CSc I6716
Spring 2011

3D Computer Vision

Introduction

Instructor: Zhigang Zhu
City College of New York
zzhu@ccny.cuny.edu

- Basic Information:
  - Course participation
  - Books, notes, etc.
  - Web page – check often!
- Homework, Assignment, Exam
  - Homework and exams
  - Grading
- *Goal*
  - What I expect from you
  - What you can expect from me
  - Resources

Book

- Textbook
  - "Introductory Techniques for 3-D Computer Vision" Trucco and Verri, 1998
- Additional readings when necessary
  - "Computer Vision – A Modern Approach" Forsyth and Ponce, 2003
  - "Three-Dimensional Computer Vision: A Geometric Viewpoint" O. Faugeras, 1998
  - "Image Processing, Analysis and Machine VIsion" Sonika, Hlavac and Boyle, 1999
- On-Line References

Prequisites

- Linear Algebra
- A little Probability and Statistics
- Programming Experience
- Reading Literature (A little)
- An Inquisitive Nature (Curiosity)
- No Fear

Course Web Page

http://www-cs.engr.ccny.cuny.edu/~zhu/CSCI6716-2011s/VisionCourse-Spring-2011.html

- Lectures available in Powerpoint format
- All homework assignments will be distributed over the web
- Additional materials and pointers to other web sites
- Course bulletin board contains last minute items, changes to assignments, etc.
  - **CHECK IT OFTEN!**
  - You are responsible for material posted there

Course Outline

- Complete syllabus on the web pages (14 meets)
- Rough Outline (3D Computer Vision):

Part 1. Vision Basics (Total 4)

1. Introduction (1)
2. Image Formation and Processing (1) (hw 1, matlab)
3-4. Features and Feature Extraction (2) (hw 2)

Part 2. 3D Vision (Total 7)

5. Camera Models (1)
6. Camera Calibration (2)(hw 3)
7. Stereo Vision (2) (project assignments)
8. Visual Motion (2) (hw 4)

Part 3. Exam and Projects (Total 3)

9. Project topics and exam review/discussion (1)
10. Midterm exam (1)
11. Student Project presentations (1)

Grading

- Homework (4): 40%
- Exam (midterm): 40%
- Course Project + Presentation: 20%
  - Groups (2-3 students) for discussions
  - Experiments – independently + collaboratively
  - Written Report - independently + collaboratively
- All homework must be yours….but you can work together until the final submission

- **Teaching Assistant:**
  - Mr. Wai L. Khoo <WKhoo@gc.cuny.edu>

3D Computer Vision
and Video Computing

C++ and Matlab

- C++
  - For some simple computation, you may use C++

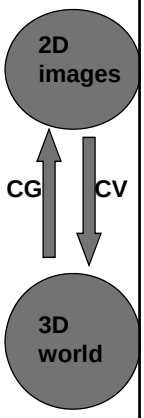
- Matlab
  - An interactive environment for numerical computation
  - Available on Computer Labs machines (both Unix and Windows)
    - Matlab primer available on line (web page)
    - Pointers to on-line manuals also available
  - Good rapid prototyping environment

- Use C++ and/or Matlab for your homework assignments and project(s); However Java will also be fine

3D Computer Vision and Video Computing

Course Goals and Questions

- What makes (3D) Computer Vision interesting ?
  - Image Modeling/Analysis/Interpretation
    - Interpretation is an Artificial Intelligence Problem
      - Sources of Knowledge in Vision
      - Levels of Abstraction
    - Interpretation often goes from 2D images to 3D structures
      - since we live in a 3D world
  - Image Rendering/Synthesis/Composition
    - Image Rendering is a Computer Graphics problem
    - Rendering is from 3D model to 2D images



3D Computer Vision and Video Computing

Related Fields

- Image Processing: image to image
- Computer Vision: Image to model
- Computer Graphics: model to image

All three are interrelated!

- Pattern Recognition: image to class
  - image data mining/ video mining
- Artificial Intelligence: machine smarts
  - Machine perception

AI

Applications

- Photogrammetry: camera geometry, 3D reconstruction
- Medical Imaging: CAT, MRI, 3D reconstruction (2nd meaning)
- Video Coding: encoding/decoding, compression, transmission

- Physics & Mathematics: basics
- Neuroscience: wetware to concept

basics

- Computer Science: programming tools and skills?

Applications

- Visual Inspection (*)
- Robotics (*)
- Intelligent Image Tools
- Image Compression (MPEG 1/2/4/7)
- Document Analysis (OCR)
- Image and Video on the Web
- Virtual Environment Construction (*)
- Environment (*)
- Media and Entertainment
- Medicine
- Astronomy
- Law Enforcement (*)
  - surveillance, security
- Traffic and Transportation (*)
- Tele-Conferencing and e-Learning (*)
- Human Computer Interaction (HCI)

3D Computer Vision
and Video Computing

Job Markets

- Homeland Security
  - Port security – cargo inspection, human ID, biometrics
  - Facility security – Embassy, Power plant, bank
  - Surveillance – military or civilian
- Media Production
  - Cartoon / movie/ TVs/ photography
  - Multimedia communication, video conferencing
- Research in image, vision, graphics, virtual reality
  - 2D image processing
  - 3D modeling, virtual walk-thorugh
- Consumer/ Medical Industries
  - Video cameras, Camcorders, Video phone
  - Medical imaging 2D -> 3D

IP vs CV

- Image processing (mainly in 2D)
  - Image to Image transformations
  - Image to Description transformations
  - Image Analysis - extracting quantitative information from images:
    - Size of a tumor
    - distance between objects
    - facial expression
  - Image restoration. Try to undo damage
    - needs a model of how the damage was made
  - Image enhancement. Try to improve the quality of an image
  - Image compression. How to convey the most amount of information with the least amount of data

What is Computer Vision?

Vision is the art of seeing things invisible.

-Jonathan Swift (1667-1745)
"Thoughts on Various Subjects"
Miscellanies in Prose and Verse
(published with Alexander Pope),
vol. 1, 1727

- Computer vision systems attempt to construct meaningful and explicit descriptions of the world depicted in an image.
- Determining from an image or image sequence:
  - The objects present in the scene
  - The relationship between the scene and the observer
  - The structure of the three dimensional (3D) space

Cues to Space and Time

Directly Measurable in an Image

- Spectral Characteristics
  - Intensity, contrast, colors and their
  - Spatial distributions
- 2D Shape of Contours
- Linear Perspective
- Highlights and Shadows
- Occlusions
- Organization
- Motion parallax and Optical Flow
- Stereopsis and sensor convergence

3D Computer Vision
and Video Computing

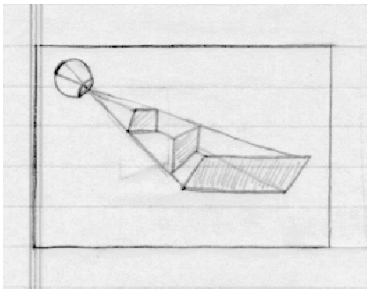
Cues to Space and Time

Inferred Properties

- Surface connectivity
- 3D Volume
- Hidden sides and parts
- Identity (Semantic category)
- Absolute Size
- Functional Properties
- Goals, Purposes, and Intents
- Organization
- Trajectories

Cues to Depth

- Question:
  - How do we perceive the three-dimensional properties of the world when the images on our retinas are only two-dimensional?
- Stereo is not the entire story!



Cues to Depth

- Monocular cues to the perception of depth in images
  - Interposition: occluding objects appear closer than occluded objects
  - Relative size: when objects have approximately the same physical size, the larger object appears closer
  - Relative height: objects lower in the image appear closer
  - Linear Perspective: objects appear smaller as they recede into the distance
    - texture gradients
  - Aerial Perspective: change in color and sharpness as object recede into the distance
  - Illumination gradients: gradients and shadow lend a sense of depth
  - Relative Motion: faster moving objects appear closer

Cues to Depth

- Physiological cues to depth:
  - Focus (accomodation): change in curvature of the lens for objects at different depths
  - Convergence: eyes turn more inward (nasal) for closer objects
  - Retinal disparity: greater for objects further away

Interposition

Interposition

Interposition

Different viewpoint

Different viewpoint

Edgar Degas: Dance Class at the Opéra, 1872

Different viewpoint

Edgar Degas:

Green Dancer,

c.1880

Different viewpoint

Edgar Degas: Frieze of Dancers, c.1895

Different viewpoint

Edgar Degas: Frieze of Dancers, c.1895

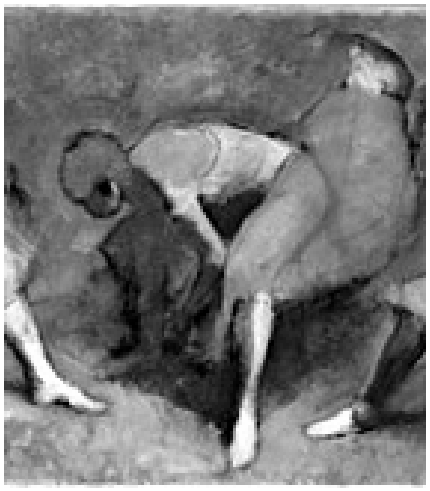
3D Computer Vision
and Video Computing

Different viewpoint

Edgar Degas: Frieze of Dancers, c.1895

3D Computer Vision and Video Computing

Different viewpoint



Edgar Degas: Frieze of Dancers, c.1895

3D Computer Vision and Video Computing

Different viewpoint

Edgar Degas: Frieze of Dancers, c.1895

3D Computer Vision and Video Computing

Different viewpoint

Edgar Degas: Frieze of Dancers, c.1895

3D Computer Vision and Video Computing

Different viewpoint

Edgar Degas: Frieze of Dancers, c.1895

Different viewpoint

Edgar Degas: Frieze of Dancers, c.1895

3D Computer Vision
and Video Computing

Different viewpoint

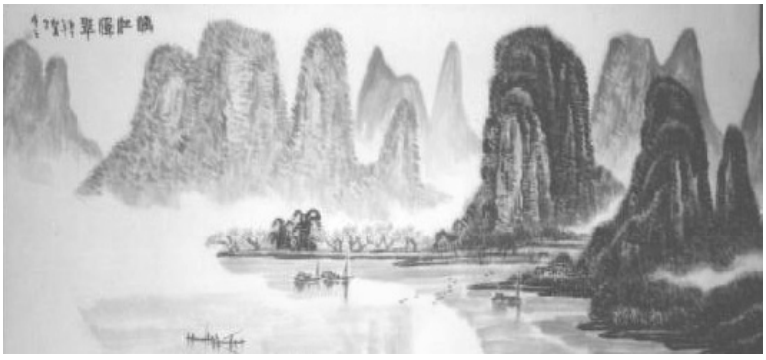
Edgar Degas: Frieze of Dancers, c.1895

Aerial Perspective

- Constable

Aerial Perspective

- Classic Chinese Paintings



Absolute Size

Relative Size

Relative Size

Absolute Size

Relative Size

Light and Surfaces

Light and Surfaces

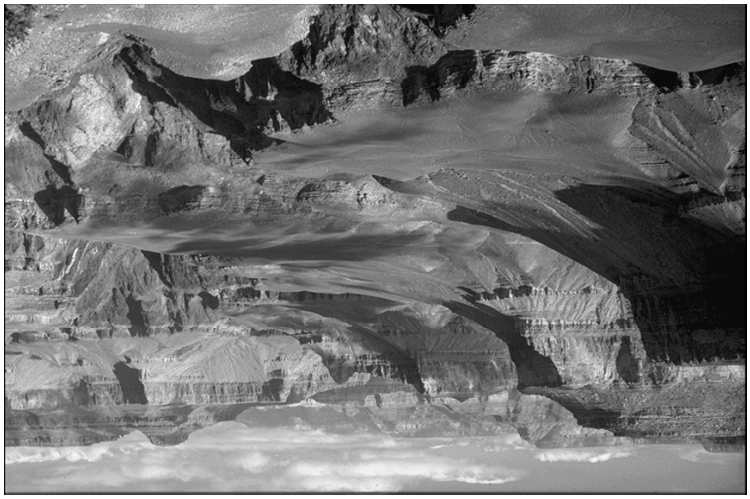
3D Computer Vision
and Video Computing

Light and Surfaces

Light and Surfaces

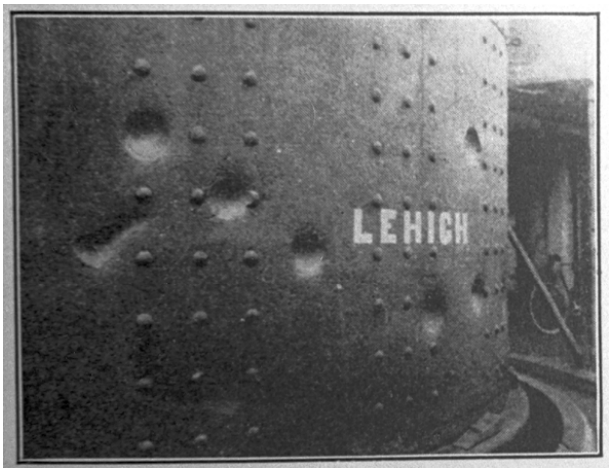
3D Computer Vision
and Video Computing

Light and Surfaces



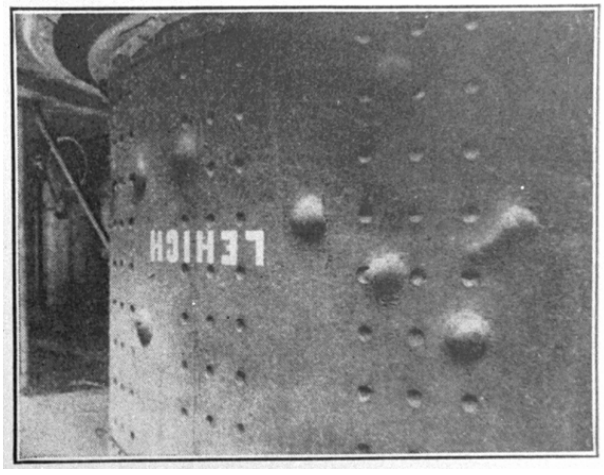
Light and Surfaces

- C. H. Stoelting Company

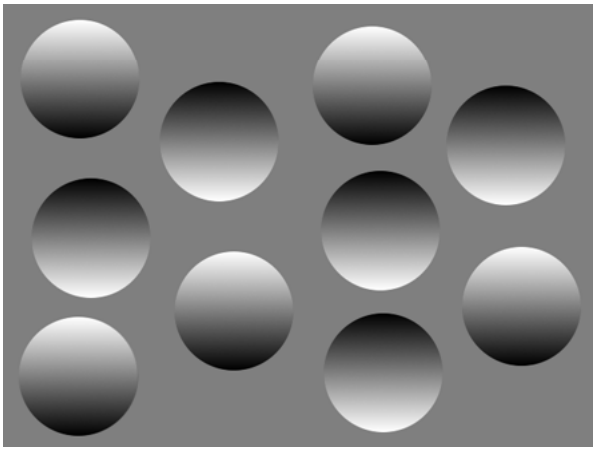


3D Computer Vision
and Video Computing

Light and Surfaces

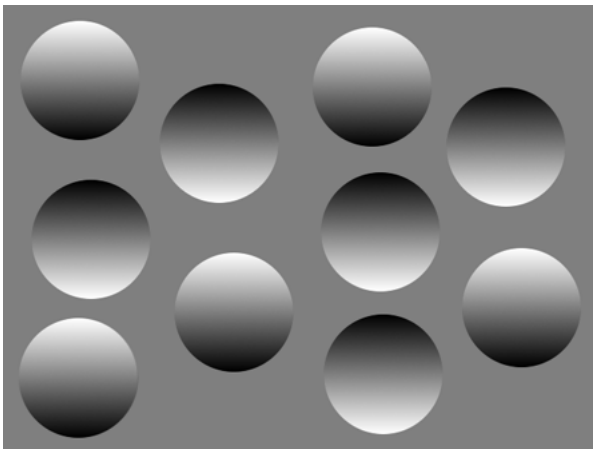


Light and Surfaces



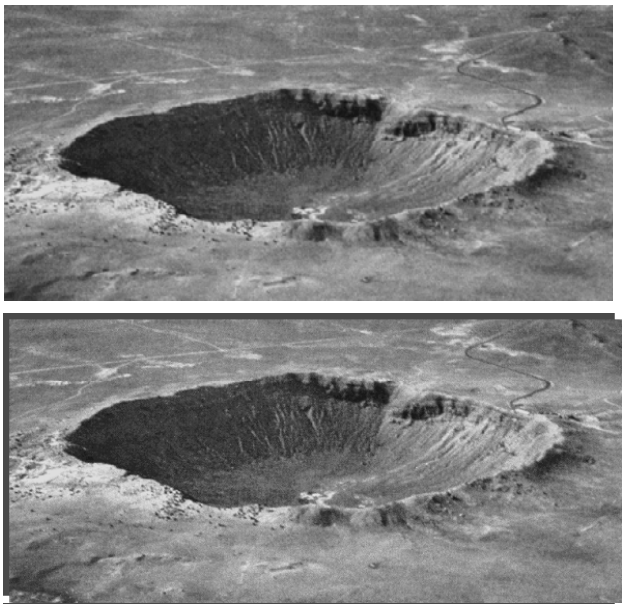
3D Computer Vision
and Video Computing

Light and Surfaces



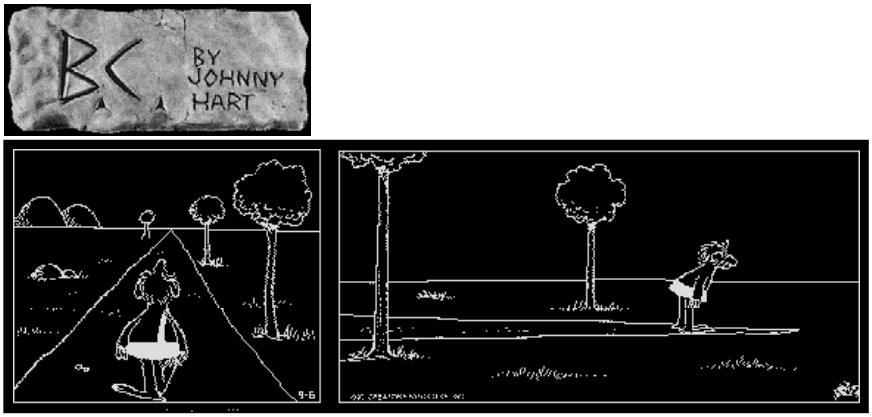
3D Computer Vision and Video Computing

Light and Surfaces



3D Computer Vision and Video Computing

The Effect of Perspective



Texture Gradient

Sunflowers in Fargo, ND
Photo by Bruce Fitz

http://www.ars.usda.gov/is/graphics/photos/

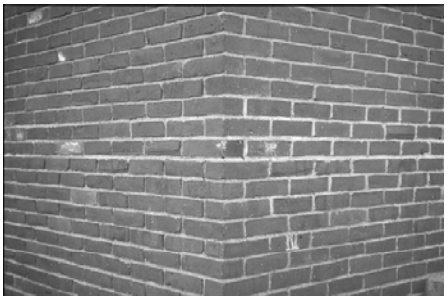
Texture Gradients

3D Computer Vision
and Video Computing

Edges

3D Computer Vision
and Video Computing

Texture Edges



Who Knows

Next

Anyone who isn't confused really doesn't understand the situation.

--Edward R. Murrow

Next:
Image Formation

Reading: Ch 1, Ch 2- Section 2.1, 2.2, 2.3, 2.5

Questions: 2.1. 2.2, 2.3, 2.5

Exercises: 2.1, 2.3, 2.4