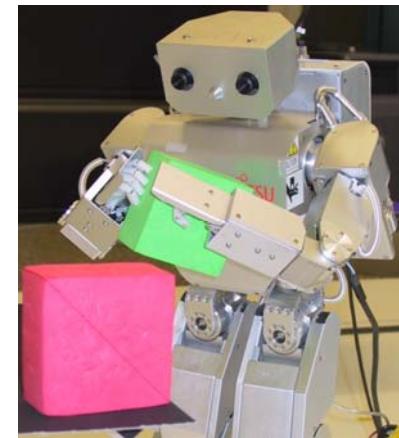


CSE 455

Computer Vision



Rajesh Rao (Instructor)
Jiun-Hung Chen (TA)

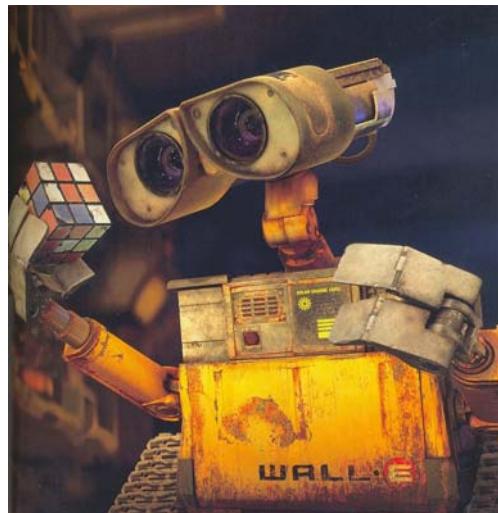
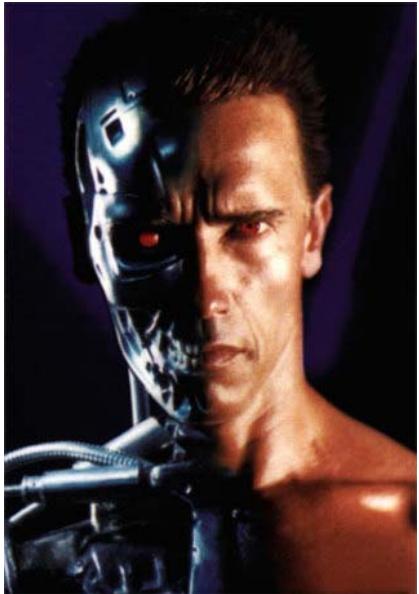
<http://www.cs.washington.edu/455>

What's on our plate today?

- What is computer vision?
- Examples of current state-of-the-art
- Goals of the course
- Logistics
- Intro to Images & Image Processing

What is computer vision?

Computer
vision
according to
Hollywood



What is computer vision?

Making useful decisions about real physical objects and scenes based on images (Shapiro & Stockman, 2001)

Extracting descriptions of the world from pictures or sequences of pictures (Forsyth & Ponce, 2003)

Analyzing images and producing descriptions that can be used to interact with the environment (Horn, 1986)

Designing representations and algorithms for relating images to models of the world (Ballard & Brown, 1982)

A picture is worth a thousand words



Can a computer infer what happened from the image?

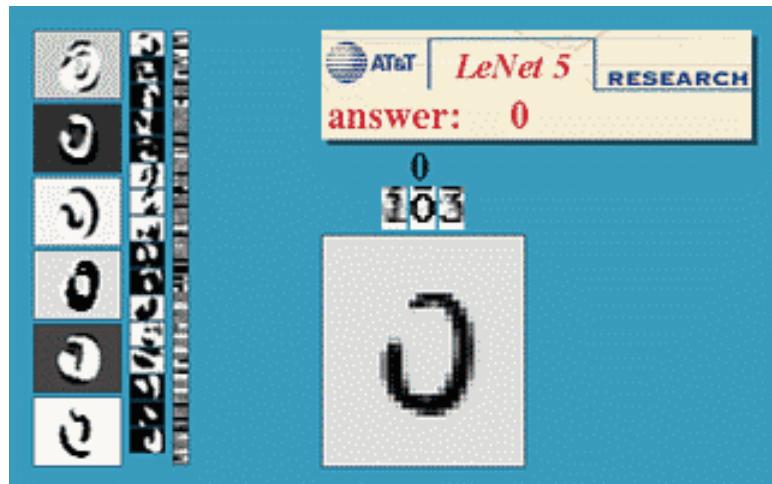
Computer Vision: Current State of the Art

The next few slides show examples of what current computer vision systems can do...

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/~yann/>

4 YCH428

4 YCH428

4 YCH428

License plate readers

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

Face Detection

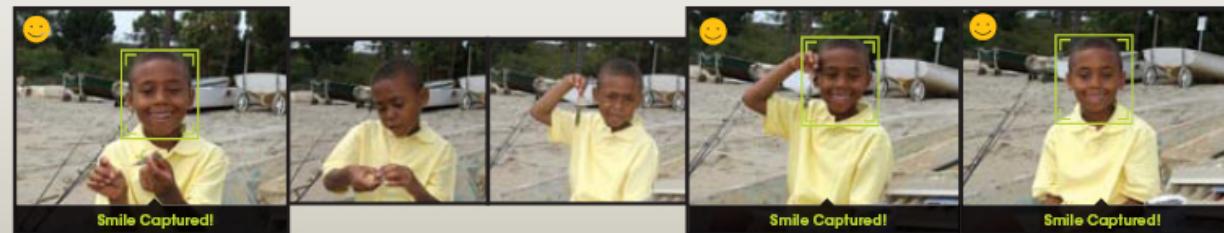
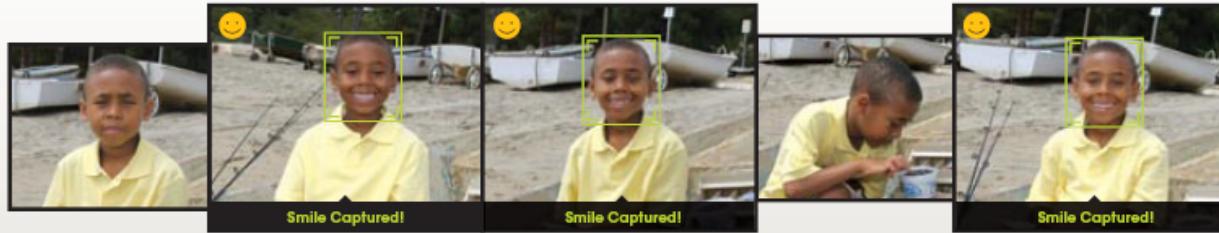


Most new digital cameras now detect faces
(sometimes badly)

Smile Detection (automatically clicks when you smile!)

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](#)

Some
unhappy
customers



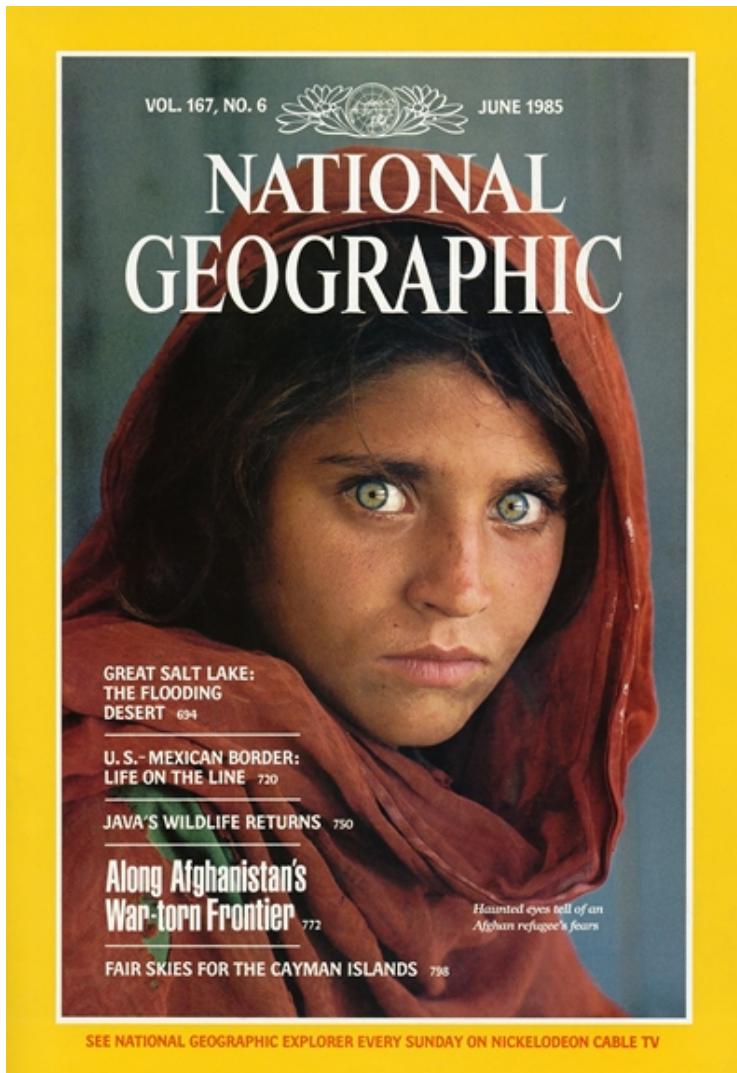
Object Recognition (in supermarkets)



LaneHawk by EvolutionRobotics

"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..."

Vision-Based Biometrics



Sharbat Gula at age 12 in an Afghan refugee camp in 1984

Traced in 2002 but is she the same person?

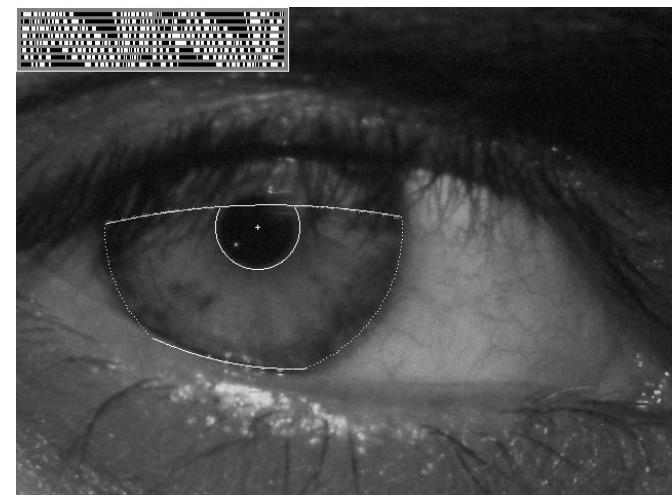
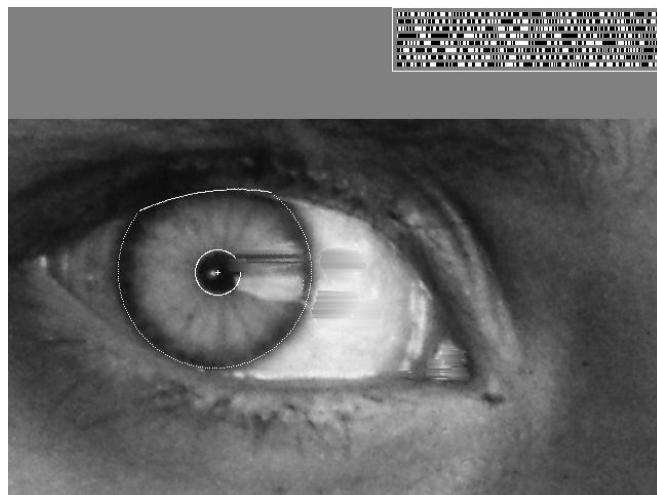
Identity verification through Iris code

1984



2002

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



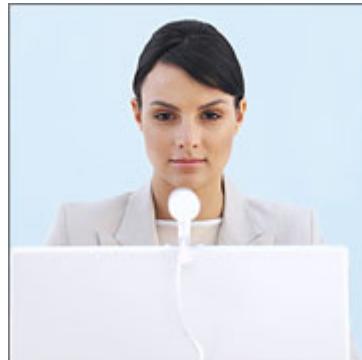
Login with your fingerprint or face



<http://www.xmicro.com>

Could be a problem if
your face changes often

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



Object recognition (in mobile phones)



This is becoming real:

- [Lincoln Microsoft Research: Mobile web search via pictures](#)
- [Nokia's Point & Find](#)

3D modeling: Earth viewers

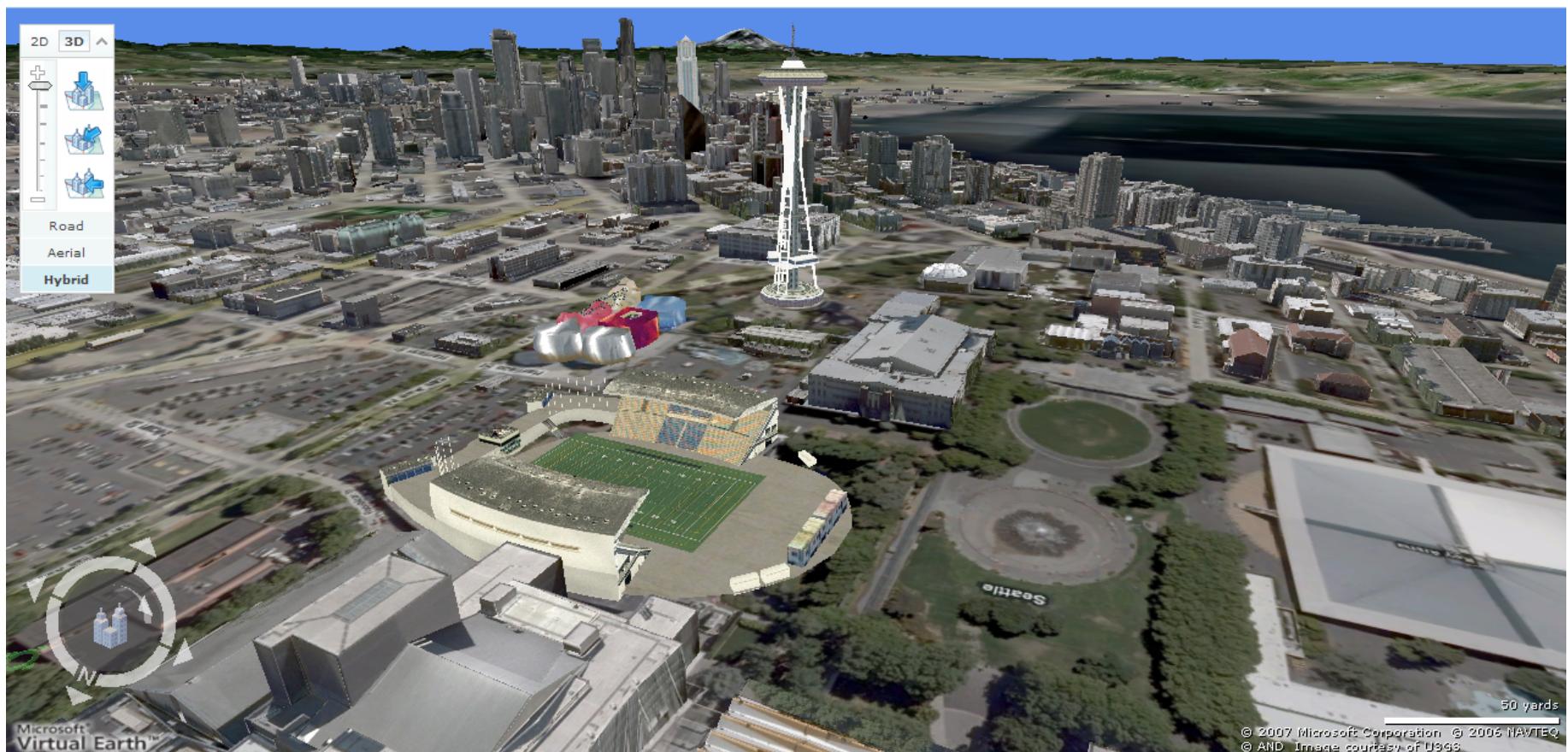


Image from Microsoft's [Virtual Earth](#)
(see also: [Google Earth](#))



- [Home](#)
- [Try it](#)
- [What is Photosynth?](#)
- [Collections](#)
- [Team blog](#)
- [Videos](#)
- [System requirements](#)
- [About us](#)
- [FAQ](#)

*"What if your photo collection was an entry point into the world,
like a wormhole that you could jump through and explore..."*

[Try it](#)



[Try the Tech Preview](#)

The **Photosynth Technology Preview** is a taste of the newest - and, we hope, most exciting - way to [view photos](#) on a computer. Our software takes a large collection of photos of a place or an object, analyzes them for similarities, and then displays the photos in a reconstructed **three-dimensional space**, showing you how each one relates to the next.

<http://photosynth.net>

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

Special effects: shape capture



The Burly Brawl scene in *The Matrix Reloaded*



Special effects: motion capture



Pirates of the Caribbean, Industrial Light and Magic
[Click here for interactive demo](#)

Sports (<http://www.sportvision.com>)



Virtual first down line
([explanation](#) on www.howstuffworks.com)



Real-time strike zone box



Ball tracking



Virtual Ads!

Smart cars

Slide content courtesy of Amnon Shashua

▶▶ manufacturer products consumer products ◀◀

Our Vision. Your Safety.

rear looking camera forward looking camera side looking camera

> EyeQ Vision on a Chip

[> read more](#)

> Vision Applications

Road, Vehicle, Pedestrian Protection and more

[> read more](#)

> AWS Advance Warning System

[> read more](#)

News

> Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System

> Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end

[> all news](#)

Events

> [Mobileye at Equip Auto, Paris, France](#)

> [Mobileye at SEMA, Las Vegas, NV](#)

[> read more](#)

Mobileye

- Vision systems currently in high-end BMW, GM, Volvo models
- By 2010: 70% of car manufacturers

Vision-based interaction and games

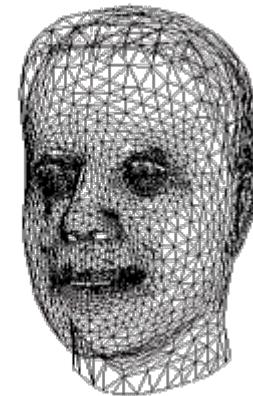


Nintendo Wii has camera-based IR tracking built in. See [Lee's work at CMU](#) on clever tricks on using it to create a [multi-touch display!](#)



["Game turns moviegoers into Human Joysticks"](#), CNET
Camera tracking a crowd, based on [this work](#).

[Digimask](#): put your face on a 3D avatar



Computer 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.

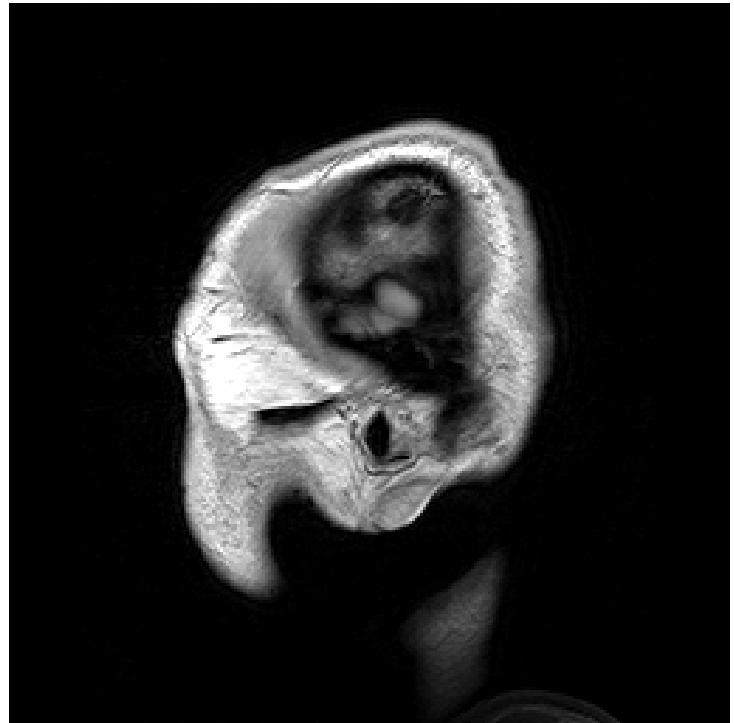
Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “[Computer Vision on Mars](#)” by Matthies et al.

Medical imaging



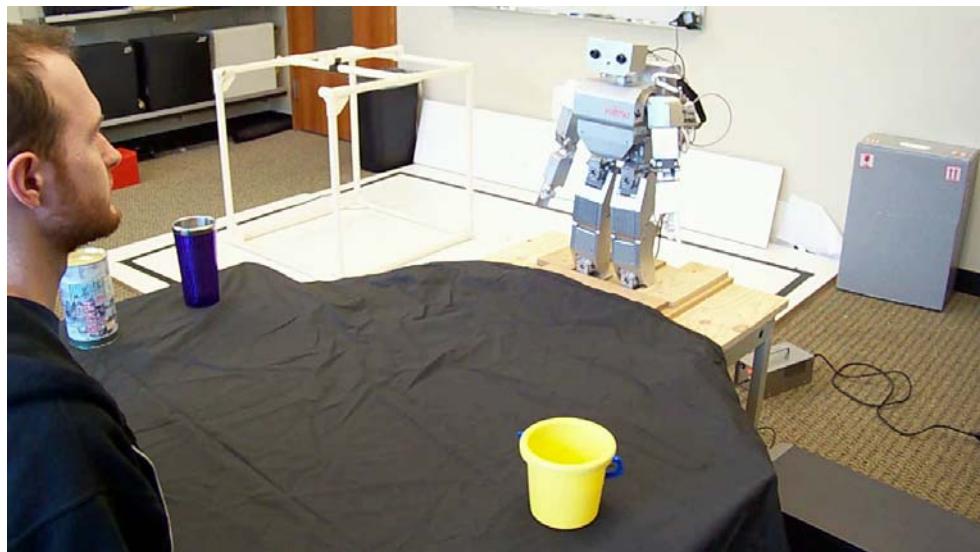
Image guided surgery
[Grimson et al., MIT](#)



3D imaging
MRI

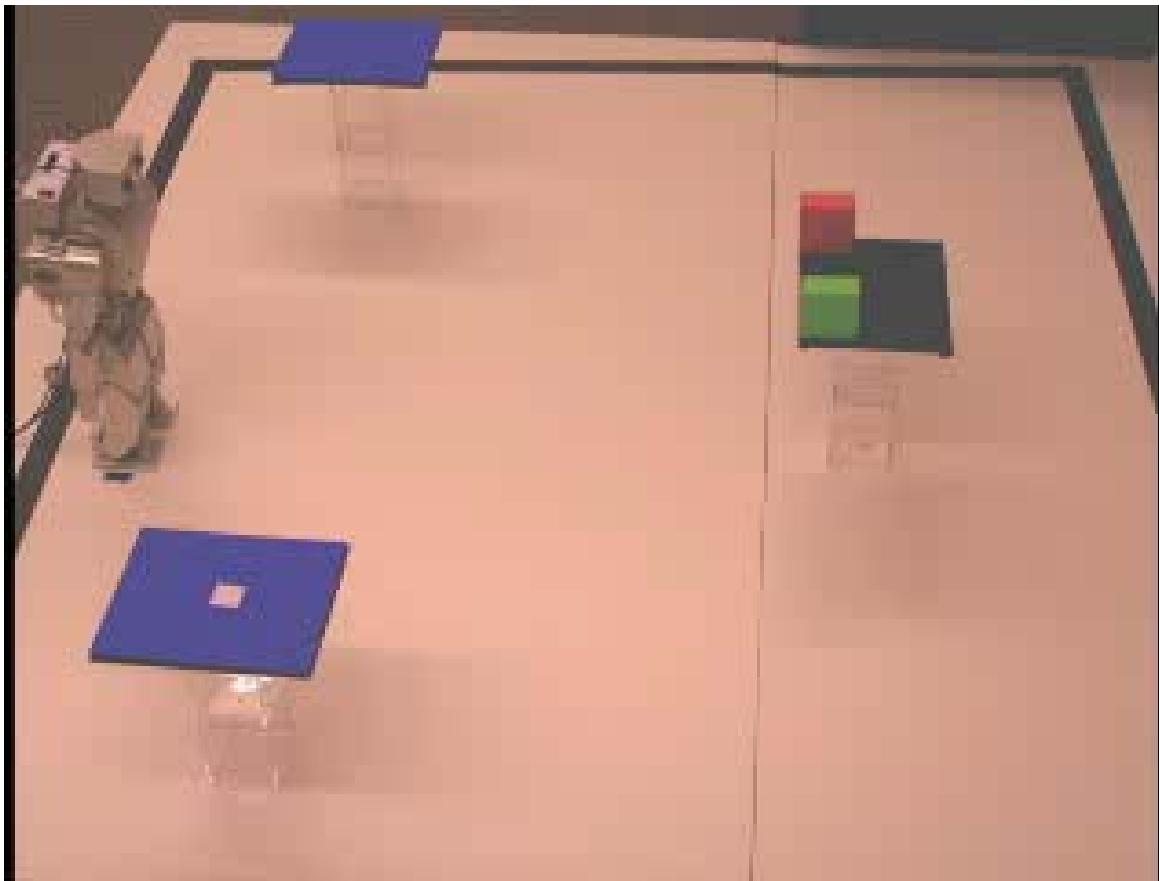
Vision-Based Robotic Learning of Language

Robot learns names for new objects through gaze following



Research done by UW CSE student Aaron Shon

Vision-Guided Brain-Robot Interfaces



[CBS News Article](#)

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

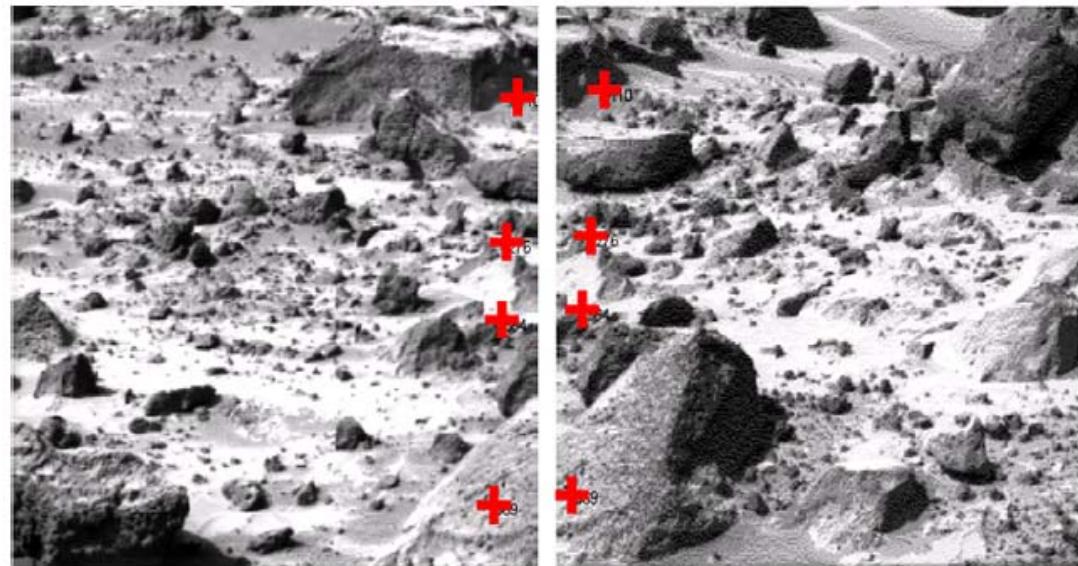
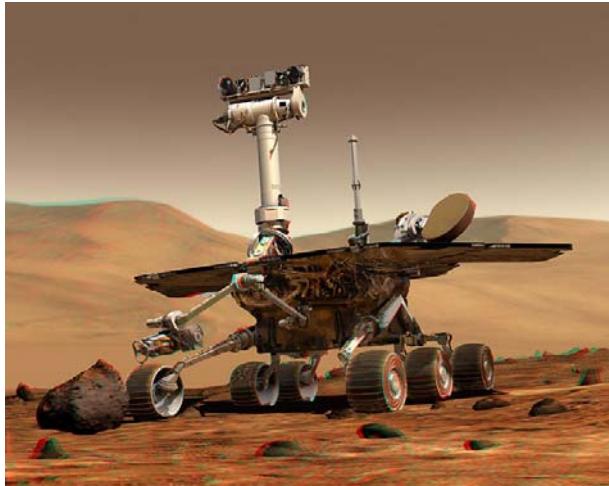
- [David Lowe](#) maintains an excellent overview of vision companies
 - <http://www.cs.ubc.ca/spider/lowe/vision.html>

Goals of the course

- Provide an introduction to computer vision
- Topics to be covered:
 - Image processing and feature detection
 - Image stitching and mosaicing
 - Human vision
 - Pattern recognition & visual learning
 - Object recognition & Image segmentation
 - Motion estimation, color & texture
 - Stereo & 3D vision
 - Applications: content-based image retrieval, tactile graphics, computer vision for Mars exploration

Invited guest lectures

- Jan 29: Prof. Clark Olson
(UW Bothell) on
“Computer vision for
Mars exploration”



Invited guest lectures

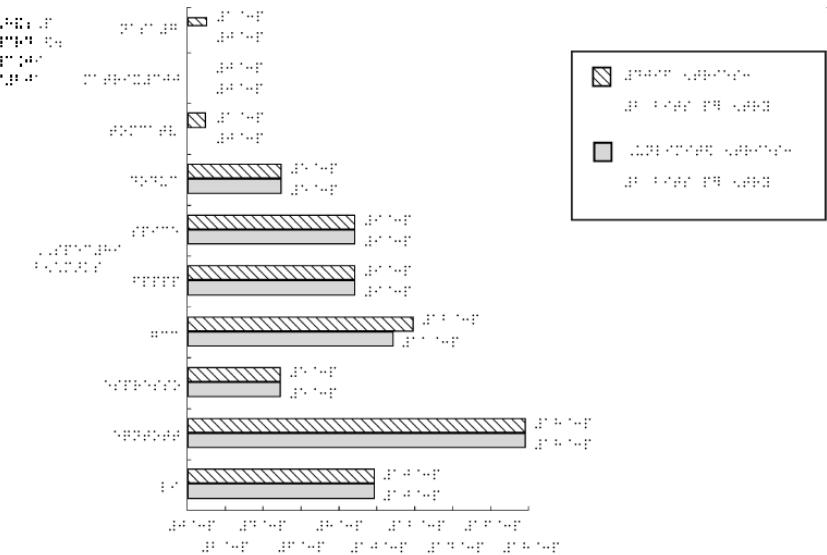
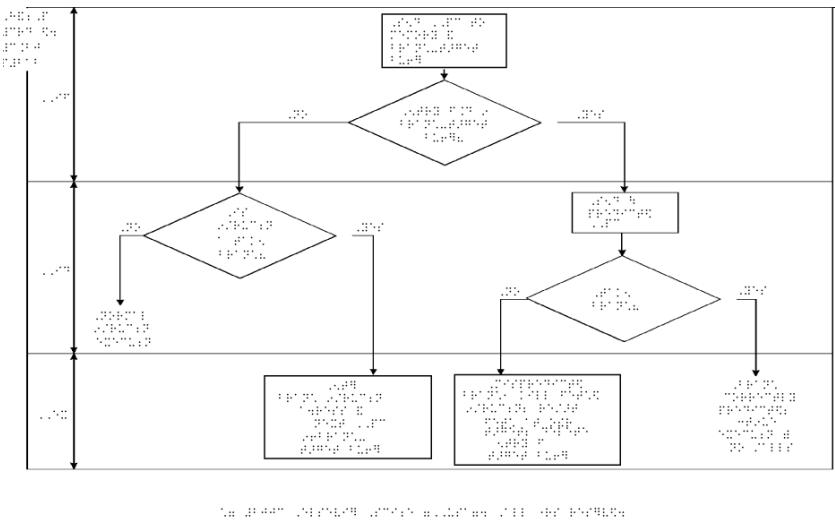
- Feb 19: Prof. Linda Shapiro
(UW Seattle) on
“Content-Based Image
Retrieval”



A screenshot of a Content-Based Image Retrieval application. At the top left, there is a small thumbnail of a cherry blossom tree with a red rectangular box drawn around its upper portion. To the right of this are three larger, more detailed versions of the same cherry blossom tree from different angles. Below these are two rows of three more images each, all showing various types of trees, likely cherry blossoms, in different stages of bloom or from different perspectives. At the bottom of the interface, there are several control buttons: a left arrow labeled "Random", a "Go" button, a "ZoomIn" button with a magnifying glass icon, and a right arrow. To the right of these buttons, the text "Found 24 matches. Displaying 1 - 6" is displayed.

Invited guest lectures

- Mar 5: Prof. Richard Ladner
(UW Seattle) on
“Tactile Graphics”



Tactile versions (with Braille) of graphical images in *Computer Architecture: A Quantitative Approach* by Hennessy and Patterson.

Projects

1. Image scissors



2. Image stitching

3. Content-based image retrieval

4. Face recognition & detection



Project 1: intelligent scissors



David Dewey, 455 02wi

Project 2: panorama stitching



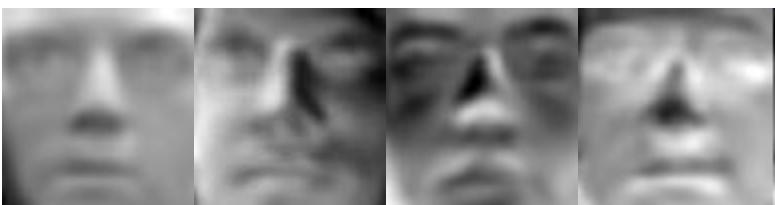
Oscar Danielsson, 455 06wi

Project 3: Content-Based Image Retrieval



Project 4: Face Recognition & Face Detection

Eigenfaces



Recognition



Detection

Grading

Programming Projects (80%)

- Image scissors (20%)
- Panoramas (20%)
- Content-based image retrieval (20%)
- Face recognition & detection (20%)

Final (20%)

Prerequisites

The following are essential!

- Data structures
- A good working knowledge of C and C++ programming
 - (or willingness/time to pick it up quickly!)
- Linear algebra
- Vector calculus

Course does **not** assume prior imaging experience

- computer vision, image processing, graphics, etc.

Okay, let's begin

What is an image?

What is an image?

Think of an **image** as a function, f , from \mathbf{R}^2 to \mathbf{R} :

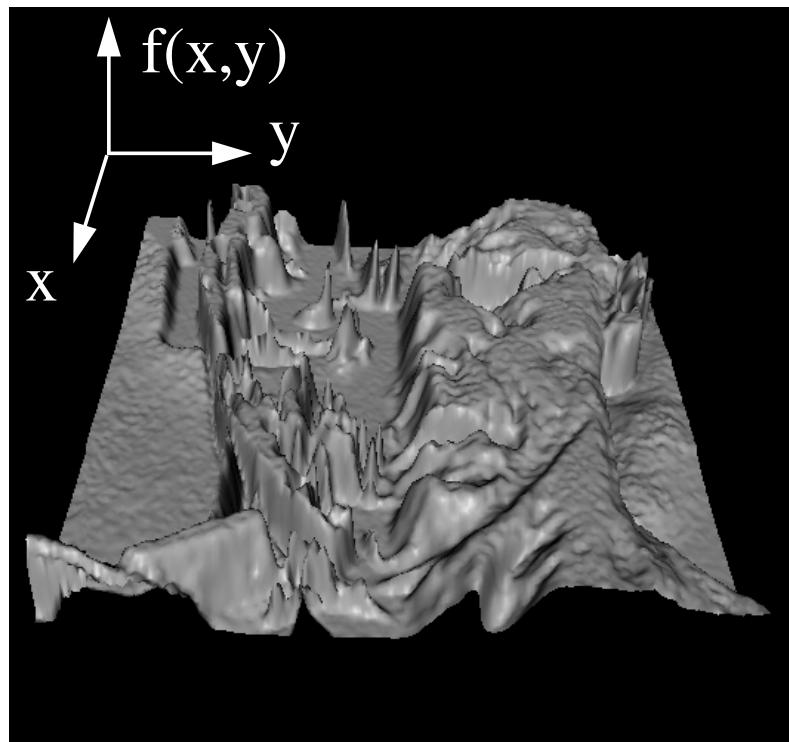
- $f(x, y)$ gives the **intensity** at position (x, y)
- Realistically, images defined over a rectangle:

$$f: [a,b] \times [c,d] \rightarrow [0,1]$$

Color image = three functions pasted together

$$f(x, y) = \begin{bmatrix} r(x, y) \\ g(x, y) \\ b(x, y) \end{bmatrix}$$

An image as a function



Bright regions are high, dark regions are low

Digital images

In computer vision we usually operate on **digital (discrete)** images:

- **Sample** the 2D space on a regular grid
- **Quantize** each sample (round to nearest integer)
- Each sample is a “**pixel**” (picture element)
- If 1 byte for each pixel, values range from 0 to 255



\xrightarrow{y}
 $\downarrow \quad x$

62	79	23	119	120	105	4	0
10	10	9	62	12	78	34	0
10	58	197	46	46	0	0	48
176	135	5	188	191	68	0	49
2	1	1	29	26	37	0	77
0	89	144	147	187	102	62	208
255	252	0	166	123	62	0	31
166	63	127	17	1	0	99	30

Image processing

An **image processing** operation converts an existing image f to a new image g

Can transform either the domain or range of f

Image processing

Range transformation: $g(x, y) = t(f(x, y))$
(What is an example?)



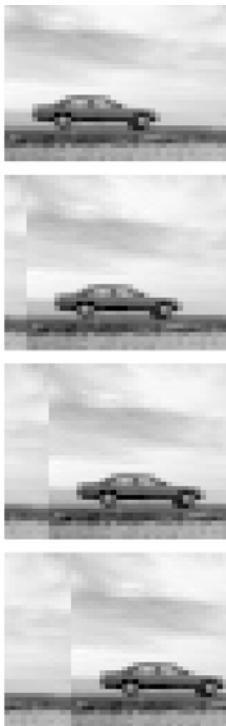
Noise filtering



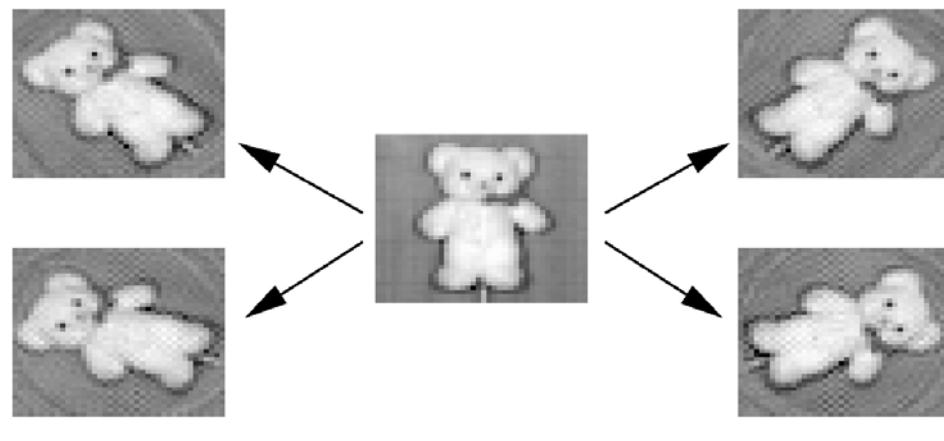
Image Processing

Domain transformation: $g(x, y) = f(t_x(x, y), t_y(x, y))$
(What is an example?)

Translation



Rotation



Next Time: Image Processing and Filtering

- Things to do:
 - Read Chap 2 & Chap 5: Sec. 5.1-5.5, 5.10
 - Browse class website
 - Mailing list: cse455@cs.washington.edu
 - Did you receive the welcome message? Otherwise, sign up
 - Brush up on C/C++ programming skills
 - Visit Vision and Graphics Lab (Sieg 327)
 - Your ID card should open Sieg 327
 - Check to make sure ASAP

