

Introduction

Introduction to Computer Vision
CSE 152
Lecture 1

CSE152, Winter 2012

Intro Computer Vision

What is computer vision?



Done?

CSE152, Winter 2012

Intro Computer Vision

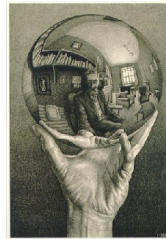
What is Computer Vision?

- Trucco and Verri (Text): Computing properties of the 3-D world from one or more digital images
- Sockman and Shapiro: To make useful decisions about real physical objects and scenes based on sensed images
- Ballard and Brown: The construction of explicit, meaningful description of physical objects from images.
- Forsyth and Ponce: Extracting descriptions of the world from pictures or sequences of pictures"

CSE152, Winter 2012

Intro 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?
4. A 1935 self portrait of Escher

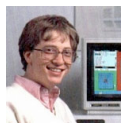
- Interpretations are ambiguous
- The forward problem (graphics) is well-posed
- The "inverse problem" (vision) is not

CSE152, Winter 2012

Intro Computer Vision

We all make mistakes

- "In 1966, Minsky hired a first-year undergraduate student and assigned him a problem to solve over the summer: connect a television camera to a computer and get the machine to describe what it sees."
– Crevier 1993, pg. 88
- "640K ought to be enough for anybody."
– Bill Gates, 1981

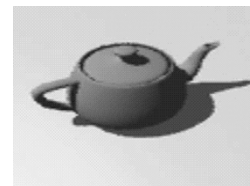


CSE152, Winter 2012

Intro Computer Vision

What do you see?

- ☐ Changing viewpoint
- ☐ Moving light source
- ☐ Deforming shape

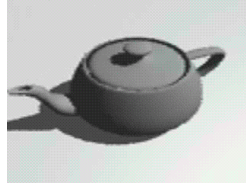


CSE152, Winter 2012

Intro Computer Vision

What was happening

- ☐ Changing viewpoint
- ☐ Moving light source
- ☒ Deforming shape



CSE152, Winter 2012

Intro Computer Vision

Should Computer Vision follow from our understanding of Human Vision?

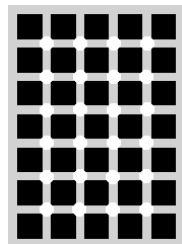
Yes & No

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 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

CSE152, Winter 2012

Intro 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.

CSE152, Winter 2012

Intro Computer Vision

How many red X's are there?

Raise your hand when you know.

CSE152, Winter 2012

Intro Computer Vision



CSE152, Winter 2012

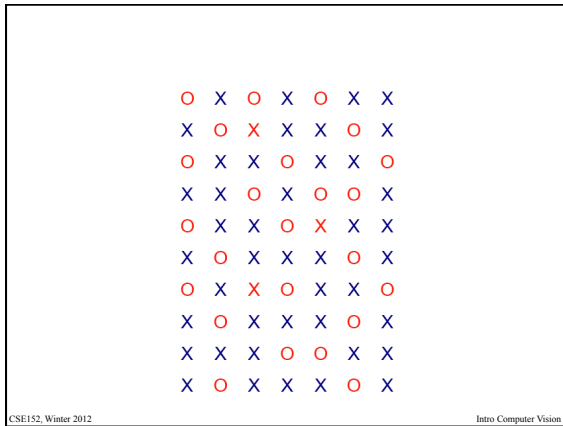
Intro Computer Vision

How many red X's are there?

Raise your hand when you know.

CSE152, Winter 2012

Intro Computer Vision



The Near Future: Ubiquitous Vision

- Digital video has become really cheap.
- More than 1 Billion cameras sold last year.
- 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!
- Oh, and computers, memory, and bandwidth have gotten pretty cheap too.

CSE152, Winter 2012

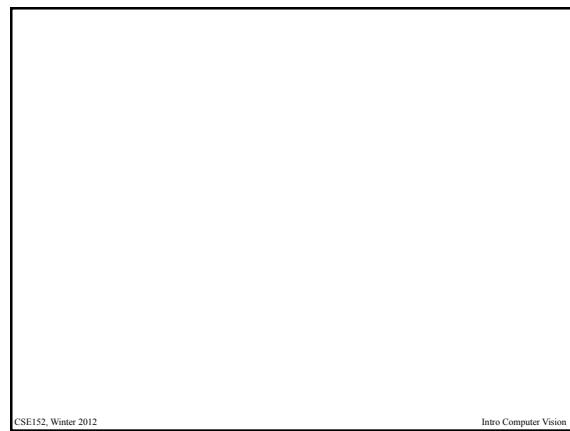
Intro Computer Vision

Applications: touching your life

- Digital photography
- Google Goggles
- Video games
- Football
- Movies
- Surveillance
- HCI – hand gestures
- Aids to the blind
- Face recognition & biometrics
- Road monitoring
- Industrial inspection
- Robotic control
- Autonomous driving
- Space: planetary exploration, docking
- Medicine – pathology, surgery, diagnosis
- Microscopy
- Military
- Remote Sensing
- Virtual Earth; street view
- Homeland security

CSE152, Winter 2012

Intro Computer Vision



Earth viewers (3D modeling)

Image from Microsoft's [Virtual Earth](#) (see also: [Google Earth](#))

CSE152, Winter 2012

Intro Computer Vision

Photosynth

http://photosynth.net

Based on [Photo Tourism technology](#) developed by Noah Snavely, Steve Seitz, and Rick Szeliski

CSE152, Winter 2012

Intro Computer Vision

MapsGL

MapsGL

CSE152, Winter 2012

Intro Computer Vision

Photo Tourism overview

Photo Tourism

Exploring photo collections in 3D

Noah Snavely Steven M. Seitz Richard Szeliski
University of Washington Microsoft Research

SIGGRAPH 2006

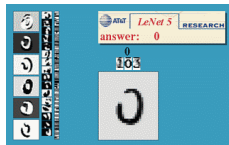
CSE152, Winter 2012

Intro 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
<http://www.research.att.com/~yanli/>



License plate readers
http://en.wikipedia.org/wiki/Summer_license_plate_recognition

CSE152, Winter 2012

Intro Computer Vision

Face detection



- Most new digital cameras now detect faces
– Canon, Sony, Fuji, ... iPhone

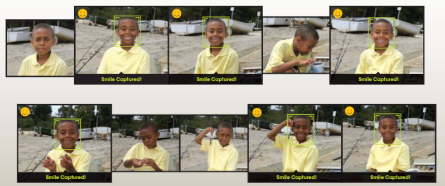
CSE152, Winter 2012

Intro 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

CSE152, Winter 2012

Intro Computer Vision

Object recognition (in supermarkets)



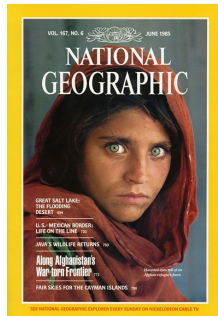
[LaneHawk by EyeQvision Robotics](#)

"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..."

CSE152, Winter 2012

Intro Computer Vision

Face recognition



Who is she?

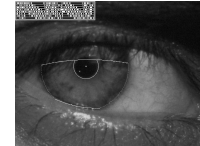
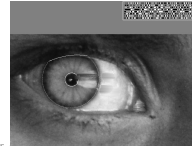
CSE152, Winter 2012

Intro Computer Vision

Vision-based biometrics



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



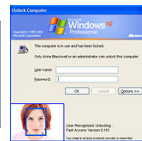
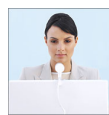
CSE152, Winter 2012

Intro Computer Vision

Login without a password...



Fingerprint scanners on many new laptops, other devices



Face recognition systems now beginning to appear more widely
<http://www.snaptellvision.com/>

CSE152, Winter 2012

Intro Computer Vision

Object recognition (in mobile phones)

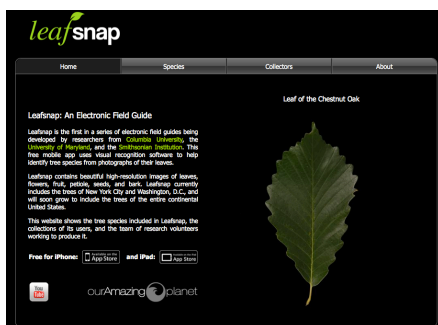


- This is becoming real:
 - [Point & Find](#), [Nokia](#)
 - SnapTell.com (now amazon)
 - Google Goggles

CSE152, Winter 2012

Intro Computer Vision

Leafsnap.com



CSE152, Winter 2012

Intro Computer Vision

Special effects: shape capture



The Matrix movies, ESC Entertainment, XYZRGB, NRC

CSE152, Winter 2012

Intro Computer Vision

Sports



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

CSE152, Winter 2012

Intro Computer Vision

Smart cars



- [Mobileye](http://www.mobileye.com)
 - Vision systems currently in high-end BMW, GM, Volvo models

CSE152, Winter 2012

Slide content courtesy of [Amnon Shashua](http://www.mobileye.com)

Smart cars



- [Mobileye](http://www.mobileye.com)
 - Vision systems currently in high-end BMW

CSE152, Winter 2012

Slide content courtesy of [Amnon Shashua](http://www.mobileye.com)

Vision-based interaction (and games)



Nintendo Wii has camera-based IR tracking built in.



[Digimask](http://www.digimask.com): put your face on a 3D avatar.



Playmotion game a Disney Epic



Xbox Kinect

CSE152, Winter 2012

Intro 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.

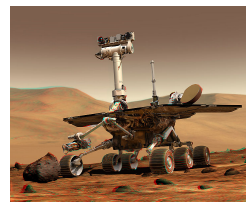
Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "[Computer Vision on Mars](http://www.computer-vision.org)" by Matthies et al.

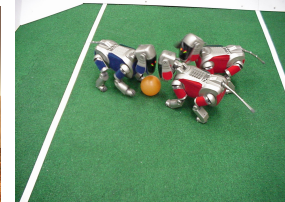
CSE152, Winter 2012

Intro Computer Vision

Robotics



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



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

CSE152, Winter 2012

Intro Computer Vision

Autonomous Vehicles

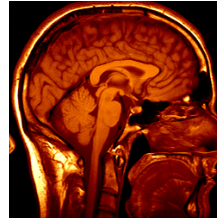


<http://www.youtube.com/watch?v=cdgQp1pUUE>

CSE152, Winter 2012

Intro Computer Vision

Medical imaging



3D imaging
MRI, CT

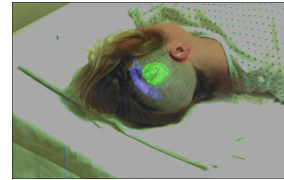
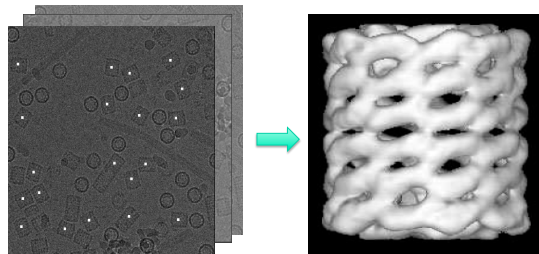


Image guided surgery
[Grimson et al., MIT](#)

CSE152, Winter 2012

Intro Computer Vision

Molecular Reconstruction from Cryo-electron Microscope Images



CSE152, Winter 2012

[Mallick, Zhu, Kriegman]

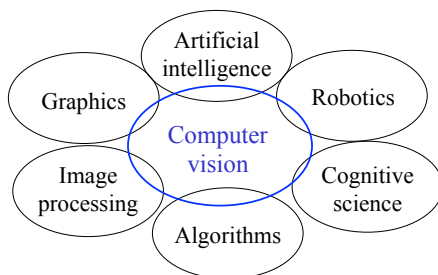
Current state of the art

- 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 apps 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>

CSE152, Winter 2012

Intro Computer Vision

Related Fields



CSE152, Winter 2012

Intro Computer Vision

What is a digital image?

- Formed by a camera (usually).
- Divided into smaller elements (pixels)
- Stored as a matrix of numbers (usually positive)
- 8-bit, 12-bit, 16-bit, float
- Color image – 3 channels



CSE152, Winter 2012

Intro Computer Vision

How do we interpret an image? Use different cues

- Variation in appearance in multiple views
 - stereo
 - motion
- Shading & highlights
- Shadows
- Contours
- Texture
- Blur
- Relationship of objects (context)
- Prior knowledge

CSE152, Winter 2012

Intro Computer Vision

An example of a cue: Shading and lighting

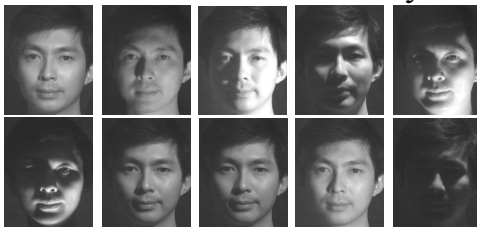
Shading as a result of differences in lighting is

1. A source of information
2. An annoyance

CSE152, Winter 2012

Intro Computer Vision

Illumination Variability



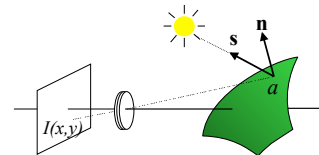
“The variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity.”

-- Moses, Adini, Ullman, ECCV '94

CSE152, Winter 2012

Intro Computer Vision

An example: Image Formation Model



Lambertian Reflectance:

At image location (x,y) the intensity of a pixel $I(x,y)$ is

$$I(x,y) = a(x,y) \mathbf{n}(x,y) \cdot \mathbf{s}$$

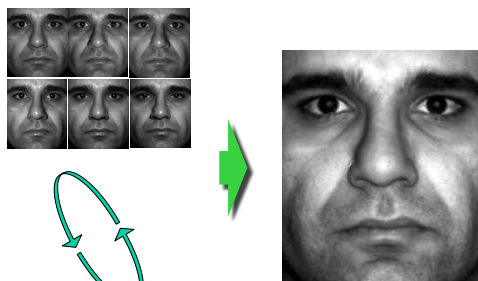
where

- $a(x,y)$ is the albedo of the surface projecting to (x,y) .
- $\mathbf{n}(x,y)$ is the unit surface normal.
- $\mathbf{s}$ is the direction and strength toward the light source.

CSE152, Winter 2012

Intro Computer Vision

An implemented algorithm: Relighting

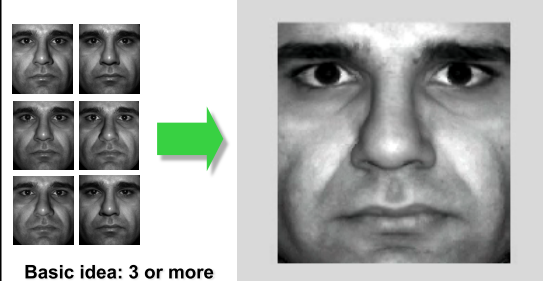


Single Light Source

CSE152, Winter 2012

Intro Computer Vision

3,4: An implemented algorithm Photometric Stereo



Basic idea: 3 or more images under slightly different lighting

CSE152, Winter 2012

Intro Computer Vision

The course

- Part 1: The Physics of Imaging
- Part 2: Early Vision (Segmentation)
- Part 3: Reconstruction (Shape-from-X)
- Part 4: Recognition

CSE152, Winter 2012

Intro Computer Vision

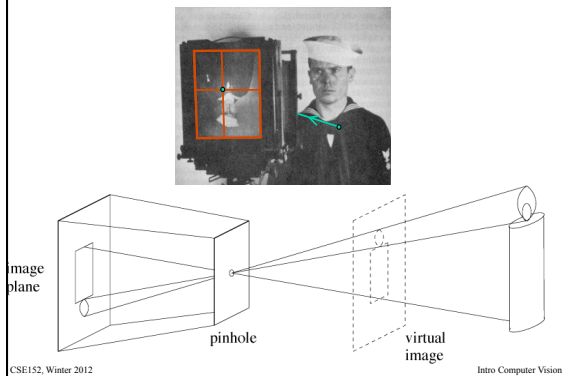
Part I of Course: The Physics of Imaging

- How images are formed
 - Cameras
 - What a camera does
 - How to tell where the camera was located
 - Light
 - How to measure light
 - What light does at surfaces
 - How the brightness values we see in cameras are determined
 - Color
 - The underlying mechanisms of color
 - How to describe it and measure it

CSE152, Winter 2012

Intro Computer Vision

A real camera ... and its model

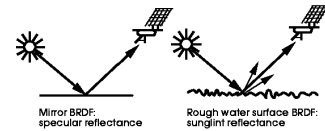


CSE152, Winter 2012

Intro Computer Vision

Lighting & Photometry

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



CSE152, Winter 2012

Intro Computer Vision

Cameras, lenses, and sensors

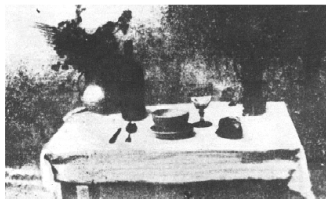


Figure 1.16 The first photograph on record, *la table servie*, obtained by Nicéphore Niépce in 1822. *Collection Harlinge-Viollet*.

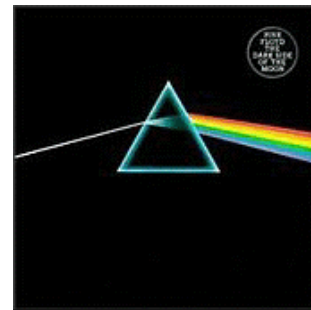
From *Computer Vision*, Forsyth and Ponce, Prentice-Hall, 2002.

CSE152, Winter 2012

Intro Computer Vision

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

Color



CSE152, Winter 2012

Intro Computer Vision

Part II: Early Vision in One Image

- Representing small patches of image
 - For three reasons
 - Sharp changes are important in practice --- known as “edges”
 - Representing texture by giving some statistics of the different kinds of small patch present in the texture.
 - Tigers have lots of bars, few spots
 - Leopards are the other way
 - We wish to establish correspondence between (say) points in different images, so we need to describe the neighborhood of the points

CSE152, Winter 2012

Intro Computer Vision

Segmentation

- Which image components “belong together”?
- Belong together $\equiv$ lie on the same object
- Cues
 - similar color
 - similar texture
 - not separated by contour
 - form a suggestive shape when assembled
- Binary Image Processing

CSE152, Winter 2012

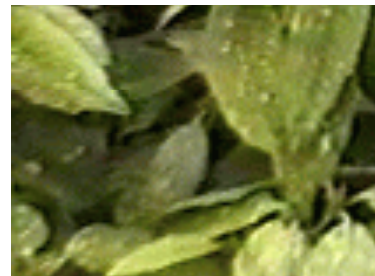
Intro Computer Vision

Boundary Detection



CSE152, Winter 2012 <http://www.robots.ox.ac.uk/~vdg/dynamics.html> Intro Computer Vision

Boundary Detection: Local cues



CSE152, Winter 2012

Intro Computer Vision

Boundary Detection: Local cues



CSE152, Winter 2012

Intro Computer Vision

Gradients



CSE152, Winter 2012

sion

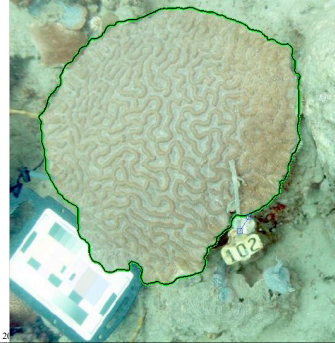
Coral Reef Health



CSE152, Winter 2012

Intro Computer Vision

Segmentation of Coral



CSE152, Winter 2012

Intro Computer Vision

Part 3: Reconstruction from Multiple Images

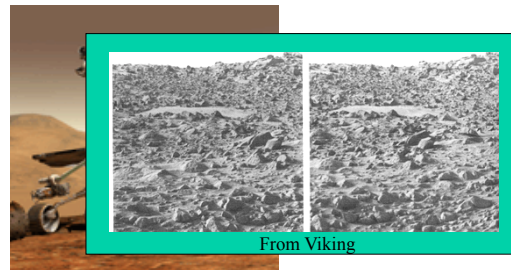
- Photometric Stereo
 - What we know about the world from lighting changes.
 - The geometry of multiple views
 - Stereopsis
 - What we know about the world from having two eyes
 - Structure from motion
 - What we know about the world from having many eyes
- or, more commonly, our eyes moving.

CSE152, Winter 2012

Intro Computer Vision

Mars Rover

Spirit



CSE152, Winter 2012

Intro Computer Vision

Xbox Kinect



CSE152, Winter 2012

Intro Computer Vision

Multi-view geometry applications

- Image stitching, mosaicing
 - <http://www.yosemite-17-gigapixels.com/>
 - 214,414 x 80,571 pixels from 2k images.
 - Pixel is less than 3 arc seconds.
- Photosynth
 - <http://photosynth.net/>
 - In Bing Maps

CSE152, Winter 2012

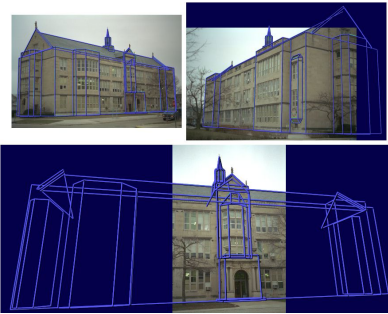
Intro Computer Vision

Images with marked features



CSE152, Winter 2012

Intro Computer Vision

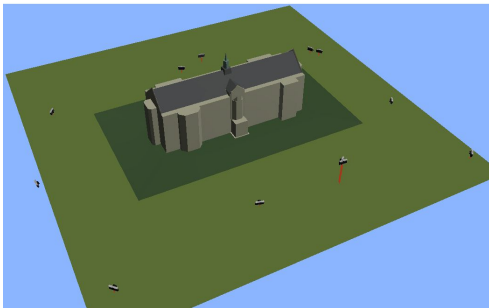


Recovered model edges reprojected through recovered camera positions into the three original images

CSE152, Winter 2012

Intro Computer Vision

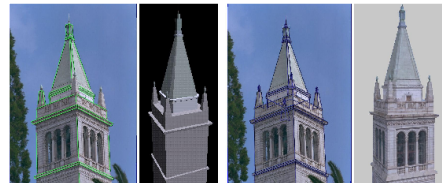
Resulting model & Camera Positions



CSE152, Winter 2012

Intro Computer Vision

Façade (Debevec, Taylor and Malik, 1996)
Reconstruction from multiple views, constraints, rendering



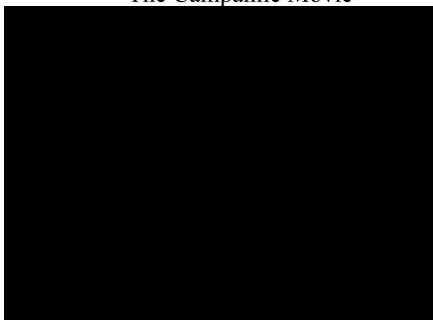
Architectural modeling:
• photogrammetry;
• view-dependent texture mapping;
• model-based stereopsis.

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.

CSE152, Winter 2012

Intro Computer Vision

Façade • The Campanile Movie



CSE152, Winter 2012

Intro Computer Vision

Video-Motion Analysis

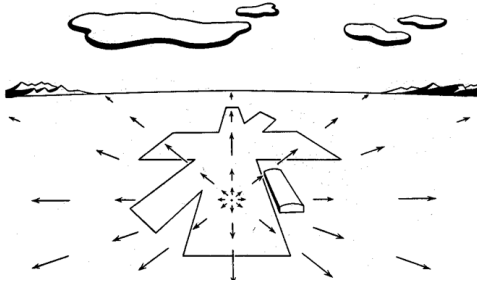


- Where "things" are moving in image – segmentation.
- Determining observer motion (egomotion)
- Determining scene structure
- Tracking objects
- Understanding activities & actions

CSE152, Winter 2012

Intro Computer Vision

Forward Translation & Focus of Expansion [Gibson, 1950]



CSE152, Winter 2012

Intro Computer Vision

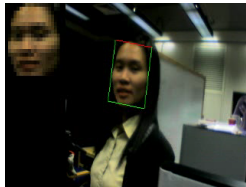
Tracking

- Use a model to predict next position and refine using next image
- Model:
 - simple dynamic models (second order dynamics)
 - kinematic models
 - etc.
- Face tracking and eye tracking now work rather well

CSE152, Winter 2012

Intro Computer Vision

Visual Tracking



Main Challenges

1. 3-D Pose Variation
2. Occlusion of the target
3. Illumination variation
4. Camera jitter
5. Expression variation etc.

[Ho, Lee, Kriegman]

CSE152, Winter 2012

Intro Computer Vision

Tracking



CSE152, Winter 2012

(www.brickstream.com)

Intro Computer Vision

Tracking



CSE152, Winter 2012

Intro Computer Vision

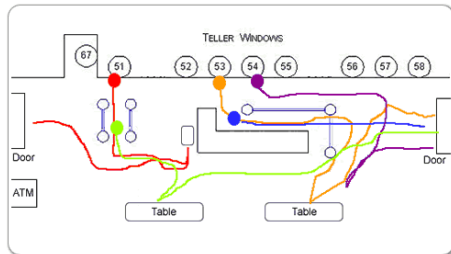
Tracking



CSE152, Winter 2012

Intro Computer Vision

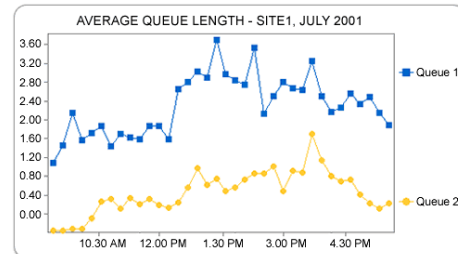
Tracking



CSE152, Winter 2012

Intro Computer Vision

Tracking



CSE152, Winter 2012

Intro 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.

CSE152, Winter 2012

Intro 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

CSE152, Winter 2012

Intro Computer Vision

Face Detection: First Step



- Digital Cameras: Face + Expression
- Street view
- Crowd counting

CSE152, Winter 2012

Intro Computer Vision

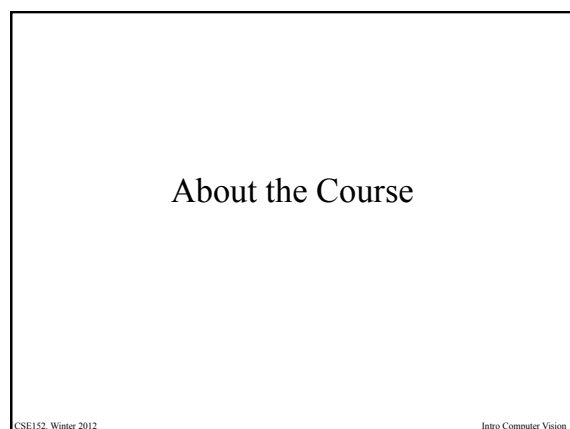
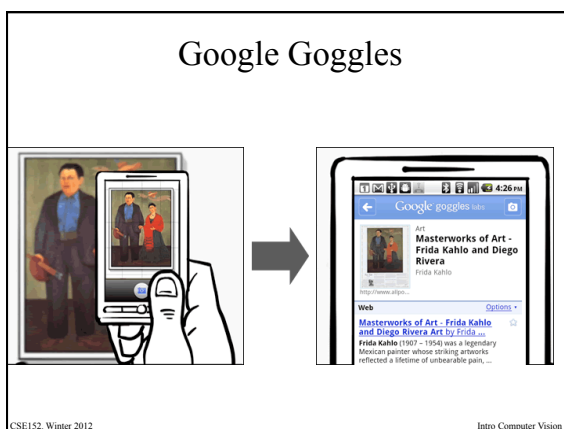
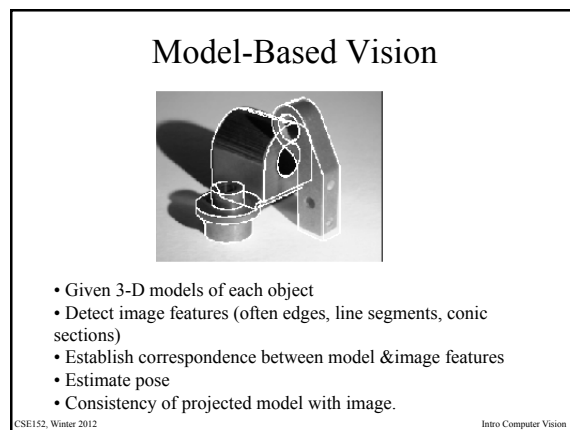
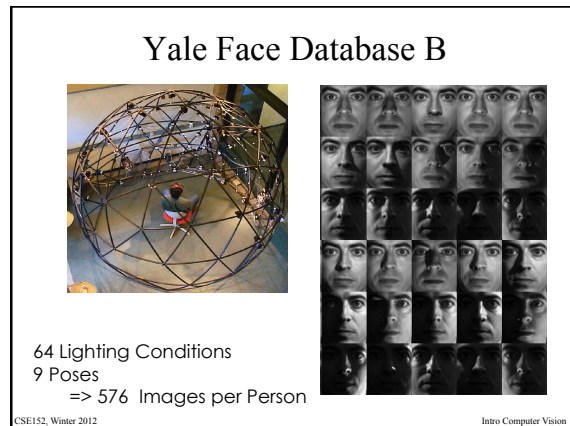
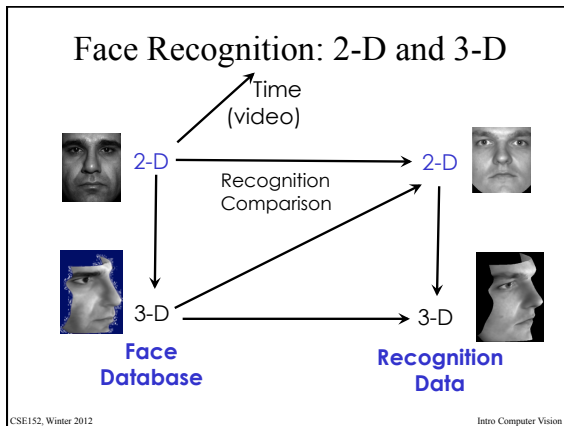
Why is Face Recognition Hard?

Many faces of Madonna



CSE152, Winter 2012

Intro Computer Vision



Primary Text



Available on Ereserves

CSE152, Winter 2012 Intro Computer Vision

Secondary Text

Computer Vision: Algorithms and Applications

© 2010 Richard Szeliski, Microsoft Research



Welcome to the Web site (<http://szeliski.org/Book>) for my computer vision textbook, which you can now purchase at a variety of locations, including [Springer](#) ([SpringerLink](#), [DOD](#), [Amazon](#), and [Barnes & Noble](#)).

This book is largely based on the computer vision courses that I have co-taught at the University of Washington (2008, 2005, 2001) and Stanford (2003) with [Steve Seitz](#) and [David Fleet](#).

You are welcome to download the PDF from this Web site for personal use, but not to repost it on any other Web site. Please post a link to this URL (<http://szeliski.org/Book>) instead. An electronic version of this manuscript will continue to be available even after the book is published. Note, however, that while the content of the electronic and hardcopy versions are the same, the page layout (pagination) is different, since the electronic version is designed for a different format.

CSE152, Winter 2012 Intro Computer Vision

The Syllabus

CSE152, Winter 2012 Intro Computer Vision

Announcements

- Class Web Page:
<http://cseweb.ucsd.edu/classes/wi13/cse152-a/>
- Assignment 0: “**Getting Started with Matlab**” (to be posted soon)
- Read Chapters 1 Trucco & Verri

CSE152, Winter 2012 Intro Computer Vision