

Introduction

Introduction to Computer Vision
CSE 152
Lecture 1

- We'll begin with some introductory material ...
- ... and end with
 - Syllabus
 - Organizational materials
 - Wait list

What is computer vision?



Done?

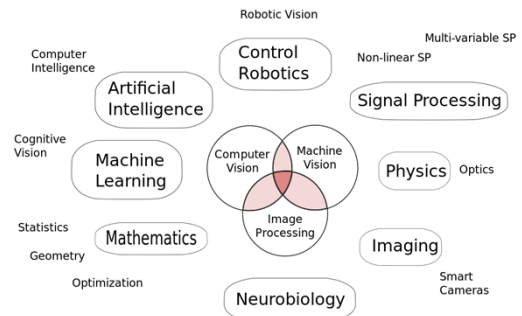
Computer Vision

- An interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or videos
- Other common definitions:
 - Computing properties of the 3-D world from one or more digital images (Trucco and Verri)
 - To make useful decisions about real physical objects and scenes based on sensed images (Stockman and Shapiro)
 - The construction of explicit, meaningful description of physical objects from images (Ballard and Brown)
 - Extracting descriptions of the world from pictures or sequences of pictures (Forsyth and Ponce)

Computer Vision

- An interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or videos
- Engineering perspective
 - Computer vision seeks to automate tasks that the human visual system can do

Related Fields



The four Rs of computer vision

- **Reprojection**
 - Rendering a scene from a different view, under different illumination, under different surface properties, etc.
- **Reconstruction**
 - Single and multiple view geometry, structure from motion, shape from X (where X is texture, shading, contour, etc.), etc.
- **Registration**
 - Tracking, alignment, optical flow, correspondence, etc.
- **Recognition**
 - Recognizing objects, scenes, events, etc.

CSE 152, Spring 2017

Others may have slightly different Rs

Introduction to Computer Vision

Rudiments: The implied fifth R

- Image filtering
- Edge detection
- Interest point detection
- Probability
- Statistics
- Linear algebra
- Projective geometry
- Optics
- Fourier analysis
- Sampling
- Algorithms
- Photometry
- Physics of color
- Human vision
- Psychophysics
- Performance evaluation

CSE 152, Spring 2017

Introduction to Computer Vision

Why is this hard?



What is in this image?

1. A hand holding a man?
2. A hand holding a mirrored sphere?
3. An Escher drawing?

- Interpretations are ambiguous
- The forward problem (graphics) is well-posed
- The “inverse problem” (vision) is not

CSE 152, Spring 2017

Introduction to Computer Vision

Underestimates

“640K ought to be enough for anybody.”

– Bill Gates, 1981

“... in three to eight years we will have a machine with the general intelligence of an average human being ... The machine will begin to educate itself with fantastic speed. In a few months it will be at genius level and a few months after that its powers will be incalculable ...”

– Marvin Minsky, LIFE Magazine, 1970

CSE 152, Spring 2017

Introduction to Computer Vision

Should Computer Vision follow from our understanding of Human Vision?

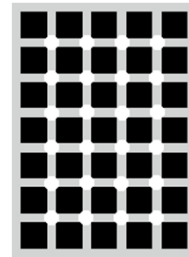
Yes & No

1. Who would ever be crazy enough to even try creating machine vision?
 2. Human vision “works”, and copying is easier than creating.
 3. Secondary benefit – in trying to mimic human vision, we learn about it.
1. Why limit oneself to human vision when there is even greater diversity in biological vision?
 2. Why limit oneself to biological vision when there may be greater diversity in sensing mechanism?
 3. Biological vision systems evolved to provide functions for “specific” tasks and “specific” environments. These may differ for machine systems.
 4. Implementation – hardware is different, and synthetic vision systems may use different techniques/methodologies that are more appropriate to computational mechanisms.

CSE 152, Spring 2017

Introduction to Computer Vision

Hermann Grid



Scan your eyes over the figure. Do you see the gray spots at the intersections? Stare at one of them and it will disappear.

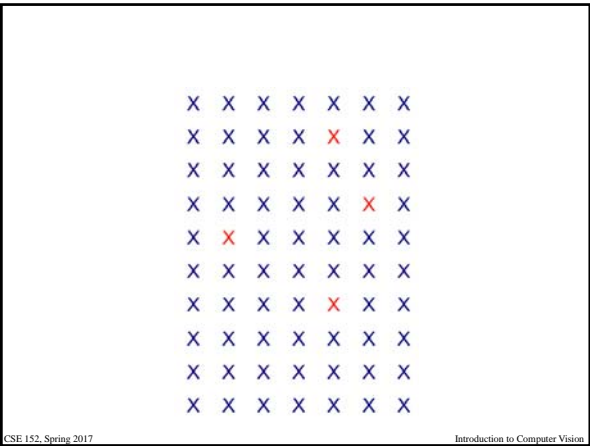
CSE 152, Spring 2017

Introduction to Computer Vision

How many red X's are there?

Raise your hand when you know.

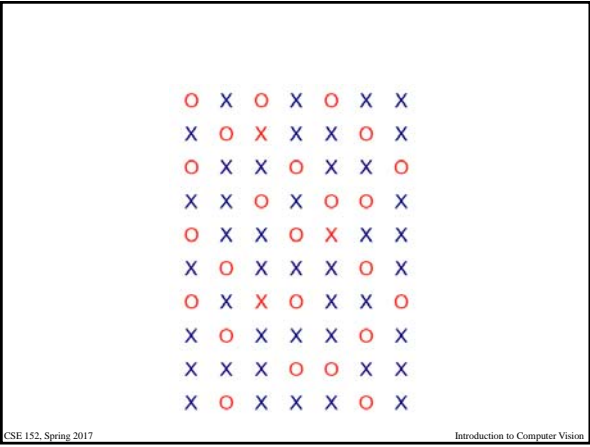
CSE 152, Spring 2017 Introduction to Computer Vision



How many red X's are there?

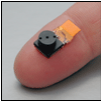
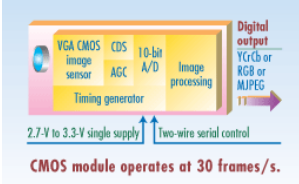
Raise your hand when you know.

CSE 152, Spring 2017 Introduction to Computer Vision



The Near Future: Ubiquitous Vision

- Digital video has become very inexpensive.
- It's widely embedded in cell phones, cars, games, etc.
- 99.9% of digitized video isn't seen by a person.
- That doesn't mean that only 0.1% is important!

CSE 152, Spring 2017 Introduction to Computer Vision

- ### Applications: touching your life
- Optical Character Recognition
 - Football
 - Movies
 - Surveillance
 - HCI – hand gestures
 - Aids to the blind
 - Face recognition & biometrics
 - Road monitoring
 - Industrial inspection
 - Virtual Earth; street view
 - Robotic control
 - Autonomous driving
 - Space: planetary exploration, docking
 - Medicine – pathology, surgery, diagnosis
 - Microscopy
 - Military
 - Remote Sensing
 - Digital photography
 - Google Goggles
 - Video games
- CSE 152, Spring 2017 Introduction to Computer Vision

Earth viewers (3D modeling)



Image from Microsoft's Virtual Earth (now Bing Maps)
(see also: Google Earth)

CSE 152, Spring 2017

Introduction to Computer Vision

Photosynth



<http://photosynth.net>
Based on [Photo Tourism technology](#) developed
by Noah Snavely, Steve Seitz, and Rick Szeliski

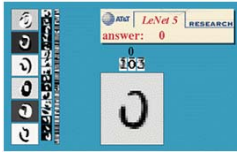
CSE 152, Spring 2017

Introduction to Computer Vision

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs



License plate readers

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

CSE 152, Spring 2017

Introduction to Computer Vision

Face detection



- Most new digital cameras now detect faces, so do smart phones...
 - Canon, Sony, Fuji, ...

CSE 152, Spring 2017

Introduction to Computer Vision

Smile detection

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



Sony Cyber-shot® T70 Digital Still Camera

CSE 152, Spring 2017

Introduction to Computer Vision

Object recognition (in supermarkets)



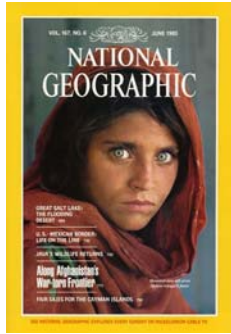
LaneHawk by EvolutionRobotics (now part of iRobot)

"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it..."

CSE 152, Spring 2017

Introduction to Computer Vision

Face recognition



Who is she?

CSE 152, Spring 2017

Introduction to Computer Vision

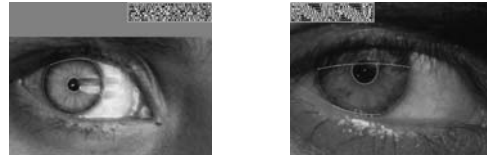
Vision-based biometrics

1984
Age 12



2002
Age 30

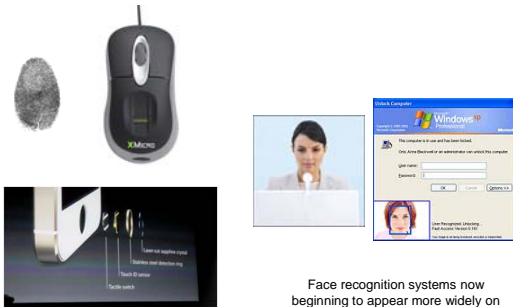
"How the Afghan Girl was Identified by Her Iris Patterns" Read the [story](#)



CSE 152, Spring 2017

Introduction to Computer Vision

Login without a password...



Fingerprint scanners on laptops, smart phones, mice, and other devices

Face recognition systems now beginning to appear more widely on computers and smart phones

CSE 152, Spring 2017

Introduction to Computer Vision

Object recognition (in mobile phones)



- Point & Find, Nokia
- SnapTell.com (now Amazon)
- Mobile Acuity
- Google Goggles

CSE 152, Spring 2017

Introduction to Computer Vision

Image-based search



CSE 152, Spring 2017

Introduction to Computer Vision

Special effects: shape capture



The Matrix movies, ESC Entertainment, XYZRGB, NRC

CSE 152, Spring 2017

Introduction to Computer Vision

Special effects: motion capture



Facial motion capture

CSE 152, Spring 2017

Introduction to Computer Vision

Light Fields

- An array of cameras



The Matrix (bullet time)

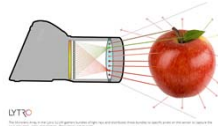
Application: view synthesis

CSE 152, Spring 2017

Introduction to Computer Vision

Computational photography

- Light-field cameras
 - A dense array of micro-lenses



Application: refocus

CSE 152, Spring 2017

Introduction to Computer Vision

Sports

- Football first down line



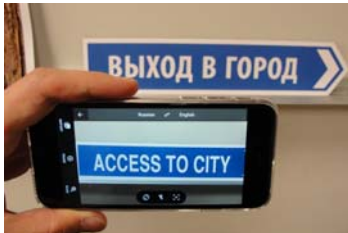
Sportvision first down line
Nice [explanation](http://www.howstuffworks.com) on www.howstuffworks.com

CSE 152, Spring 2017

Introduction to Computer Vision

Augmented reality

- Text detection, localization, and translation, then render with similar font



CSE 152, Spring 2017

Introduction to Computer Vision

Augmented reality

- Pokémon Go






CSE 152, Spring 2017

Introduction to Computer Vision

Augmented reality


- Microsoft HoloLens


Simultaneous localization and mapping (SLAM)

CSE 152, Spring 2017 Introduction to Computer Vision


Vision-based interaction (and games)




Nintendo Wii has camera-based IR tracking built in



Your face on a 3D avatar



Playmotion game a Disney Epicot



Xbox Kinect


CSE 152, Spring 2017 Introduction to Computer Vision

Structured light-based sensor



CSE 152, Spring 2017 Introduction to Computer Vision

Vision in space




NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks


- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

CSE 152, Spring 2017 Introduction to Computer Vision

Robotics




NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



<http://www.roboocup.org/>

CSE 152, Spring 2017 Introduction to Computer Vision

Smart cars



Our Vision. Your Safety.

- EyeQ Vision on a Chip
- Vision Applications: Road, Vehicle, Pedestrian Protection and more
- AWS Advance Warning System

• Mobileye
– Vision systems currently in high-end BMW, GM, Volvo models

Slide content courtesy of Amnon Shashua

CSE 152, Spring 2017 Introduction to Computer Vision

Autonomous Vehicles



CSE 152, Spring 2017

Introduction to Computer Vision

First person vision

Google Glass



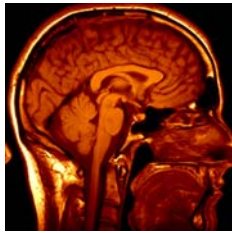
Oracam



CSE 152, Spring 2017

Introduction to Computer Vision

Medical imaging



3D imaging
MRI, CT

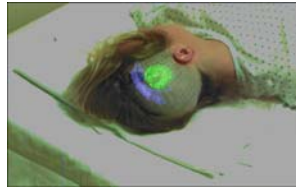


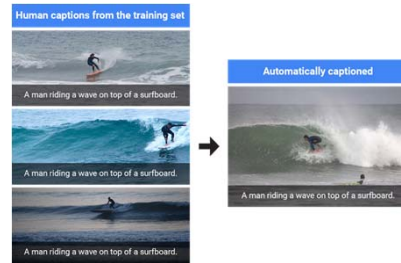
Image guided surgery
[Grimson et al., MIT](#)

CSE 152, Spring 2017

Introduction to Computer Vision

Automatic image captioning

- Deep learning

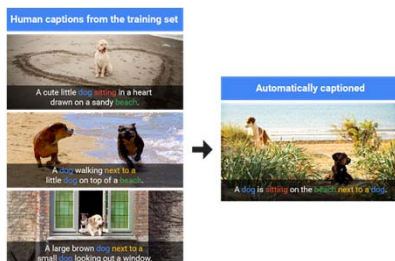


CSE 152, Spring 2017

Introduction to Computer Vision

Automatic image captioning

- Deep learning

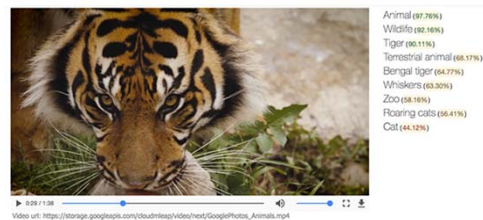


CSE 152, Spring 2017

Introduction to Computer Vision

Video understanding

- Deep learning



CSE 152, Spring 2017

Introduction to Computer Vision

Current state of the art

- You just saw examples of current systems.
 - Many of these are less than 5 years old
- This is a very active research area, and rapidly changing
 - Many new applications in the next 5 years
- To learn more about vision applications and companies
 - David Lowe maintains an excellent overview of vision companies
 - <http://www.cs.ubc.ca/spider/lowe/vision.html>

CSE 152, Spring 2017

Introduction to Computer Vision

Image Interpretation - Cues

- Variation in appearance in multiple views
 - stereo
 - motion
- Shading & highlights
- Shadows
- Contours
- Texture
- Blur
- Geometric constraints
- Prior knowledge

CSE 152, Spring 2017

Introduction to Computer Vision

An example of a cue: Shading and lighting

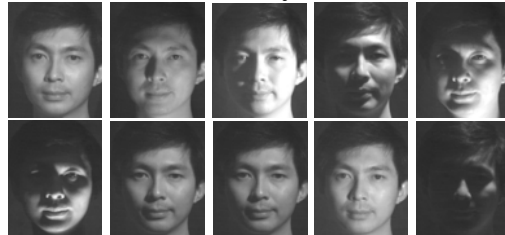
Shading as a result of differences in lighting is

1. A source of information
2. An annoyance

CSE 152, Spring 2017

Introduction to Computer Vision

Illumination Variability An annoyance



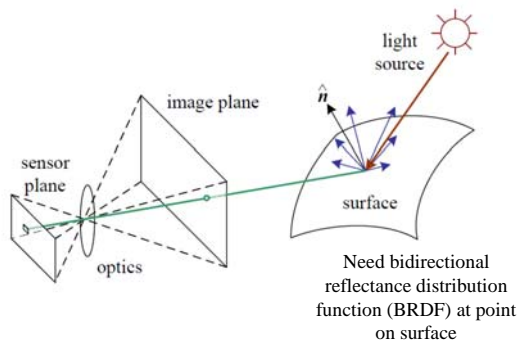
“The variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity.”

-- Moses, Adini, Ullman, ECCV '94

CSE 152, Spring 2017

Introduction to Computer Vision

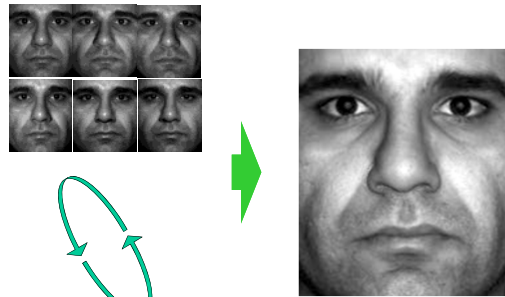
Photometric image formation



CSE 152, Spring 2017

Introduction to Computer Vision

An implemented algorithm: Relighting



Single Light Source

CSE 152, Spring 2017

Introduction to Computer Vision

The course

- Part 1: The Physics of Imaging
- Part 2: Early Vision
- Part 3: Reconstruction
- Part 4: Recognition

CSE 152, Spring 2017

Introduction to Computer Vision

Part 1 of Course: The Physics of Imaging

- How images are formed
 - Cameras
 - What a camera does
 - Projection models (projective spaces, etc.)
 - How to tell where the camera was located
 - Light
 - How to measure light
 - What happens to light at surfaces
 - How the brightness values we see in images are determined
 - Color
 - The underlying mechanisms of color
- How to describe it and measure it

CSE 152, Spring 2017

Introduction to Computer Vision

Cameras, lenses, and sensors



- Pinhole cameras
- Lenses
- Projection models
- Geometric camera parameters

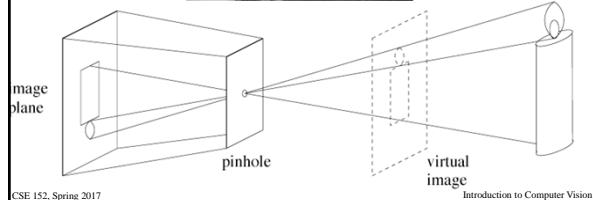
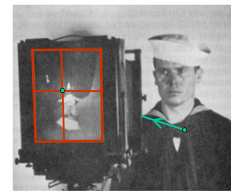
Figure 1.16 The first photograph on record, *la table servie*, obtained by Nicéphore Niepce in 1822. *Collection Harlinge-Viollet*.

From *Computer Vision*, Forsyth and Ponce, Prentice-Hall, 2002.

CSE 152, Spring 2017

Introduction to Computer Vision

A real camera ... and its model

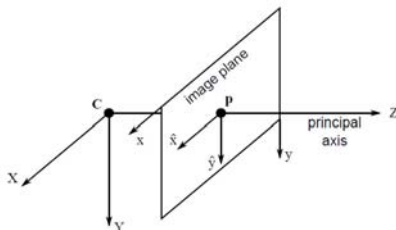


CSE 152, Spring 2017

Introduction to Computer Vision

Geometry

- How do 3D world points project to 2D image points?

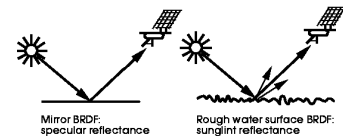


CSE 152, Spring 2017

Introduction to Computer Vision

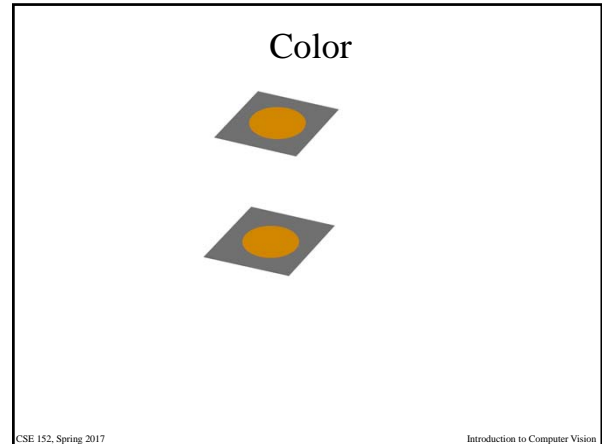
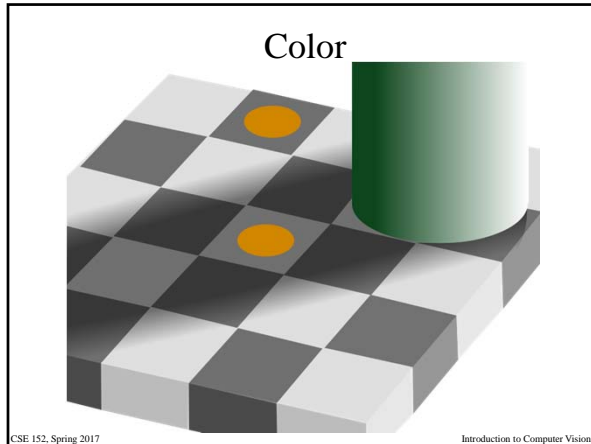
Lighting & Photometry

- How does measurement relate to light energy?
- Sensor response
- Light sources
- Reflectance



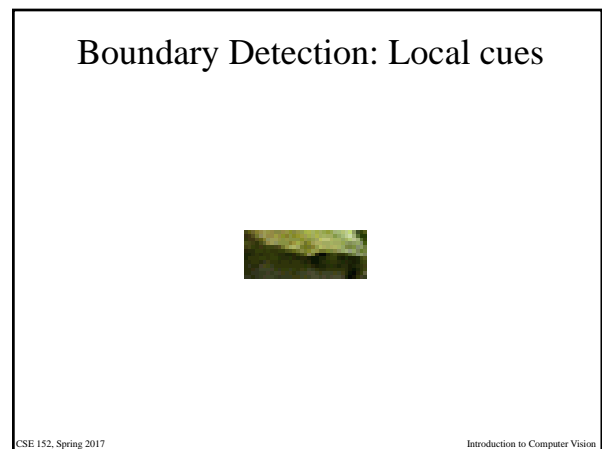
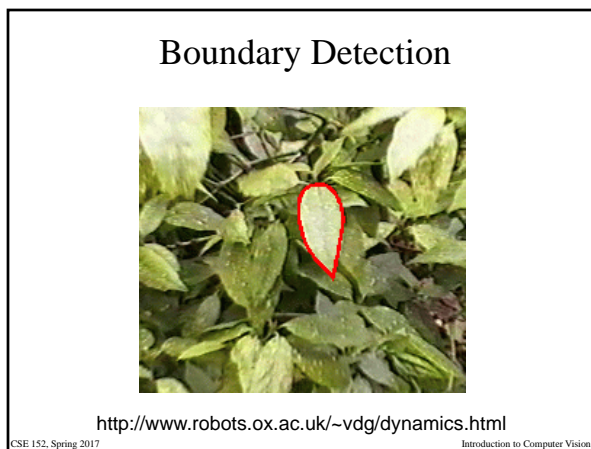
CSE 152, Spring 2017

Introduction to Computer Vision

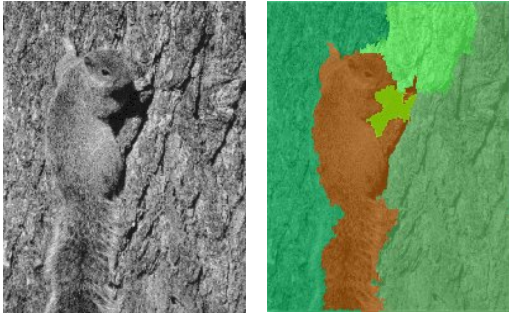


- ### Part 2: Early Vision in One Image
- Representing small patches of image
 - Noise
 - Filtering
 - Edge detection
 - Corner detection
 - Texture
 - Segmentation
- CSE 152, Spring 2017 Introduction to Computer Vision

- ### Segmentation
- Which image components “belong together”?
 - Belong together \cong lie on the same object
 - Cues
 - Similar color
 - Similar texture
 - Not separated by contour
 - Form a suggestive shape when assembled
- CSE 152, Spring 2017 Introduction to Computer Vision



Boundary Detection



(Sharon, Balun, Brandt, Basri)

CSE 152, Spring 2017

Introduction to Computer Vision

Part 3: Reconstruction from Multiple Images

- Photometric Stereo
 - What we know about the world from lighting changes
- The geometry of multiple views
 - Stereopsis
 - What we know about the world from having two eyes
 - Structure from motion
 - What we know about the world from having many eyes (or, more commonly, our eyes moving)

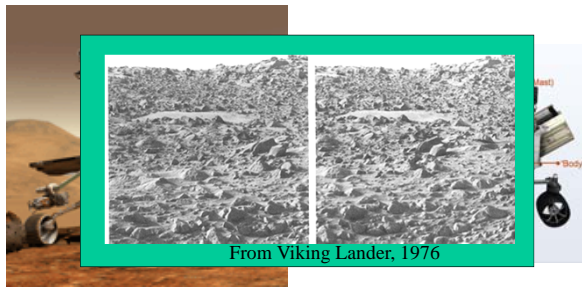
CSE 152, Spring 2017

Introduction to Computer Vision

Mars Rover

Spirit

Curiosity



From Viking Lander, 1976

CSE 152, Spring 2017

Introduction to Computer Vision

Façade (Debevec, Taylor and Malik, 1996)

Reconstruction from multiple views, constraints, rendering



Architectural modeling:

- photogrammetry;
- view-dependent texture mapping;
- model-based stereopsis.

Reprinted from "Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach,"
By P. Debevec, C.J. Taylor, and J. Malik,
Proc. SIGGRAPH (1996). © 1996 ACM, Inc.
Included here by permission.

CSE 152, Spring 2017

Introduction to Computer Vision

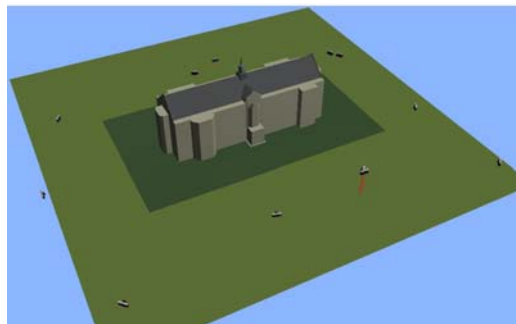
Images with marked features



CSE 152, Spring 2017

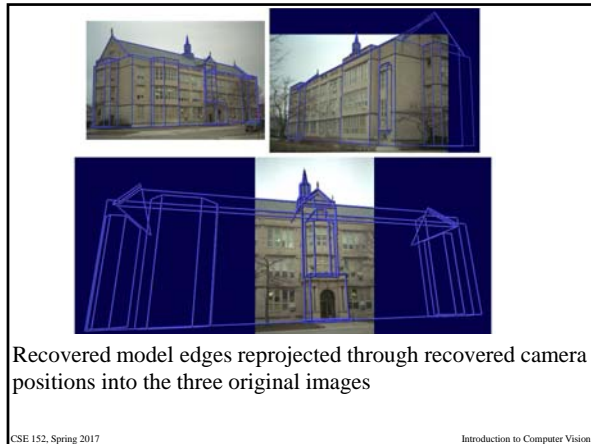
Introduction to Computer Vision

Resulting model & Camera Positions



CSE 152, Spring 2017

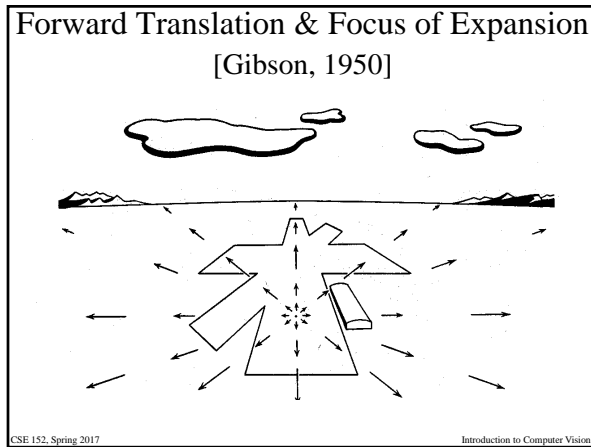
Introduction to Computer Vision



Video-Motion Analysis

- Where “things” are moving in image – segmentation
- Determining observer motion (egomotion)
- Determining scene structure
- Tracking objects
- Understanding activities & actions

CSE 152, Spring 2017 Introduction to Computer Vision



Part 4: Recognition

Given a database of objects and an image determine what, if any of the objects are present in the image

CSE 152, Spring 2017 Introduction to Computer Vision

Recognition Challenges

- Within-class variability
 - Different objects within the class have different shapes or different material characteristics
 - Deformable
 - Articulated
 - Compositional
- Pose variability:
 - 2-D image transformation (translation, rotation, scale)
 - 3-D pose variability (perspective, orthographic projection)
- Lighting
 - Direction (multiple sources & type)
 - Color
 - Shadows
- Occlusion – partial
- Clutter in background -> false positives

CSE 152, Spring 2017 Introduction to Computer Vision

Face Detection (not face recognition):

Classify face vs. non-face

CSE 152, Spring 2017 Introduction to Computer Vision

Why is Face Recognition Hard? Many faces of Madonna



CSE 152, Spring 2017

Introduction to Computer Vision

Syllabus

- Instructor: Ben Ochoa
 - Office hours: Mondays, 8:00 PM-9:00 PM
- TAs: Lenord Melvix, Jean Choi, and Daniel Peroni
 - Office hours: Mondays, 12:00 noon-2:00 PM; Tuesdays, 1:00 PM-2:00 PM; Wednesdays, 12:00 noon-2:00 PM
- Course website
 - <http://cseweb.ucsd.edu/classes/sp17/cse152-a/>
- 19 lecture meetings
 - No meeting on Memorial Day Observance (Monday, May 29)
- Weekly discussion section (optional)
- Class discussion
 - Piazza

CSE 152, Spring 2017

Introduction to Computer Vision

Syllabus

- Grading
 - 5 homework assignments
 - By hand and programming
 - Late submissions are not accepted
 - Piazza
 - Ask (and answer) questions using Piazza, not email
 - Good participation could raise your grade (e.g., raise a B+ to an A-)
 - No midterm exams
 - No final exam*

CSE 152, Spring 2017

Introduction to Computer Vision

Textbook (optional)

- Computer Vision: A Modern Approach, second edition (do not use first edition)
 - David A. Forsyth and Jean Ponce



CSE 152, Spring 2017

Introduction to Computer Vision

Academic Integrity Policy

Integrity of scholarship is essential for an academic community. The University expects that both faculty and students will honor this principle and in so doing protect the validity of University intellectual work. For students, this means that all academic work will be done by the individual to whom it is assigned, without unauthorized aid of any kind.

CSE 152, Spring 2017

Introduction to Computer Vision

Collaboration Policy

It is expected that you complete your academic assignments on your own and in your own words and code. The assignments have been developed by the instructor to facilitate your learning and to provide a method for fairly evaluating your knowledge and abilities (not the knowledge and abilities of others). So, to facilitate learning, you are authorized to discuss assignments with others; however, to ensure fair evaluations, you are not authorized to use the answers developed by another, copy the work completed by others in the past or present, or write your academic assignments in collaboration with another person. If the work you submit is determined to be other than your own, you will be reported to the Academic Integrity Office for violating UCSD's Policy on Integrity of Scholarship.

CSE 152, Spring 2017

Introduction to Computer Vision

Wait List

- Number of enrolled students is limited by
 - Size of lecture and discussion rooms
 - Number of TAs
- General advice
 - Wait for as long as you can
- Concurrent enrollment (Extension) students have lowest priority