scientist at Adobe Systems, working in the Photoshop group. He received his PhD in Physics from Southern Illinois University in 1996. He concentrates on applications of mathematical methods taken from theoretical physics to image processing, graphics and vision. He is the author of the Healing Brush tool in Photoshop (2002), the method better known as Poisson image editing. He has published several articles on applications of the mathematics of covariant derivatives in image processing and vision. He is also interested in a wide range of theoretical and practical aspects of optics, light field cameras and capture/manipulations of the optical field. This naturally leads to his recent works on light field camera designs, view interpolation and vision techniques. He has several papers and patents in the related areas.

http://www.tgeorgiev.net
Ramesh Raskar

Ramesh Raskar is a Senior Research Scientist at MERL. His research interests include projector-based graphics, computational photography and non-photorealistic rendering. He has published several articles on imaging and photography including multi-flash photography for depth edge detection, image fusion, gradient-domain imaging and projector-camera systems. His papers have appeared in SIGGRAPH, EuroGraphics, IEEE Visualization, CVPR and many other graphics and vision conferences. He was a course organizer at Siggraph 2002 through 2005. He was the panel organizer at the Symposium on Computational Photography and Video in Cambridge, MA in May 2005 and taught a graduate level class on Computational Photography at Northeastern University, Fall 2005. He is a member of the ACM and IEEE.

http://www.merl.com/people/raskar/

Jack Tumblin

Jack Tumblin is an assistant professor of computer science at Northwestern University. His interests include novel photographic sensors and lighting devices to assist museum curators in historical preservation, computer graphics and visual appearance, and image-based modeling and rendering. During his doctoral studies at Georgia Tech and post-doc at Cornell, he investigated tone-mapping methods to depict high-contrast scenes. His MS in Electrical Engineering (December 1990) and BSEE (1978), also from Georgia Tech, bracketed his work as co-founder of IVEX Corp., (>45 people as of 1990) where his flight simulator design work was granted 5 US Patents. He was co-organizer of a course on Computational Photography at Siggraph 2005. He is an Associate Editor of ACM Transactions on Graphics.

http://www.cs.northwestern.edu/~jet
Copyright Information

The following copyright information appears on the first page of each corresponding paper.

© 2007 IEEE. Reprinted, with permission, from (…).

© 2007 ACM Reprinted, with permission, from (…).

All trademarks are the property of their respective owners. Papers and Photos are copyrighted and shall not be reproduced without permission.
Welcome

- Understanding Film-like Photography
  - Parameters, Nonlinearities, Ray-based concepts
- Image Processing and Reconstruction Tools
  - Multi-image Fusion, Gradient domain, Graph Cuts
- Improving Camera Performance
  - Better dynamic range, focus, frame rate, resolution
- Future Directions
  - HDR cameras, Gradient sensing, Smart optics/lighting

Goals

- Review of 30+ recent papers
- Understand computational aspects of cameras
  - Discuss issues in traditional cameras
  - Explore alternative imaging methods
  - Learn vision and optics techniques
- Discuss image processing and reconstruction tools
- What we will not cover
  - Film Cameras, Novel view rendering (IBR), Color issues, Traditional image processing/editing

Speaker: Marc Levoy

Marc Levoy is an Associate Professor of Computer Science and Electrical Engineering at Stanford University. He received his PhD in Computer Science from the University of North Carolina in 1989. In the 1970's Levoy worked on computer animation, developing an early computer-aided cartoon animation system. In the 1980's Levoy worked on volume rendering, a family of techniques for displaying sampled three-dimensional functions, such as CT and MRI data. In the 1990's he worked on technology and algorithms for 3D scanning. This led to the Digital Michelangelo Project, in which he and a team of researchers spent a year in Italy digitizing the statues of Michelangelo using laser rangefinders. His current interests include light field sensing and display, computational imaging, and digital photography. Levoy received the NSF Presidential Young Investigator Award in 1991 and the SIGGRAPH Computer Graphics Achievement Award in 1996 for his work in volume rendering.

http://graphics.stanford.edu/~levoy/

Speaker: Shree Nayar

Shree K. Nayar is a Professor at Columbia University. He received his PhD degree in Electrical and Computer Engineering from the Robotics Institute at Carnegie Mellon University in 1990. He heads the Columbia Automated Vision Environment (CAVE), which is dedicated to the development of advanced computer vision systems. His research is focused on three areas: the creation of novel vision sensors, the design of physics-based models for vision, and the development of algorithms for scene understanding. His work is motivated by applications in the fields of digital imaging, computer graphics, and robotics. Professor Nayar has received best paper awards at ICCV 1995, ICPR 1994, CVPR 1994, ICCV 1993, CVPR 2000 and CVPR 2004. He is the recipient of the David and Lucile Packard Fellowship (1992), the NTT Distinguished Scientific Achievement Award (1994), and the Keck Foundation Award for Excellence in Teaching (1996). He has published over 100 scientific papers and has several patents on inventions related to vision and robotics.

http://www.cs.columbia.edu/~nayar/
Ramesh Raskar is a Senior Research Scientist at MERL. His research interests include projector-based graphics, computational photography and non-photorealistic rendering. He has published several articles on imaging and photography including multi-flash photography for depth edge detection, image fusion, gradient-domain imaging and projector-camera systems. His papers have appeared in SIGGRAPH, EuroGraphics, IEEE Visualization, CVPR and many other graphics and vision conferences. He was a course organiser at Siggraph 2002 through 2005. He was the panel organizer of the Symposium on Computational Photography and Video in Cambridge, MA in May 2005 and taught a graduate level class on Computational Photography at Northeastern University, Fall 2005. He is a member of the ACM and IEEE.

http://www.merl.com/people/raskar/raskar.html

Jack Tumblin is an Assistant Professor of Computer Science at Northwestern University. His interests include novel photographic sensors to assist museum curators in historical preservation, computer graphics and visual appearance, and image-based modeling and rendering. During his doctoral studies at Georgia Tech and post-doc at Cornell, he investigated tone-mapping methods to depict high-contrast scenes. He MS in Electrical Engineering (December 1995) and BSEE (1978), also from Georgia Tech, bracketed his work as co-founder of IVEX Corp., (>45 people as of 1990) where his flight simulator design work was granted 5 US Patents. He is an Associate Editor of ACM Transactions on Graphics, was a member of the SIGGRAPH Papers Committee (2003, 2004), and in 2001 was a Guest Editor of IEEE Computer Graphics and Applications.

http://www.cs.northwestern.edu/~jet

Professor Andrew Davidhazy is a professor in the Imaging and Photographic Technology Department of the School of Photographic Arts and Sciences at the Rochester Institute of Technology (RIT) in Rochester, NY. He is a teacher with over 30 years of experience and while specializing in the technical aspects of photography he is almost equally active in the application of technical imaging concepts to aesthetic purposes. He was a NASA/ASEE Research Fellow in 1994 at NASA Langley Research Center, VA, and the inaugural Kodak Visiting Professor to RMIT in Australia. He was the recipient of the Eisenhart Award for Outstanding Teaching at RIT and the 1995 Professor Raymond C. Bowman Award from the Society for Imaging Science and Technology and was awarded the Society Fellowship in 2001. He is a Fellow of SPIE, the International Society for Optical Engineering and an Associate member of the Royal Photographic Society.

http://www.rit.edu/~andpph

Todor Georgiev is a Senior Research Scientist at Adobe Systems, working in the Photoshop group. He received his PhD in Physics from Southern Illinois University in 1996. He concentrates on applications of mathematical methods taken from theoretical physics to image processing, graphics and vision. He is the author of the Healing Brush tool in Photoshop (2002), the method better known as Poisson image editing. He has published several articles on applications of the mathematics of covariant derivatives in image processing and vision. He is also interested in a wide range of theoretical and practical aspects of optics, light field cameras and capture/manipulations of the optical field. This naturally leads to his recent work on light field camera design, view interpolation and vision techniques. He has several papers and patents in the related areas.

http://www.tgeorgiev.net

Opportunities

- Unlocking Photography
  - How to expand camera capabilities
  - Digital photography that goes beyond film-like photography
- Think beyond post-capture image processing
  - Computation well before image processing and editing
- Learn how to build your own camera-toys
- Review of 30+ recent papers
- What we will not cover
  - Film Cameras, Novel view rendering (IBR), Color issues, Traditional image processing/editing

Traditional Photography

Detector

Lens

Pixels

Image

Courtesy: Shree Nayar
Radial Stereoscopic Imaging

Dual photography from diffuse reflections

Fluttered Shutter Photography

IEEE Computer Special Issue on Computational Photography

Computational Photography

Siggraph 2006

Hybrid Images
- Chou et al. (MIT)
- Drag-and-Follow Pasting
- Xu et al. (MIT)

Two-tone Tone Management for Photographs
- Agarwala et al. (Georgia Tech)
- Greenspan et al. (MIT)

Interactive Local Adjustment of Tonal Values
- Lischinski et al. (Tel Aviv)

Image-Based Material Editing
- Wan et al. (Princeton)

Flash Matting
- Sun et al. (Microsoft Research Asia)

Natural Video Matting Using Camera Arrays
- Judd et al. (UCSD / MERL)

Removing Camera Shake From a Single Photograph
- Fergus (MIT)

Coded Exposure Photography: Motion Deblurring
- Raskar et al. (MERL)

Photo Tourism: Exploring Photo Collections in 3D
- Snavely et al. (Stanford)

AutoCollage
- Rother et al. (Microsoft Research Cambridge)

Photographing Long Scenes with Multiple Viewpoint Cameras
- Agarwala et al. (University of Washington)

Projection Refocus Analysis for Scene Capture and Image Display
- Zhang et al. (Columbia University)

Multi-view Radial Catadioptric Imaging for Scene Capture
- Kuthirummal et al. (Columbia University)

Light Field Microscopy (Project)
- Levoy et al. (Stanford University)

Fast Separation of Direct and Global Components of a Scene Using High Frequency Illumination
- Nayar et al. (Columbia University)
A2: What is the Ideal Photographic Signal?

Jack Tumblin
Northwestern University

What is Photography?

Safe answer:

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

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

- A ‘bucket’ word: a neat container for messy notions (e.g. aviation, music, comprehension)

- A record of what we see, or would like to see, in tangible form.

- Does photography always capture it? no.

- So, what do we see?

‘Film-Like’ Photography

Film Camera designs still dominate:
- ‘Instantaneous’ light measurement…
- Of focal plane image behind a lens.
- Reproduce those amounts of light;
- EXACT MATCH!

Implied:
“What we see is \( \cong \) focal-plane intensities.” well, no…
Why we like Photography

**PHYSICAL**
- 3D Scene: light sources, BRDFs, shapes, positions, movements, …
- Eyepoint: position, movement, projection, …

**Exposure**
- Light & Optics
- Exposure Control, tone map

**Display**
- RGB(x,y,t)

**Tangible Record**
- Editable, storable as Film or Pixels

**PERCEIVED**
- Scene: light sources, BRDFs, shapes, positions, movements, …
- Eyepoint: position, movement, projection, …

Vision

What’s Beyond Film-Like Photography?

Thought Experiment:

• COMPARE:
  - Digital Camera result.
  - Digitized (Scanned) Film result.

? Can we See more, Do more, Feel more?

? Has photography really *changed* yet?
What other ways better reveal appearance to human viewers? (Without direct shape measurement?)

Can you understand this shape better?

Goals for a New Photography

Vision

Sensor(s)

Computing

Light & Optics

Tangible Record

estimates we can capture, edit, store, display...

3D Scene

Light sources, BRDFs, shapes, positions, movements...

Eyepoint...

projection...

Meaning...

PHYSICAL

PERCEIVED or UNDERSTOOD

missing: Dynamic Display, Interactive...

What other ways better reveal appearance to human viewers? (Without direct shape measurement?)

Can you understand this shape better?


Can you understand this shape better?

What other ways better reveal appearance to human viewers? (Without direct shape measurement?)

Can you understand this shape better?
Missing: **More Revealing Sets of Rays**

*Multiple-Center-of-Projection Images* Rademacher, P, Bishop, G., SIGGRAPH '98

---

**Taking Images versus Taking Pictures**

**Image:** A copy of light intensities.

(Just *one kind* of picture, made by copying a scaled map of scene light intensities, as a lens might)

**Visual Appearance:** What we *think* we see.

(Consciously-available estimates of our surroundings, made from the light reaching our eyes)

**Picture:** A ‘container’ for visual appearance.

(something we make to hold what we see, or what would like to see. Includes non-photorealistic drawings)
Photographic Signal: Pixels, Rays

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

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

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

The Photographic Signal Path

Computing can improve every component:

- Light Sources
- Sensors
- Data Types, Processing
- Display

“Scene”
Review: How many Rays are there?

4-D set; infinitesimal members. Imagine:
- Convex Enclosure of a 3D scene
- Inward-facing ray camera at every surface point
- Pick the rays you need for ANY camera outside.
- 2D surface of cameras,
  + 2D ray set for each camera,
  $\rightarrow$ 4D set of rays.

4-D Light Field / Lumigraph

Measure all the \textit{outgoing} light rays.
4-D Illumination Field

Same Idea: Measure all the *incoming* light rays

4D x 4D = 8-D Reflectance Field

**Ratio:** \( R_{ij} = \frac{\text{outgoing ray}_i}{\text{incoming ray}_j} \)
Is a 4-D Light Source Required?

- Multiple dynamic Virtual Viewpoints
- Efficient Bandwidth usage: ‘send only what you see’
- Yang, et al 2002
- 64 tightly packed commodity CMOS webcams
- 30 Hz, Scaleable, Real-time:

Is A 4D Camera Required?

e.g. MIT Dynamic Light Field Camera 2002

Is this the whole answer?

or is it just "more film-like cameras, but now with computers!"?
Or do Ray Changes Convey Appearance?

5 ray sets → explicit geometric occlusion boundaries

• These rays + all these rays give me…

• MANY more useful details I can examine…
**Mild Viewing & Lighting Changes; Are these Enough?**

Convicing visual appearance:
Is Accurate Depth really necessary?

a few good 2-D images may be enough...

---

**‘The Ideal Photographic Signal’**

I CLAIM IT IS:

All Rays? Some Rays? Changes in Rays

Photographic ray space is vast and boring.
>8 dimensions: 4D view, 4D light, time, λ,
Gather only ‘visually significant’ ray changes

? What rays should we measure?
? How should we combine them?
? How should we display them?
Future Photography:

Novel Cameras:
- Generalized Sensor
  - Ray Reconstructor
- Generalized Processing
  - 4D Ray Sampler
- General Optics:
  - 4D Ray Benders

Novel Displays:
- Generalized Display
  - Recreated 4D Light field

Novel Illuminators:
- Lights
  - 4D Modulators
- General Optics:
  - 4D Ray Benders
  - 4D Incident Lighting
  - Scene: 8D Ray Modulator

Beyond ‘Film-Like’ Photography

‘Computational Photography’;
To make ‘meaningful ray changes’ tangible,

- Sensors can do more…
- Displays can do more…
- Light Sources can do more…
- Optics can do more…
- Ray Descriptors can do more…

by applying low-cost storage, computation, and control.
Course: Computational Photography

A3: Film-like Photography: The Ray-Optics Model

Jack Tumblin
Northwestern University

A3: Film-like Photography

Film Camera designs still dominate:

- ‘Instantaneous’ light measurement...
- Of focal plane image behind a lens.
- Reproduce those amounts of light;
- Display ‘exactly matches’ the scene

Implied:

“What we see is \( \cong \) focal-plane intensities.”

well, no...
‘Film-Like Photography’: Ray Model

Light + 3D Scene: Image:
Illumination, shape, movement, surface BRDF,… Planar 2D map of light intensities

Image Plane
I(x,y)

‘Center of Projection’ (P^3 or P^2 Origin)

Light + 3D Scene: Image:
Illumination, shape, movement, surface BRDF,… Planar 2D map of light intensities

Image Plane
I(x,y)

‘Center of Projection’ (P^3 or P^2 Origin)

Film-Like Photography

- **Lighting:** Ray Sources (external)
- **Scene:** Ray Modulator (external)
- **Optics:** Ray Benders Thin Lens Approx.
- **Sensors:** Ray Bundle Measurement Sensor Irradiance E(x,y)
- **Processing** Ray Normalized E(x,y)
- **Display** Recreate Rays Normalized E(x,y)
**Film-like Optics: ‘Thin Lens Law’**

- Focal length, where parallel rays converge
- Object at distance $S_1$ forms image at $S_2$
- Focus at infinity: Adjust for $S_2 = f$
  
  
  \[
  \frac{1}{S_1} + \frac{1}{S_2} = \frac{1}{f}
  \]

  
  Larger $S_2$ for closer focus

**Rays Are Doubly Differential**

- Lens Systems: approximate rays with bundles
- Finite angle, not rays (lens aperture)
- Finite area, not points (circle of confusion)
Ray BUNDLES approximate Rays

• **Rays are doubly infinitesimal!**
  – A ‘ray’ leaves a span of infinitesimal area $0^+$
  – And covers a span of infinitesimal directions $0^+$

• Ray Bundles:
  Finite, measurable power from combined rays
  – A finite span of SOLID ANGLE, *and*
  – A finite span of SURFACE AREA

---

**EXAMPLE:**

- Power from 1 point on a spherical lamp?
  
  $0^+ = \frac{60\text{Watts}}{\infty \text{points}}$
  
  - BUT has a finite, measurable ratio: (flux/area)
  
    $\frac{60\text{Watts}}{30 \text{cm}^2 \text{area}} = 2 \text{W/cm}^2$
Ray BUNDLES approximate Rays

• Rays are **doubly** infinitesimal!
  A ‘ray’ Leaves $0^+$ area in $0^+$ directions

**EXAMPLE:**
• Power from spherical lamp in just 1 direction?
  $0^+ = \frac{60\text{Watts}}{\infty \text{ directions}}$
  – BUT has finite ratio:
  $(60\text{Watts} / 4\pi \text{ steradians}) = 4.77 \text{ W/cm}^2$

Ray Measurement: Radiance $L$

• Incoming light directions form hemisphere $\Omega$; Ray == one point on the hemisphere

THUS
‘Incident Rays’ measured in **Radiance Units** $L$:
Irradiance per unit solid angle
$L = \frac{\text{(watts / area)}}{\text{sr}}$
(sr = steradians; solid angle; $\Omega = \text{surface area on unit sphere}$)
Ray ‘Bundles’

- Rays have no surface area (just a point)
- Rays have no solid angle (just a point)

THUS:
- A Ray carries \textit{infinitesimal} power (0^+ \text{ Watts}).
- Only BUNDLES of rays are measurable!

? How can estimate the ‘Photographic Signal’ when we can’t directly measure it?

Lens Flaws: Depth of Focus

For the \textit{same} focal length:
- \textbf{Larger lens}
  - Gathers a wider ray bundle:
  - More light: brighter image
  - Shallower depth-of-focus
- \textbf{Smaller lens}
  - Gathers a narrower ray bundle:
  - Less light: dimmer image
  - Deeper depth-of-focus
Lens Flaws: Geometric Aberration

- **Aberrations:** Real lenses don’t converge rays perfectly
  - **Spherical:** edge rays ≠ center rays
  - **Coma:** diagonal rays focus deeper at edge

Radial Distortion
(e.g. ‘Barrel’ and ‘pin-cushion’)

straight lines curve around the image center
Lens Flaws: Chromatic Aberration

- Dispersion: wavelength-dependent refractive index
  - (enables prism to spread white light beam into rainbow)
- Modifies ray-bending and lens focal length: \( f(\lambda) \)

- color fringes near edges of image
- Corrections: add ‘doublet’ lens of flint glass, etc.

http://www.swgc.mun.ca/physics/physlets/opticalbench.htm

Lens Flaws: Chromatic Aberration

- **Lens Design Fix:** Multi-element lenses
  - Complex, expensive, many tradeoffs!
- **Computed Fix:** Geometric warp for R,G,B.

Near Lens Center  Near Lens Outer Edge
Lens Flaws: Intensity Aberrations

Image ‘Vignette’: bright at center, dark at edges.

Several compounded causes:
- Internal shadowing—angle-dependent Ray bundles
- Longer paths for off-axis Rays; Dark Glass
- Planar detector: outer pixels span greater angle

Compensation:
- Use anti-vignetting filters, (darkest at center)
- OR Position-dependent pixel-detector sensitivity.
http://homepage.ntlworld.com/j.houghton/vignette.htm

Polarization

Sunlit haze is often strongly polarized. Polarization filter yields much richer sky colors.
Film-like Color Sensing

- Visible Light: narrow band of emag spectrum
- $\lambda \approx 400-700 \text{ nm} \, (\text{nm} = 10-9 \text{ meter wavelength})$
- (humans:<1 octave ↔ honey bees: 3-4 ‘octaves do honey bees sense harmonics, see color ‘chords’?)
Color Sensing

- 3-chip vs. 1-chip: quality vs. cost

Practical Color Sensing: Bayer Grid

- Estimate RGB at 'G' cells from neighboring values
Conclusions

- Film-like photography methods limit digital photography to film-like results or less.

- Broaden, unlock our views of photography:

- 4D, 8D, even 10D Ray Space holds the photographic signal. Look for new solutions by creating, gathering, processing RAYS, not focal-plane intensities.

- Choose the best, most expressive sets of rays, THEN find the best way to measure them.

Useful links:

- Interactive Thin Lens Demo (or search ‘physlet optical bench’)  
  www.swgc.mun.ca/physics/physlets/opticalbench.html

For more about color:
- Prev. SIGGRAPH courses (Stone et al.)
- Good: www.cs.rit.edu/~ncs/color/a_spectr.html
- Good: www.colourware.co.uk/cpfaq.htm
- Good: www.yorku.ca/eye/toc.htm
Image Processing and Reconstructions Tools

Ramesh Raskar
Mitsubishi Electric Research Labs
Cambridge, MA

Image Tools

- Gradient domain operations,
  - Applications in tone mapping, fusion and matting

- Graph cuts,
  - Applications in segmentation and mosaicing

- Bilateral and Trilateral filters,
  - Applications in image enhancement

Intensity Gradient in 1D

Gradient at x, $G(x) = I(x+1) - I(x)$
Forward Difference

Reconstruction from Gradients

For $n$ intensity values, about $n$ gradients

Reconstruction from Gradients

1D Integration
$I(x) = I(x-1) + G(x)$
Cumulative sum

Intensity Gradient in 2D

Gradient at x,y as Forward Differences

$G_x(x,y) = I(x+1, y) - I(x,y)$
$G_y(x,y) = I(x, y+1) - I(x,y)$
$G(x,y) = (G_x, G_y)$
Intensity Gradient Vectors in Images

Reconstruction from Gradients

Given \( G(x,y) = (G_x, G_y) \)

How to compute \( I(x,y) \) for the image?

For \( n^2 \) image pixels, \( 2n^2 \) gradients!

Intensity Gradient in 2D

Recovering Original Image

Grad X

Grad Y

2D Integration

Recovering Manipulated Image

New Grad X

New Grad Y

Gradient Processing

2D Integration

Intensity Gradient Manipulation

Recovering Original Image

Recovering Manipulated Image

Intesity Gradient Manipulation

Gradient Processing

2D Integration
Intensity Gradient Manipulation

A Common Pipeline

Reconstruction from Gradients

• Look for image \( I \) with gradient closest to \( G \) in the least squares sense.

\[
F(\nabla I, G) = \|\nabla I - G\|^2 = \left( \frac{\partial I}{\partial x} - G_x \right)^2 + \left( \frac{\partial I}{\partial y} - G_y \right)^2
\]

Euler-Lagrange Equation

• \( I \) must satisfy:

\[
\frac{\partial F}{\partial I} - \frac{d}{dx} \frac{\partial F}{\partial I_x} - \frac{d}{dy} \frac{\partial F}{\partial I_y} = 0
\]

• Substituting \( F \) we get:

\[
\nabla^2 I = \text{div } G
\]

Application: Compressing Dynamic Range

How could you put all this information into one Image?

Attenuate High Gradients

Maintain local detail at the cost of global range

Fattal et al Siggraph 2002

Basic Assumptions

• The eye responds more to local intensity differences (ratios) than global illumination

• A HDR image must have some large magnitude gradients

• Fine details consist only of smaller magnitude gradients
Gradient Compression in 1D

Basic Method

- Take the log of the luminances
- Calculate the gradient at each point
- Scale the magnitudes of the gradients with a progressive scaling function (Large magnitudes are scaled down more than small magnitudes)
- Re-integrate the gradients and invert the log to get the final image

Summary: Intensity Gradient Manipulation

Graph and Images

Credits: Jianbo Shi
Graph Based Image Segmentation
Segmentation = Graph partition

\[ G = \{V, E\} \]

- \( V \): graph node
- \( E \): edges connecting nodes
- \( W_{ij} \): Edge weight

Image objective

Source images → Brush strokes → Computed labeling → Composite

Minimum Cost Cuts in a graph

Cut: Set of edges whose removal makes a graph disconnected

\[ S_{ij} \]: Similarity between pixel \( i \) and pixel \( j \)

Cost of a cut,

\[ \text{cut}(A, \bar{A}) = \sum_{i \in \lambda, j \in \bar{\lambda}} S_{ij} \]
Problem with min cuts

Min. cuts favors isolated clusters

Normalize cuts in a graph

\[ N_{\text{cut}}(A, B) = \text{cut}(A, B) \left( \frac{1}{\text{vol}(A)} + \frac{1}{\text{vol}(B)} \right) \]

Graph Cuts for Segmentation and Mosaicing

Cut ~ String on a height field

Brush strokes \rightarrow Computed labeling

Bilateral and Trilateral Filter

Input \rightarrow Bilateral \rightarrow Trilateral

Bilateral and Trilateral Filtering

Outline

- Unilateral filtering
  - Smoothing using filtering
- Bilateral filtering
  - Strength and 3 weaknesses
- Trilateral filtering
  - Key ideas
- Application in tone mapping
  - Detail preserving contrast reduction

‘Unilateral’ Filter

Traditional, linear, FIR filters

Key Idea: Convolution

- Output(x) = local weighted avg. of inputs.
  - Weights vary within a ‘window’ of nearby x
Smoothes away details, **BUT** blurs result

Note that weights always sum to 1.0
**‘Unilateral’ Filter**

Forces a Tradeoff:
- Broad window: better detail removal
  - OR -
- Narrow window: better large structure

But we want BOTH...

---

**Bilateral Filter**

A 2-D filter window: weights vary with intensity

Further Analysis: [Black99] [Elad02] [Durand&Dorsey02], ...

<table>
<thead>
<tr>
<th>Domain</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(x)$</td>
<td>$x$</td>
</tr>
</tbody>
</table>

2 Gaussian Weights: product = ellipsoidal footprint

Why it works: graceful segmentation
- Filtering in one region ignores filtering in another
- Gaussian $s$ acts as a ‘filtered region’ finder

---

**Bilateral Filter: Strengths**

Piecewise smooth result:
- averages local small details, ignores outliers
- preserves steps, large-scale ramps, and curves,...
- Equivalent to anisotropic diffusion and robust statistics
- Simple & Fast (esp. w/ [Durand02] FFT-based speedup)

---

**Bilateral Filter: 3 Difficulties**

- Poor Smoothing in High Gradient Regions
- Smooths and blunts cliffs, valleys & ridges
- Can combine disjoint signal regions

Output at $f(x)$ is average of a tiny region
**Bilateral Filter: 3 Difficulties**

- Poor Smoothing in High Gradient Regions
- Smoothes and blunts cliffs, valleys & ridges
- Can combine disjoint signal regions

**New Solution: "Trilateral" Filter**

- Keep best features of bilateral, adds more
- Corner sharpening resembles PDE shocks
- User sets 1 parameter (good defaults for 7 internals)

**Bilateral→Trilateral Filter**

**Key Ideas:**

- **Tilt** the filter window according to bilaterally-smoothed gradients
- **Limit** the filter window to connected regions of similar smoothed gradient.
- **Adjust** Parameters from measurements of the windowed signal

**Three Key Ideas:**

- **Tilt** the filter window according to bilaterally-smoothed gradients
- **Limit** the filter window to connected regions of similar smoothed gradient.
- **Adjust** Parameters from measurements of the windowed signal
Application: Tone Mapping

- Filter removes details.
- Goal: Detail-Preserving Contrast Reduction
  - in log domain, difference == contrast
  - remove details, compress contrast, replace details

\[
\begin{align*}
\log_{10} & \quad \text{In} \\
\text{Detail-Removing Filter} & \quad \text{Details} \\
\text{Base} & \quad \text{Out} \\
\gamma & \quad \langle 1: \text{compression} \rangle \\
\end{align*}
\]

More Trilateral Results

Comparable to Gradient Attenuation
[Fattal et al 2002]

Similar to LCIS
[Tumblin&Turk'99]
[Bertozzi'03]

Simple, Robust
Course 15: Computational Photography

**B1: Reconstruction**

Ramesh Raskar
Mitsubishi Electric Research Labs

Jack Tumblin
Northwestern University

Course WebPage: http://www.merl.com/people/raskar/photo

---

**Image Fusion and Reconstruction**

- Epsilon Photography
  - Vary time, view
  - Vary focus, exposure polarization, illumination
  - Better than any one photo
- Achieve effects via multi-image fusion
- Understand computer vision methods
- Exploit lighting

---

**Time-Lapse**

- Duchamp
  - Nude Descending a Staircase

Richard Hundley 2001

---

**Shape Time Photography**

Freeman et al 2003

---

**Varying Focus: Extended depth-of-field**

Agrawala et al, Digital Photomontage. Siggraph 2004
Computer Vision Techniques

- **Photometric Stereo**
  - Varying light source positions
  - Estimate surface normal from shading
  - Diffuse objects: minimum 3 lights

- **Depth from Defocus**
  - Varying focus

- **Defogging**
  - Varying time and polarization

Varying Focus: Depth from Defocus

\[
\frac{1}{f} + \frac{1}{i} = \frac{1}{o}
\]

Previous Work:
Pentland 87, Subbarao 88, Nayar 89.

Real Time Defocus Depth Camera (Movies)

Performance: 512 x 480 Depth map at 30 frames per sec.
Varying Polarization
Yoav Y. Schechner, Nir Karpel 2005

Best polarization state
Worst polarization state

The recovered image is much clearer, especially at distant objects, than the raw image.

Varying Polarization
- Schechner, Narasimhan, Nayar
- Instant dehazing of images using polarization

Varying Wavelength: Multispectral Fusion
Vegetation Mapping of the Forest

SAR + Optical Landsat =

- Mountain forest
- Mixed forest
- Okume forest
- Grassland Savannah
- Fern Savanna
- Burnt Savanna

Varying IR Wavelength Image Fusion

Non-photorealistic Camera: Depth Edge Detection and Stylized Rendering
using Multi-Flash Imaging

Ramesh Raskar, Karhan Tan, Rogerio Feris, Jingyi Yu, Matthew Turk
Mitsubishi Electric Research Labs (MERL), Cambridge, MA
U of California at Santa Barbara
U of North Carolina at Chapel Hill
Image Fusion and Reconstruction

- Epsilon Photography
  - Vary focus, exposure polarization, illumination
  - Vary time, view
  - Better than any one photo
- Achieve effects via multi-image fusion
- Understand computer vision methods
- Exploit lighting

Improving FILM-LIKE Camera Performance

What would make it ‘perfect’?

- Dynamic Range

Film-Style Camera: Dynamic Range Limits

Under-Exposure
- Highlight details: Captured
- Shadow details: Lost

Over-Exposure
- Highlight details: Lost
- Shadow details: Captured

Problem: Map Scene to Display

Domain of Human Vision:
from $-10^{-8}$ to $-10^{8}$ cd/m²

Range of Typical Displays:
from $-1$ to $-100$ cd/m²
High dynamic range capture (HDR)

- overcomes one of photography’s key limitations
  - negative film = 250:1 (8 stops)
  - paper prints = 50:1
  - [Debevec97] = 250,000:1 (18 stops)
  - hot topic at recent SIGGRAPHs

Debevec'97 (see www.HDRshop.com)

- number the images ‘i’
- pick fixed spots \((x_j, y_j)\) that sample scene’s radiance values \(\log L_i\) well:

STEP 1:

- Collect pixel values \(Z_{ij}\) (from image \(i\), location \(j\))
- (All of them sample the response curve \(f(\log L)\))

Debevec'97 (see www.HDRshop.com)

- Use the multiple samples to reconstruct the response curve;

Debevec'97 (see www.HDRshop.com)

- Then use the inverse response curve to reconstruct the intensities that caused the responses
**HDR Direct Sensing?**

- An open problem! (esp. for video...)

- A direct (and expensive) solution:
  - Flying Spot Radiometer: brute force instrument, costly, slow, delicate

- Some Other Novel Image Sensors:
  - line-scan cameras (e.g. Spheron: multi-detector)
  - logarithmic CMOS circuits (e.g. Fraunhofer Inst)
  - Self-resetting pixels (e.g. sMal: Cypress Semi)
  - Gradient detectors (CVPR 2005 Tumblin, Raskar et al)

**HDR From Multiple Measurements**

**MANY ways to make multiple exposure measurements**

**Sequential Exposure Change:**
- Ginosar et al '92, Burt & Kolczynski '93, Shaalan '93, Tsai '94, Saito '95, Mann '95, Debevec & Malik '97, Mitsunaga & Nayar '99, Kato et al '03

**Mosaicing with Spatially Varying Filters:**
- Schechner and Nayar '01, Aggarwal and Ahuja '01

**Multiple Image Detectors:**
- Nayar '92, Saito '95, Mann '95, Kato et al '03, Nayar and Narasimhan '02

**Assorted-Pixel Camera Prototype**

( Courtesy: Sony Kihara Research Lab )

- Digital Still Camera
- Camera with Assorted Pixels

**Another Approach: Locally Adjusted Sensor Sensitivity**

**Computational Pixels:**
- (pixel sensitivity set by its illumination)

- No Gradient Camera: Ramesh Has It
**Sensor: LCD Adaptive Light Attenuator**

LCD Light Attenuator limits image intensity reaching 8-bit sensor

**Unprotected 8-bit Sensor Output:**

**Attenuator-Protected 8-bit Sensor Output**

---

**Improving FILM-LIKE Camera Performance**

- Vary Focus Point-by-Point

---

**High depth-of-field**

Levoy et al., SIGG2005

- adjacent views use different focus settings
- for each pixel, select sharpest view

- close focus
- distant focus
- composite

---

**Single-Axis Multi-Parameter Camera (SAMP)**

Idea: Cameras + Beamsplitters Place MANY (8) cameras at same virtual location

- Morgan McGuire (Brown), Wojciech Matusik (MERL), Hanspeter Pfister (MERL), Fredo Durand (MIT), John Hughes (Brown), Shree Nayar (Columbia)

---

**SAMP Prototype System (Layout)**

---

**Multiple Simultaneous Focus Depths**

- Strongly desired in microscopy, too. see
  [http://www.micrographia.com/article/artmicgr/mesci000.htm](http://www.micrographia.com/article/artmicgr/mesci000.htm)
Long-range synthetic aperture photography

Focus Adjustment: Sum of Bundles

Improving FILM-LIKE Camera Performance

• Field of view vs. Resolution?

Are we done?
  • Almost EVERY digital camera has panoramic stitching.

No; Much more is possible:

A tiled camera array

• 12 × 8 array of VGA cameras
  • abutted: 7680 × 3840 pixels
  • overlapped 50%: half of this
  • total field of view = 29° wide
  • seamless mosaicing isn’t hard
  • cameras individually metered
  • Approx same center-of-proj.

Tiled panoramic image
(before geometric or color calibration)

Tiled panoramic image
(after calibration and blending)
Improving FILM-LIKE Camera Performance

- Exposure time and Frame rate

High Speed Video

Say you want 120 frame per second (fps) video.
- You could get one camera that runs at 120 fps
- Or…

High Speed Video

Say you want 120 frame per second (fps) video.
- You could get one camera that runs at 120 fps
- Or… get 4 cameras running at 30 fps.

Conclusions

- Multiple measurements:
  - Multi-camera, multi-sensor, multi-optics, multi-lighting
- Intrinsic limits seem to require it
  - lens diffraction limits, noise, available light power.
- Are we eligible for Moore’s law? Or will lens making, mechanics limit us?

52 Camera Cluster, 1560 FPS

Levoy et al., SIGG2005
Computational Cameras: Convergence of Optics and Software

Shree K. Nayar

Computer Science
Columbia University

http://www.cs.columbia.edu/CAVE/

Support:
NSF, ONR, Packard Foundation
T. C. Chang Endowed Chair

Traditional Camera
Computational Cameras

Detector

Computations

New Optics

Vision

Wide Angle Imaging

Multiple Cameras

Examples: Disney 55, McCutchen 91, Nalwa 96, Swaminathan & Nayar 99, Cutler et al. 02

Catadioptric Imaging

Examples: Rees 70, Charles 87, Nayar 88, Yagi 90, Hong 91, Yamazawa 95, Bogner 95, Nalwa 96, Nayar 97, Chahl & Srinivasan 97

© Shree Nayar, Columbia University
What’s the Mirror’s Shape?

(With Simon Baker, ICCV 98)

Complete Class of Mirrors

\[
\left( z - \frac{c}{2} \right)^2 - r^4 \left( \frac{k}{2} - 1 \right) = \frac{c^4}{4} \left( \frac{k - 2}{k} \right) \quad (k > 0)
\]

\[
\left( z - \frac{c}{2} \right)^2 - r^4 \left( 1 + \frac{c^2}{2k} \right) = \frac{2k - c^2}{4} \quad (k \geq 2)
\]

OneShot 360 by RemoteReality

(Nayar 97)

4 Megapixel (2000 x 2000)
360 degree still camera
Video Conferencing

(with Venkat Peri 96)
Omnidirectional Periscope

Wide Area Surveillance

Vehicle Navigation

Perimeter Monitoring

Commercial Security

Near Vehicle Awareness

(Courtesy: RemoteReality Inc.)

Radial Stereoscopic Imaging

(with Sujit Kuthirummal, SIGGRAPH '06)

© Shree Nayar, Columbia University
Mosaicing

......Redundant Measurements

© Shree Nayar, Columbia University
Generalized Mosaicing

(Schechner and Nayar, ICCV 2001)

Field of View
Dynamic Range
Spectrum
Depth of Field
Polarization

© Shree Nayar, Columbia University
Multispectral Mosaicing

Spectral

Multispectral Mosaic

© Shree Nayar, Columbia University
High Dynamic Range Imaging: Assorted Pixels

Spatially Varying Exposures (SVE)

© Shree Nayar, Columbia University (with Tomoo Mitsunaga, CVPR 2000)
Fundamental Trade-Off in Imaging

Temporal Resolution (fps)

<table>
<thead>
<tr>
<th>Spatial Resolution (pixels)</th>
<th>High Resolution Camera</th>
<th>Low Resolution Camera</th>
</tr>
</thead>
<tbody>
<tr>
<td>3M 2048x1536</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75K 320x240</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Deblurring Approach: Hybrid Imaging

Low-Res. Camera → Motion Analysis → PSF Estimation → Deconvolution → High-Res. Camera

Same Time Period

(with Moshe Ben-Ezra, CVPR 2003)
Hybrid Imaging System: Prototype

Primary Detector (2048x1536)

Secondary Detector (360x240)

Resolution Ratio of 1 : 36

Example: Blurred High Resolution Image

f = 633mm, Exp. Time 1 Sec (> -9 stops)
Example: PSF Estimation from Motion

Low Resolution Video

Blurred Image

Deblurred image

Tripod image (Ground Truth)
Super-Resolution using Jitter Video

Conventional Video

Jitter Video

(With Moshe Ben-Ezra and Assaf Zomet, CVPR 2004)

Jitter Camera

Lens

Detector

Micro-Actuator

Jitter is Instantaneous and Synchronized
Jitter is Instantaneous and Synchronized

© Shree Nayar, Columbia University
Super-Resolution using Jitter Video

1 (out of 4) Jitter Camera Image

Super-Resolution

Video Super-Resolution Using Controlled Sub-Pixel Detector Shifts

© Shree Nayar, Columbia University
Imaging Through Micro-Mirrors

Geometry: Ray Orientation

\[ o_i (x_i) = G (n_i) \]

Photometry: Ray Attenuation

\[ a_i (x_i) = \frac{t (n_i)}{t (n_i) + t (n_b)} \]

Digital Micromirror Device (DMD)

DMD Array:

Micromirror Architecture:

DMD for Imaging:

(Malbet et al. 95, Kearney et al. 98, Castracane et al. 99, Christensen et al. 02)
Programmable Imaging System

Imaging Lens  DMD Electronics
Camera Electronics  Tilted CCD  Lens Focused on DMD

Modulation: Examples

Scene  DMD Image  Camera Image

*  =  
*  =  
*  =  

© Shree Nayar, Columbia University
Adaptive Dynamic Range Imaging (ADR)

Normal (Constant Exposure) Video

Pixel-wise Dynamic Range Control
(Nayar & Branzoi 03)
(Christensen et al. 02)

ADR Video

DMD Control Video

Camera with a Lens

© Shree Nayar, Columbia University
Lensless Camera with Volumetric Aperture

Volumetric Aperture

Image Detector

Scene

Volumetric Aperture

Scene

Volumetric Aperture

Scene

Single Layer Aperture

Image Detector

Pixel Brightness:

\[ I(x, y) = \int \int S(u, v)T(x - fu, y - fv)\, du\, dv \]

© Shree Nayar, Columbia University
**Initial Implementation: LCD Attenuator**

- Camera without Lens
- LCD Aperture
- LCD Controller

© Shree Nayar, Columbia University

---

**Panning without Moving Parts**

- LCD Attenuator
- Image Detector
- Captured Video

© Shree Nayar, Columbia University
Multiple Aperture Layers

\[ u = \tan(\alpha) \]

\[ S(u,v) \]

Multi-Layered Aperture

\[ \alpha \]

Multi-Layered Aperture

\[ j=1 \]

\[ j=2 \]

\[ \cdots \]

Image Detector

Pixel Brightness:

\[ I(x,y) = \int \int S(u,v) \prod_{j=1}^{N} T_j(x - f_j u, y - f_j v) \, du \, dv \]

Scene Transmittance Functions

© Shree Nayar, Columbia University
Split Field of View using Multiple Layers

Fov 1  Fov 2  Fov 3

Pinholes

Attenuating Layers

Image Detector

Split Field of View

Lens Camera

Lensless Camera

© Shree Nayar, Columbia University
Computational Cameras

Detector

Computations

Pixels

New Optics

Vision

---

Programmable Imaging

Detector

Computations

Pixels

New Optics

Programmable Controller

Vision

© Shree Nayar, Columbia University
Scanning Photography
…where the 2nd dimension is TIME

Andrew Davidhazy
Imaging and Photographic Technology
School of Photographic Arts and Sciences
Rochester Institute of Technology

GEORGE SILK
ALFRED EISENSTADT
GJON MILI
NEIL LEIFER
JOHN ZIMMERMAN
SAM BREITMAN
BOB GALBRAITH
JERRY DANTZIC
WILLIAM LARSON
ETC.
Operation of an “instantaneous” shutter

Operation of a scanning focal plane shutter (slit-scan)
PHOTOGRAPHY: THEORY AND PRACTICE

Fig. 17.43. Image Distortion with Focal Plane Shutter
RESULT

RESULT
Simultaneity and duration timing
Ping Pong ball accelerating down inclined ramp
RESULT
remember …

SCANNING

TIME

Andrew Davidhazy
Imaging and Photographic Technology
School of Photographic Arts and Sciences
Rochester Institute of Technology

http://www.rit.edu/~andpph
Undistorted Reproduction achieved when

\[ \text{aspect ratio of original} = \text{aspect ratio of reproduction} \]

Height of Reproduction determined by Optical Magnification

Length of Reproduction determined by CCD Readout Rate

( for a given Image Rate of Motion )
Light field photography and microscopy

Marc Levoy

Computer Science Department
Stanford University

34:15 total + 30% = ~45 minutes
The light field
[Gershun 1936]

Radiance as a function of position and direction

- for general scenes
  - the “plenoptic function”
  - five-dimensional
  - \( L(x, y, z, \theta, \phi) \) (w/m\(^2\)sr)

- in free space
  - four-dimensional
  - \( L(u, v, s, t) \)

• light fields
  • from Parry Moon (1940’s)
  • he included wavelength

• \( L(x,y,z,\theta,\phi) \)
  • a.k.a. plenoptic function
    (Adelson & Bergen 1991)

• free space
  • no occluders or participating media

• 4D function
  • How to parameterize it?
    i.e. \( L \) (what?)

0:45

We’ll come back to this...
Devices for recording light fields

- big scenes
  - handheld camera [Buehler 2001]
  - array of cameras [Wilburn 2005]
  - plenoptic camera [Ng 2005]
  - light field microscope [Levoy 2006]

- small scenes

- 0:45
  - using geometrical optics
    - i.e. omitting holography
  - as you’ve probably guessed
    - I’ve organized these devices roughly according to the smallest-size scene it can a record a light field of
  - Is the scale of the scene important?
    - when it reaches the micron scale, yes...
High performance imaging using large camera arrays

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

(Proc. SIGGRAPH 2005)

• 0:30
• 16:30 for this third
• 44:30 for entire talk (fat chance)
Stanford multi-camera array

- 640 × 480 pixels × 30 fps × 128 cameras
- synchronized timing
- continuous streaming
- flexible arrangement

Time = 0:45
Ways to use large camera arrays

• widely spaced → light field capture

0:30
In order that you understand the video,
– I need to explain a few things first
Ways to use large camera arrays

• widely spaced ➔ light field capture
• tightly packed ➔ high-performance imaging

•0:30
•In order that you understand the video,
  – I need to explain a few things first
Ways to use large camera arrays

• widely spaced → light field capture
• tightly packed → high-performance imaging
• intermediate spacing → synthetic aperture photography

• 0:30
• In order that you understand the video,
  – I need to explain a few things first
Synthetic aperture photography

• 2:00, 10:15 for SAP
Synthetic aperture photography
Synthetic aperture photography
Synthetic aperture photography
Synthetic aperture photography
Synthetic aperture photography
Example using 45 cameras
[Vaish CVPR 2004]

•0:30
Light field photography using a handheld plenoptic camera

Ren Ng, Marc Levoy, Mathieu Brédif, Gene Duval, Mark Horowitz and Pat Hanrahan

(Proc. SIGGRAPH 2005 and TR 2005-02)

• 0:30
• 14:15 for this third
• Light field capture not using an array of cameras, but
  – using a single, handheld camera
• What we’ll do with these light fields is not seeing through crowds, but
  – refocusing a picture after we take it, and
  – moving the observer (slightly) after we take the picture
Conventional versus plenoptic camera

• 3:15 for camera
• 0:45
• plenoptic camera
  – on which our microscope is based
Conventional versus plenoptic camera

Subject

Main lens

Photosensor

uv-plane

st-plane

Main lens
Prototype camera

Contax medium format camera

Kodak 16-megapixel sensor

Adaptive Optics microlens array

125μ square-sided microlenses

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

•0:30
Digital refocusing

- refocusing = summing windows extracted from several microlenses
A digital refocusing theorem

• an $f/N$ light field camera, with $P \times P$ pixels under each microlens, can produce views as sharp as an $f/(N \times P)$ conventional camera

— or —

• it can produce views with a shallow depth of field ($f/N$) focused anywhere within the depth of field of an $f/(N \times P)$ camera

•0:30
Example of digital refocusing

• 1:15
Refocusing portraits
Action photography

Focusing through a splash of water
Extending the depth of field

- Conventional photograph, main lens at $f/4$
- Conventional photograph, main lens at $f/22$
- Light field, main lens at $f/4$, after all-focus algorithm [Agarwala 2004]

- 1:00
- Main lens at $f/22$
  - Captured with light field camera and f/4 lens,
  - Computed by extracting only the middle pixel of that image
  - Would be the same image if no microlenses and larger pixels
Macrophotography

•0:15
Digitally moving the observer

• moving the observer = moving the window we extract from the microlenses

© 2007 Marc Levy
Example of moving the observer
Moving backward and forward
Implications

• cuts the unwanted link between exposure (due to the aperture) and depth of field

• trades off (excess) spatial resolution for ability to refocus and adjust the perspective

• sensor pixels should be made even smaller, subject to the diffraction limit

  \[
  \frac{36 \text{mm} \times 24 \text{mm}}{2 \mu \text{pixels}} = 216 \text{ megapixels}
  \]

  \[
  18 \times 12 \text{K pixels}
  \]

  \[
  1800 \times 1200 \text{ pixels} \times 10 \times 10 \text{ rays per pixel}
  \]

• 1:45
Light Field Microscopy

Marc Levoy, Ren Ng, Andrew Adams, Matthew Footer, and Mark Horowitz

(Proc. SIGGRAPH 2006)
A traditional microscope

- eyepiece
- intermediate image plane
- objective
- specimen
A light field microscope (LFM)

- 40x / 0.95NA objective
  ↓
  0.26μ spot on specimen
  × 40x = 10.4μ on sensor
  ↓
  2400 spots over 25mm field

- 125^2-micron microlenses
  ↓
  200 × 200 microlenses with 12 × 12 spots per microlens

→ reduced lateral resolution on specimen
  = 0.26μ × 12 spots = 3.1μ

- 0.95NA
  - captures rays up to 70° away from the optical axis
  - no photographic lens comes close to this
  - this would be an f/0.5 lens!

- 2400 spots
  - limited by the wave optics of light
  - regardless of the number of pixels in our sensor
A light field microscope (LFM)

• we don’t have a sensor 2.5mm above the microlens array
  – in this prototype, we use a camera and 1:1 relay lens instead
Example light field micrograph

- orange fluorescent crayon
- mercury-arc source + blue dichroic filter
- 16x / 0.5NA (dry) objective
- f/20 microlens array
- 65mm f/2.8 macro lens at 1:1
- Canon 20D digital camera
The geometry of the light field in a microscope

- Microscopes make orthographic views
- Translating the stage in X or Y provides no parallax on the specimen
- Out-of-plane features don’t shift position when they come into focus
- Front lens element size = aperture width + field width
- PSF for 3D deconvolution microscopy is shift-invariant (i.e. doesn’t change across the field of view)

Objective lenses are telecentric
Panning and focusing

panning sequence
focal stack
Real-time viewer
Mouse embryo lung
(16x / 0.5NA water immersion)

•0:30
Zebradish optic tectum
(63x / 0.9NA water immersion)
Axial resolution (a.k.a. depth of field)

• wave term + geometrical optics term

\[ DOF_{tot} = \frac{\lambda n}{NA^2} + \frac{n}{M \times NA} \times e \]

• ordinary microscope (16x/0.4NA (dry), \( e = 0 \))

\[ = \frac{0.535 \times 1}{0.4^2} = 3.3\mu \quad \text{(wave optics dominates)} \]

• with microlens array \( (e = 125\mu) \)

\[ = \frac{0.535 \times 1}{0.4^2} + \frac{1}{16 \times 0.4} \times 125\mu = 3.3\mu + 19.5\mu = 22.8\mu \quad \text{(geometrical optics dominates)} \]

• stopped down to one pixel per microlens

\[ = 3.3\mu + 19.5\mu \times 12 \text{ spots} = 237\mu \]

→ number of slices in focal stack = 12

© 2007 Marc Levy
From light fields to volumes

- 4D light field → digital refocusing →
  3D focal stack → deconvolution microscopy →
  3D volume data

(DeltaVision)

Time = 1:45
3D deconvolution

[McNally 1999]

focus stack of a point in 3-space is the 3D PSF of that imaging system

- object * PSF → focus stack
- $F \{\text{object}\} \times F \{\text{PSF}\} \rightarrow F \{\text{focus stack}\}$
- $F \{\text{focus stack}\} \div F \{\text{PSF}\} \rightarrow F \{\text{object}\}$
- spectrum contains zeros, due to missing rays
- imaging noise is amplified by division by ~zeros
- reduce by regularization (smoothing) or completion of spectrum
- improve convergence using constraints, e.g. object $> 0$

-2:30
- object * PSF
  - assumes object is semi-transparent, hence scattering is minimal
  - each pixel is a line integral of attenuation (if backlit) or emission (if frontlit and object is fluorescent)
- missing rays
  - not as bad as a photographic lens, because microscope lenses have a very wide aperture,
  - for example, 0.95NA = about 70 degrees each side of the optical axis, as shown in the PSF above
- Ref:
Silkworm mouth
(40x / 1.3NA oil immersion)

- slice of focal stack
- slice of volume
- volume rendering

• 0:30
Insect legs
(16x / 0.4NA dry)

volume rendering

all-focus image

[Agarwala 2004]
3D reconstruction (revisited)

- 4D light field $\rightarrow$ digital refocusing $\rightarrow$
  3D focal stack $\rightarrow$ deconvolution microscopy $\rightarrow$
  3D volume data

- 4D light field $\rightarrow$ tomographic reconstruction $\rightarrow$
  3D volume data

(DeltaVision)

(from Kak & Slaney)
Implications of this equivalence

• light fields of minimally scattering volumes contain only 3D worth of information, not 4D
• the extra dimension serves to reduce noise; could it be used to restore lost spatial resolution?

Optical Projection Tomography [Sharpe 2002]
Extensions

- extending the field of view by correcting digitally for objective aberrations

Nikon 40x 0.95NA (dry) Plan-Apo
Extensions

- extending the field of view by correcting digitally for objective aberrations

  correcting for aberrations caused by imaging through thick specimens whose index of refraction doesn’t match that of the immersion medium
Extensions

- extending the field of view by correcting digitally for objective aberrations
- gradient filters at the aperture plane
Extensions

• extending the field of view by correcting digitally for objective aberrations
• gradient filters at the aperture plane
Extensions

• extending the field of view by correcting digitally for objective aberrations
• gradient filters at the aperture plane

... or polarization direction
... or ???

• gives up digital refocusing?
Extensions

• extending the field of view by correcting digitally for objective aberrations
• gradient filters at the aperture plane
• photographing the aperture plane
  – transposes the light field
Using lenses to transpose a light field

\[
\text{st} \quad \text{uv} \quad \text{st'} \quad \text{uv'}
\]

[Images of light field transposition with scale bars: 200μ and 10°]
Extensions

• extending the field of view by correcting digitally for objective aberrations
• gradient filters at the aperture plane
• photographing the aperture plane
  – transposes the light field
  – high angular resolution × low spatial resolution
  → an imaging microscope scatterometer
    » would need to control 4D illumination
Reflectance field of single squid skin iridiphore
Other applications of light field illumination:
“4D designer lighting”
Spatioangular Resolution Tradeoff in Integral Photography

Todor Georgiev

Early work on integral photography

F. Ives (1903)

G. Lippmann (1908)
Display – Camera duality

(From Zwicker et al.)

(From Ng et al.)
Low angular resolution is typical:

Small number of cameras (100) or
Small number of pixels behind each micro lens (100)

Both possible due to Lambertian surface of observed object
(slow change in radiance in angular directions at surface).
Very sparse sampling is OK with view morphing.

Big number of pixels in each camera (100, 000)

Capture 4D radiance with 2D sensor

Two ways of multiplexing:
(1) big array of small angular images
(2) small array of big spatial images

We want to trade angular for spatial resolution.
At low angular resolution case (1) has
significant problem at boundary pixels.

We have chosen to work on optical design (2).
Design 2

Optics

Light field transform at a lens
\[
\begin{pmatrix}
x' \\
\theta'
\end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{T} & 1 \end{pmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix}
\]

Light field traveling distance T
\[
\begin{pmatrix}
x' \\
\theta'
\end{pmatrix} = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix}
\]

Light field transform at a prism (affine transform)
\[
\begin{pmatrix}
x' \\
\theta'
\end{pmatrix} = \begin{pmatrix} x \\ \theta \end{pmatrix} + \begin{pmatrix} 0 \\ \alpha \end{pmatrix}
\]
Light field transform at a shifted lens

\[
\begin{pmatrix}
  x' \\
  \theta'
\end{pmatrix} = \begin{pmatrix}
  x \\
  \theta
\end{pmatrix} - \begin{pmatrix}
  s \\
  0
\end{pmatrix}
\]

\[
\begin{pmatrix}
  x'' \\
  \theta''
\end{pmatrix} = \begin{pmatrix}
  1 & 0 \\
  -\frac{1}{f} & 1
\end{pmatrix} \begin{pmatrix}
  x-s \\
  \theta
\end{pmatrix}
\]

\[
\begin{pmatrix}
  x''' \\
  \theta'''
\end{pmatrix} = \begin{pmatrix}
  1 & 0 \\
  -\frac{1}{f} & 1
\end{pmatrix} \begin{pmatrix}
  x-s \\
  \theta
\end{pmatrix} + \begin{pmatrix}
  s \\
  0
\end{pmatrix}
\]

\[
\begin{pmatrix}
  x'''' \\
  \theta'''
\end{pmatrix} = \begin{pmatrix}
  1 & 0 \\
  -\frac{1}{f} & 1
\end{pmatrix} \begin{pmatrix}
  x \\
  \theta
\end{pmatrix} + \begin{pmatrix}
  0 \\
  \frac{s}{f}
\end{pmatrix}
\]

A shifted lens is equivalent to a lens-prism pair

Designs
Designs

a)

b)

c)

d)

e)

f)
3-View Morphing

Refocusing
Conclusion:

The way to increase spatial resolution with a fixed sensor is to trade angular for spatial resolution. Then, view-morph.

The Plenoptic (Adelson–Wang, Ng et al.) design (1) has difficulties with low angular resolution. That’s why we chose the other design (2).

We showed optical light field transforms and 5 new camera designs. Lens-prism pairs.
Computational Illumination

Ramesh Raskar
Mitsubishi Electric Research Labs

Course WebPage:
http://www.merl.com/people/raskar/photo/course/

Computational Illumination

• Presence or Absence
  – Flash/No-flash

• Light position
  – Multi-flash for depth edges
  – Programmable dome (image re-lighting and matting)

• Light color/wavelength

• Spatial Modulation
  – Synthetic Aperture Illumination

• Temporal Modulation
  – TV remote, Motion Tracking, Sony ID-cam, RFfG

• General lighting condition
  – Day/Night

Novel Cameras

Generalized Sensor

Processing

Generalized Optics

Display

Remote 4D Lightfield

Recreate 4D Lightfield

Programmable 4D Illumination Field + Time + Wavelength

Computational Photography

Novel Illumination

Light Sources

Computational Illumination:

Programmable 4D Illumination Field + Time + Wavelength

• Presence or Absence
  – Flash/No-flash

• Light position
  – Multi-flash for depth edges
  – Programmable dome (image re-lighting and matting)

• Light color/wavelength

• Spatial Modulation
  – Synthetic Aperture Illumination

• Temporal Modulation
  – TV remote, Motion Tracking, Sony ID-cam, RFfG

• Exploiting (uncontrolled) natural lighting condition
  – Day/Night Fusion

Flash and Ambient Images

(Agrawal, Raskar, Nayar, Li Siggraph05)
Denoising Challenging Images

Available light:
+ nice lighting
- noise/blurriness
- color

Flash:
+ details
+ color
- flat/artificial

Computational Illumination

- Presence or Absence
  - Flash/No-flash
- Light position
  - Multi-flash for depth edges
  - Programmable dome (image re-lighting and matting)
- Light color/wavelength
- Spatial Modulation
  - Synthetic Aperture Illumination
- Temporal Modulation
  - TV remote, Motion Tracking, Sony ID-cam, RFIG
- General lighting condition
  - Day/Night
Synthetic Lighting
Paul Haeberli, Jan 1992

Non-photorealistic Camera:
Depth Edge Detection and Stylized Rendering using Multi-Flash Imaging

Ramesh Raskar, Karhan Tan, Rogerio Feris, Jingyi Yu, Matthew Turk
Mitsubishi Electric Research Labs (MERL), Cambridge, MA
U of California at Santa Barbara
U of North Carolina at Chapel Hill

Depth Edge Camera

Mitsubishi Electric Research Labs
MultiFlash NPR Camera
Raskar, Tan, Feris, Yu, Turk

Depth Discontinuities

Internal and external Shape boundaries, Occluding contour, Silhouettes

Our method captures shape edges
Imaging Geometry

- Shadow lies along epipolar ray,
- Epipole and Shadow are on opposite sides of the edge

Depth Edge Camera

- Light epipolar rays are horizontal or vertical
Negative transition along epipolar ray is depth edge

% Max composite
maximg = max( left, right, top, bottom);

% Normalize by computing ratio images
r1 = left ./ maximg;  r2 = top ./ maximg;
  r3 = right ./ maximg;  r4 = bottom ./ maximg;

% Compute confidence map
v = fspecial('sobel'); h = v';
d1 = imfilter(r1, v);  d3 = imfilter(r3, v);  % vertical sobel
    d2 = imfilter(r2, h);  d4 = imfilter(r4, h);  % horizontal sobel

% Keep only negative transitions
silhouette1  = d1 .* (d1>0);
silhouette2 = abs( d2 .* (d2<0) );  silhouette3 = abs( d3 .* (d3<0) );
silhouette4  = d4 .* (d4>0);

% Pick max confidence in each
confidence = max( silhouette1, silhouette2, silhouette3, silhouette4);
imwrite( confidence, 'confidence.bmp');

Negative transition along epipolar ray is depth edge

Debevec et al. 2002: ‘Light Stage 3’
Computational Illumination

- Presence or Absence
  - Flash/No-flash
- Light position
  - Multi-flash for depth edges
  - Programmable dome (image re-lighting and matting)
- Light color/wavelength

- Spatial Modulation
  - Synthetic Aperture Illumination
- Temporal Modulation
  - TV remote, Motion Tracking, Sony ID-cam, RFiG
- General lighting condition
  - Day/Night

“Light Waving”
Tech Sketch (Winnemoller, Mohan, Tumblin, Gooch)

Synthetic Aperture Illumination: Comparison with Long-range synthetic aperture photography

- width of aperture 6'
- number of cameras 45
- spacing between cameras 5”
- camera’s field of view 4.5°
The scene

- distance to occluder: 110'
- distance to targets: 125'
- field of view at target: 10'

Synthetic aperture photography
using an array of mirrors

- 11-megapixel camera (4064 x 2047 pixels)
- 18 x 12 inch effective aperture, 9 feet to scene
- 22 mirrors, tilted inwards → 22 views, each 750 x 500 pixels

Synthetic aperture illumination

- technologies
  - array of projectors
  - array of microprojectors
  - single projector + array of mirrors

What does synthetic aperture illumination look like?

What are good patterns?

<table>
<thead>
<tr>
<th>pattern</th>
<th>one trial</th>
<th>16 trials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Underwater confocal imaging with and without SAP

6-D Methods and beyond...

Relighting with 4D Incident Light Fields Vincent Masselus, Pieter Peers, Philip Dutre and Yves D. Willems SIGG2003

Computational Illumination

- Presence or Absence
  - Flash/No-flash
- Light position
  - Multi-flash for depth edges
  - Programmable dome (image re-lighting and matting)
- Light color/wavelength
- Spatial Modulation
  - Synthetic Aperture Illumination
- Temporal Modulation
  - TV remote, Motion Tracking, Sony ID-cam, RF/G
- General lighting condition
  - Day/Night

Demodulating Cameras

- Simultaneously decode signals from blinking LEDs and get an image
  - Sony ID Cam
  - Phoci
- Motion Capture Cameras
  - Visualeyz™ VZ4000 Tracking System
  - PhaseSpace motion digitizer

Demodulating Cameras

- Decode signals from blinking LEDs + image
  - Sony ID Cam
  - Phoci
RFGLamps:
Interacting with a Self-describing World via Photosensing Wireless Tags and Projectors

Ramesh Raskar, Paul Beardsley, Jeroen van Baar, Yao Wang, Paul Dietz, Johnny Lee, Darren Leigh, Thomas Willwacher
Mitsubishi Electric Research Labs (MERL), Cambridge, MA

RFGLamps
Radio Frequency Identification Tags (RFID)
No batteries,
Small size,
Cost few cents

Radio Frequency Identification Tags (RFID)

Conventional Passive RFID

Warehousing
Routing
Livestock tracking
Library
Baggage handling
Currency

Conventional Passive RFID

Tagged Books in a Library

✓ Id
Easy to get list of books in RF range

✗ No Precise Location Data
Difficult to find if the books in sorted order?
Which book is upside down?

Where are boxes with
Products close to Expiry Date?
**RFID**

(Radio Frequency Identification)

**RFIG**

(Radio Frequency Id and Geometry)

Photosensing RFID tags are queried via RF.

Tags respond via RF, with date and precise (x,y) pixel location. Projector beams 'O' or 'X' at that location for visual feedback.

Multiple users can simultaneously work from a distance without RF collision.

Projector beams a time-varying pattern unique for each (x,y) pixel which is decoded by tags.

Photosensor and Light compatible with RFID size and power needs.

Directional transfer, AR with Image overlay.

Multiple users can simultaneously work from a distance without RF collision.

Photo-sensing RF tag

Conventional RF tag

RFIG Lamps

Prototype Tag

Prototype Tag

RF tag + photosensor

RF Transponder

Photo Sensor
Projected Sequential Frames

- Handheld Projector beams binary coded stripes
- Tags decode temporal code

For each tag
a. From light sequence, decode x and y coordinate
b. Transmit back to RF reader (Id, x, y)

\[ X=12 \]
Visual feedback of 2D position

a. Receive via RF \((x_1, y_1), (x_2, y_2), \ldots\) pixels

b. Illuminate those positions

Computational Illumination

- Presence or Absence
  - Flash/No-flash

- Light position
  - Multi-flash for depth edges
  - Programmable dome (image re-lighting and matting)

- Light color/wavelength

- Spatial Modulation
  - Synthetic Aperture Illumination

- Temporal Modulation
  - TV remote, Motion Tracking, Sony ID-cam, RFIG

- Natural lighting condition
  - Day/Night Fusion

A Night Time Scene:
Objects are Difficult to Understand due to Lack of Context

Dark Bldgs
Reflections on bldgs
Unknown shapes

Enhanced Context:
All features from night scene are preserved, but background in clear

Well-lit Bldgs
Reflections in bldgs windows
Tree, Street shapes

Background is captured from day-time scene using the same fixed camera

Night Image
Day Image

Result: Enhanced Image

Mask is automatically computed from scene contrast

Ramesh Raskar, Computational Illumination
But, Simple Pixel Blending Creates Ugly Artifacts

Our Method: Integration of blended Gradients

Reconstruction from Gradient Field

- Problem: minimize error $|\nabla I' - G|$
- Estimate $I'$ so that $G = \nabla I'$
- Poisson equation $\nabla^2 I' = \text{div} G$
- Full multigrid solver

Video Enhancement using Fusion

- Video from fixed cameras
  - Improve low quality InfraRed video using high-quality visible video
  - Fill in dark areas, enhance change in intensity
  - Output style: better context
- Current Demo
  - Fusion of Night video and Daytime image

Details

- Combine day and night time images
  - Night videos have low contrast, areas with no detail
  - Same camera during day can capture static information
  - Dark areas of night video are replaced to provide context
  - Moving object (from night) + Static scene (from day)
Video Enhancement

Overview of Process

Original night-time traffic camera 320x240 video
Daytime image: By averaging 5 seconds of day video

Enhanced Video
Note: exit ramp, lane dividers, buildings not visible in original night video, but clearly seen here.

Algorithm

Frame N
Gradient field
Temporal averaged importance mask
Mixed gradient field
Processed binary mask
Final result
Daytime image
Computational imaging in the sciences

Marc Levoy

Due to copyright restrictions, some images have been removed from this version of the slides. To see presentation with these images intact, go to:
http://graphics.stanford.edu/courses/cs448a-06-winter/
and look under the heading “Lectures for SIGGRAPH 2007 course on Computational Photography”.

Computer Science Department
Stanford University

34:15 total + 30% = ~45 minutes
What’s going on in the basic sciences?

• new instruments ⇒ scientific discoveries
• most important new instrument in the last 50 years: the digital computer
• computers + digital sensors = computational imaging
  Def: imaging methods in which computation is inherent in image formation.
    – B.K. Horn
• the revolution in medical imaging (CT, MR, PET, etc.) is now happening all across the basic sciences

•1:15
• great source for volume and point data
  – and more importantly...
  – they need our help – lack of good visualization tools is holding them back
  – all the arguments made for ViSC (Visualization in Scientific Computing) in the 1980s now apply to computational imaging
Examples of computational imaging in the sciences

• medical imaging
  – rebinning
  – transmission tomography
  – reflection tomography (for ultrasound)

• geophysics
  ✓ – borehole tomography
  – seismic reflection surveying

• applied physics
  ✓ – diffuse optical tomography
  ✓ – diffraction tomography
  ✓ – scattering and inverse scattering

•30:45 up to molecular probes
•1:45
• biology
  ✔ – confocal microscopy
  ✔ – deconvolution microscopy
• astronomy
  ✔ – coded-aperture imaging
    – interferometric imaging
• airborne sensing
  – multi-perspective panoramas
  – synthetic aperture radar

applicable at macro scale too
• optics
  – holography
  – wavefront coding
Computational imaging technologies used in neuroscience

- Magnetic Resonance Imaging (MRI)
- Positron Emission Tomography (PET)
- Magnetoencephalography (MEG)
- Electroencephalography (EEG)
- Intrinsic Optical Signal (IOS)
- In Vivo Two-Photon (IVTP) Microscopy
- Microendoscopy
- Luminescence Tomography
- New Neuroanatomical Methods (3DEM, 3DLM)
The Fourier projection-slice theorem
(a.k.a. the central section theorem)  [Bracewell 1956]

- $P_\square(t)$ is the integral of $g(x,y)$ in the direction $\square$
- $G(u,v)$ is the 2D Fourier transform of $g(x,y)$
- $G_\square(\omega)$ is a 1D slice of this transform taken at $\square$
- $\mathcal{F}^{-1} \{ G_\square(\omega) \} = P_\square(t)$!

• 1:30
• integrals
  – of attenuation (e.g. of X-rays or visible light), of emission (e.g. fluorescence), etc.
• Ref:
  – Bracewell, R. N. Strip Integration in Radio Astronomy, 
Reconstruction of $g(x,y)$ from its projections

- add slices $G(\omega)$ into $u,v$ at all angles and inverse transform to yield $g(x,y)$, or
- add 2D backprojections $P(t, s)$ into $x,y$ at all angles

$P_\theta(t) = \int_{-\infty}^{+\infty} P_\theta(t, s) \, ds$

$G(\omega) = G(u,v) |_\theta$

0:45
The need for filtering before (or after) backprojection

- sum of slices would create $1/\omega$ hot spot at origin
- correct by multiplying each slice by $|\omega|$, or
- convolve $P(t)$ by $F^{-1}\{|\omega|\}$ before backprojecting
- this is called filtered backprojection

- 1:30
- omega
  - distance from the origin of frequency space
- convolving by $F^{-1}(|w|)$
  - is similar to taking a 2\textsuperscript{nd} derivative
Summing filtered backprojections

Image removed due to copyright restrictions

(from Kak)
Example of reconstruction by filtered backprojection

- X-ray
- sinugram
- filtered sinugram
- reconstruction

Image removed due to copyright restrictions

Ref:
• 1:15
• occlusions
  – violate the assumption that pixels are line integrals, so reconstruction fails
  – thus, you cannot reconstruct volumetric models of opaque scenes using digital photography followed by tomography (although many have tried)
• 0:45
• here’s another violation of the assumptions on tomography, but one that is more easily addressed
• artifacts
  – elongation in direction of available projections
  – object-dependent sinusoidal variations

• Ref:
Reconstruction using the Algebraic Reconstruction Technique (ART)

- applicable when projection angles are limited or non-uniformly distributed around the object
- can be under- or over-constrained, depending on N and M

\[ p_i = \sum_{j=1}^{N} w_{ij} f_j, \quad i = 1, 2, \ldots, M \]

- M projection rays
- N image cells along a ray
- \( p_i \) = projection along ray \( i \)
- \( f_j \) = value of image cell \( j \) (\( n^2 \) cells)
- \( w_{ij} \) = contribution by cell \( j \) to ray \( i \)

(a.k.a. resampling filter)

•1:15
Procedure
• make an initial guess, e.g. assign zeros to all cells
• project onto p₁ by increasing cells along ray 1 until \( \Sigma = p₁ \)
• project onto p₂ by modifying cells along ray 2 until \( \Sigma = p₂ \), etc.
• to reduce noise, reduce by \( \alpha \Delta \tilde{f}^{(k)} \) for \( \alpha < 1 \)

\[
\tilde{f}^{(k)} = \tilde{f}^{(k-1)} - \frac{\tilde{f}^{(k-1)} \cdot (\overline{w}_i - p_i)}{\overline{w}_i \cdot \overline{w}_i} \\
\tilde{f}^{(k)} = k^{th} \text{ estimate of all cells} \\
\overline{w}_i = \text{weights (} w_{i1}, w_{i2}, \ldots, w_{iN} \text{) along ray } i \\
\]

• 1:15
• Formula is derived in Kak, chapter 7, p. 278
• linear system, but big, sparse, and noisy
• ART is solution by method of projections [Kaczmarz 1937]
• to increase angle between successive hyperplanes, jump by 90°
• SART modifies all cells using $f^{(k-1)}$, then increments $k$
• overdetermined if $M > N$, underdetermined if missing rays
• optional additional constraints:
  • $f > 0$ everywhere (positivity)
  • $f = 0$ outside a certain area

Procedure
• make an initial guess, e.g. assign zeros to all cells
• project onto $p_1$ by increasing cells along ray 1 until $\Sigma = p_1$
• project onto $p_2$ by modifying cells along ray 2 until $\Sigma = p_2$, etc.
• to reduce noise, reduce by $\alpha \Delta f^{(k)}$ for $\alpha < 1$

• SIRT = Simultaneous Iterative Reconstruction Technique
• SART = Simultaneous ART
• linear system, but big, sparse, and noisy
• ART is solution by *method of projections* [Kaczmarz 1937]
• to increase angle between successive hyperplanes, jump by 90°
• SART modifies all cells using $f^{(k-1)}$, then increments $k$
• overdetermined if $M > N$, underdetermined if missing rays
• optional additional constraints:
  • $f > 0$ everywhere (positivity)
  • $f = 0$ outside a certain area

• Ref:
  – Olson, T., A stabilized inversion for limited angle tomography.
    Manuscript.
  – 35 degrees missing
• Nonlinear constraints
  \[-f = 0 \text{ outside of circle (oval?)}\]
Borehole tomography

• receivers measure end-to-end travel time
• reconstruct to find velocities in intervening cells
• must use limited-angle reconstruction methods (like ART)

• 0:45
• what are other sciences that use tomography?
  – A. geophysics

• Ref:
Applications

mapping a *seismosaurus* in sandstone using microphones in 4 boreholes and explosions along radial lines

mapping ancient Rome using explosions in the subways and microphones along the streets?

- 1:00
- obvious applications:
  - looking for **oil**
- Left picture is from Reynolds, right picture is from Stanford’s Forma Urbis Romae project
- no dinosaur fossils were harmed in the acquisition of this data ;-(

*Time* =
Optical diffraction tomography (ODT)

- for weakly refractive media and coherent plane illumination
- if you record amplitude and phase of forward scattered field
- then the Fourier Diffraction Theorem says \( \mathcal{F} \{ \text{scattered field} \} = \text{arc in } \mathcal{F} \{ \text{object} \} \) as shown above, where radius of arc depends on wavelength \( \lambda \)
- repeat for multiple wavelengths, then take \( \mathcal{F} \) \(^{-1} \) to create volume dataset
- equivalent to saying that a broadband hologram records 3D structure

- 2:00
- diffraction tomography
  - like any tomographic technique, it is applicable only to semi-transparent objects
  - useful when the object refracts, or diffracts, making Fourier Projection Slice Theorem (and backprojection) invalid
- forward scattered field
  - the shape taken by the initially planar waves after refraction or diffraction
  - which is equivalent to measuring their amplitude and phase
- volume dataset
  - index of refraction as a function of \( x \) and \( y \)
- broadband hologram
  - of a semi-transparent object
  - alternatively, you could hold wavelength constant and vary incident direction, thus filling frequency space with arcs of the same radius

- Ref:
• for weakly refractive media and coherent plane illumination
• if you record amplitude and phase of forward scattered field
• then the Fourier Diffraction Theorem says \( \mathcal{F} \{ \text{scattered field} \} = \text{arc in } \mathcal{F} \{ \text{object} \} \) as shown above, where radius of arc depends on wavelength \( \lambda \)
• repeat for multiple wavelengths, then take \( \mathcal{F}^{-1} \) to create volume dataset
• equivalent to saying that a broadband hologram records 3D structure

1:00

• example (from Devaney)
  – PSF of a **single point scatterer** (real part, as a function of x and y)
  – **measuring phase** typically requires a reference beam and interference between it and the main beam, i.e. a holographic procedure
  – arc is sometimes called Ewald circle

• Ref:
  – Devaney, A., Inverse scattering and optical diffraction tomography, Powerpoint presentation.
• for weakly refractive media and coherent plane illumination
• if you record amplitude and phase of forward scattered field
• then the Fourier Diffraction Theorem says \( \mathcal{F} \{\text{scattered field}\} = \text{arc in } \mathcal{F} \{\text{object}\} \) as shown above, where radius of arc depends on wavelength \( \lambda \)
• repeat for multiple wavelengths, then take \( \mathcal{F}^{-1} \) to create volume dataset
• equivalent to saying that a broadband hologram records 3D structure
Inversion by filtered backpropagation

• depth-variant filter, so more expensive than tomographic backprojection, also more expensive than Fourier method
• applications in medical imaging, geophysics, optics

Diffuse optical tomography (DOT)

- assumes light propagation by multiple scattering
- model as diffusion process

[Arridge 2003]

1:30

so what happens if the object is strongly scattering?

- then it becomes a problem in inverse scattering

Refs:

- Arridge, S.R., Methods for the Inverse Problem in Optical Tomography, 

- Schweiger, M., Gibson, A., Arridge, S.R.,
  “Computational Aspects of Diffuse Optical Tomography,”

- Jensen, H.W., Marschner, S., Levoy, M., Hanrahan, P., A Practical Model for Subsurface Light Transport,
Diffuse optical tomography

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

**acquisition**

- **81 source positions, 81 detector positions**
  - for each source position, measure light at all detector positions
  - use time-of-flight measurement to estimate **initial guess** for absorption, to reduce cross-talk between absorption and scattering
From microscope light fields to volumes

- 4D light field $\rightarrow$ *digital refocusing* $\rightarrow$ 3D focal stack $\rightarrow$ *deconvolution microscopy* $\rightarrow$ 3D volume data

–1:00
–let’s switch gears completely,
  »and talk about something *besides tomography*
–in the paper I’m presenting Wednesday,
  »I’ll be talking about *3D deconvolution*
  »which produces *volume data*
3D deconvolution

focus stack of a point in 3-space is the 3D PSF of that imaging system

- object * PSF $\rightarrow$ focus stack
- $\mathcal{F} \{\text{object}\} \times \mathcal{F} \{\text{PSF}\} \rightarrow \mathcal{F} \{\text{focus stack}\}$
- $\mathcal{F} \{\text{focus stack}\} \odot \mathcal{F} \{\text{PSF}\} \rightarrow \mathcal{F} \{\text{object}\}$
- spectrum contains zeros, due to missing rays
- imaging noise is amplified by division by ~zeros
- reduce by regularization (smoothing) or completion of spectrum
- improve convergence using constraints, e.g. object $> 0$

- 2:30
- object * PSF
  - assumes object is semi-transparent, hence scattering is minimal
  - each pixel is a line integral of attenuation (if backlit) or emission (if frontlit and object is fluorescent)
- missing rays
  - not as bad as a photographic lens, because microscope lenses have a very wide aperture,
  - for example, 0.95NA = about 70 degrees each side of the optical axis, as shown in the PSF above
- Ref:
Silkworm mouth
(40x / 1.3NA oil immersion)

slice of focal stack  slice of volume  volume rendering

• 0:45
From microscope light fields to volumes

- 4D light field \( \rightarrow \) digital refocusing \( \rightarrow \)
  3D focal stack \( \rightarrow \) deconvolution microscopy \( \rightarrow \)
  3D volume data

\[ \text{(DeltaVision)} \]

- 4D light field \( \rightarrow \) tomographic reconstruction \( \rightarrow \)
  3D volume data

\[ \text{(from Kak)} \]

- 1:15
- tomographic reconstruction
  - backprojection doesn’t work well, because we don’t have all rays
  - but we could use ART

- it turns out that tomographic reconstruction using \text{SART with a positivity constraint}
  - is the same as digital refocusing followed by \text{deconvolution microscopy with the same positivity constraint}
  - we prove this in our SIGGRAPH 2006 paper, Light Field Microscopy
Optical Projection Tomography (OPT)

- 1:15
- OPT
  - reconstruction of semi-transparent objects from digital photographs
  - object must be immersed in refraction index-matched fluid
- James Sharpe, Science, 2002
  - object was fluorescent, so each pixel is a line integral of emission
  - and microscopes are orthographic,
  - hence tomography applies
- Trifonov EGSR 2006
  - object was fully transparent, immersed in colored liquid and backlit, so each pixel is a line integral of attenuation
  - digital cameras are perspective,
  - hence he had to use ART
- if this works, then what good is the extra dimension in the light field?
  - come to my talk on Wednesday

Time = 31
Coded aperture imaging

• optics cannot bend X-rays, so they cannot be focused
• pinhole imaging needs no optics, but collects too little light
• use multiple pinholes and a single sensor
• produces superimposed shifted copies of source

•1:30
•Ref:

Reconstruction by backprojection

- backproject each detected pixel through each hole in mask
- superimposition of projections reconstructs source + a bias
- essentially a cross correlation of detected image with mask
- also works for non-infinite sources; use voxel grid
- assumes non-occluding source

\[\text{(from Zand)}\]

• assumes non-occluding source
  – otherwise it’s the **voxel coloring problem**
A simple example:

- **cross correlation is just convolution** (of detected image by mask) without first reversing detected image in x and y
- conversion of blacks to -1’s in “decoding matrix” just serves to avoid normalization of resulting reconstruction (to remove the bias)
- performing this on an **image of gumballs**, rather than a 3D gumball scene, is equivalent to assuming the **gumballs cover the sky at infinity**, i.e. they are an angular function

- I would love to try this in my lab
  - using a mask with holes, a screen behind it, and a camera to photograph the image falling on the screen

**Ref:**

- Paul Carlisle,
  Coded Aperture Imaging,
  [http://www.paulcarlisle.net/old/codedaperture.html](http://www.paulcarlisle.net/old/codedaperture.html)
Computational imaging below the diffraction limit

- Molecular probes
  - 50-nm fluorescent microspheres
  - smaller than the wavelength of light
  - in fact their size controls their color
  - too small to image, you say?
- Gustafsson
  - uses structured illumination and non-linear (saturation) imaging to beat the diffraction limit by 10x (in X and Y)!
- Ref:
  - Gustafsson, M.G.L.,
  Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution,
  Proc. National Academy of Sciences (PNAS),
Interesting techniques
I didn’t have time to cover

• reflection tomography
• synthetic aperture radar & sonar
• holography
• wavefront coding
Smart Optics, Modern Sensors and Future Cameras

Ramesh Raskar
Mitsubishi Electric Research Labs

Jack Tumblin
Northwestern University

Course WebPage:
http://www.merl.com/people/raskar/photo

Computational Photography

Novel Illumination
Light Sources

Generalized Sensor

Processing

Generalized Optics

Display

Recreate 4D Lightfield

Scene: 8D Ray Modulator
Future Directions

- Smart Lighting
  - Light stages, Domes, Light waving, Towards 8D
- Computational Imaging outside Photography
  - Tomography, Coded Aperture Imaging
- Smart Optics
  - Handheld Light field camera, Programmable imaging/aperture
- Smart Sensors
  - HDR Cameras, Gradient Sensing, Line-scan Cameras, Demodulators
- Speculations

Wavefront Coding: 10X Depth of Field

- In-focus: small ‘Circle of Confusion’
- Out-of-focus: LARGE “circle of confusion”
- Coma-like distortion: Make Circle MOVE as focus changes:

http://www.cdm-optics.com/site/extended_dof.php
Wavefront Coding: 10X Depth of Field

- In-focus: small ‘Circle of Confusion’
- Out-of-focus: LARGE “circle of confusion”
- Coma-like distortion allows us to De-convolve, sharpen out-of-focus items

http://www.cdm-optics.com/site/extended_dof.php

Light field photography using a handheld plenoptic camera

Ren Ng, Marc Levoy, Mathieu Brédif, Gene Duval, Mark Horowitz and Pat Hanrahan
Conventional versus light field camera

Subject → Main lens → Photosensor

Subject → Micro lens array → Main lens → Photosensor

Conventional versus light field camera

Subject → Main lens → Photosensor

uv-plane → st-plane → Main lens
Conventional versus light field camera

Prototype camera

Contax medium format camera
Kodak 16-megapixel sensor

Adaptive Optics microlens array
125μ square-sided microlenses

4000 × 4000 pixels ÷ 292 × 292 lenses = 14 × 14 pixels
Digital refocusing

refocusing = summing windows extracted from several microlenses
Example of digital refocusing

Extending the depth of field

conventional photograph, main lens at $f/4$
conventional photograph, main lens at $f/22$
light field, main lens at $f/4$, after all-focus algorithm [Agarwala 2004]
Programmable Imaging

Vision → Programmable Controller → New Optics → Detection

Computations

Geometry: Ray Orientation
\[ o_i (x_i) = G(n_i) \]

Photometry: Ray Attenuation
\[ a_i (x_i) = \frac{t(n_i)}{t(n_i) + t(n_b)} \]

Imaging Through Micro-Mirrors (Nayar, Branzoi and Boult, 2004)
Digital Micromirror Device (DMD)
(by Texas Instruments)

DMD Array:

Micromirror Architecture:

-10° to 10°

DMD for Imaging:
(Malbet et al. 95, Kearney et al. 98, Castracane et al. 99, Christensen et al. 02)

Programmable Imaging System

Imaging Lens
DMD Electronics

Camera Electronics
Tilted CCD
Lens Focused on DMD
### Modulation: Examples

<table>
<thead>
<tr>
<th>Scene</th>
<th>DMD Image</th>
<th>Camera Image</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Scene Image" /></td>
<td><img src="image2.png" alt="DMD Image" /></td>
<td><img src="image3.png" alt="Camera Image" /></td>
</tr>
<tr>
<td><img src="image4.png" alt="Scene Image" /></td>
<td><img src="image5.png" alt="DMD Image" /></td>
<td><img src="image6.png" alt="Camera Image" /></td>
</tr>
<tr>
<td><img src="image7.png" alt="Scene Image" /></td>
<td><img src="image8.png" alt="DMD Image" /></td>
<td><img src="image9.png" alt="Camera Image" /></td>
</tr>
</tbody>
</table>

### Optical Intra-Pixel Feature Detection

**Laplacian**

\[ f \ast g = f \ast g^+ - f \ast g^- \]

**Laplacian Image:**

<table>
<thead>
<tr>
<th>a b c</th>
<th>d e f</th>
<th>g h i</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 4 1</td>
<td>4 -20 4</td>
<td>1 4 1</td>
</tr>
<tr>
<td>4 0 4</td>
<td>4 0 4</td>
<td>1 4 1</td>
</tr>
<tr>
<td>1 4 1</td>
<td>1 4 1</td>
<td>0 0 0</td>
</tr>
<tr>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
</tbody>
</table>

**Lapacian Image:**

<table>
<thead>
<tr>
<th>a b c</th>
<th>d e f</th>
<th>g h i</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 4 1</td>
<td>4 0 4</td>
<td>1 4 1</td>
</tr>
<tr>
<td>4 0 4</td>
<td>4 0 4</td>
<td>1 4 1</td>
</tr>
<tr>
<td>1 4 1</td>
<td>1 4 1</td>
<td>0 0 0</td>
</tr>
<tr>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
</tbody>
</table>

**Lapacian Image:**

<table>
<thead>
<tr>
<th>a b c</th>
<th>d e f</th>
<th>g h i</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 4 1</td>
<td>4 0 4</td>
<td>1 4 1</td>
</tr>
<tr>
<td>4 0 4</td>
<td>4 0 4</td>
<td>1 4 1</td>
</tr>
<tr>
<td>1 4 1</td>
<td>1 4 1</td>
<td>0 0 0</td>
</tr>
<tr>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
</tbody>
</table>
Optical Edge Detection

Scene Video

Edge Video

Generalized Optics and Sensors

- Smart Optics
  - Handheld Light field camera,
  - Programmable imaging/aperture

- Smart Sensors
  - HDR Cameras,
  - Gradient Sensing,
  - Line-scan Cameras,
  - Demodulators
Future Directions

• Smart Lighting
  – Light stages, Domes, Light waving, Towards 8D
• Computational Imaging outside Photography
  – Tomography, Coded Aperture Imaging
• Smart Optics
  – Handheld Light field camera, Programmable imaging/aperture
• Smart Sensors
  – HDR Cameras, Gradient Sensing, Line-scan Cameras, Demodulators
• Speculations

Foveon: Thick Sensor

COLOR FILM contains three layers of emulsion which directly record red, green, and blue light.

TYPICAL DIGITAL SENSORS have just one layer of pixels and capture only part of the color.

Foveon's Foveon XL sensors have three layers of pixels which directly capture all of the color.
Pixim

Fuji's SuperCCD S3 Pro camera has a chip with high and low sensitivity sensors per pixel location to increase dynamic range
Gradient Camera

Sensing Pixel Intensity Difference with Locally Adaptive Gain

Ramesh Raskar, MERL
Work with Jack Tumblin, Northwestern U,
Amit Agrawal, U of Maryland

Natural Scene Properties

Intensity vs x

Intensity Histogram

Gradient vs x

Gradient Histogram
Intensity values ranging from 0 to 1800
Intensity ramp plus low contrast logo

8 bit camera for 1:1000 range
Problem: saturation at high intensity regions

8 bit log for 1:10^6 range
Problem: Visible quantization effects at high intensities

Locally Adaptive Gain
Pixel divided by the average of local neighborhood.
Thus the low frequency contents are lost and only detail remains.

Gradient Camera Image
In proposed method, we sense intensity differences. We use a 8 bit A/D with resolution of log(1.02) to capture 2% contrast change between adjacent pixels. Notice that the details at both high and low intensities are captured.

Gradient Camera

- Two main features
  1. Sense difference between neighboring pixel intensity
     At each pixel, measure \((\nabla_x, \nabla_y)\), \nabla_x = I_{x+1, y} - I_{x, y}   ,   \nabla_y = I_{x, y+1} - I_{x, y}
  2. With locally adaptive gain

- Gradient camera is very similar to locally adaptive gain camera
- Locally Adaptive Gain Camera
  - Gain is different for each pixel
  - Problem: Loses low frequency detail and preserves only high frequency features (edges)
- Gradient Camera
  - The gain is same for four adjacent pixels
  - Difference between two pixels is measured with same gain on both pixels
  - Reconstruct original image in software from pixel differences by solving a linear system (solving Poisson Equation)
**Camera Pipeline**

On-board Hardware

- Log
- Gain
- A/D
- Diff

4 Pixel Cliques

Difference between pixels

Local gain adaptive to difference

Software

- 2D Integration to reconstruct the image

**Detail Preserving**

Intensity Camera  Log Intensity Camera  Gradient Camera

*Intensity cameras capture detail but lose range*

*Log cameras capture range but lose detail*
Quantization

Original Image
Uniform quantization 3 bits
Log Uniform quantization 3 bits
Log Uniform gradients quantization 3 bits
GradCam requires fewer bits
In the reconstructed image, error is pushed to high gradient pixel positions which is visually imperceptible

High Dynamic Range Images

Scene
Intensity camera saturation map
Gradient camera saturation map

Intensity camera fail to capture range
Gradients saturate at very few isolated pixels
3D Cameras

- **Time of flight**
  - ZCam (Shuttered Light Pulse)

- **Phase Decoding of modulated illumination**
  - Canesta (Phase comparison)
  - Phase difference = depth
  - Magnitude = reflectance

- **Structured Light**
  - Binary coded light and triangulation

---

**ZCam (3Dvsystems), Shuttered Light Pulse**

Resolution: 1cm for 2-7 meters
Graphics can be inserted behind and between characters.

**Figure 2. Time of flight measurement.**

**Figure 3. A method of phase/amplitude calculation.**

Objects that lay outside the designated range (both near the camera or further away in the background) are not included in the key.
Demodulating Cameras

- Motion Capture Cameras
  - Visualeyez™ VZ4000 Tracking System
  - PhaseSpace motion digitizer
Demodulating Cameras

- Decode signals from blinking LEDs + image
  - Sony ID Cam
  - Phoci

Fluttered Shutter Camera

Raskar, Agrawal, Tumblin Siggraph2006
Figure 2 results

Input Image

Rectified Image to make motion lines parallel to scan lines.
Approximate cutout of the blurred image containing the taxi (vignetting on left edge). Exact alignment of cutout with taxi extent is not required.

Image Deblurred by solving a linear system. No post-processing.

Coded Exposure Photography
Raskar, Agrawal, Tumblin Siggraph2006

<table>
<thead>
<tr>
<th>Short Exposure</th>
<th>Traditional</th>
<th>MURA</th>
<th>Coded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Coded Exposure Photography

Discrete Fourier Transform of Convolving Filter

Converting Deblurring into a Well-posed Problem
Novel Sensors

- Gradient sensing
- HDR Camera, Log sensing
- Line-scan Camera
- Demodulating
- Motion Capture
- Fluttered Shutter
- 3D

Fantasy Configurations

- ‘Cloth-cam’: ‘Wallpaper-cam’ elements 4D light emission and 4D capture in the surface of a cloth...

- Floating Cam: ad-hoc wireless networks form camera arrays in environment...

- Other ray sets: Multilinear cameras, canonical ‘basis’ cameras (linear combination of 8 types) McMillan’04, ‘05
Dream of A New Photography

<table>
<thead>
<tr>
<th></th>
<th>Old</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>People and Time</td>
<td>~Cheap</td>
<td>Precious</td>
</tr>
<tr>
<td>Each photo</td>
<td>Precious</td>
<td>Free</td>
</tr>
<tr>
<td>Lighting</td>
<td>Critical</td>
<td>Automated?</td>
</tr>
<tr>
<td>External Sensors</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>‘Stills / Video’</td>
<td>Disjoint</td>
<td>Merged</td>
</tr>
<tr>
<td>Exposure Settings</td>
<td>Pre-select</td>
<td>Post-Process</td>
</tr>
<tr>
<td>Exposure Time</td>
<td>Pre-select</td>
<td>Post-Process</td>
</tr>
<tr>
<td>Resolution/noise</td>
<td>Pre-select</td>
<td>Post-Process</td>
</tr>
<tr>
<td>‘HDR’ range</td>
<td>Pre-select</td>
<td>Post-Process</td>
</tr>
</tbody>
</table>

Computational Photography

Novel Illumination
- Light Sources
- Modulators
- Generalized Optics
- 4D Incident Lighting
- Scene: 8D Ray Modulator

Novel Cameras
- Generalized Sensor
- Processing
  - Ray Reconstruction
  - Upto 4D Ray Sampler
- Generalized Optics
- 4D Ray Bender
- 4D Light Sampler
- 4D Light Field
- Display
- Recreate 4D Lightfield
**Digital Photography**

- Image processing applied to captured images to produce "better" images.

**Computational Processing**

- Processing of a set of captured images to create "new" images.
  - Examples: Mosaic, Matting, Super-Resolution, Multi-Exposure HDR, Light field from Multiple View, Structure from Motion, Shape from X.

**Computational Imaging/Optics**

- Capture of optically coded images and computational decoding to produce "new" images.
  - Examples: Coded Aperture, Optical Tomography, Diaphanography, SA Microscopy, Integral Imaging, Assorted Pixels, Catadioptric Imaging, Holographic Imaging.

**Computational Sensor**

- Detectors that combine sensing and processing to create "smart" pixels.
  - Examples: Artificial Retina, Retinex Sensors, Adaptive Dynamic Range Sensors, Edge Detect Chips, Focus of Expansion Chips, Motion Sensors.

**Computational Illumination**

- Adapting and controlling illumination to create 'revealing' image.
  - Examples: Flash/no flash, Lighting domes, Multi-flash for depth edges, Dual Photos, Polynomial texture Maps, 4D light source.

---

**Acknowledgements**

- MERL, Northwestern Graphics Group
- Amit Agrawal
- Shree Nayar
- Marc Levoy
- Jinbo Shi
- Ankit Mohan, Holger Winnemoller

**Image Credits**

- Ren Ng, Vaibhav Vaish, William Bennet
- Fredo Durand, Aseem Agrawala
- Morgan McGuire, Paul Debevec
- And more
Table-top Computed Lighting for Practical Digital Photography

Ankit Mohan\textsuperscript{1}, Jack Tumblin\textsuperscript{1}, Bobby Bodenheimer\textsuperscript{2}, Cindy Grimm\textsuperscript{3}, Reynold Bailey\textsuperscript{3}

\textsuperscript{1}Northwestern University, \textsuperscript{2}Vanderbilt University, \textsuperscript{3}Washington University in St. Louis

Abstract

We apply simplified image-based lighting methods to reduce the equipment, cost, time, and specialized skills required for high-quality photographic lighting of desktop-sized static objects such as museum artifacts. We place the object and a computer-steered moving-head spotlight inside a simple foam-core enclosure, and use a camera to quickly record low-resolution photos as the light scans the box interior. Optimization guided by interactive user sketching selects a small set of frames whose weighted sum best matches the target image. The system then repeats the lighting used in each of these frames, and constructs a high resolution result from re-photographed basis images. Unlike previous image-based relighting efforts, our method requires only one light source, yet can achieve high resolution light positioning to avoid multiple sharp shadows. A reduced version uses only a hand-held light, and may be suitable for battery-powered, field photography equipment that fits in a backpack.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture I.3.3 [Computer Graphics]: Picture/Image Generation

1. Introduction

Modern digital cameras have made picture-taking much easier and more interactive. However, lighting a scene for good photography is still difficult, and practical methods to achieve good lighting have scarcely changed at all. We show that sketch-guided optimization and simplified forms of image-based lighting can substantially reduce the cost, equipment, skill, and patience required for small-scale studio-quality lighting.

Good studio lighting is difficult because it is a 4D inverse problem that photographers must solve by making successive approximations guided by years of experience. For non-experts, good studio lighting can be surprisingly frustrating. Most people can specify the lighting they want in screen space (e.g., “get rid of this obscuring highlight, make some shadows to reveal rough texture here, but fill in the shadows there”), but determining what kind of lights to use, where to place them, and how to orient them is never easy.

We are especially interested in camera-assisted lighting for human-scale, desktop-sized static objects. We want lighting that accurately reveals the shape, texture, materials, and most visually meaningful features of the photographed item. In particular, we seek a method to help museum curators as they gather digital photographic archives of their vast collections of items.

Pioneering work in image-based lighting [DHT\textsuperscript{00}, HCD\textsuperscript{01}, DWT\textsuperscript{02}, MPDW\textsuperscript{03}] offers promising approaches that can help with the photographic lighting problem. Unfortunately, most require too many precise measurements and adjustments for day-to-day use outside the laboratory. Precision is required to address more ambitious goals such as recovering shape, BRDF, and appearance under arbitrary viewing and lighting conditions. For the much smaller, yet more widespread problem of photographic lighting, we need a method that requires less time, expense, and complexity, yet allows users who are not lighting experts to quickly find the lighting they want.

This paper offers three contributions. We extend existing...
image-based lighting ideas to reduce the required equipment to a single light source and single camera; we replace trial-and-error light repositioning with optimization and on-screen painting; and we reduce the need for high dynamic range photography, thus reducing the capture time. The result is a novel and inexpensive system that a novice can use to intuitively describe and obtain the desired lighting for a photograph.

2. Related Work

Lighting has long been recognized as a hard problem in computer graphics and many papers have explored optimization for light placement and other parameters [SDS93, KPC93, PF95, CSF99, SL01]. Some of these systems used painting interfaces to specify desired lighting in a 3D scene [SDS93, PF95, PRJ97], and we use a similar approach to make lighting for photography more intuitive. The system by Shacked et al. [SL01] was even able to provide fully automatic lighting by applying image quality metrics. Marschner et al. [MG97] used inverse rendering techniques to estimate and alter the directional distribution of incident light in a photograph. However, all these systems require 3D information unavailable in our photographic application.

Several commercial photographic products have also used lighting enclosures similar to ours, but they achieve very soft lighting with limited user controls. Moreover, they do not help users solve light placement problems. These systems include diffusive tents [Pho], photo-boxes [MK] and translucent back-lit platforms with an array of individually dimmed light sources [Ast].

Image-based methods have also been used to permit arbitrary relighting of well-measured objects. Most methods, including ours, perform relighting using a weighted sum of differently lit basis images, done first by [NSD94]. However, prior efforts used more elaborate and expensive equipment because their goals were different from ours. These include measurement of a 4D slice of the reflectance field of the human face [DHT00], museum artifacts measured by a rotating-arm light stage [HCD01], an ingenious but extensive system by Debevec et al. [DWT02] for real-time video playback and measurement of light fields, a dome of electronic flashes for real time image relighting [MGW01], a free form light stage to enable portable gathering of light-field data with some calibration [MDA02], and full 4D incident light measurements by Masselus et al. [MPDW03]. In all of these cases, data-gathering required either customized equipment or collection times much longer than would be practical for photographic lighting.

Three recent systems also offered novel sketch guided relighting from basis images. Akers et al. [ALK03] used a robotic light-positioning gantry to gather precisely lit images, and like us, provided a painting interface to guide relighting. But unlike us they used spatially varying weights that could produce physically impossible lighting. Digital Photomontage [ADA04] used sketch guided graph-cut segmentation coupled with gradient domain fusion to seamlessly merge several photographs. They demonstrated merging differently lit photographs to create novel illumination conditions. Though their interaction scheme worked well for a small number of images (~10), it may be impractical for the hundreds of images required for complete control over lighting directions. Also, their system does nothing to help the user with light placement, and may produce physically unrealizable results. Anrys and Dutre [AD04] used a Debevec-style light stage with around 40 fixed, low powered light sources and a painting interface to guide lighting. Their optimization only found light intensities, and light placement was still left up to the user. Also, their point light sources could cause multiple shadows and highlights which might be undesirable for archival purposes. The data capture time was high since they captured high-dynamic-range (HDR) photos for every light location.

Unlike previous attempts, our system does not require users to decide on correct or complete light source placement. This is possible because our capture process is significantly dif-
different, and better suited for the task of photography. We require less than five minutes to capture the initial image and a few more minutes to get the final result. The equipment required is minimal and portable, and our handheld version can be carried in a backpack. Also, HDR capture is reduced to a minimum in our system.

3. Simplifications: HDR and 2D lighting

Our goal is to do what a good photographer does, but with computational help. We want to light a scene for a particular photograph, not build a calibrated 4D data set to reconstruct every possible form of illumination. Photographers make consistent choices about which types of lights to use, how to adjust them, and where to place them. We will show how our streamlined image-based method follows these same choices.

Like most previous image-based lighting methods, we apply the observations formalized by Nimeroff [NSD94] that lights and materials interact linearly. If a fixed camera makes an observation, then the lights are perceived as a single light source. Multiple light sources result in multiple hard shadows, while overlapping area light sources can be used to simulate a larger light source.

Figure 2: All possible lighting angles parameterized by light position $(\theta_p, \phi_p)$ and direction $(\theta_a, \phi_a)$. Point light sources (on the left side of the hemisphere) result in multiple hard shadows, while overlapping area (on the right) light sources can be used to simulate a larger light source.

Formally, arbitrary external illumination is four-dimensional for a desktop scene: $L(\theta_p, \phi_p, \theta_a, \phi_a) = L(\Theta)$. Suppose that the photographed object receives all its light from a hemisphere of tiny, invisible, inward-pointing video projectors, each at a distance $r$ from the object. Each projector’s position in desktop polar coordinates is $(\theta_p, \phi_p)$. Each projector’s centermost pixel $P(\theta_p, \phi_p)$ forms a ray that illuminates the center point of our desktop, and in the projector’s polar coordinates the other pixels are $P(\theta_a, \phi_a)$, as shown in Figure 2. All projectors’ light output is the 4-D incident light field, and describes all possible lighting. To simulate all possible lighting, we would need a new image $I$ to capture object. Distance to the light affects foreshortening of shadow shapes, but these effects are subtle and rarely noticed. Third, they adjust lights to control shadow softness versus sharpness. Light sources (or more accurately, the shadows they form) become ‘softer’ by increasing the angular extent as measured from the lit object. Fourth, they seek out lighting arrangements that produce a simple set of shadows and highlights that best reveal the object’s shape, position, and surface qualities. They avoid complex overlapping shadows, lack of shadows due to overly-soft light, and contrast extremes due to large specular highlights or very dark shadows. Simpler shadows usually mean fewer lights, and thus fewer basis images.

Accordingly, we use commercially available light sources instead of custom or special-purpose devices. We place light sources at a moderate distance (typically around 1 meter) from the object. We use small-to-moderate area ‘soft’ light sources instead of the much sharper point-like sources often used in earlier approaches. Overlapped soft shadows blend far less noticeably than sharp shadows from the same light positions (as shown in Figure 2), thus requiring fewer images to avoid multiple shadow artifacts. Also, overlapping area light sources can be combined to produce a larger area light source.

Note that we do not need to know the light positions or their absolute intensities for our images; we select weights $w_i$ and images $I_i$ by their ability to match the lighting target images a user sketches for us. Instead of calibration, we only need consistency in the aiming direction of a single, commercially available steerable light, and consistency in the light response curve of a commercially available digital camera.

We also avoid the use of HDR photographs where possible, as these typically require multiple calibrated exposures and computation to merge them [DM97]. Instead, we rely on the camera’s automatic exposure adjustments to capture what we call light-aiming images suitable for interactive lighting design. We photograph high resolution basis images afterwards, for construction of the output image, and only resort to HDR capture methods for a basis image with large overexposed regions. Under-exposed regions can be ignored, as their contributions are already invisible, and are further reduced as their weights are less than one ($w_i \leq 1$).

Formally, arbitrary external illumination is four-dimensional for a desktop scene: $L(\theta_p, \phi_p, \theta_a, \phi_a) = L(\Theta)$. Suppose that the photographed object receives all its light from a hemisphere of tiny, invisible, inward-pointing video projectors, each at a distance $r$ from the object. Each projector’s position in desktop polar coordinates is $(\theta_p, \phi_p)$. Each projector’s centermost pixel $P(\theta_p, \phi_p)$ forms a ray that illuminates the center point of our desktop, and in the projector’s polar coordinates the other pixels are $P(\theta_a, \phi_a)$, as shown in Figure 2. All projectors’ light output is the 4-D incident light field, and describes all possible lighting. To simulate all possible lighting, we would need a new image $I$ to capture
light from each pixel of each video projector! Instead, we use only broad beams of light \( P(\theta_a, \phi_a) \propto \cos(\theta_a) \cos(\phi_a) \), regular sampling of light placement angles \((\theta_p, \phi_p)\), and specify ‘softer’ to ‘sharper’ shadows by varying the angular extent \((\theta_p, \phi_p)\) as measured from the lit object. This angular extent should not be confused with the lamp’s beam width \((\theta_a, \phi_a)\); in our ‘hemisphere of video projectors’ analogy, beam width sets the image from a projector, but angular extent sets the number of adjacent projectors that emit this same image.

In summary, rather than recreate arbitrary 4D incident light fields, we use weighted sums of basis images that represent the type of lighting used by professional photographers. This method is much more practical and efficient, with little, if any, loss of useful generality.

4. Method

We construct a high quality user-guided picture in three steps. First the system automatically captures low-resolution light-aiming photos for densely sampled lighting angles around the photographed object. These quick photos are used only to guide the lighting design, not to form the final output. Second, the user iteratively paints the desired lighting by simple lighten-darken operations to generate a target image. The system finds weights \( w_i \) for each light-aiming photo such that their weighted sum matches the target image in the least-squares sense. Finally the system takes a few selected high resolution basis images by relighting the scene from light source positions that have weights \( w_i \) greater than a threshold. A weighted sum of these high resolution images gives the final result. If the result is not satisfactory, the user can sketch on the current result for use as the next iteration’s target image.

4.1. Enclosed Light Source & Aiming Images

Freed from photometric and angular calibration requirements as discussed in Section 3, we are able to build a much simpler and cost-effective controlled light source. We place the object and a gimbal-mounted moving-head spotlight inside an enclosure of almost any convenient size, shape and material. The powerful computer-airmed light pivots to any desired pan and tilt angle with good repeatability \((\leq \pm 0.5^\circ)\) to light any desired spot inside our enclosure. The enclosure acts as a reflector, and effectively provides a controllable 2D area light source around the object. The size and shape of the enclosure is almost irrelevant as long as the light is close enough to the object to keep parallax low, and the light is powerful enough for the camera to get a reasonable exposure.

We built a 1 \( \times \) 1 \( \times \) 1.5 \( m^3 \) sized box of white 1/2” foam-core board as our enclosure, and chose an inexpensive moving-head spotlight. The 150-watt American DJ Auto Spot 150 disco-light, shown in Figure 1 can tilt 270°, pan 540°, and includes 9 color filters, gobos and several other fun features.

Computer control by the DMX512 protocol is easy to program with the SoundLight USB DMX controller. Our foam-core enclosure resembles a hemi-cube around a pair of tables. We place the gimbal light on a small table that lowers its rotation center to the plane of an adjacent taller table holding the photographed object, as shown in Figure 3. Using adjacent but separate tables reduces vibration, permits gimbal angles to approximate hemisphere angles, and separates the object from the swiveling lamp. We place the camera behind a small opening cut in the enclosure wall on the end farthest from the light source.

The system gathers aiming images rapidly and automatically. Through the DMX512 controller we direct the gimbal light to scan the upper hemisphere of light aiming directions in equal-angle increments as we record low-resolution aiming images, either by collecting viewfinder video \((320 \times 240@10Hz)\) or by individual computer-triggered photographs using auto-exposure. We are able to record hundreds of individual aiming images per minute, and can complete all the data gathering in less than five minutes using a Pentium 2GHz computer, and a Canon Powershot G3 camera.

To the best of our knowledge, no other image-based lighting work exploits these movable and controllable lights. Enclosed pivoting lights retain many advantages of the more sophisticated lighting systems, avoid multiple sharp shadows, can offer variable ‘softness’ by spot size adjustment, and are much simpler and cheaper to construct. Of course, they do not easily provide accurate lighting direction calibration or point-light illumination, but these features are not needed for our goals.

After recording, we linearize each captured frame (RGB) by applying the camera’s inverse response curve, recovered by the method of Debevec et al. [DM97], and converted to luminance values. Linear response ensures weighted sums of
whole images are accurate representations of physically realizable lighting. We then down-sample the linearized aiming image dataset to 64 × 64 for use as the aiming basis set for the following optimization step.

### 4.2. Sketch-Guided Lighting Optimization

After gathering aiming images, users can interactively specify and refine lighting by sketching the desired intensity on a target intensity image. This grey-scale image (examples in Figure 5) approximates the final output image the user would like to see. For editing the target image, the user starts off either with a simple grey wash (such as uniform grey, or light grey fading to dark grey across the image, etc.), or the previous iteration’s result. The user then carries out a series of lighten and darken operations in the different regions of the image to approximate the desired results. The process is extremely simple and intuitive, and takes a few minutes at most.

Given a target image, the optimization finds weights \( w_i \) for each aiming image that produces the best match to the target image. We take a constrained least-squares approach, solving for weights \( w_i \) for each of the small, luminance-only aiming basis images. Let \( N \) be the number of images in the aiming image set, each of size \( m \times n \). We formulate the optimization problem as follows:

\[
\min_w |Aw - t|^2
\]

subject to \( 0 \leq w_i \leq 1 \) \( \forall i \in (1 \ldots N) \)

where \( w \) is the \( N \)-dimensional vector of weights, \( A \) is an \( (m \times n) \times N \) matrix of basis images (that is, each basis image is treated as a vector), \( t \) is the \( (m \times n) \) vector representing the target image painted by the user, and \( | . | \) represents the \( L^2 \) norm of the vector. We solve this bound constrained quadratic optimization problem using an active set method [NW99]. The optimization is quite fast and takes around 1-2 minutes on a 2GHz Pentium 4 desktop machine.

The result is a least-squares optimal match to the supplied target image. As the objective function is quadratic, weights for images with weak contributions are rapidly driven to zero. In our experience, the number of significant nonzero weights is consistently small (5 – 15). This greatly reduces the number of images needed for the final lighting solution.

After finding the \( w_i \) weights, we apply them to the linearized color aiming images, then re-apply the camera response function to display a preview of the output image. The user then has the option of replacing the target with a grayscale version of this result and can repeat the sketching and optimization cycle until satisfied with the color preview of the output image.

### 4.3. Output Assembly

The user now has the desired visually pleasing, but low-resolution, image that is a weighted sum of a small subset of the linearized aiming images. For high-quality results, we wish to replace each of these aiming images with an image taken at the maximum resolution available from the camera. We re-take just those photos that correspond to the aiming images with significant weights \( w_i \), again using auto-exposure on the camera, and record a set of high-resolution photos called basis images. Recall that we can exactly replicate the lighting using the gimballed spotlight; the only things that change are the camera settings.

We capture HDR photographs for images that contain large over-exposed regions as a result of the camera’s autoexposure. As discussed in Section 3, under-exposed regions do not require HDR photos. We then linearize each basis image to remove effects of the camera response curve. As before, we construct a linear output image as a weighted sum of basis images, using the weights determined by the optimization to match the target image. Finally, we re-apply the camera’s response function to the linear output image to get the desired high resolution result.

### 5. Portable, Hand-held Method

Even a foam-core box and a moving-head spotlight are impractical to carry around everywhere. However, the “Free-form light-stage” [MDA02] showed that it is possible to gather calibrated image sets suitable for 2D relighting with nothing more than four small light-probe-like spheres, a digital camera on a tripod, a hand-held point-light source, possibly battery-powered, and approximately 30 minutes of time to take several hundred digital photographs. Pang et al. [PWH04] also used a similar approach by mounting a
camera on the light source and used camera calibration techniques to estimate lighting directions with reasonable accuracy. While these methods try to meet the ambitious goal of incident light field capture, they would tax anyone’s patience to record more than just a few items. We present a faster and simpler variant that serves our purposes better.

In the method of Section 4, we required repeatable light source positioning. However, if we record all of our ‘aiming images’ at the final output resolution, and if we either ignore over-exposed specular highlights or record high dynamic range images when needed, then repeatability is not needed. This allows us to use a hand-held light source instead. As shown in Figure 4, we use a small 250W hand-held light intended for television news cameras, attached to a diffuse reflector (foam core again), and limit the beam width with barn-doors to form a well-defined area light source.

To gather all photos, we hold the light outstretched and “dance” (see video). We sample the hemisphere of lighting directions by a polar-coordinate scan in $\phi$-major order as the camera takes sequential photographs. A Nikon D70 camera, takes a steady stream of photos at about 3 frames per second using autoexposure for each frame. The user stands facing the object, and holds the light at arms’ length while moving the lamp in an arc that passes directly over the object. The user moves the lamp from one side of the table to the other, scanning by $\pi$ radians in $\theta$ axis with constant $\phi$, and the natural alignment of their shoulders helps aim the light’s centerline directly at the object. After each pass over the object with the light, the user steps sideways to change the $\phi$ angle for the next scan, and makes enough of these passes to cover $0 \leq \phi < \pi$ radians. In practice the user can be more careless with the light, as long as the hemisphere of light directions is well-sampled and the images are not over-exposed. After the image capture dance is complete, we downsample all images to construct aiming photos, and proceed with the sketch guided lighting design as before.

We find this process is quite simple and pleasing, and in under three minutes we can gather around 150 high-quality aiming/basis photos. An experienced user might not need to scan the whole hemisphere, but can quickly illuminate just from the useful and interesting lighting directions.

6. Results

Images in Figure 5 show results from our sketch guided lighting system. Both the moving-head light and the hand-held methods are equally successful at creating arbitrary cleanly-lit images of desktop-sized objects. The data sets gathered by either method is sufficiently dense to allow easy lighting design. Additionally, our system yields reasonable results even when presented with unrealistic targets or highly reflective objects.

Figure 5(a), demonstrates a user interaction sequence with the system. Starting from a uniform grayscale image as the target, the user guides the optimization, iteratively improving the target until she gets the desired output. Figure 5(b) shows how simple approximate sketching on the target image can give an interesting sidelighting effect. Figure 5(c) shows how the highlight can bring out the underlying texture in a surface.

Figure 5(d) shows lighting for a highly specular object. Good lighting for such smooth, highly reflective objects is always difficult, as the light source itself is visible in the reflection. Our system produces results similar to the target image without large, objectionable saturated regions. In future systems we may hide the enclosure seams by constructing wide smooth rounded corners resembling a photographer’s ‘cyc’.

Figure 5(f) shows results from the handheld method of Section 5. The data gathering time was under 3 minutes, and the results are comparable to the moving-head light method. While the handheld method is not practical for photographing a large collection of objects, it can be an invaluable tool for well-lit photography in the field.

7. Discussion and Future Work

The ability to have large area light sources is crucial for photographing highly specular objects. Light source size also affects the sharpness of shadows and highlights. Our system has a unique advantage in that larger area light sources can be simulated by combining pictures illuminated with overlapping light sources. We could extend our optimization to penalize each distinct light source cluster, thus preventing disjoint highlights. The softness of the light can also be controlled by varying the beam width between a point-source and a large area source as it quickly sweeps over the hemisphere of lighting directions. More advanced moving-head spotlights usually provide controllable spot sizes suitable for this purpose.

Even though our system is aimed primarily at non-professional photographers, a few simple additions can make it a flexible tool for a creative expert to experiment with different lighting designs more easily. For example, the user might specify a simple weighting mask to set the importance of different image regions and influence the optimization process. While weighting masks would make the system more flexible, they would complicate the target sketching process. We do not know yet if the results would warrant the increase in complexity. Also, tools to directly tweak the light position and size on a virtual hemisphere around the object might also aid expert users.

There are several possible ways of dealing with the ambient light in the reflective enclosure. Underexposing all images using exposure compensation on the camera, using a larger enclosure or one made of materials with special reflective properties would greatly minimize the ambient component. Finally, it might also be possible to explicitly subtract the ambient term from the basis images.

© The Eurographics Association 2005.
This paper takes the problem of good lighting for desktop photography and finds a simple and practical solution using image-based relighting techniques. More sophisticated image-based measurements might also be achievable while maintaining the simplicity and elegance of the system. For example, we could estimate the incoming light direction by calibrating the ad-hoc enclosure setup with a light-probe, or by using dimensionality reduction [WMTG05] for the handheld case. Combined with surface normals, such calibration might suffice for image-based estimates of BRDF.

8. Acknowledgements

We would like to thank the computer graphics group at Northwestern University for their support and feedback and the anonymous reviewers for their helpful comments and suggestions. Thanks to Jingjing Meng and Xing Hu for help in early stages of the project, Kelli Johnson and Nathan Matsuda for help with the construction of the enclosure, Vincent Masselus and Amy Gooch for help with Figure 2, and Holger Winnemöller for help with the video. This research was funded in part by NSF grants 0238062 and 0237621.

References


(a) Sequence showing successive sketching/optimization iterations to get the desired lighting. The first result uses a constant grayscale target, while the others use previous results as starting points for the target image.

(b) Strategic placement of highlights in the target result in an interesting side-lit image.

(c) Positioning of highlights reveals underlying texture in the surface.

(d) Lighting a highly specular object by forcing the background to be dark.

(e) Target results in image suggesting illumination from the right.

(f) Data captured by the handheld method. Image on the left uses a smooth grayscale gradient as the target image.

Figure 5: Sample target images and lit photographs.

© The Eurographics Association 2005.
Assorted Pixels:
Multi-Sampled Imaging with Structural Models

Shree K. Nayar and Srinivasa G. Narasimhan

Department of Computer Science, Columbia University
New York, NY 10027, U.S.A.
{nayar, srinivas}@cs.columbia.edu

Abstract. Multi-sampled imaging is a general framework for using pixels on an image detector to simultaneously sample multiple dimensions of imaging (space, time, spectrum, brightness, polarization, etc.). The mosaic of red, green and blue spectral filters found in most solid-state color cameras is one example of multi-sampled imaging. We briefly describe how multi-sampling can be used to explore other dimensions of imaging. Once such an image is captured, smooth reconstructions along the individual dimensions can be obtained using standard interpolation algorithms. Typically, this results in a substantial reduction of resolution (and hence image quality). One can extract significantly greater resolution in each dimension by noting that the light fields associated with real scenes have enormous redundancies within them, causing different dimensions to be highly correlated. Hence, multi-sampled images can be better interpolated using local structural models that are learned offline from a diverse set of training images. The specific type of structural models we use are based on polynomial functions of measured image intensities. They are very effective as well as computationally efficient. We demonstrate the benefits of structural interpolation using three specific applications. These are (a) traditional color imaging with a mosaic of color filters, (b) high dynamic range monochrome imaging using a mosaic of exposure filters, and (c) high dynamic range color imaging using a mosaic of overlapping color and exposure filters.

1 Multi-Sampled Imaging

Currently, vision algorithms rely on images with 8 bits of brightness or color at each pixel. Images of such quality are simply inadequate for many real-world applications. Significant advances in imaging can be made by exploring the fundamental trade-offs that exist between various dimensions of imaging (see Figure 1). The relative importances of these dimensions clearly depend on the application at hand. In any practical scenario, however, we are given a finite number of pixels (residing on one or more detectors) to sample the imaging dimensions. Therefore, it is beneficial to view imaging as the judicious assignment of resources (pixels) to the dimensions of imaging that are relevant to the application.

Different pixel assignments can be viewed as different types of samplings of the imaging dimensions. In all cases, however, more than one dimension is simultaneously sampled. In the simplest case of a gray-scale image, image brightness and image space are sampled, simultaneously. More interesting examples result from using image detectors made of an assortment of pixels, as shown in Figure
Figure 1. A few dimensions of imaging. Pixels on an image detector may be assigned to multiple dimensions in a variety of ways depending on the needs of the application.

2. Figure 2(a) shows the popular Bayer mosaic [Bay76] of red, green and blue spectral filters placed adjacent to pixels on a detector. Since multiple color measurements cannot be captured simultaneously at a pixel, the pixels are assigned to specific colors to trade-off spatial resolution for spectral resolution. Over the last three decades various color mosaics have been suggested, each one resulting in a different trade-off (see [Dil77], [Dil78], [MOS83], [Par85], [KM85]).

Historically, multi-sampled imaging has only been used in the form of color mosaics. Only recently has the approach been used to explore other imaging dimensions. Figure 2(b) shows the mosaic of neutral density filters with different transmittances used in [NM00] to enhance an image detector’s dynamic range.

In this case, spatial resolution is traded-off for brightness resolution (dynamic range). In [SN01], spatially varying transmittance and spectral filters were used with regular wide FOV mosaicing to yield high dynamic range and multi-spectral mosaics. Figure 2(c) shows how space, dynamic range and color can be sampled simultaneously by using a mosaic of filters with different spectral responses and transmittances. This type of multi-sampling is novel and, as we shall show, results in high dynamic range color images. Another example of assorted pixels was proposed in [BE00], where a mosaic of polarization filters with different orientations is used to estimate the polarization parameters of light reflected by scene points. This idea can be used in conjunction with a spectral mosaic, as shown in Figure 2(d), to achieve simultaneous capture of polarization and color.

Multi-sampled imaging can be exploited in many other ways. Figures 2(e) shows how temporal sampling can be used with exposure sampling. This example is related to the idea of sequential exposure change proposed in [MP95], [DM97] and [MN99] to enhance dynamic range. However, it is different in that the exposure is varied as a periodic function of time, enabling the generation of high dynamic range, high framerate video. The closest implementation appears to be the one described in [GHZ92] where the electronic gain of the camera is varied periodically to achieve the same effect. A more sophisticated implementation may sample space, time, exposure and spectrum, simultaneously, as shown in Figure 2(f).

The above examples illustrate that multi-sampling provides a general framework for designing imaging systems that extract information that is most per-
Fig. 2. A few examples of multi-sampled imaging using assorted pixels. (a) A color mosaic. Such mosaics are widely used in solid-state color cameras. (b) An exposure mosaic. (c) A mosaic that includes different colors and exposures. (d) A mosaic using color and polarization filters. (e), (f) Multi-sampling can also involve varying exposure and/or color over space and time.
tinent to the application. Though our focus is on the visible light spectrum, multi-sampling is, in principle, applicable to any form of electromagnetic radiation. Therefore, the pixel assortments and reconstruction methods we describe in this paper are also relevant to other imaging modalities such as X-ray, magnetic resonance (MR) and infra-red (IR). Furthermore, the examples we discuss are two-dimensional but the methods we propose are directly applicable to higher-dimensional imaging problems such as ones found in tomography and microscopy.

2 Learned Structural Models for Reconstruction
How do we reconstruct the desired image from a captured multi-sampled one? Nyquist’s theory [Bra65] tells us that for a continuous signal to be perfectly reconstructed from its discrete samples, the sampling frequency must be at least twice the largest frequency in the signal. In the case of an image of a scene, the optical image is sampled at a frequency determined by the size of the detector and the number of pixels on it. In general, there is no guarantee that this sampling frequency satisfies Nyquist’s criterion. Therefore, when a traditional interpolation technique is used to enhance spatial resolution, it is bound to introduce errors in the form of blurring and/or aliasing. In the case of multi-sampled images (see Figure 2), the assignment of pixels to multiple dimensions causes further undersampling of scene radiance along at least some dimensions. As a result, conventional interpolation methods are even less effective.

Our objective is to overcome the limits imposed by Nyquist’s theory by using prior models that capture redundancies inherent in images. The physical structures of real-world objects, their reflectances and illuminations impose strong constraints on the light fields of scenes. This causes different imaging dimensions to be highly correlated with each other. Therefore, a local mapping function can be learned from a set of multi-sampled images and their corresponding correct (high quality) images. As we shall see, it is often beneficial to use multiple mapping functions. Then, given a novel multi-sampled image, these mapping functions can be used to reconstruct an image that has enhanced resolution in each of the dimensions of interest. We refer to these learned mapping functions as local structural models.

The general idea of learning interpolation functions is not new. In [FP99], a probabilistic Markov network is trained to learn the relationship between sharp and blurred images, and then used to increase spatial resolution of an image. In [BK00], a linear system of equations is solved to estimate a high resolution image from a sequence of low resolution images wherein the object of interest is in motion. Note that both these algorithms are developed to improve spatial resolution, while our interest is in resolution enhancement along multiple imaging dimensions.

Learning based algorithms have also been applied to the problem of interpolating images captured using color mosaics. The most relevant among these is the work of Wober and Soini [WS95] that estimates an interpolation kernel from training data (high quality color images of test patterns and their corresponding color mosaic images). The same problem was addressed in [Bra94] using a Bayesian method.
We are interested in a general method that can interpolate not just color mosaic images but any type of multi-sampled data. For this, we propose the use of a structural model where each reconstructed value is a polynomial function of the image brightnesses measured within a local neighborhood. The size of the neighborhood and the degree of the polynomial vary with the type of multi-sampled data being processed. It turns out that the model of Wober and Soini [WS95] is a special instance of our model as it is a first-order polynomial applied to the specific case of color mosaic images. As we shall see, our polynomial model produces excellent results for a variety of multi-sampled images. Since it uses polynomials, our method is very efficient and can be easily implemented in hardware. In short, it is simple enough to be incorporated into any imaging device (digital still or video camera, for instance).

3 Training Using High Quality Images

Since we wish to learn our model parameters, we need a set of high quality training images. These could be real images of scenes, synthetic images generated by rendering, or some combination of the two. Real images can be acquired using professional grade cameras whose performance we wish to emulate using lower quality multi-sampling systems. Since we want our model to be general, the set of training images must adequately represent a wide range of scene features. For instance, images of urban settings, landscapes and indoor spaces may be included. Rotated and magnified versions of the images can be used to capture the effects of scale and orientation. In addition, the images may span the gamut of illumination conditions encountered in practice, varying from indoor lighting to overcast and sunny conditions outdoor. Synthetic images are useful as one can easily include in them specific features that are relevant to the application.

Some of the 50 high quality images we have used in our experiments are shown in Figure 3. Each of these is a 2000 x 2000 color (red, green, blue) image with 12 bits of information in each color channel. These images were captured using a film camera and scanned using a 12-bit slide scanner. Though the total...
number of training images is small they include a sufficiently large number of
local (say, $7 \times 7$ pixels) appearances for training our structural models.

Given such high quality images, it is easy to generate a corresponding set
of low-quality multi-sampled images. For instance, given a $2000 \times 2000$ RGB
image with 12-bits per pixel, per color channel, simple downsampling in space,
color, and brightness results in a $1000 \times 1000$, 8 bits per pixel multi-sampled
image with the sampling pattern shown in Figure 2(c). We refer to this process
of generating multi-sampled images from high quality images as *downgrading*.

With the high quality images and their corresponding (downgraded) multi-
sampled images in place, we can learn the parameters of our structural model. A
structural model is a function $f$ that relates measured data $M(x, y)$ in a multi-
sampled image to a desired value $H(i, j)$ in the high quality training image:

$$H(i, j) = f(M(1, 1), ..., M(x, y), ..., M(X, Y))$$

(1)

where, $X$ and $Y$ define some neighborhood of measured data around, or close
to, the high quality value $H(i, j)$. Since our structural model is a polynomial, it
is linear in its coefficients. Therefore, the coefficients can be efficiently computed
from training data using linear regression.

Note that a single structural model may be inadequate. If we set aside the
measured data and focus on the type of multi-sampling used (see Figure 2), we
see that pixels can have different types of neighborhood sampling patterns. If
we want our models to be compact (small number of coefficients) and effective
we cannot expect them to capture variations in scenes as well as changes in the
sampling pattern. Hence, we use a single structural model for each type of local
sampling pattern. Since our imaging dimensions are sampled in a uniform man-
ner, in all cases we have a small number of local sampling patterns. Therefore,
only a small number of structural models are needed. During reconstruction,
given a pixel of interest, the appropriate structural model is invoked based on
the pixel’s known neighborhood sampling pattern.

4 Spatially Varying Color (SVC)

Most color cameras have a single image detector with a mosaic of red, green and
blue spectral filters on it. The resulting image is hence a widely used type of
multi-sampled image. We refer to it as a spatially varying color (SVC) image.
When one uses an NTSC color camera, the output of the camera is nothing but
an interpolated SVC image. Color cameras are notorious for producing inade-
quate spatial resolution and this is exactly the problem we seek to overcome
using structural models. Since this is our first example, we will use it to describe
some of the general aspects of our approach.

4.1 Bayer Color Mosaic

Several types of color mosaics have been implemented in the past [Bay76], [Dil77],
[Dil78], [MOS83], [Par85], [KM85]. However, the most popular of these is the
Bayer pattern [Bay76] shown in Figure 4. Since the human eye is more sensitive
to the green channel, the Bayer pattern uses more green filters than it does red
and blue ones. Specifically, the spatial resolutions of green, red and blue are 50%, 25% and 25%, respectively. Note that the entire mosaic is obtained by repeating the $2 \times 2$ pattern shown on the right in Figure 4. Therefore, given a neighborhood size, all neighborhoods in a Bayer mosaic must have one of four possible sampling patterns. If the neighborhood is of size $3 \times 3$, the resulting patterns are $p_1, p_2, p_3$ and $p_4$ shown in Figure 4.

![Fig. 4. Spatially varying color (SVC) pattern on a Bayer mosaic. Given a neighborhood size, all possible sampling patterns in the mosaic must be one of four types. In the case of a $3 \times 3$ neighborhood, these patterns are $p_1, p_2, p_3$ and $p_4$.](image)

### 4.2 SVC Structural Model

From the measured SVC image, we wish to compute three color values (red, green and blue) at each pixel, even though each pixel in the SVC image provides a single color measurement. Let the measured SVC image be denoted by $M$ and the desired high quality color image by $H$. A structural model relates each color value in $H$ to the measured data within a small neighborhood in $M$. This neighborhood includes measurements of different colors and hence the model implicitly accounts for correlations between different color channels.

As shown in Figure 5, let $M_p$ be the measured data in a neighborhood with sampling pattern $p$, and $H_p(i + 0.5, j + 0.5, \lambda)$ be the high quality color value at the center of the neighborhood. (The center is off-grid because the neighborhood is an even number of pixels in width and height.) Then, a polynomial structural model can be written as:

$$H_p(i + 0.5, j + 0.5, \lambda) = \sum_{(x, y) \in S_p(i, j)} \sum_{(k \neq x, l \neq y) \in S_p(i, j)} \sum_{n=0}^{N_p} \sum_{q=0}^{N_p-n} C_p(a, b, c, d, \lambda, n) M_p^n(x, y) M_p^q(k, l).$$  \hspace{1cm} (2)

$S_p(i, j)$ is the neighborhood of pixel $(i, j)$, $N_p$ is the order of the polynomial and $C_p$ are the polynomial coefficients for the pattern $p$. The coefficient indices $(a, b, c, d)$ are equal to $(x - i, y - j, k - i, l - j)$.

The product $M_p(x, y) M_p(k, l)$ explicitly represents the correlation between different pixels in the neighborhood. For efficiency, we have not used these cross-terms in our implementations. We found that very good results are obtained
Fig. 5. The measured data $M_p$ in the neighborhood $S_p(i,j)$ around pixel $(i,j)$ are related to the high quality color values $H_p(i + 0.5, j + 0.5, \lambda)$ via a polynomial with coefficients $C_p$.

$$H_p(i + 0.5, j + 0.5, \lambda) = \sum_{(x,y) \in S_p(i,j)} \sum_{n=0}^{N_p} C_p(a, b, \lambda, n)M_p^n(x, y). \tag{3}$$

The mapping function (3), for each color $\lambda$ and each local pattern type $p$, can be conveniently rewritten using matrices and vectors, as shown in Figure 6:

$$H_p(\lambda) = A_p C_p(\lambda). \tag{4}$$

For a given pattern type $p$ and color $\lambda$, $A_p$ is the measurement matrix. The rows of $A_p$ correspond to the different neighborhoods in the image that have the pattern $p$. Each row includes all the relevant powers (up to $N_p$) of the measured data $M_p$ within the neighborhood. The vector $C_p(\lambda)$ includes the coefficients of the polynomial mapping function and the vector $H_p(\lambda)$ includes the desired high quality values at the off-grid neighborhood centers. The estimation of the model parameters $C_p$ can then be posed as a least squares problem:

$$C_p(\lambda) = (A_p^T A_p)^{-1}A_p^T H_p(\lambda), \tag{5}$$

When the signal-to-noise characteristics of the image detector are known, (5) can be rewritten using weighted least squares to achieve greater accuracy [Aut01].
4.3 Total Number of Coefficients

The number of coefficients in the model (3) can be calculated as follows. Let the neighborhood size be \( u \times v \), and the polynomial order corresponding to each pattern \( p \) be \( N_p \). Let the number of distinct local patterns in the SVC image be \( P \) and the number of color channels be \( \Lambda \). Then, the total number of coefficients needed for structural interpolation is:

\[
|C| = \left( P + u \times v \times \sum_{p=1}^{P} N_p \right) \times \Lambda.
\]

(6)

For the Bayer mosaic, \( P = 4 \) and \( \Lambda = 3 \) \((R,G,B)\). If we use \( N_p = 2 \) and \( u = v = 6 \), the total number of coefficients is 876. Since these coefficients are learned from real data, they yield greater precision during interpolation than standard interpolation kernels. In addition, they are very efficient to apply. Since there are \( P = 4 \) distinct patterns, only 219 (a quarter) of the coefficients are used for computing the three color values at a pixel. Note that the polynomial model is linear in the coefficients. Hence, structural interpolation can be implemented in real-time using a set of linear filters that act on the captured image and its powers (up to \( N_p \)).

4.4 Experiments

A total of 30 high quality training images (see Figure 3) were used to compute the structural model for SVC image interpolation. Each image is downgraded to obtain a corresponding Bayer-type SVC image. For each of the four sampling patterns in the Bayer mosaic, and for each of the three colors, the appropriate image neighborhoods were used to compute the measurement matrix \( A_p \) and the reconstruction vector \( H_p(\lambda) \). While computing these, one additional step was taken; each measurement is normalized by the energy within its neighborhood to make the structural model insensitive to changes in illumination intensity and camera gain. The resulting \( A_p \) and \( H_p(\lambda) \) are used to find the coefficient vector \( C_p(\lambda) \) using linear regression (see (5)). In our implementation, we used the parameter values \( P = 4 \) (Bayer), \( N_p = 2 \), \( u = v = 6 \) and \( \Lambda = 3 \) \((R,G,B)\), to get a total of 876 coefficients.

The above structural model was used to interpolate 20 test SVC images that are different from the ones used for training. In Figure 7(a), a high quality (8-bits per color channel) image is shown. Figure 7(b) shows the corresponding (downgraded) SVC image. This is really a single channel 8-bit image and its pixels are shown in color only to illustrate the Bayer pattern. Figure 7(c) shows a color image computed from the SVC image using bi-cubic interpolation. As is usually done, the three channels are interpolated separately using their respective data in the SVC image. The magnified image region clearly shows that bi-cubic interpolation results in a loss of high frequencies; the edges of the tree branches and the squirrels are severely blurred. Figure 7(d) shows the result of applying structural interpolation. Note that the result is of high quality with minimal loss of details.
Fig. 7. (a) Original (high quality) color image with 8-bits per color channel. (b) SVC image obtained by downgrading the original image. The pixels in this image are shown in color only to illustrate the Bayer mosaic. Color image computed from the SVC image using (c) bi-cubic interpolation and (d) structural interpolation. (e) Histograms of luminance error (averaged over 20 test images). The RMS error is 6.12 gray levels for bi-cubic interpolation and 3.27 gray levels for structural interpolation.
We have quantitatively verified our results. Figure 7(e) shows histograms of the luminance error for bi-cubic and structural interpolation. These histograms are computed using all 20 test images (not just the one in Figure 7). The difference between the two histograms may appear to be small but is significant because a large fraction of the pixels in the 20 images belong to "flat" image regions that are easy to interpolate for both methods. The RMS errors (computed over all 20 images) are 6.12 and 3.27 gray levels for bi-cubic and structural interpolation, respectively.

5 Spatially Varying Exposures (SVE)

In [NM00], it was shown that the dynamic range of a gray-scale image detector can be significantly enhanced by assigning different exposures (neutral density filters) to pixels, as shown in Figure 8. This is yet another example of a multi-sampled image and is referred to as a spatially varying exposure (SVE) image. In [NM00], standard bi-cubic interpolation was used to reconstruct a high dynamic range gray-scale image from the captured SVE image; first, saturated and dark pixels are eliminated, then all remaining measurements are normalized by their exposure values, and finally bi-cubic interpolation is used to find the brightness values at the saturated and dark pixels. As expected, the resulting image has enhanced dynamic range but lower spatial resolution. In this section, we apply structural interpolation to SVE images and show how it outperforms bi-cubic interpolation.

![Fig. 8. The dynamic range of an image detector can be improved by assigning different exposures to pixels. In this special case of 4 exposures, any 6 × 6 neighborhood in the image must belong to one of four possible sampling patterns shown as p1...p4.](image)

5.1 SVE Structural Model

As in the SVC case, let the measured SVE data be \( M \) and the corresponding high dynamic range data be \( H \). If the SVE detector uses only four discrete exposures (see Figure 8), it is easy to see that a neighborhood of any given size can have only one of four different sampling patterns \( P = 4 \). Therefore, for each sampling pattern \( p \), a polynomial structural model is used that relates the captured data \( M_p \) within the neighborhood to the high dynamic range value \( H_p \) at the center of the neighborhood:

\[
H_p(i + 0.5, j + 0.5) = \sum_{(x,y)\in S_p(i,j)} \sum_{n=0}^{N_p} C_p(a, b, n) \ M_p^n(x, y),
\]
where, as before, \((a, b) = (x - i, y - j), S_p(i, j)\) is the neighborhood of pixel \((i, j)\), \(N_p\) is the order of the polynomial mapping, and \(C_p\) are the polynomial coefficients for the pattern \(p\). Note that there is only one channel in this case (gray-scale) and hence the parameter \(\lambda\) is omitted. The above model is rewritten in terms of a measurement matrix \(A_p\) and a reconstruction vector \(H_p\), and the coefficients \(C_p\) are found using (5). The number of coefficients in the SVE structural model is determined as:

\[
|C| = P + u * v * \sum_{p=1}^{P} N_p .
\]  

(8)

In our implementation, we have used \(P = 4, N_p = 2\) and \(u = v = 6\), which given a total of \(292\) coefficients. Since \(P = 4\), only \(73\) coefficients are needed for reconstructing each pixel in the image.

5.2 Experiments

The SVE structural model was trained using 12-bit gray-scale versions of 6 of the images shown in Figure 3 and their corresponding 8-bit SVE images. Each SVE image was obtained by applying the exposure pattern shown in Figure 8 (with \(e_4 = 4e_3 = 16e_2 = 64e_1\)) to the original image, followed by a downgrade from 12 bits to 8 bits. The structural model was tested using 6 test images, one of which is shown in Figure 9. Figure 9(a) shows the original 12-bit image, Figure 9(b) shows the downgraded 8-bit SVE image, Figure 9(c) shows a 12-bit image obtained by bi-cubic interpolation of the SVE image, and Figure 9(d) shows the 12-bit image obtained by structural interpolation. The magnified images shown on the right are histogram equalized to bring out the details (in the clouds and walls) that are lost during bi-cubic interpolation but extracted by structural interpolation. Figure 9(e) compares the error histograms (computed using all 6 test images) for the two cases. The RMS errors were found to be 33.4 and 25.5 gray levels (in a 12-bit range) for bi-cubic and structural interpolations, respectively. Note that even though a very small number (6) of images were used for training, our method outperforms bi-cubic interpolation.

6 Spatially Varying Exposure and Color (SVEC)

Since we are able to extract high spatial and spectral resolution from SVC images and high spatial and brightness resolution from SVE images, it is natural to explore how these two types of multi-sampling can be combined into one. The result is the simultaneous sampling of space, color and exposure (see Figure 10). We refer to an image obtained in this manner as a spatially varying exposure and color (SVEC) image. If the SVEC image has 8-bits at each pixel, we would like to compute at each pixel three color values, each with 12 bits of precision. Since the same number of pixels on a detector are now being used to sample three different dimensions, it should be obvious that this is a truly challenging interpolation problem. We will see that structural interpolation does remarkably well at extracting the desired information.
Fig. 9. (a) Original 12-bit gray scale image. (b) 8-bit SVE image. (c) 12-bit (high dynamic range) image computed using bi-cubic interpolation. (d) 12-bit (high dynamic range) image computed using structural models. (d) Error histograms for the two case (averaged over 6 test images). The RMS error for the 6 images are 33.4 and 25.5 gray levels (on a 12-bit scale) for bi-cubic and structural interpolation, respectively. The magnified image regions on the right are histogram equalized.
Fig. 10. A combined mosaic of 3 spectral and 4 neutral density filters used to simultaneously sample space, color and exposures using a single image detector. The captured 8-bit SVEC image can be used to compute a 12-bit (per channel) color image by structural interpolation. For this mosaic, for any neighborhood size, there are only 16 distinct sampling patterns. For a $4 \times 4$ neighborhood size, the patterns are $p_1 \ldots p_{16}$.

Color and exposure filters can be used to construct an SVEC sensor in many ways. All possible mosaics that can be constructed from $A$ colors and $E$ exposures are derived in [Aut01]. Here, we will focus on the mosaic shown in Figure 10, where 3 colors and 4 exposures are used. For any given neighborhood size, it is shown in [Aut01] that only 16 different sampling patterns exist (see Figure 10).

6.1 SVEC Structural Model

The polynomial structural model used in the SVEC case is the same as the one used for SVC, and is given by (3). The only caveat is that in the SVEC case we need to ensure that the neighborhood size used is large enough to adequately sample all the colors and exposures. That is, the neighborhood size is chosen such that it includes all colors and all exposures of each color.

The total number of polynomial coefficients needed is computed the same way as in the SVC case and is given by (6). In our experiments, we have used the mosaic shown in Figure 10. Therefore, $P = 16$, $A = 3$ ($R$, $G$, and $B$), $N_p = 2$ for each of the 16 patterns, and $u = v = 6$, to give a total of 3504 coefficients. Once again, at each pixel, for each color, only $3504/48 = 73$ coefficients are used. Therefore, even for this complex type of multi-sampling, our structural models can be applied to images in real-time using a set of linear filters.

6.2 Experiments

The SVEC structural model was trained using 6 of the images in Figure 3. In this case, the 12-bit color images in the training set were downgraded to 8-bit SVEC images. The original and SVEC images were used to compute the 3504 coefficients of the model. The model was then used to map 10 different test SVEC images to 12-bit color images. One of these results is shown in Figure 11. The original 12-bit image shown in Figure 11(a) was downgraded to obtain the 8-bit SVEC image shown in Figure 11(b). This image has a single channel
Fig. 11. (a) Original 12-bit color image. (b) 8-bit SVEC Image. 12-bit color images reconstructed using (c) bi-cubic interpolation and (d) structural interpolation. (e) Luminance error histogram computed using 10 test images. RMS luminance errors were found to be 118 and 80 (on a 12-bit scale) for bi-cubic and structural interpolation, respectively. The magnified images on the right are histogram equalized.
and is shown here in color only to illustrate the effect of simultaneous color and exposure sampling. Figures 11(c) and (d) show the results of applying bi-cubic and structural interpolation, respectively. It is evident from the magnified images on the right that structural interpolation yields greater spectral and spatial resolution. The two interpolation techniques are compared in Figure 11(e) which shows error histograms computed using all 10 test images. The RMS luminance errors were found to be 118 gray-levels and 80 gray-levels (on a 12 bit scale) for bi-cubic and structural interpolations, respectively.

Acknowledgements: This research was supported in parts by a ONR/ARPA MURI Grant (N00014-95-1-0601) and an NSF ITR Award (IIS-00-85864).

References


Camera for Conical Peripheral and Panoramic Photography

Andrew Davidhazy
Imaging and Photographic Technology Department
School of Photographic Arts and Sciences
Rochester Institute of Technology
One Lomb Memorial Drive, Rochester, New York 14623

ABSTRACT

This paper is a report on the development, operating principles and applications of a camera capable of making distortion free peripheral reproductions of conical objects. In addition, the parameters governing its use for the generation of conical panoramic photographs are also described. Examples of both types of applications are included. Suggestions for other applications are given. A brief review of the operating principles of standard peripheral, panoramic and linear type strip cameras is included in the introduction to the paper.

1. INTRODUCTION

The principles governing the operation of panoramic cameras which are capable of making extremely wide angle photographs by rotating roughly about the rear nodal point of their lenses while moving film behind an open slit located just in front of the film surface are well known. Basic relationships are presented below in general terms to establish a common level of understanding so that the rest of the material can be presented in the context of where it fits in with established camera systems.

Among the earliest panoramic cameras of the type described above is the Cirkut camera, introduced by Eastman Kodak in the late 1800's and which was available in a wide variety of film sizes. Modern cameras of similar design include the Hulcherama, the Alpa Rotocamera, the Panalux or Roundshot and the Globuscope.

It is fundamental to fully understand the method by which these cameras operate to realize that they record subject DISTANCE, or image position, along the slit and TIME at right angles to the slit. Since in all of the above applications the velocity at which the film moves is the same at the top and bottom of the slit, the elapsed time for any given length of film which passes though the camera is the same at the top and bottom margins of the film.

Panoramic records made with "cirkut" type cameras are made by rotating the camera about a vertical axis thus scanning the scene in front of the camera in sequential fashion. The film accumulates the changing image information presented to it by the scanning camera and eventually builds up a record on a length of film containing a view of any desired angle even up to (or beyond) 360 degrees. For the record to look sharp the film must move at the same velocity as the image. This is accomplished by moving a length of film equal to 2pi times the lens focal length past the open slit during the time for one revolution of the camera. In actuality it is the rear nodal point to film distance that matters but since in most applications the object distance is large, the lens focal length is usually a close enough approximation to this measure.

Related to the panoramic camera, the peripheral camera is not as well recognized in either the literature or in terms of working models. The cause for this lack of recognition for peripheral
photography is probably due to the fact that this application is somewhat more specialized. It is interesting to note, however, that the making of peripheral photographs, sometimes called cyclographs, of Greek and Mayan pottery, has been practiced since the late 1800's in a number of major museums and also in industrial situations.

Peripheral records of the surface of cylindrical subjects can be easily made by rotating the subject in front of the camera. The film in the camera is made to move at right angles to the axis of rotation of the subject and in the same direction as it's image. The slit restricts the angle of view along the axis parallel to film motion to a very narrow angle, thus it encompasses only a small portion of the cylindrical subject's surface. At the slit, then, the image of the subject's surface appears to pass by in linear fashion, as in racetrack photography. Again, the film velocity is set to match the value given by the subject's surface velocity multiplied times magnification.

Note then, that peripheral records of subjects which vary in circumference can only be reproduced properly at those points where their image velocity happens to match the film velocity.

Since the film can only match one particular image velocity, it follows that image velocities which are slower than the film velocity will produce records which are stretched out and those which move faster will produce compressed reproductions. That is, since the lengths of the images recorded behind any given point along the slit are all the same, this results in invariable distortion of all areas which did not move at the same velocity as the film.

Finally, the other two variations on the above themes, that of racetrack photofinish and synchroballistic cameras and that of aerial strip cameras, apparently complete the applications circle of those cameras which make more or less realistic records by moving a length of film past a slit and capturing the image of the subject by making it's image travel across this slit at the same speed as the film. In effect, in these cameras the moving image scans itself onto the passing film by virtue of its motion.

In these "linear" type cameras, the film is simply made to move at the expected velocity of the image. In racetracks the camera is fixed and the image of the racers passes over the slit in the camera. The film velocity is adjusted to approximately match the expected velocity of the images of the subjects. In aerial cameras, the plane moving with respect to the ground below causes a moving image to pass by the slit of the camera. The film velocity is again adjusted to match that of the passing image. Film velocity can be adjusted as the plane changes velocity by making a visual comparison between the motion of the ground's image with the motion of a chain which is matched mechanically to the velocity of the film. Alternately, it is possible to make the comparison electronically. When neither is practical, then camera operators must take into account the operating magnification of the camera and adjust the film velocity to the apparent subject velocity multiplied by the camera magnification.
Generically all of these cameras or systems can be labeled as variants of "strip" recording cameras and can be called simply "strip cameras". Examples of images made by each of these approaches are shown in Figure 1. At this point it is assumed that the operating principle of these cameras has been discussed in sufficient detail and the development of the present camera will be presented next.

Figure 1. a) 360 degree panoramic photograph with characteristic distortion; b) partial peripheral photograph of automobile tire surface; c) typical photofinish photograph where print is a line as a function of time; d) moving strip camera sees subjects head-on and thus building sides are not visible.

2. PERIPHERAL PHOTOGRAPHY

2.1 Development of the camera

As stated above, with standard strip cameras peripheral reproductions of subjects of irregular diameter can not be produced without distortion because the film can only match one particular image velocity at a time. For distortion free records of these subjects the film must be able to move at various velocities simultaneously, moving different amounts of film in equal time increments.

A special case of a subject of varying diameters is a one whose diameter changes in straight linear fashion. The development of the present camera is a direct result of a request for a photographic peripheral record of a subject which had just such a shape. In brief, the subject had a conical, rather than a cylindrical, shape.

To recap then, while the surface velocity of a rotating cylinder is uniform and thus its image velocity can be matched everywhere by a strip of film moving behind the slit of the camera, the surface velocity of a rotating cone varies, leading to various image velocities at the slit which can not be matched by the film in a standard strip camera except at one point along the slit. In order to cope with such a conical subject what is needed is either some optical technique which can nullify differences in subject surface velocity or the development of a peripheral camera capable of moving the film at a different velocity at one end of the slit as contrasted to that at the other end. While a variety of optical compensation methods were tried in order to equalize the image velocity from subject areas moving at unequal velocities, none succeeded in establishing a 1:1 image to film velocity relationship when the subject was conical in nature.
which were attempted dealt with rotating the subject about a tilted axis and with using the tilts of a moveable front and rear standard to attempt to introduce differential magnification to compensate for the change in subject circumference.

Eventually, it was a modification of the camera itself that resulted in the development of a peripheral camera which could deal effectively with conical subjects. It neatly solved the problem of differential image velocity along the slit by moving the film in a circular, rather than linear, fashion, as is the practice in all other "strip" cameras.

The basis for the design was the realization that on a turntable the surface velocity is a function of the distance from the center of rotation. When a slit is extended from the center of a turntable, a piece of film attached to the turntable moves past this slit at increasing velocities with increasing distance from the turntable's center. At a later date I made another connection with an existing imaging system when I noticed that the manner in which a conventional strip peripheral camera delivers undistorted records of cylindrical objects is similar to the operation of a printing press. Here, a cylinder with information on it's surface, transfers onto the support a series of perfect rectangles. That is, the length of the transferred images per revolution of the impression cylinder is the same at one end as at the other since the circumference of the original cylinder is also uniform from one end to the other.

When one tries the same procedure with a conical subject by rolling it along a surface, it becomes obvious that the surface which is generated is a circular one. If the cone has an apex, then the apex will be located at the center of the circle. The number of degrees out of a complete circle which the cone describes during one revolution is a function of the steepness of the angle of it's side. The steeper the angle, or more pointy the cone, the smaller part of a complete circle which will be produced by making the cone complete one revolution. If it is so steep that it is, in fact, a cylinder, then the length of the image along the circle will be reduced to the width of a single line. In the case that the cone is so flat that it is actually a flat circle, then the transferred record will also be a full 360 degree circle. These conditions are obviously extremes but it helps to think of these extremes to relate them to standard strip cameras and to appreciate how these systems work.

Anyway, once I realized the operating principle of a camera which could transport the film past a slit at various velocities, it took me very little time to build a prototype model. It is shown in Figure 2. The camera was designed to accept 4" diameter discs of film cut from regular sheet film. These discs were held on the surface of the rotating circular film holder by means of two-sided adhesive tape. The rotation rate of the film holder could be varied by changing the voltage to the DC gearhead motor to which it was directly attached.

There was provision built into the camera lens mount to allow it to move along the length of the slit so that an image could be formed at any given distance less than 50mm away from the center of the film disc.
Movement of the lens along the slit, or radius of the film disc, is necessary because the lens position needs to be adjusted depending on the characteristics of the cone being photographed. Also, this capability allows the choice of the largest possible magnification for a given situation. This ensures that the sharpness and grainlessness of the records is as great as possible since the peripheral images could always be made as large as the limits of the 4" diameter film disc would allow for a given application.

Figure 2. Strip camera with adjustable lens positioning track and variable rotation rate film stage driven by DC gearhead motor attached to it through camera back.

2.2 Setting up for peripheral photography of cones

The operational features of the camera permit it to effectively deal with conical subjects of a wide variety of steepness to their sides. To set up for operation the characteristics of the subject cone are first determined. These include the measurement of the inside angle of the side at the base of the cone, the circumference of the base of the cone and the length of the side.

Knowledge of the inside side angle allows one to compute the part of the circular film in degrees which will be taken up by the image of the cone's surface, A:

\[ A = 360 \text{ degrees} \times \cos \text{inside angle} \]

The inside side angle also determines time relationship between one revolution of the subject \( R(s) \) and one revolution of the film bearing turntable, \( R(f) \).

\[ R(s) = R(f) \times \cos \text{inside angle} \]

Under normal conditions, this means that if the time for one revolution of the film disc is fixed, the time for one revolution of the subject will be a fraction of the time for one revolution of the film and a number of peripheral records can be produced on a single sheet of film.

The location of the image of the subject along the radius of the film bearing turntable, or along the shutter slit, determines the operating magnification of the camera. It is advantageous to operate at the maximum possible magnification to minimize graininess. Once the magnification is known, then the camera position is adjusted to make the image formed by the lens the right size for proper reproduction.

I have found that the following procedure is convenient for setting up the camera. First, the distance from center the base of the cone will be placed at needs to be decided. To work at the maximum magnification, the base of the image of the subject must be placed at the largest possible distance from center of the film disc. In the camera described here the maximum useful distance from the center is about 45 mm.

The operating magnification, \( M \), is determined by finding the required image base circumference, \( C(i) \), and dividing by the base circumference of the real cone, \( C(c) \). Image base circumference, \( C(i) \), is found by mutiplying the circumference of the circle of which the base of the image is a part of the outside circumference of the inside angle of the cone. Thus,
\[ M = \frac{C(i)-C(c)}{C(s)} \]

Alternately, M can also be found by dividing the circumference of the film disc, \( C(f) \), along the circle on which the base of the image will be placed, by the circumference of the circle of which the base of the subject cone is a part of, \( C(s) \).

\[ M = \frac{C(f)}{C(s)} \]

The denominator of the fraction above, \( C(s) \), is given by dividing subject base circumference by the cosine of the inside angle. The magnification, M, determines what the image height of the cone's known side must be at the film plane when the base of the image of the cone is placed on the circle of chosen radius. Given a particular lens focal length, the camera distance is then adjusted so that with the base of the cone located at the chosen distance from the center of the film, the length of the image of the side of the cone, \( I \), along the slit, with \( S \) being the length of the real cone's side, is:

\[ I = M \times S \]

For distortion free reproduction, the optical axis of the camera must be adjusted so that it is perpendicular to the surface of the cone. The film, then, is parallel to the angled side of the cone. Then, the camera should be adjusted in elevation so that the rotation axis of the film is aimed at the center of the side of the subject. Small departures from this location are tolerable. With the camera at the proper distance for the required magnification, the distance of the lens axis from the center of rotation of the film must be adjusted to place the base of the image of the subject at the desired radius. If the subject cone has an apex, then this should now fall at the center of rotation of the film. If the subject is a truncated cone, then the point in space where the virtual apex is located must be placed at the center of rotation of the film. If the image of the side of the cone which the lens forms along the slit does not cover from the base to the apex of the cone, the result will be an unexposed band around the center of the image.

The exposure time, \( ET \), is basically the length of time it takes a given point on the film to pass by the open slit. Since in this camera the slit is cut so that it is wedge shaped, the exposure time along the slit does not vary even though the film is traveling more slowly towards the center of the film disc than along the edges. Exposure time in seconds, \( ET \), is determined by the angular size of the slit, \( SW \) (deg.), and the rotation rate of the film disc, \( R \) deg./sec, as follows:

\[ ET = \frac{SW(\text{deg.})}{R \text{ deg./sec.}} \]

Once the exposure time is determined, the aperture is set so that with the film loaded in the camera and the available lighting level a proper exposure will result.

The present camera can achieve film rotation rates of about 10 degrees per second, and the slit is about 1 degree in size so that minimum exposure times are in the order of 1/10 th of a second. The film can be slowed down to about 1 degree per second by simply decreasing the voltage to the DC gearhead motor. The camera is shown in a typical application in Figure 3. A peripheral negative and enlargement generated by the camera are shown in Figure 4 and 5 respectively.
3. CONICAL PANORAMIC PHOTOGRAPHY

3.1 Development of conical panoramic application

Once the camera served its function of making the desired distortion free peripheral record of a conical subject, I investigated applications in panoramic photography.

By way of introduction to this application of the present camera it may be appropriate to state that it is a well established operational fact among panoramic photographers that panoramic cameras must rotate about a vertical axis unless one is willing to accept horizon lines that wander up and down along the panoramic image. A variation on this theme is one in which photographers have tilted their cameras down while still keeping the axis of rotation vertical so that the horizon line effectively remains level. Attachments are available, particularly for the Cirkut-type cameras, which allow the cameras to point up, or down, while still keeping the axis of rotation of the camera vertical. The idea is that if one can raise or lower the angle of view and still keeping the axis of rotation vertical, one can lower or raise the horizon line while still keeping it parallel to the film edges along the panorama.

The difficulty which photographers who have investigated the use of these camera tilting attachments have found is that their photographs are no longer sharp from top to bottom, although indeed the horizon line is parallel to the sides of the film from end to end. The reason for this lack of sharpness is that in this mode the slit and film plane in their basic strip panoramic camera no longer describes a cylindrical path but rather a conical one. The result is that while the camera views equiangular rates of change along the slit these do not encompass equidistant displacements in the subject. In fact, if the camera could be tilted so far down or up that the point about which the camera rotates were included on the film, this point would be standing still. Therefore, the conical surface which the slit of the tilted camera describes results in uneven image velocity along the slit.

Yet, as discussed earlier, the film, moving in linear fashion, can only produce a cylindrical record since the film velocity is constant along the slit of the camera. To solve this particular problem and to produce conical panoramic images, the camera which I designed and constructed to make conical peripheral photographs can itself be mounted on a rotating turntable to produce conical panoramic photographs. See Fig. 6.
3.2 Setting up for conical panoramic photography

The operating procedures for using this camera in the panoramic mode depend on the lens which will be used and the dimensions and characteristics of the cone which one wishes to make. Once the lens focal length which will be used is chosen and the desired side angle of the cone is fixed, these two parameters determine where the lens must be placed with respect to the center of the film disc and the relationship between the time for one revolution of the film disc in relationship to the time for one revolution of the camera.

The inside side angle which is chosen must be the angle that the film disc surface must maintain with respect to the horizontal as the camera rotates about a vertical axis. The camera may be pointed upward or downward. In the former case the slit must be located above the axis of rotation of the film disc, while in the second case the slit must be located below the axis. Alternately, the camera needs to be merely inverted when it is pointed downwards. The reason for the location of the slit above or below the disc rotation axis is that the center of the disc must record those areas of the subject which are nearest the apex of the cone which the rotating camera is describing. In both instances the film disc must be turned in such a direction that it matches the direction in which the image moves with respect to the slit in the camera.

The location of the lens from the center of rotation of the film disc, \( D(L) \), is a function of the lens focal length, \( F(L) \), and the side angle. More precisely, the rear nodal point of the lens must always be located directly above or below the center of rotation of the film disc. This displacement of the lens axis from the center of the disc is a function of the tangent of the inside angle and the lens focal length:

\[
D(L) = F(L) \times \tan(\text{inside angle})
\]

This relationship fixes that the lens position, when the side angle is 0 degrees, equaling the inside angle of a flat "cone", must be such that the lens axis is directly above the center of rotation of the disc. In this case the camera is pointed straight up and only half of the image circle produced by the lens falls on the slit. Conversely, when the inside angle approaches 90 degrees, the lens must be located at a great distance from the center of the film disc in order for it to be above the center of the disc. At a side angle of 90 degrees the lens will be infinitely far away. At such a distance, as far as the lens is concerned, the film will move in linear fashion rather than circular fashion. At this extreme the length of film required to cover a 360 degree panorama will be equal to 2\( \pi \) times the lens focal length. In fact, this extreme is a special case of the "conical" camera, and is exemplified by the traditional cirkut type panoramic camera!

The above factors determine whether the film disc available in a given camera is large enough to accommodate the chosen lens at a desired side angle. For any given diameter film disc, use of long focal length lenses will restrict the camera to the production of shallow cones, while short focal length lenses will allow cones of steep inside angle although at reduced image sizes. The four inch diameter disc which this camera uses can make cones of up to about 65-70 degrees side angle (average for lamp shade use) with lens focal lengths of about 20mm.

Unlike cylindrical panoramic cameras, where the lens can be raised or lowered to alter the position of the horizon line, in conical cameras, the placement of the lens at other than the one position determined from the factors named above will produce blurring along the height of the panorama or radius of the circle.
The relationship between the time for the film disc making one revolution, \( R(f) \), and the time taken by the camera to scan 360 degrees, \( R(c) \), are given by:

\[
R(c) = R(f) \times \cos \text{side angle}
\]

This means that under normal conditions, more than one 360 degree panorama can be included on one disc of film. In fact, when the inside angle of the cone is 60 degrees, exactly two 360 panoramas can be recorded because the cosine of this angle is .5.

The angular rotation rate of the film determines the exposure time for a given angular slit width. Exposure time is determined through exactly the same procedures described above when they were applied to the making of peripheral photographs. Figures 6 and 7 illustrate the general set-up of the camera for making a conical panoramic record and the resulting negative.

Figure 6. (on right) Camera mounted on motorized, tiltable, tripod head.

Figure 7. (on left) Two full 360 degree panoramas are recorded on less than 360 of film.

3.3 Applications for conical panoramic images

The images produced this way could find direct application in transfer to such items as lampshades, see Figure 8, novelty hats, and decorative purposes for any number of items which have a more or less conical shape. Presently, adaptation of standard images to the surface of objects of conical section requires digital analysis and manipulation which is not an impossible task but not readily available to photographers used to more standard recording techniques.

A further application of the present camera is that with a slight loss in sharpness it could be used to distort conventional cylindrical panoramic or other images so that they could be bent into, or adapted to fit, conical shapes. This loss in sharpness is associated with the mismatch in
image vs. film velocities resulting from the alteration of the aspect ratio of the original as the conical image is made. The blurring effects of this differential movement of the image with respect to the film can be minimized by making the exposing slit very narrow, thus limiting the exposure time during which blurring can affect the recorded image.

4. CONCLUSION

In summary, the development and operating characteristics of a camera suitable for making distortion free peripheral reproductions of the surface detail appearing on subjects of conical shape have been described. The application of the camera for making panoramic photographs and their subsequent display as conical objects has also been reported. When the operating limits of the circularly moving film camera design, as exemplified by the present camera, were investigated it was established that strip cameras in which the film moves in linear fashion are special cases of the design and operating principles of the circularly moving film approach described above.

While the number of instances where this camera would be found useful is probably quite low, significant improvement over the use of standard strip cameras for peripheral photography of objects with a tapered shape has been demonstrated. Further, when applied to panoramic work, this camera is able to produce photographs which are suitable for a variety of utilitarian purposes without additional manipulation.

Figure 8. The George Eastman House and adjacent grounds made into a 360 degree panoramic conical lampshade.

References for additional information:

Research Engineers Corp. of Great Britain used to manufacture an attachment for 4x5 view cameras which allowed the user to produce peripheral photographs onto sheet film. Along with the attachment they also sold a precision turntable and centering device. The Charles Hulcher Co. and Robot Inc. both offer 35mm image motion or strip cameras which are quite suitable for peripheral photographs when coupled with user supplied turntables. An advantage of the Hulcher camera is that it can also easily make 360 degree wide angle or panoramic photographs. The Globuscope panoramic camera can also be used for peripheral photography.
although its usefulness is limited due to the high rates of film transport it is designed to deliver. The Sugawara Co. offers the Film Streak V, an attachment for standard 35mm cameras to allow them fairly smooth film transport in the rewind mode as described above.

With the development of linear array digital scanning camera backs the possibilities of peripheral photography are being re-invented by a group of people who suddenly realize the potential for imaging subjects in the scanning, rather than instantaneous, mode. Something that panoramic and peripheral and photofinish and aerial photographers using strip cameras have known since at least the start of the 20th century!

Should you have a need for precision peripheral photography services, these are also not readily available. The author of this paper is in a position to help with a limited number of projects through the cooperation of the Imaging and Photographic Technology department at the Rochester Institute of Technology. He welcomes inquiries regarding this process from individuals desiring help with specific problems they may have as they become involved in peripheral recording techniques.

Andrew Davidhazy is a Professor in the Imaging and Photographic Technology department of the College of Graphic Arts and Photography at the Rochester Institute of Technology. His teaching centers on instruction related to the use of photography a a tool of measurement and visualization for researchers, scientists, engineers and technicians. Among the topics included in his laboratory's activities are infrared and ultraviolet photography, high speed and time lapse, panoramic and peripheral photography, low level aerial photography and close range photogrammetry, thermography, close up and long range photography, photographic documentation, etc..

PERTINENT REFERENCE MATERIAL

Nicholas Helmut has been doing rollout (peripheral) photographs of Mayan vases and has several examples of his work on the web. Take a peek at: Maya Vases and his brief history of rollout photography page.


Invention of peripheral photography may be attributed to Cyril Smith or to Arthur Murray Smith under the name of "cyclographs"...maybe.

On the other hand, another account has it that Arthur Hamilton Smith, Keeper of the Greek and Roman Antiquities at the British Museum designed a peripheral instrument with the assistance of the optical firm of Ross for the photography of ancient pottery. This was reported in the Journal of the Royal Photographic Society volume 19, May, 1895.

Andrew Oliver, Jr., Associate Curator, Greek and Roman Art, The Metropolitan Museum of Art, New York, NY 10028 may know something of application and/or history of technique.

Raymond Davis designed a camera for photographing short lengths of corroded pipe and described it for the American Bureau of Standards No.517, Vol. 20, December 1925.

Emil D'Isoz of the Budapest Museum seems to have developed a camera also.

The British Iron and Steel Research Association made the Evolute Camera around the early 1950's or so.

The Shell Development Co. reported in the S.A.E. Journal a camera for the photography of pistons in Vol. 51, No. 2, February 1943.

ADDITIONAL RELATED MATERIAL:

Makers of Peripheral Cameras:

L.F. Deardorff & Sons, Inc. 11 S. Des Plaines St., Chicago, IL 60606
Research Engineers, Ltd., Orsamn Road., London, ENGLAND N1 5RD
Charles Hulcher Inc., "G" Street, Hampton, VA
Hermann Seitz, Switzerland

articles, etc:

Kodak Job Sheet Number 8, Peripheral Photography, Kodak Pamphlet No. P-100-8 Published by the Professional, Commercial and Industrial Photography Division of the Eastman Kodak Company, 9-67


Photography Handbook Number 2, p. 102, 1938. "Flat pictures of round objects" published by Fawcett Publications, Greenwich, CT


D'Isoz, Emil, Budapest Museum
British Iron and Steel Research Association - "The Evolute Camera"


related articles below by Davidhazy, A.


"Looking at Life Through a Slit", MILWAUKEE JOURNAL, June 27, 1971.


"Principles of Peripheral Photography", Fall 1988 issue of the POLAROID PhotoEducation NEWSLETTER FOR PHOTOGRAPHIC EDUCATORS, pp 6-8.

"Forenklade system for Svep/spalt och scanningfotografering", pp 16-21 of #2/1992 issue of the Tidskrift for MEDICINSK OCH TEKNISK FOTOGRAFI".

"Camera for Conical Peripheral and Panoramic Photography" in the Conference on Current Developments in Optical Engineering and Commercial Optics, Part of 33rd Annual SPIE Conference held in San Diego, CA in 1989. Published in the proceedings of this conference.

The following patents may have related information concerning the peripheral reproduction of conical subjects: United States Patent Number

2,617,337 Snyder Reproduction of Conical Forms Jan. 19, 1949
1,001,549 Mertens Aug. 22, 1911
1,176,384 Lotka Mar. 16, 1916
1,456,954 Von Lucken May 29, 1923
1,738,095 Carleton Dec. 3, 1929
1,844,162 Hirsch Feb. 9, 1932
1,904,672 Berthon Apr. 18, 1933
2,066,782 Heymer Jan. 5, 1937
2,286,880 Weber Jun. 16, 1942
2,288,352 Henderson Jun. 30, 1942
Abstract

Photographs imply that they are representations of a particular scene in terms height, width and an instant in time. There are cameras that display time itself as a dimension of the final record. These are sometimes called "streak" or “strip” cameras. These cameras can be thought of as strip chart recorders where the subject information is gathered optically. This makes streak cameras powerful tools for non-contact measurement of subject changes over time. But they can also be used for other than purely technical applications.

In this presentation several improvised cameras of this type based on film and CCD or solid-state technology are presented and illustrated with applications. based the application of a linear CCD removed from an inexpensive hand-scanner and installed in the back of a 35mm camera body. I’ve used them to demonstrate a variety of applications where quantitative data about subject performance is desired and have also applied the camera for more aesthetically oriented purposes such as peripheral and panoramic photography. The cameras and their applications will be described in this presentation.

1. Scanning approach to photography

The process of scanning is becoming almost a household word in the field of imaging and indeed photography in general. This is due to the introduction of electronic scanning devices that approach imaging in a fundamentally different mode than the one normally associated with that of our own visual system.

We look at the world around us in an instantaneous fashion and perceive our surroundings all at the same time although we may concentrate on some aspect of a scene more closely than others. Standard photographic methods essentially duplicate our personal experience with remembering a slice of time, an instantaneous reality that presented itself to our eyes at some specific time.

Sometimes referred to or classified as 2-dimensional imaging, photography has assigned to it the characteristic of being a “witness of reality” because it captures an instant in time and makes a faithful record of whatever appeared in front of the camera at the selected time. We assume that photographs are to a large extent a reflection of reality. However, there are a host of film-based as well as digital image acquisition processes that operate in a fundamentally different manner than standard camera and when the images produced by these systems are seen under conventional conditions our perception and understanding of these images becomes confused and our imaginations often stretch to the limit.
At the heart of this alternative approach to image making is the focal plane shutter. This consists of a narrow slit or slot that travels across the film plane of a camera during the course of making a record of the scene it is recording on film (or a digital array). When the rate at which this slot travels across the image plane is short in comparison to the rate that images might move at during the exposure process the final photograph looks very much like the original scene. However, if the images move significantly during the time the slot travels across the image then various types of distortion are introduced in the final photograph. This effect is appreciated by photographers as “focal plane shutter distortion”. It is a type of distortion that is actually seldom seen. However, as shown in Figure 1., it can be made visible by altering the mechanical characteristics of a shutter or by chance as Henri Lartigue’s famous photograph of a race car appearing to lean forward in its direction of motion while the background shows bystanders and telephone poles leaning in the opposite direction demonstrates.

Generally we assume that there is no such thing as focal-plane shutter distortion. We assume that photographs are made instantaneously all over the surface of the film or of a solid state sensor in a digital camera. However, such instantaneity of recording, quite possible with standard cameras, is very difficult to achieve with electronic cameras when a truly high quality image are required. Instead, simultaneity of capture is sacrificed and a one dimensional array of CCDs (some call them pixels) is used and it is sequentially exposed to an image. Because the sensor is a 1-dimensional, not a 2-dimensional, sensor it is much cheaper to manufacture. In order to capture the full detail in a 3-dimensional scene and reproduce it as a 2-dimensional reproduction (as regular photographs do) the a single row of sensors is made to scan the image plane by mechanically moving it across the image. This means that large amounts of image bearing data can be extracted and stored. This leads to extremely high quality digital images.

This process of scanning a scene or the image of a subject has been known in earlier times and has been exploited by various specialized photographic technologies but has never gained the widespread acceptance that the electronic derivatives of these techniques will probably achieve.

The images recorded by slits or linear arrays can themselves be stationary in which case the slit or array is moved or the slit or array may remain stationary and the image moved past it. In this manner current electronic imaging systems are related to traditional scanning techniques associated with strip cameras and to traditional cameras equipped with focal plane shutters.

### 2. Streak photographs made with stationary slits or CCD arrays

If we turn the “slit-scan” or focal plane shutter method of scanning images around so that instead of moving the slit or the array, instead we move the film or simply “dump” data from a stationary CCD linear array into computer memory, we have an imaging system often described as a “streak” camera system. These cameras, while not often used to capture real looking reproductions of subjects, provide time or timing information that is often crucial for a quantitative analysis of a particular subject.

For example, if two or more light sources are placed along the array and turned on and off at slightly different times the final record will be as shown in Fig. 2. The streaks associated with the lamp will differ slightly from each other in terms of length and position on the record. The difference in length of each record can be traced back to differences in “on” time and differences in placement on the record (along the “time” dimension) can be assigned to variations in the
time at which one lamp was turned on with respect to the other.

As shown in Fig. 3, a rolling ball whose image travels along the slit, or array, over time will leave a curved path on the final record in which the slope of the curve at any point can be assigned a particular velocity and the curve itself is a visual representation of acceleration.

Streak cameras are what might be called “magnificent time machines”. They allow the visualization of time itself as an integral component of a photograph.

3. Applications of strip or scanning photography

When a camera equipped with a slit shutter past which film moves and over which an image moves or a CCD linear array across which an image of a subject moves, it is possible to make more or less faithful reproductions of the original subjects by matching the film velocity to that of the image or adjusting the sampling rate of the CCD so that proper aspect ratio of the image is maintained in the final record. When images move across the slit or the CCD array they are called generically “strip” cameras. They also have a wide variety of applications.

Listed below are but a few of the most basic ones that have enabled photographers to solve visual and technical problems that standard camera systems are simply incapable of dealing with effectively.

2.1 Linear strip photography. Linear strip photography comes in two different "flavors". In the most common example, the strip camera, fitted with a narrow slit past which film can be moved, is aimed at a subject that moves across the field of view of the camera. The image of the subject passes over the slit in the camera at about the same speed as the film moves under the slit.

Two widespread applications of this technique are the production of photofinish pictures at races, such as Fig. 4, and the making of synchroballistic photographs in the analysis of rocket performance.

Conversely, the camera can be placed in a moving vehicle such as a plane or a car and moved with respect to a stationary subject. In this way a long swath of subject can be recorded over time. One of the foremost applications of this technique is its use as an aerial camera system in photogrammetry, reconnaissance and remote sensing.

2.2 Peripheral photography. A variation of the above methods is to rotate a subject in front of the camera. In this case the slit is typically lined up with the center of rotation of the subject and with each revolution a complete 360 degree view of the subject's surface features is recorded onto the moving film. The process is called peripheral photography.
2.3 Panoramic photography. The last common example of the application of these cameras is in the use of the scanning principle for the making of ultra-wide angle panoramic photographs. Scanning panoramic cameras able to cover angles of view extending to a full 360 degree record. This is accomplished by rotating the strip camera. The film moving beneath the slit (or across the CCD array) then records various locations around the camera at different times and eventually a complete photograph of the scene surrounding the camera is secured.

2.4 Conical Peripheral and Panoramic Photography. A seldom seen development in the area of peripheral and panoramic photography is the use of film moving in circular rather than linear fashion behind the slit of a specially constructed strip camera. The design allows for the making of undistorted reproductions of subjects that have a conical shape and for the making of panoramic photographs that resemble conical projection techniques used in cartography.
The process of making images by the scanning method is full of potential for a myriad of applications. Hopefully, with the development of electronic imaging processes and greater access to linear arrays this method of image making will flourish and become an integral imaging process in art, science and technology and achieve a level of recognition in the photographic and imaging community in general that the forerunners of these devices never accomplished.

References

Code A: Matlab Code for Poisson Image Reconstruction from Image Gradients

% Read Input Gray Image
imgstr = 'test.png'; disp(sprintf('Reading Image %s',imgstr));
img = imread(imgstr);
[H,W,C] = size(img);
img = double(img);

% Find gradients
gx = zeros(H,W);
gy = zeros(H,W);
j = 1:H-1;
k = 1:W-1;
gx(j,k) = (img(j,k+1) - img(j,k));
gy(j,k) = (img(j+1,k) - img(j,k));

% Reconstruct image from gradients for verification
img_rec = poisson_solver_function(gx,gy,img);

figure;imagesc(img);colormap gray;colorbar;title('Image');
figure;imagesc(img_rec);colormap gray;colorbar;title('Reconstructed');
figure;imagesc(abs(img_rec-img));colormap gray;colorbar;title('Abs error');

function [img_direct] = poisson_solver_function(gx,gy,boundary_image);
% function [img_direct] = poisson_solver_function(gx,gy,boundary_image)
% Inputs; Gx and Gy -> Gradients
% Boundary Image -> Boundary image intensities
% Gx Gy and boundary image should be of same size
[H,W] = size(boundary_image);
gxx = zeros(H,W);
gyy = zeros(H,W);
f = zeros(H,W);
j = 1:H-1;
k = 1:W-1;

% Laplacian
gyy(j+1,k) = gy(j+1,k) - gy(j,k);
gxx(j,k+1) = gx(j,k+1) - gx(j,k);
f = gxx + gyy;

% boundary image contains image intensities at boundaries
boundary_image(2:end-1,2:end-1) = 0;
disp('Solving Poisson Equation Using DST'); tic
j = 2:H-1;
k = 2:W-1;
f_bp = zeros(H,W);
f_bp(j,k) = -4*boundary_image(j,k) + boundary_image(j,k+1) + boundary_image(j,k-1) + boundary_image(j-1,k) + boundary_image(j+1,k);
clear j k
f1 = f - reshape(f_bp,H,W);% subtract boundary points contribution
clear f_bp f

% DST Sine Transform algo starts here
f2 = f1(2:end-1,2:end-1);
%compute sine transform
tt = dst(f2);
f2sin = dst(tt)';
clear f2

% compute Eigen Values
[x,y] = meshgrid(1:W-2,1:H-2);
denom = (2*cos(pi*x/(W-1))-2) + (2*cos(pi*y/(H-1)) - 2);

%divide
f3 = f2sin./denom;
clear f2sin x y

%compute Inverse Sine Transform
tt = idst(f3);
clear f3;
img_tt = idst(tt)';
clear tt
time_used = toc;
disp(sprintf('Time for Poisson Reconstruction = %f secs',time_used));

% put solution in inner points; outer points obtained from boundary image
img_direct = boundary_image;
img_direct(2:end-1,2:end-1) = 0;
img_direct(2:end-1,2:end-1) = img_tt;
function b=dst(a,n)
%DST    Discrete sine transform    (Used in Poisson reconstruction)
%      Y = DST(X) returns the discrete sine transform of X.
%      The vector Y is the same size as X and contains the
%      discrete sine transform coefficients.
%      Y = DST(X,N) pads or truncates the vector X to length N
%      before transforming.
%      If X is a matrix, the DST operation is applied to each
%      column. This transform can be inverted using IDST.

error(nargchk(1,2,nargin));

if min(size(a))==1
   if size(a,2)>1
      do_trans = 1;
   else
      do_trans = 0;
   end
   a = a(:);
else
   do_trans = 0;
end
if nargin==1,   n = size(a,1); end
m = size(a,2);

% Pad or truncate a if necessary
if size(a,1)<n,
   aa = zeros(n,m);
   aa(1:size(a,1),:) = a;
else
   aa = a(1:n,:);
end

y=zeros(2*(n+1),m); y(2:n+1,:) = aa;
   y(n+3:2*(n+1),:) = -flipud(aa);
yy=fft(y);
   b=yy(2:n+1,:)/(-2*sqrt(-1));

if isreal(a), b = real(b); end
if do_trans, b = b.'; end

function b=idst(a,n)
%IDST   Inverse discrete sine transform    (Used in Poisson reconstruction)
%
%      X = IDST(Y) inverts the DST transform, returning the
%      original vector if Y was obtained using Y = DST(X).
%      X = IDST(Y,N) pads or truncates the vector Y to length N
%      before transforming.
%      If Y is a matrix, the IDST operation is applied to
%      each column.

if nargin==1
   if min(size(a)) ==1
      n=length(a);
   else
      n=size(a,1);
   end
end

nn=n+1;  b=2/nn*dst(a,n);
Code B: Matlab Code for Graph Cuts on Images

% Read gray scale image
I = imread('test.png'); [H,W,C] = size(I);

% Find graph cut
Ncut = graphcuts(I,33,255);

figure;imagesc(I);colormap gray;title('Image');
figure;imagesc(Ncut);colormap gray;title('Segmentation');

function [Ncut] = graphcuts(I,pad,MAXVAL)
% function [Ncut] = graphcuts(I)
% Input: I image
% pad: spatial connectivity; eg. 3
% MAXVAL: maximum image value
% Output: Ncut: Binary map 0 or 1 corresponding to image segmentation

I = double(I); [H,W] = size(I);

% Find weights between nodes I1 and I2, w = exp(a*abs(I1-I2));
% Set a to have a weight of 0.01 for diff = MAXVAL
a = log(0.01)/MAXVAL; x = [0:MAXVAL/100:MAXVAL]'; y = exp(a*x);
figure;plot(x,y);xlabel('intensity diff');ylabel('weights'); title('weights')

ws = 2*pad + 1;
if(ws <= 3)       ws = 3;  end

% Build the weight matrix
disp('Building Weight Matrix'); close all; tic
WM = zeros(H*W,H*W); countWM = 0;
for kk = 1:W
    for jj = 1:H
        mask = logical(zeros(H,W));
        cs = kk-pad;        ce = kk+pad;        rs = jj-pad;        re = jj+pad;
        if(cs<1)            cs = 1;  end;
        if(ce>W)            ce = W;         end;
        if(rs<1)            rs = 1;         end;
        if(re>H)            re = H;         end;
        mask(rs:re,cs:ce) = 1;
        idx = find(mask==1);
        p = abs(I(idx) - I(jj,kk));        p = exp(a*p);
        countWM = countWM + 1;             WM(countWM,idx) = p(:)';
    end
end
ttime = toc; disp(sprintf('Time for generating weight matrix = %f',ttime)); clear countWM

% Weight between a node and itself is 0
for jj = 1:H*W
    WM(jj,jj) = 0;
end
WM = sparse(WM);

% Shi and Malik Algorithm: second smallest eigen vector
disp('Finding Eigen Vector');
d = sum(WM,2);      D = diag(d);      tic
B = (D-WM);        B = (B+B')/2;    OPTS.disp = 0;
[v,d,flag] = eigs(B,D,2,'SA',OPTS);   ttime = toc;
disp(sprintf('Time for finding eigen vector = %f',ttime)); clear OPTS
y = v(:,2);
Ncut = reshape(y,H,W);
Ncut = Ncut > 0;
function [img1] = bilateral_filtering(img,winsize,sigma)

% Bilateral Filtering(img,winsize,sigma)
% Input     -> Image img
%      -> winsize: spatial filter width
%      -> sigma for intensity diff gaussian filter
%      -> sigma for spatial filter = winsize/6
% Output    -> Filtered Image
% Author: Amit Agrawal, 2004

disp('Bilateral Filtering');

[H,W] = size(img);

% Gaussian spatial filter

G_filter = fspecial('gaussian',winsize,winsize/6);
padnum = (winsize-1)/2;
A = padarray(img, [padnum padnum], 'replicate', 'both');
img1 = zeros(size(img));

for jj = padnum+1:(padnum+1+H-1)
    for kk = padnum+1:(padnum+1+W-1)
        % Get a local neighborhood
        imgwin = A(jj-padnum:jj+padnum,kk-padnum:kk+padnum);

        % Find weights according to intensity diffs
        Wwin = exp(-abs(imgwin - imgwin(padnum+1,padnum+1))/sigma^2);

        % Find composite filter
        newW = Wwin.*G_filter;

        t = sum(sum(newW));
        if(t>0)
            newW = newW/t;
        end

        img1(jj-padnum,kk-padnum) = sum(sum(imgwin.*newW));
    end
end
Bibliography

(We will update this bibliography as we refresh our course material)

Fusion of Images Taken by Varying Camera and Scene Parameters

General


Time


Exposure


Focus


Illumination


Passive Illumination


Polorization


Wavelength


Location


Matting


Techniques

General


Graph Cuts


**Gradient Domain**


**Smoothing, Bilateral and Trilateral Filter**


**Video Textures**


**Feature Extraction and Scene Understanding**

**Shape/Material/Illumination, Surface normals**

BASRI, R. JACOBS, D. Photometric stereo with general, unknown lighting, Computer Vision and Pattern Recognition, 2001


ANDREAS WENGER, A GARDNER, CHRIS TCHOU, J UNGER, T HAWKINS, P DEBEVEC, Postproduction Relighting and Reflectance Transformation With Time-Multiplexed Illumination, SIGGRAPH 2005

**Depth edges**

Ramesh RASKAR, Karhan TAN, Rogerio FERIS, Jingyi YU, Matthew TURK, Non-photorealistic Camera: Depth Edge Detection and Stylized Rendering Using a Multi-Flash Camera, SIGGRAPH 2004
S. K. NAYAR and Y. NAKAGAWA,, Shape from Focus, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, No. 8, pp. 824-831,

**Motion**


**Transfer and denoising**

**Flash to no-flash**

Elmar EISEMANN and Frédö DURAND, Flash Photography Enhancement Via Intrinsic Relighting, SIGGRAPH 2004


**Noise**


Eric P. BENNETT  and Leonard McMILLAN  "Video Enhancement Using Per-Pixel Virtual Exposures" SIGGRAPH 2005

**Geometric Operations**

**Panorama**


Synthetic Aperture and Multi-views


REN NG, Fourier Slice Photography, SIGGRAPH 2005


A. STERN and B. JAVIDI, "3-D computational synthetic aperture integral imaging (COMPSAI)," Opt. Express 11, 2446-2451 (2003),


Deblurring and Superresolution


Dynamic Range


**Active Illumination**


**Smart, Unconventional Cameras**

**MEMS Technology**


**High Speed Imaging**


X.Q. LIU AND ABBAS EL GAMAL. Synthesis of High Dynamic Range Motion Blur Free Image From Multiple Captures, In IEEE Transactions on circuits and systems (TCASI), VOL. 50, NO. 4, pp 530-539, APRIL 2003


**Programmable SIMD**

Advanced, Programmable, Demodulating Cameras and Temporal Correlation

CANESTA Inc, 2004

PIXIM Inc, 2004

FOVEON Inc, 2004

JENOPTIK Inc, 2004

IVP Inc, Ranger Camera, 2004


Others


TUMBLIN, J., AGRAWAL, A. AND RASKAR, R. Why I want a Gradient Camera, IEEE CVPR 2005

Optics


Zand, J., Coded aperture imaging in high energy astronomy, NASA Laboratory for High Energy Astrophysics (LHEA) at NASA's GSFC, 1996.

Light field optics


http://www.tgeorgiev.net/Spatioangular.pdf

**Fourier optics**


Computational Photography

SIGGRAPH 2007 Course 1 Notes

Organizers

Ramesh Raskar
Senior Research Scientist
MERL - Mitsubishi Electric Research Labs
201 Broadway
Cambridge, MA 02139
Email: raskar@merl.com
http://www.merl.com/people/raskar/

Jack Tumblin
Assistant Professor
Computer Science Dept., Northwestern University
1890 Maple Avenue, Evanston IL 60201,
Email: jet@cs.northwestern.edu
http://www.cs.northwestern.edu/~jet

Course Web Page

http://www.merl.com/people/raskar/photo/