

# CS6550 Computer Vision

- **Class Meeting:**  
M7M8 (3:30pm – 5:20pm), R6 (2:20pm – 3:10pm).  
Delta 106 (M7M8), 台達館106室  
Delta 102 (R6), 台達館102室
- **Instructor:** Prof. Shang-Hong Lai, Delta 636,  
賴尚宏, 台達館636室,  
Tel: ext. 42958, Email: [lai@cs.nthu.edu.tw](mailto:lai@cs.nthu.edu.tw),  
URL: <http://www.cs.nthu.edu.tw/~lai>
- **Office Hours:** M9F2 or by appointment

# Prerequisite

- Linear Algebra
- Probability and Statistics
- Basic Programming

羅大釗

# Course Description

- This course is to provide an introductory background in computer vision for graduate students to start research in this field.
- We will focus on teaching representative computer vision algorithms in class.
- You will need to implement some algorithms with computer programs for the homeworks and the final project.

# Course Contents

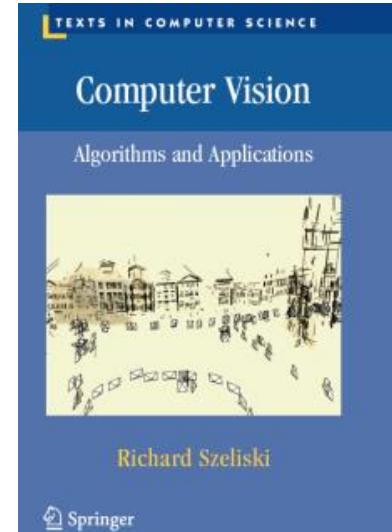
1. Image Formation (1 week)
2. Image Features (2 weeks)
3. Camera Calibration (1 week)
4. Two-View Geometry (1 week)
5. Image Segmentation (2 weeks)
6. Stereo Reconstruction (1 week)
7. Image Matching (1 week)
8. Object Recognition (1 week)
9. Deep Learning (2 weeks)
10. Final Project Presentation

# Textbooks

- *Primary:*

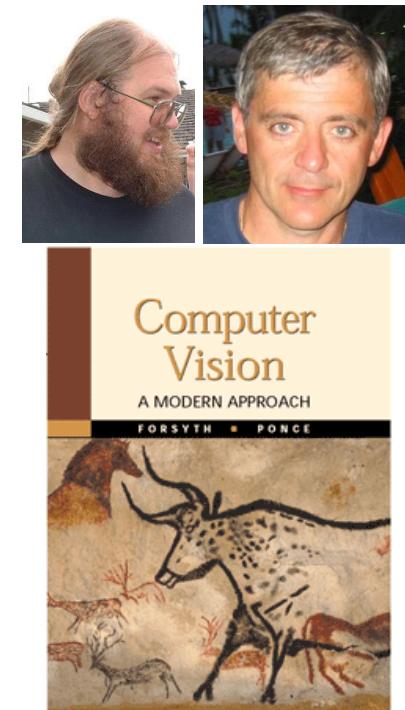
Computer Vision: Algorithms and Applications,  
by Richard Szeliski, Springer-Verlag, 2011.

<http://szeliski.org/Book/>



- *Secondary:*

- *Computer Vision: A Modern Approach*, by D. Forsyth and J. Ponce, 2<sup>nd</sup> Edition, Prentice Hall, 2012.
- *Image Processing, Analysis, and Machine Vision*, by M. Sonka, V. Hlavac, R. Boyle, Thomson Engineering, 3rd Edition, 2007.

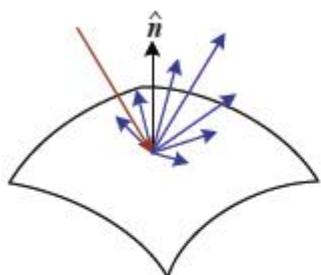


- Lecture slides distributed in class.



Richard Szeliski

Springer



2. Image Formation



3. Image Processing



4. Features



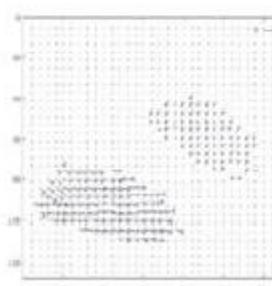
11. Stereo



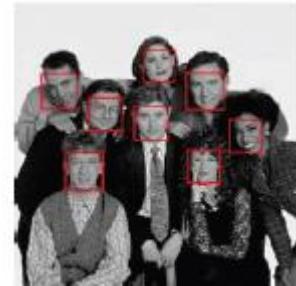
5. Segmentation



6-7. Structure from Motion



8. Motion



14. Recognition

# Grading

Midterm Exam. (11/21)	25%
Final Project	20%
Homeworks (4)	40%
Class Participation(Questions)	5%
Class Attendance	5%
Quizzes	5%

# Homework Policy

- Discussion of homework is encouraged, but you have to write your own. **No copying** is strictly enforced.
- Homework should be delivered before the announced due time, normally before the lecture. Late homework will be degraded by **20%** per day.
- No make-up homeworks 5 days after the deadline.

# iLMS Learning System

- <http://lms.nthu.edu.tw/>
- iLMS contains the course slides, basic course information, and class announcement.
- Course webpage: <http://cv.cs.nthu.edu.tw/courses.php>

# Class Participation

- Class attendance is required and treated as the basic requirement for class participation. (5%)
- Asking questions is strongly encouraged. (5%)

# CS 6550 Classroom Rule

- No eating is permitted in class.
- No sleeping is allowed in class.
- Disturbance to others in class should be minimized.
- Cell phone should be turned off during the class.

# Computer Vision

Make computers understand images and video.



What kind of scene?

Where are the cars?

How far is the  
building?

...

# What is Computer Vision?

- To extract useful information about real physical objects and scenes from sensed images/video.
  - 3D reconstruction from images
  - Object detection/recognition
- Automatic understanding of images and video
  - Computing properties of the 3D world from visual data (*measurement*)
  - Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. (*perception and interpretation*)

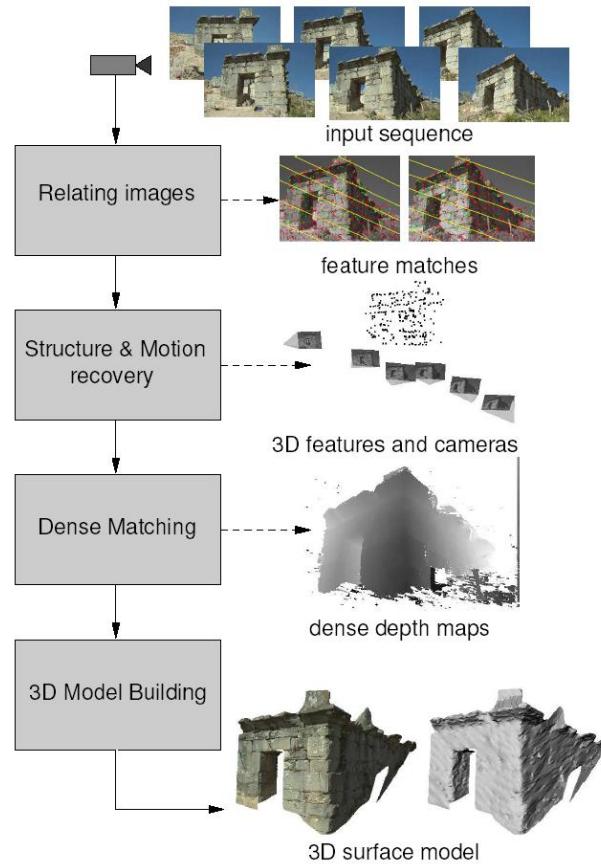
# Vision for measurement

## Real-time stereo



Pollefeys et al.

## Structure from motion

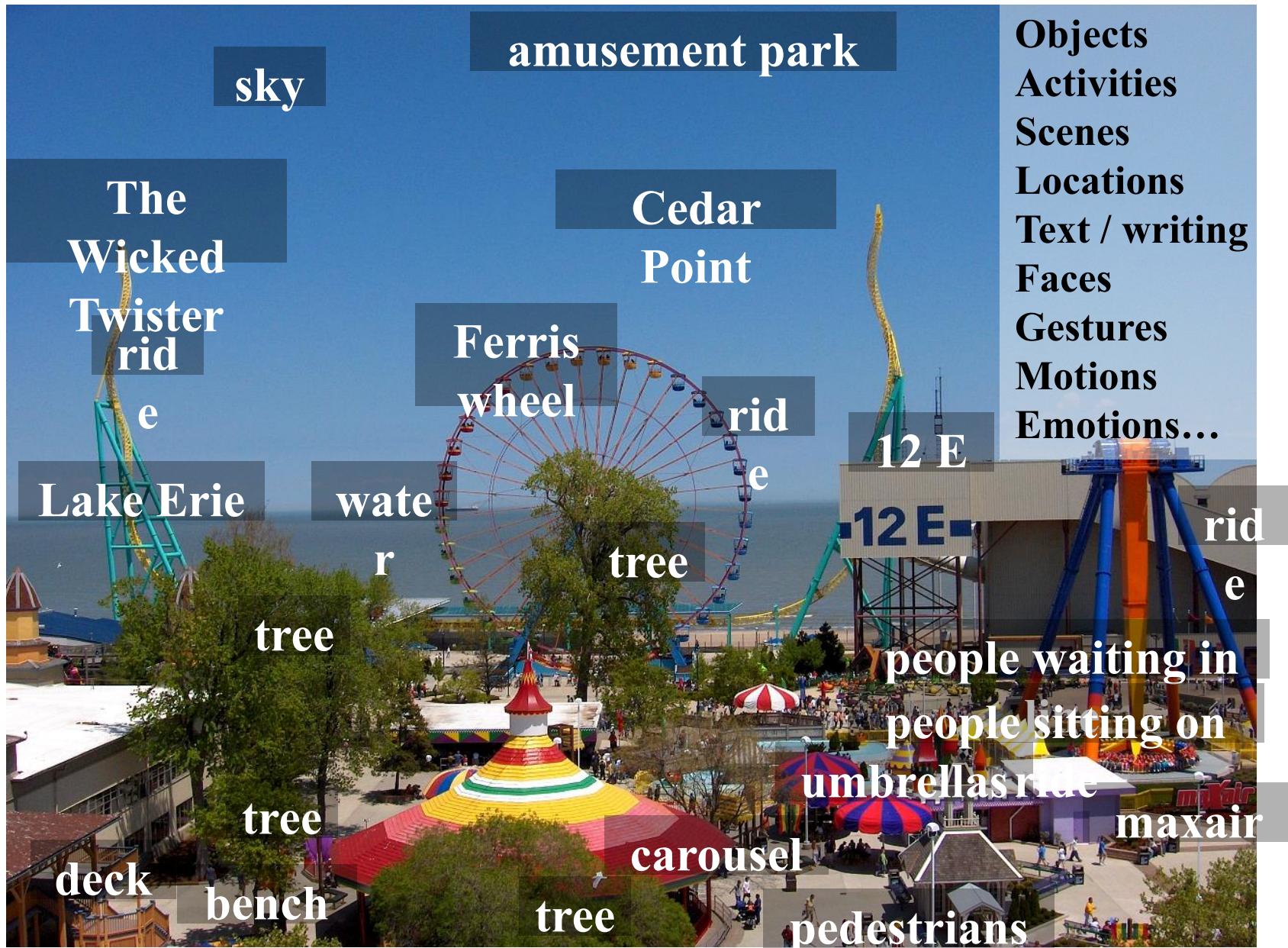


## Multi-view stereo for community photo collections

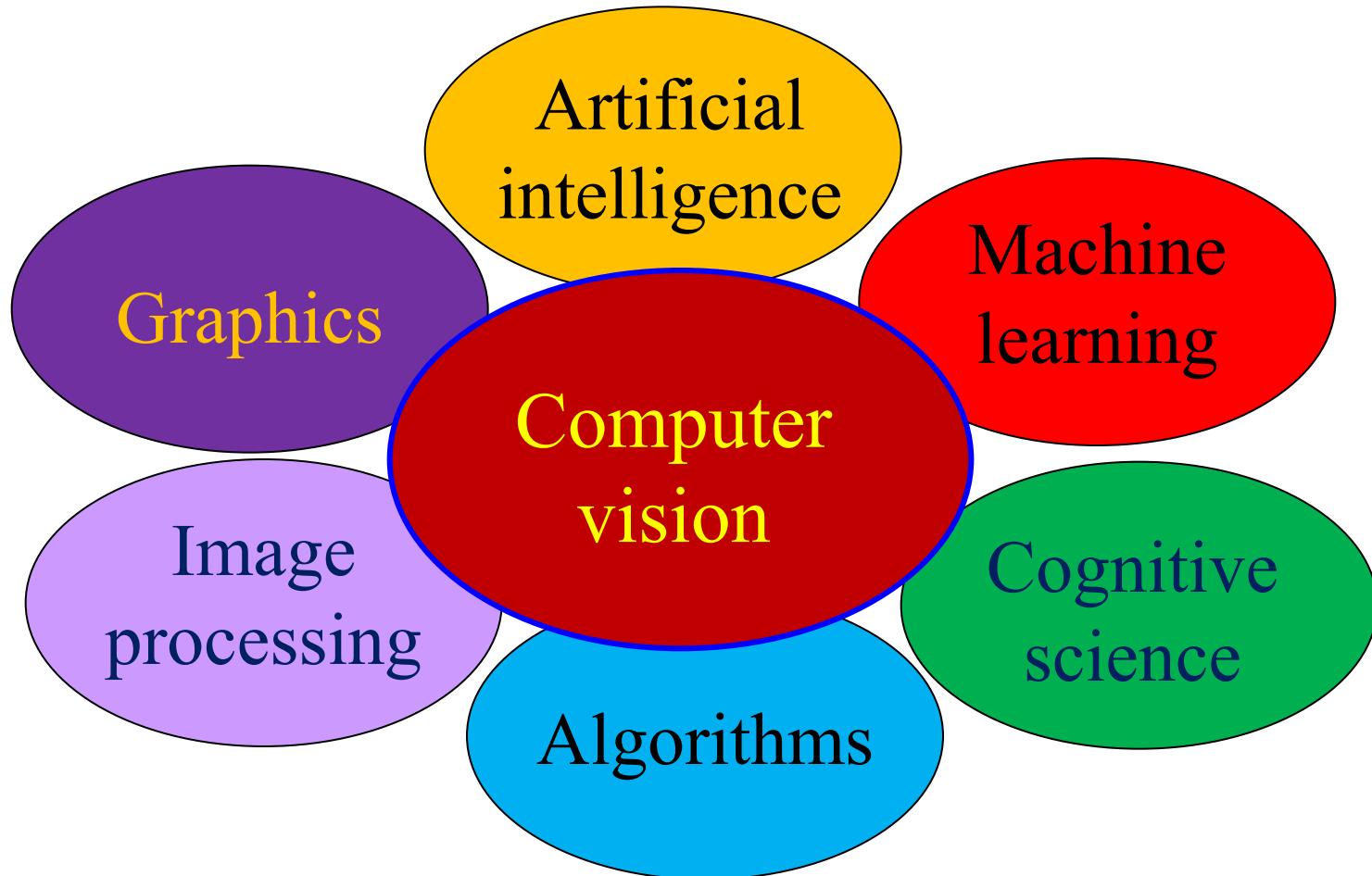


Goesele et al.

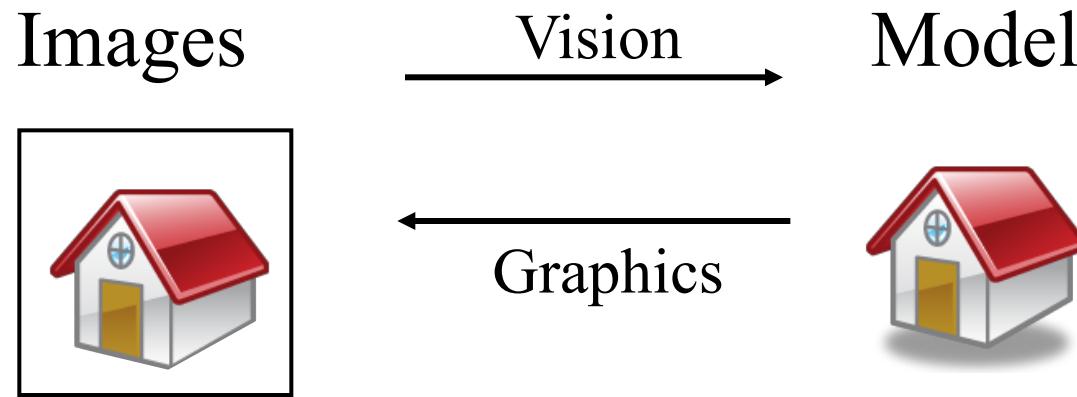
# Vision for perception, interpretation



# Related Disciplines



# Vision and Graphics



Inverse problems: analysis and synthesis.

# Why computer vision?

- As image sources multiply, so do applications
  - Relieve humans of boring, easy tasks
  - Enhance human abilities: human-computer interaction, visualization
  - Perception for robotics / autonomous agents
  - Organize and give access to visual content

# Why computer vision?

- Images and videos are everywhere!



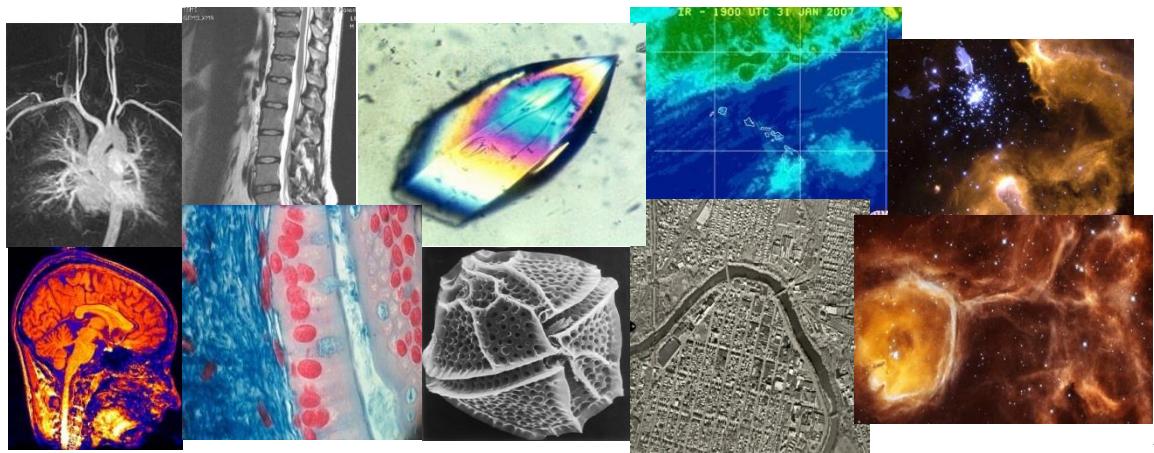
Personal photo albums



Movies, news, sports



Surveillance and security

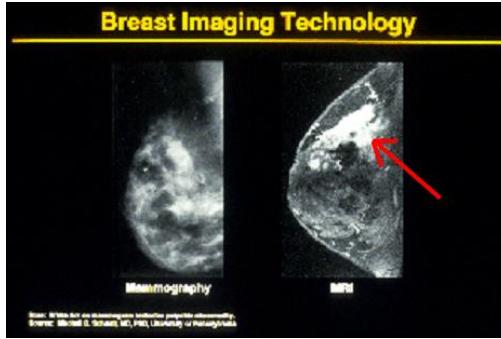


Medical and scientific images

# Why computer vision matters?



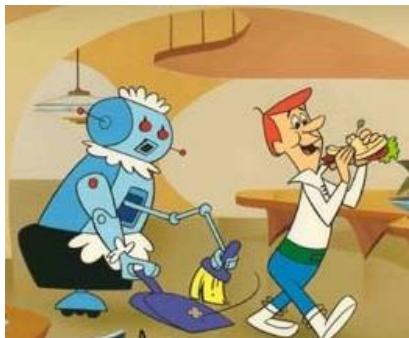
Safety



Health



Security



Comfort



Fun



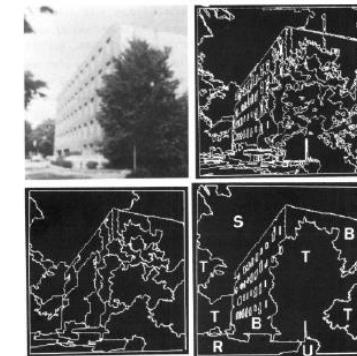
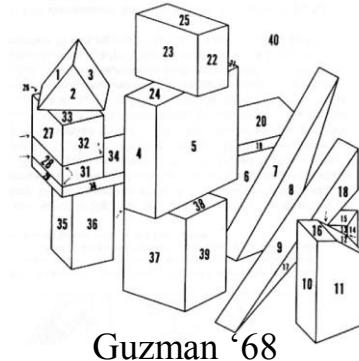
Access

# Again, what is computer vision?

- Mathematics of geometry of image formation?
- Statistics of the natural world?
- Models for neuroscience?
- Engineering methods for matching images?
- Science Fiction?

# Very brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; computational photography starts



Ohta Kanade '78



Turk and Pentland '91

# Applications of Computer Vision

- Robot Vision / Autonomous Vehicles
- Biometric Identification / Recognition
- Industrial Inspection
- Video Surveillance
- Digital Camera
- Medical Image Analysis/Processing
- Remote Sensing
- Multimedia Retrieval
- Augmented Reality
- .....

# Consumer Applications



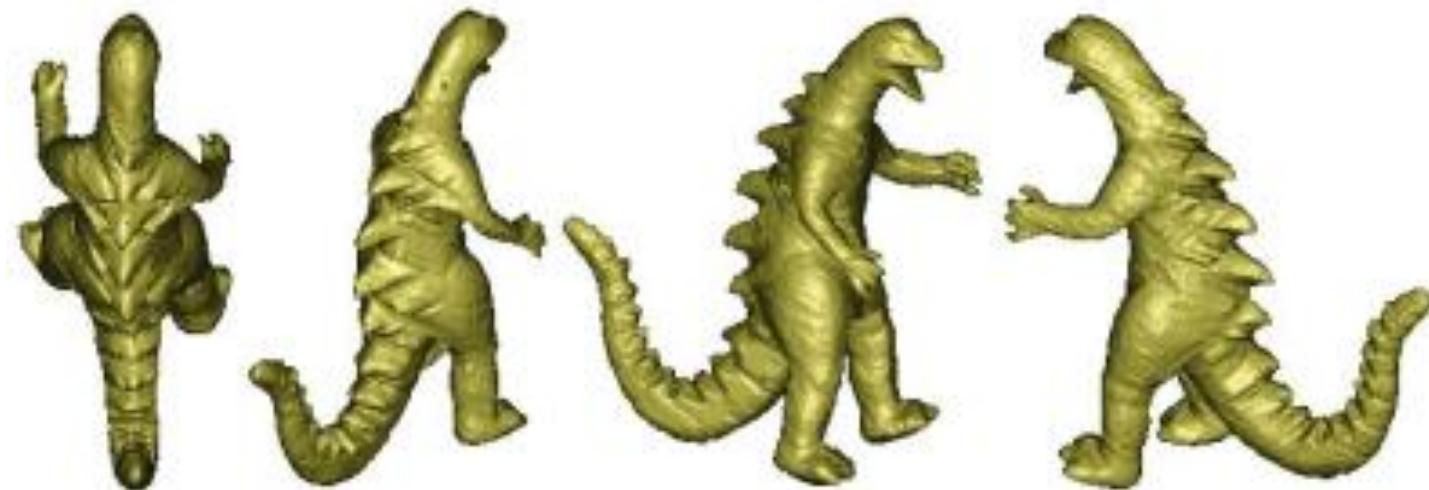
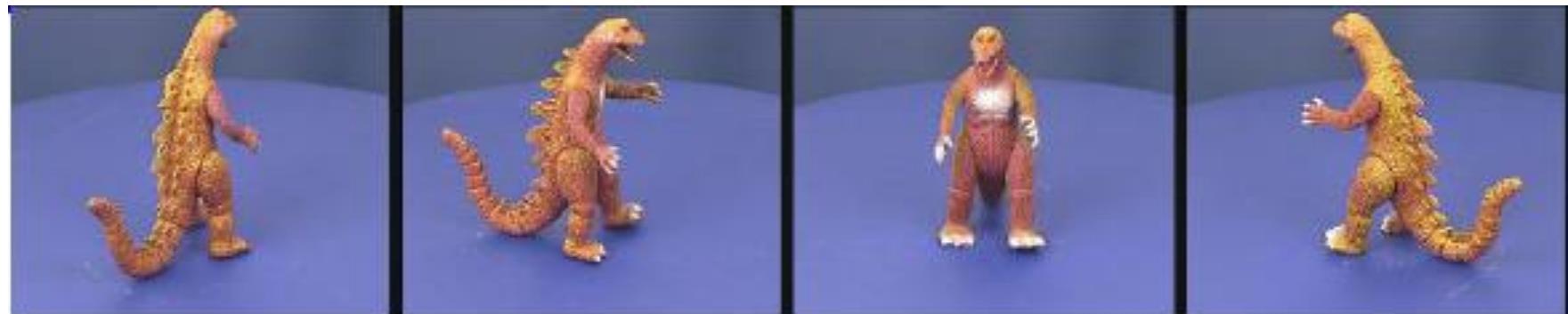
(a)



(b)

(a) image stitching: merging different views (Szeliski and Shum 1997) (b) exposure bracketing: merging different exposures.

# 3D Reconstruction from Images



# Earth viewers (3D modeling)

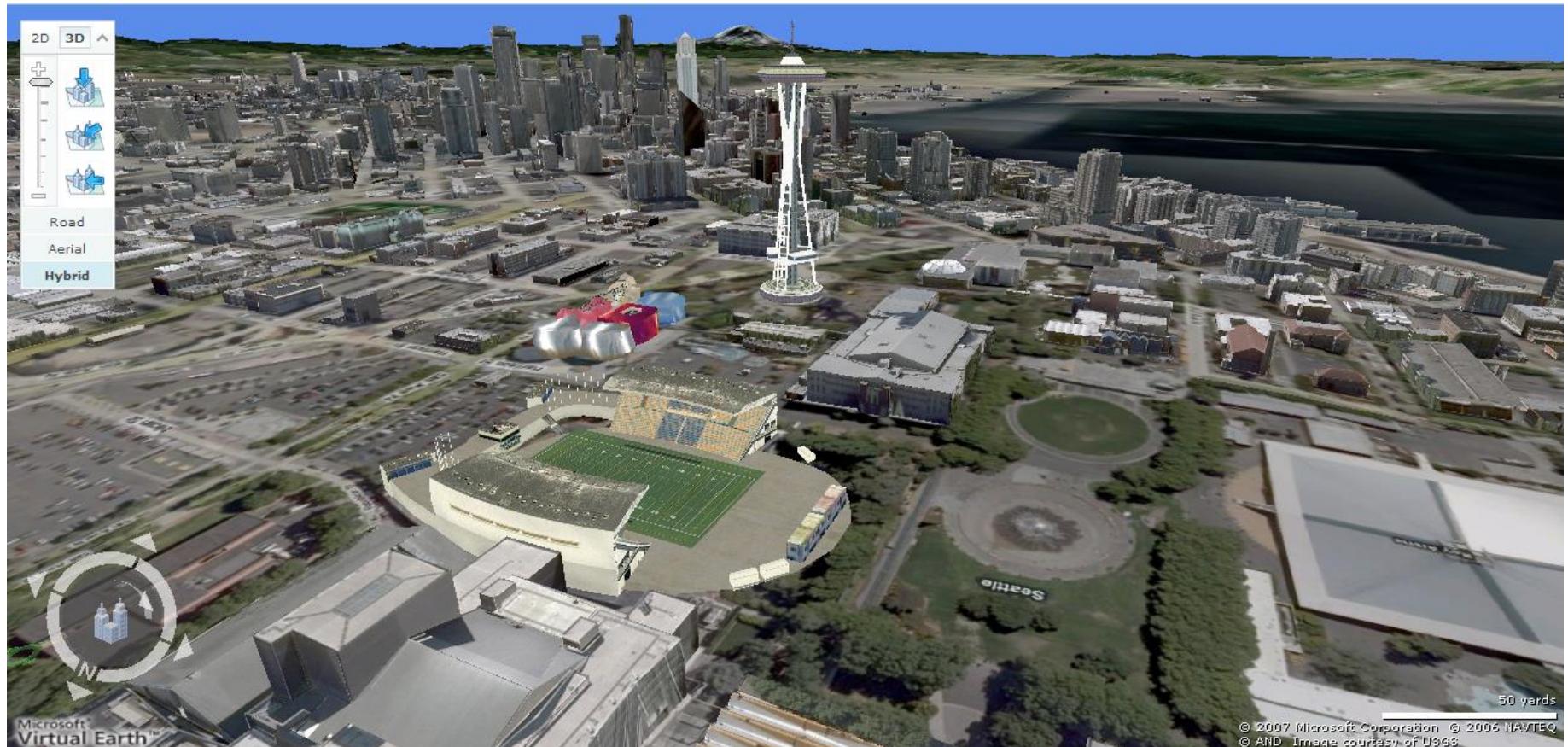
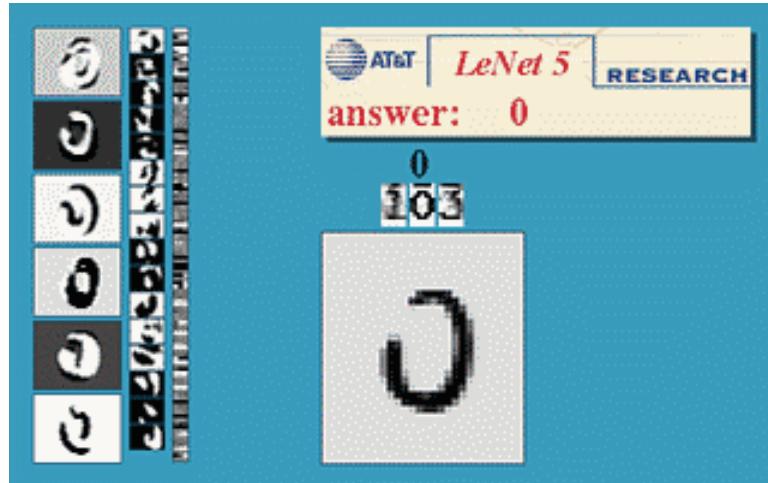


Image from Microsoft's [Virtual Earth](#)  
(see also: [Google Earth](#))

# 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/>



License plate readers  
[http://en.wikipedia.org/wiki/Automatic\\_number\\_plate\\_recognition](http://en.wikipedia.org/wiki/Automatic_number_plate_recognition)

# Face Detection

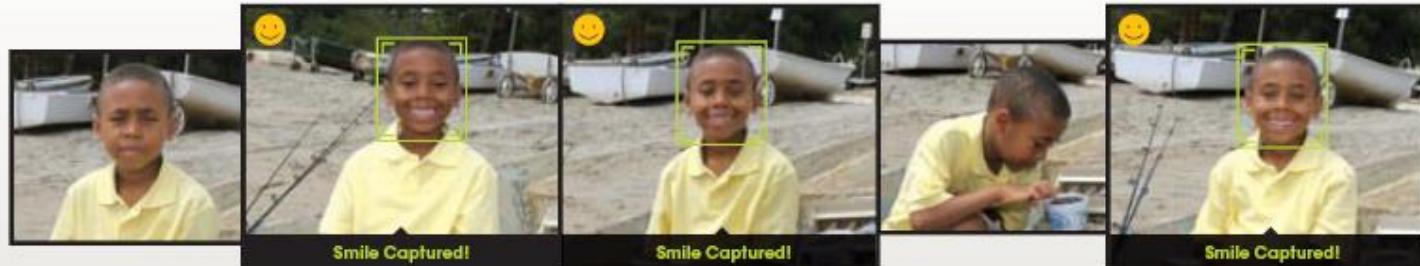


- Most digital cameras now detect faces
  - Auto-focus
  - Auto white-balance

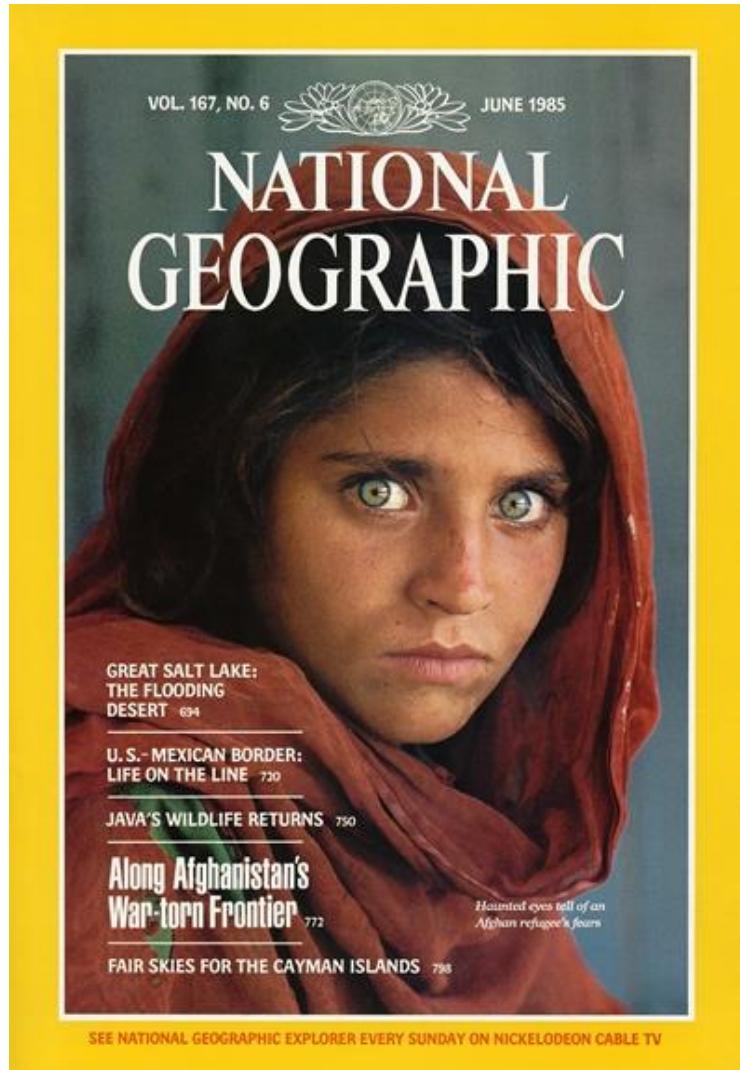
# Smile detection?

## 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.



# Biometric Recognition

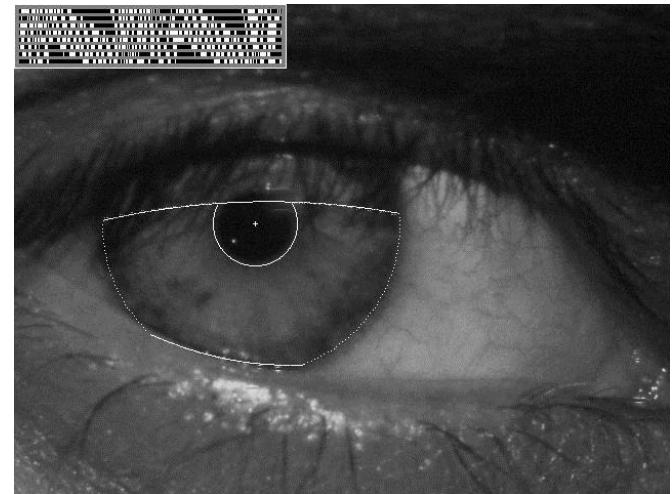
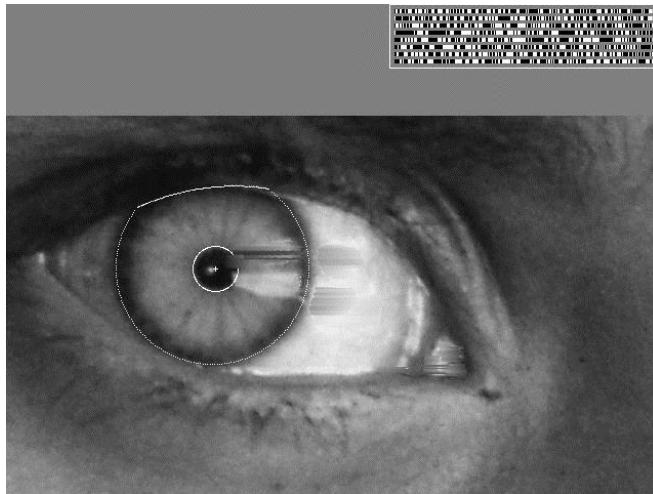


Who is she?

# Vision-based Biometrics



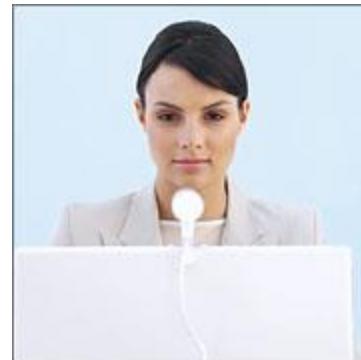
*How the Afghan Girl was Identified by Her Iris Patterns?*  
<http://www.cl.cam.ac.uk/~jgd1000/afghan.html>



# Biometric Identification

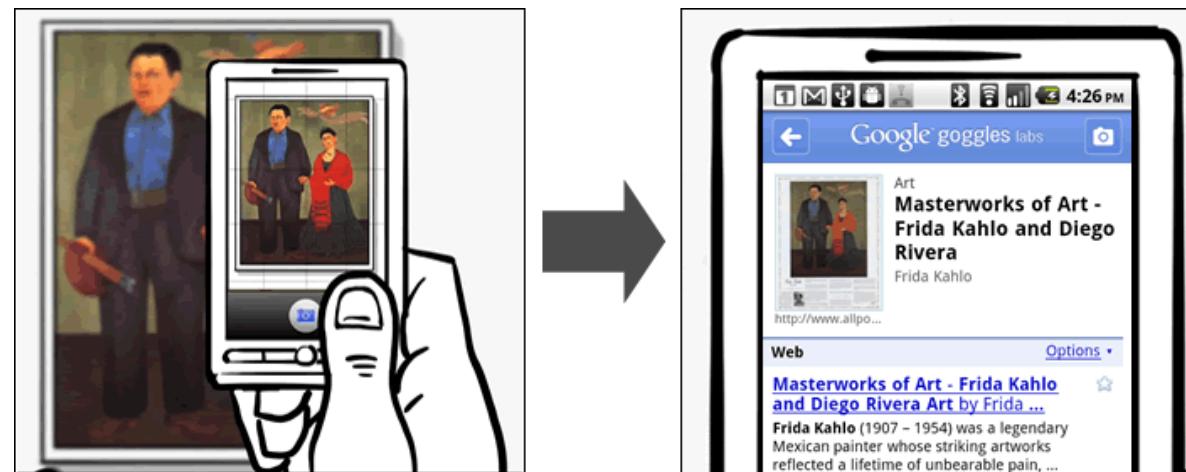
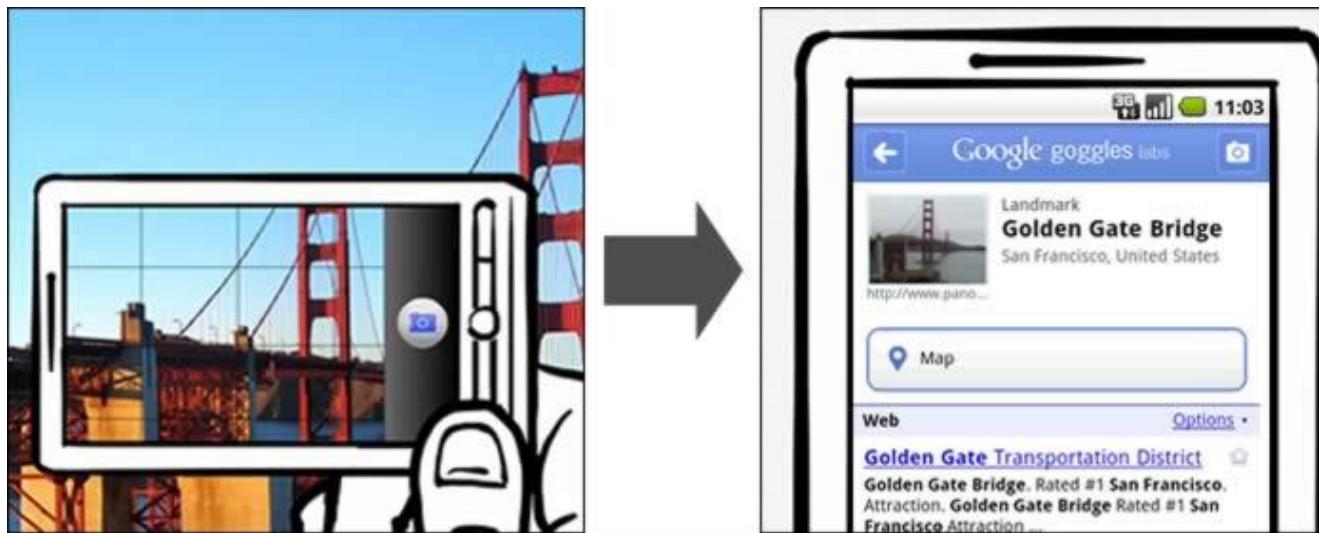


Fingerprint scanners on  
many new laptops,  
other devices

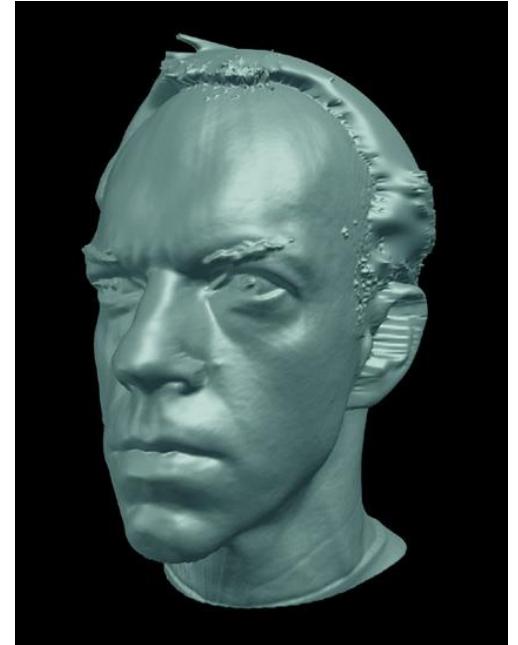


Face recognition systems now  
beginning to appear more widely  
<http://www.sensiblevision.com/>

# Mobile Visual Recognition

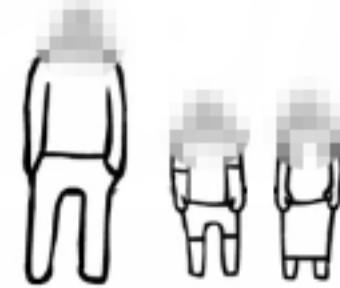


# Special effects: shape capture



*The Matrix* movies, ESC Entertainment, XYZRGB, NRC

# Google street view



Shopping Mail more ▾

philipp.lenssen@gmail.com | My Profile | Help | Web History | My Account | Sign out

e.g., "10 market st, san francisco" or "hotels near lax"

1504 broadway, nyc

Search Maps Show search options

Search the map Find businesses Get directions

Search Results My Maps

1504 Broadway  
New York, NY 10036  
Make this my default location

New! Explore this area »

Photos

User-Created Maps

Cheap Eats in NYC  
North American New Humanist Forum  
Barrel to Bottles

More photos, videos, and user-created maps »

Street View Traffic More... Map Satellite Terrain

Street View Help Full-screen

1508 Broadway Address is approximate

Zoom Out © 2008 Google

35

# Google street view



# Advanced Driver Assistance Systems (ADAS)

►► manufacturer products consumer products ◀◀

## Our Vision. Your Safety.

rear looking camera      forward looking camera  
side looking camera

**EyeQ** Vision on a Chip

Road, Vehicle, Pedestrian Protection and more

**AWS** Advance Warning System

Events

- > [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](#)

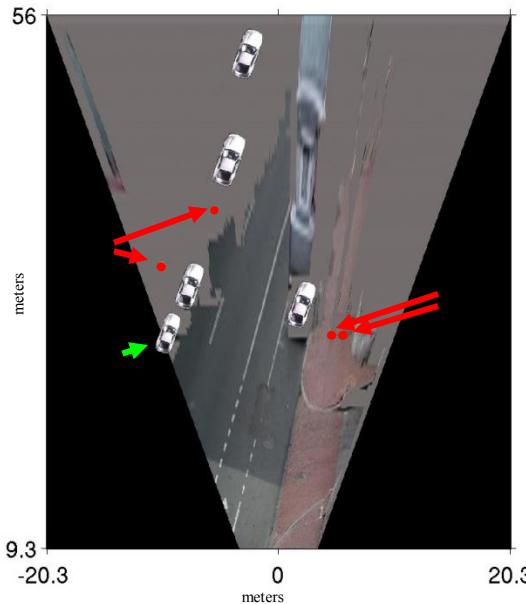
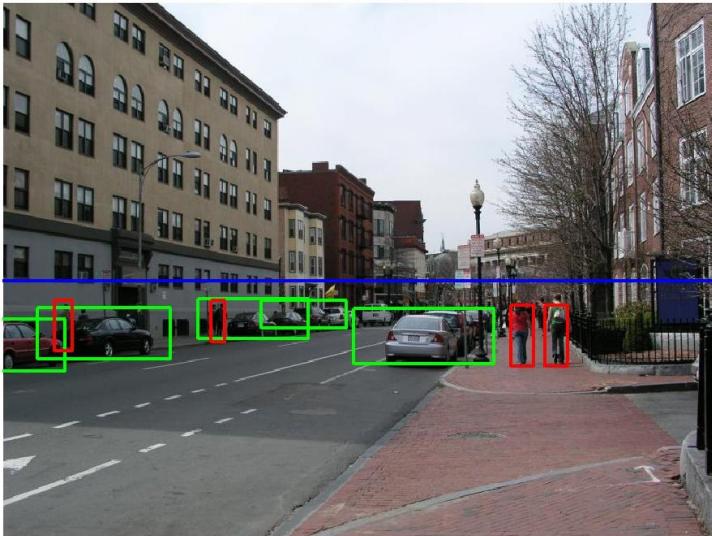
> all news

> read more

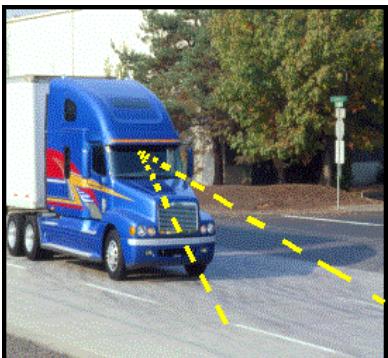
- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models
  - Lane detection, vehicle detection, pedestrian detection, traffic sign detection/recognition, surround view system

# Assisted or Autonomous Driving

Pedestrian and car detection



Lane detection



- Collision warning systems with adaptive cruise control,
- Lane departure warning systems,
- Rear object detection systems,

# Vision-based Interaction



ASUS



Nintendo Wii has camera-based IR tracking built in.

Control games with your own body motion/gesture and create immersive experiences by combining 3D personal image into the game scene

# Vision in Space

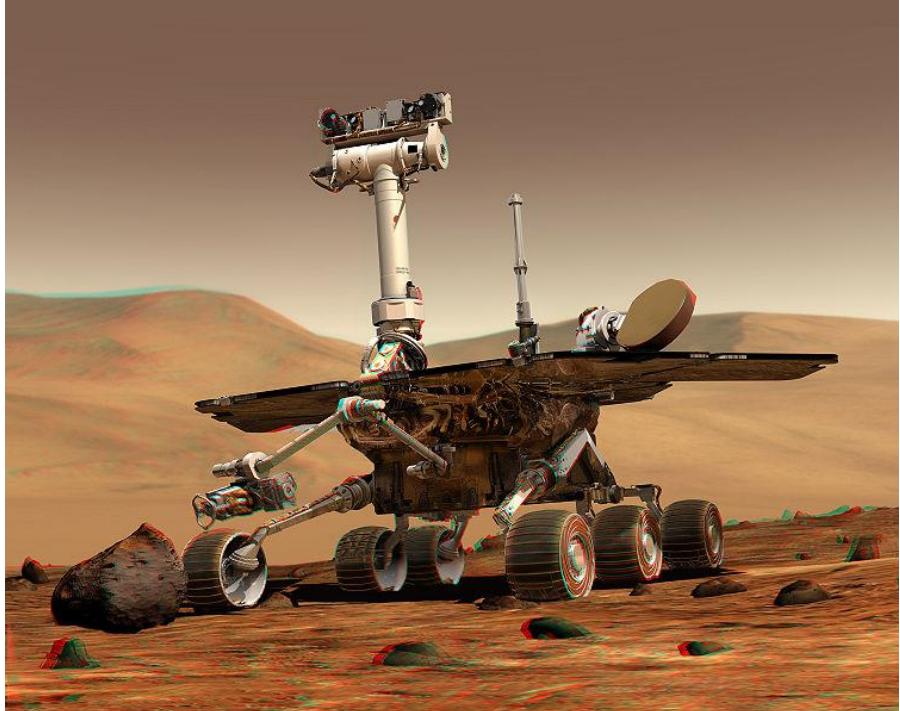


[NASA'S Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

## Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking

# Robotics

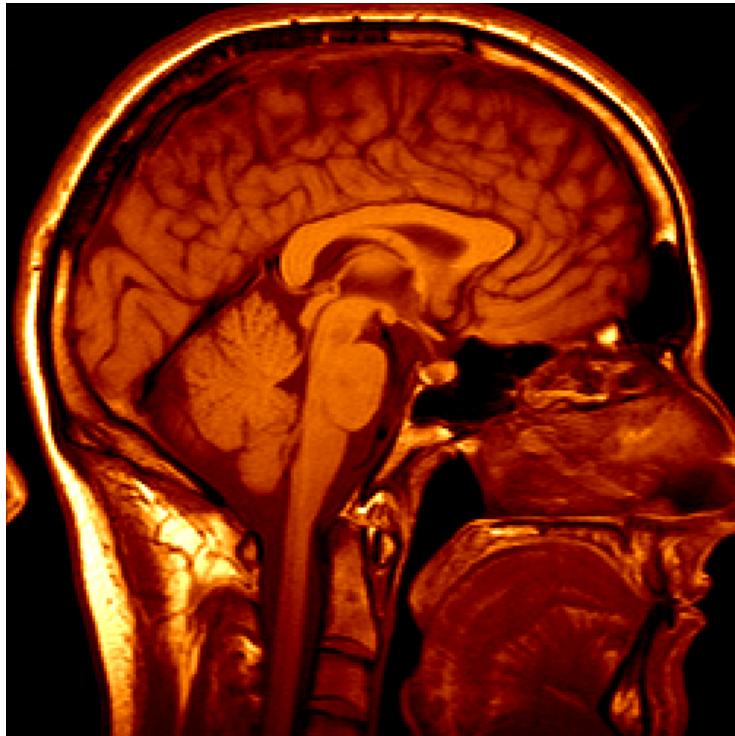


NASA's Mars Spirit Rover  
[http://en.wikipedia.org/wiki/Spirit\\_rover](http://en.wikipedia.org/wiki/Spirit_rover)



Darpa's Robotics Challenge  
[http://www.darpa.mil/Our\\_Work/TTO/Programs/DARPA\\_Robotics\\_Challenge.aspx](http://www.darpa.mil/Our_Work/TTO/Programs/DARPA_Robotics_Challenge.aspx)

# Medical Imaging

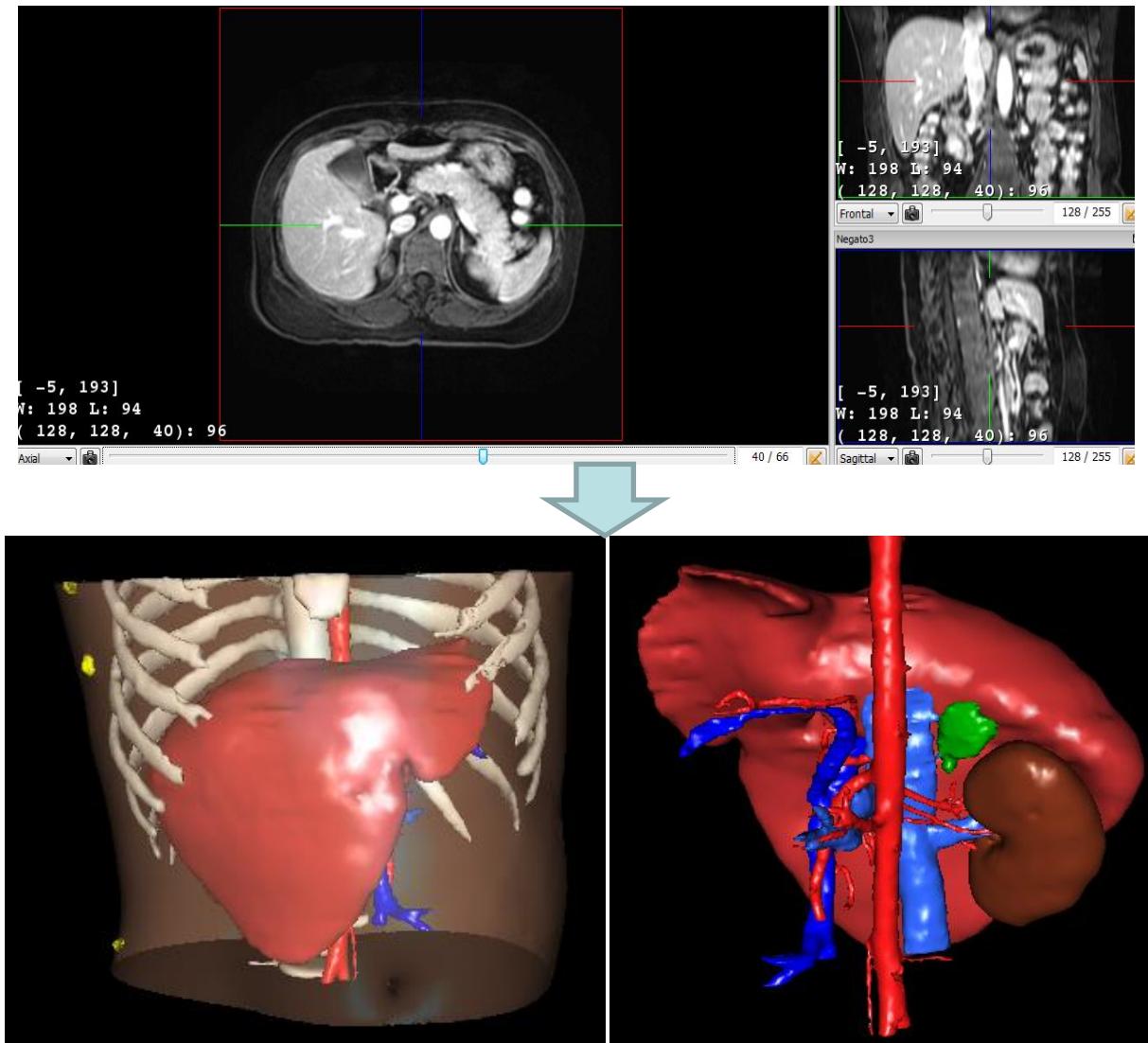


3D imaging  
MRI, CT



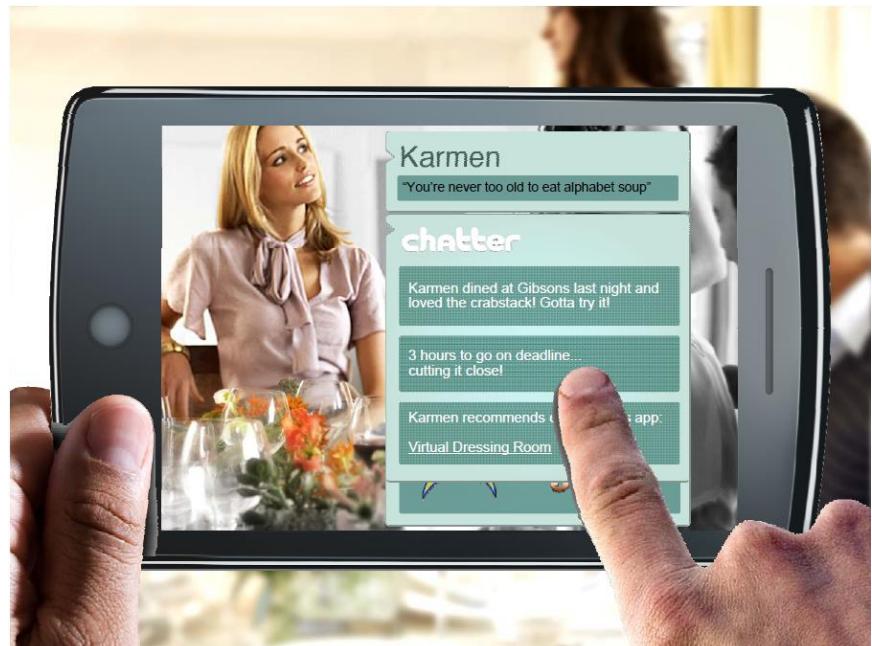
Image guided surgery  
Grimson et al., MIT

# 3D Model Reconstruction from Medical Images



# Augmented Reality

- AR allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. Therefore, AR supplements reality, rather than completely replacing it.
- Google Glasses is a research and development program to develop an augmented reality head-mounted display (HMD).



# Things to Do

- Read Chap. 1 & 2 (Szeliski)
- Next classes
  - Image formation (Chap. 2, Szeliski)