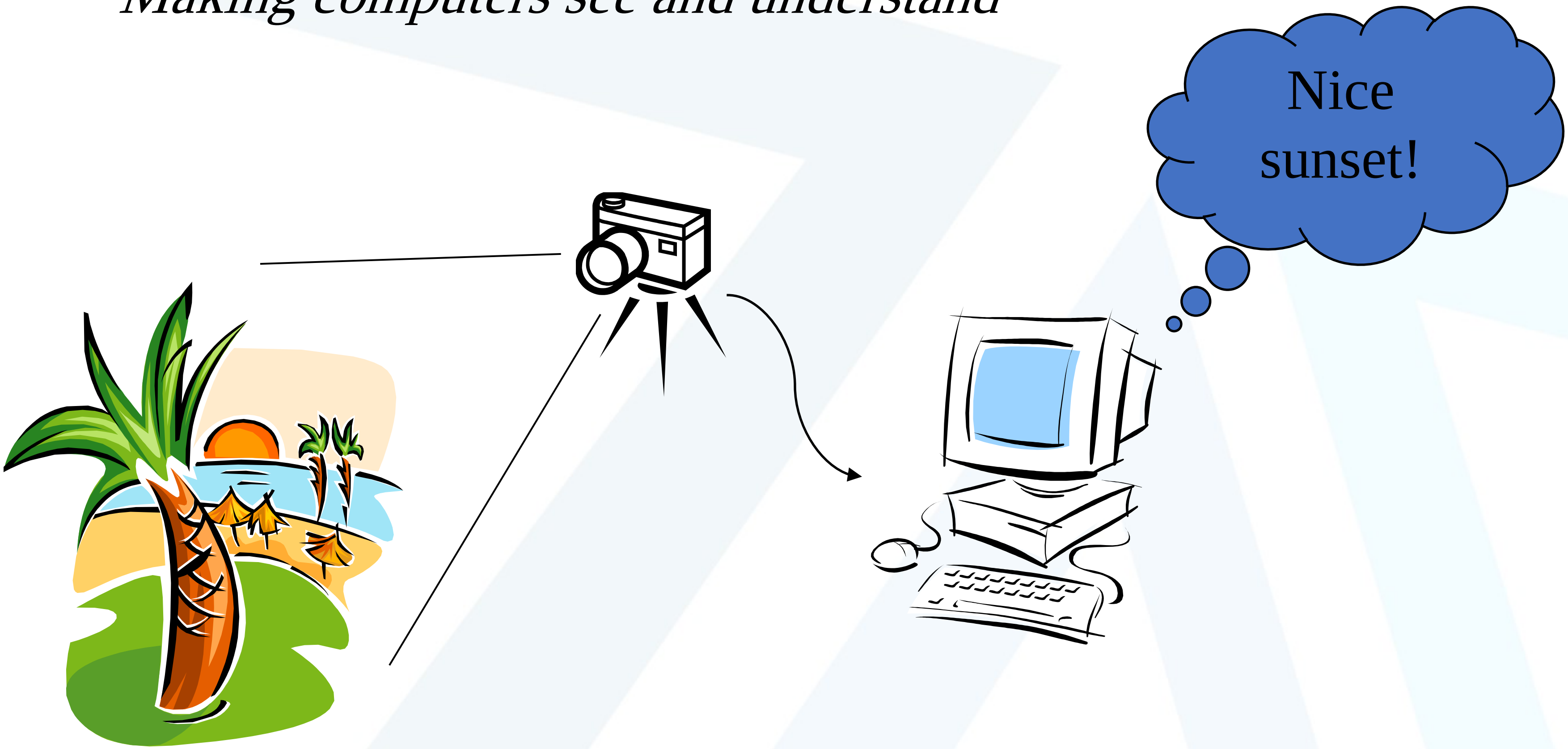
Introduction to Computer Vision

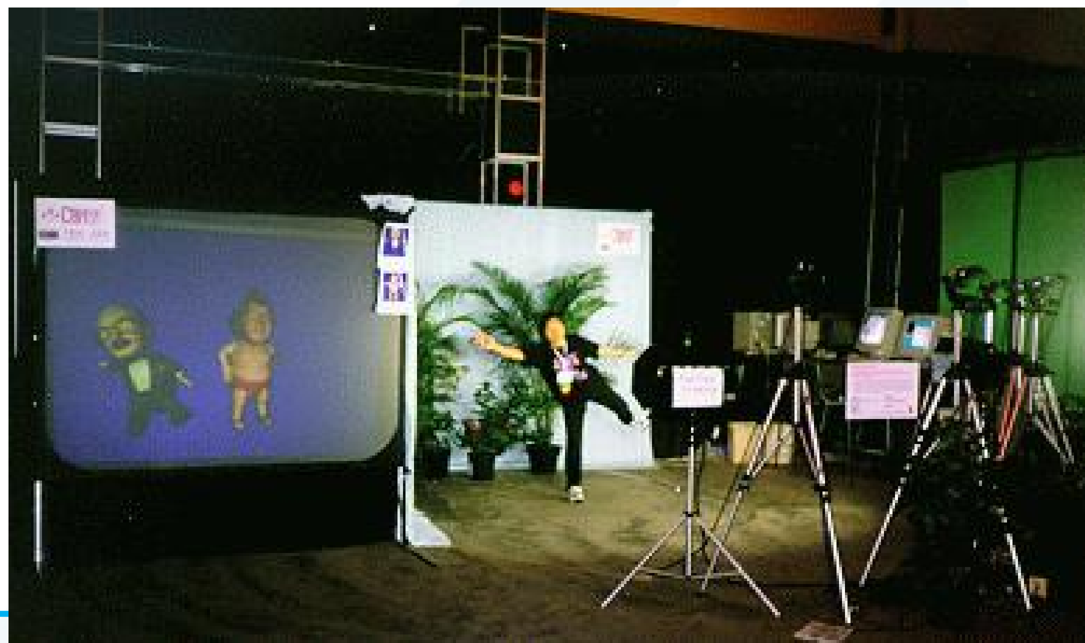
What is Computer Vision?

"Making computers see and understand"

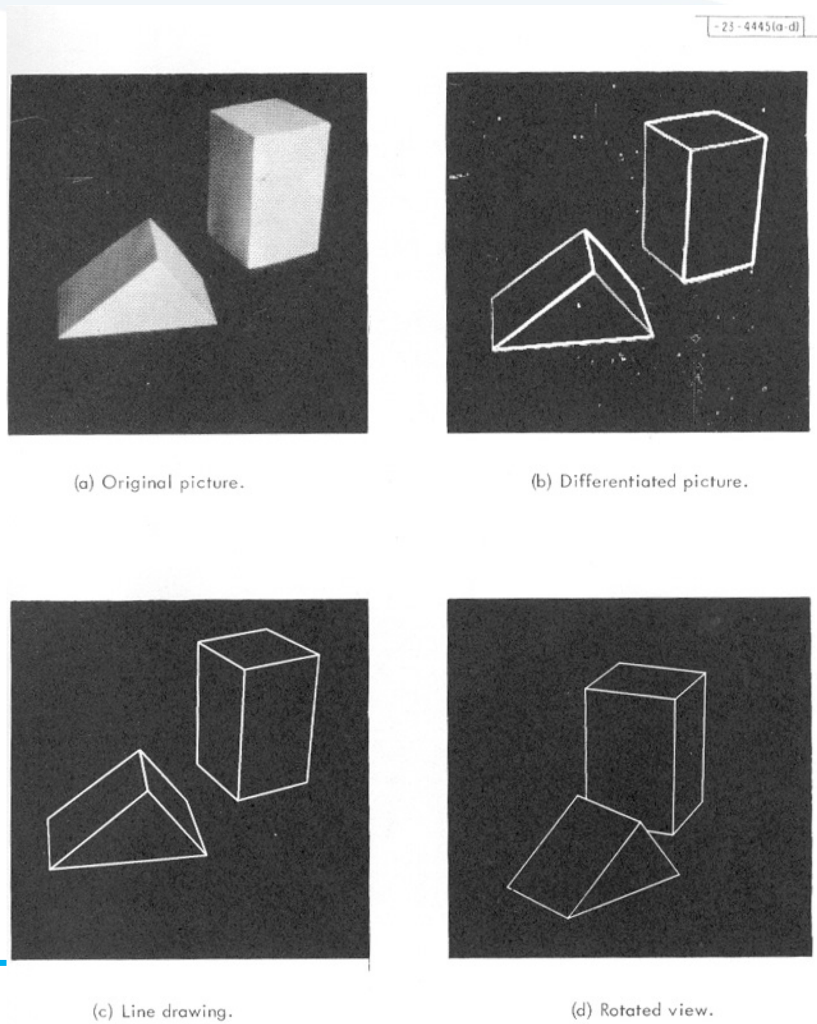


Main Objectives: Theory + Algorithms

- Development of the **theoretical** and **algorithmic** basis by which useful information about the 3D world can be automatically extracted and analyzed from a **single** or **multiple** 2D images of the world.



Origins of computer vision

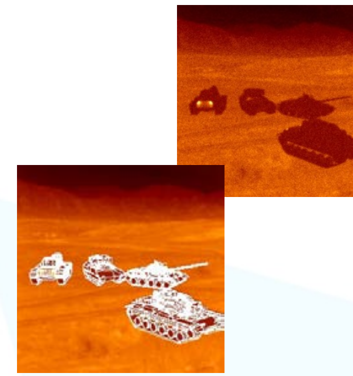


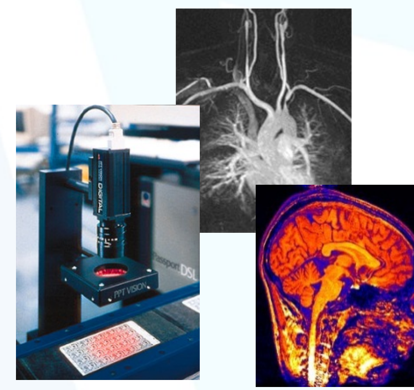
L. G. Roberts, *Machine Perception of Three Dimensional Solids*

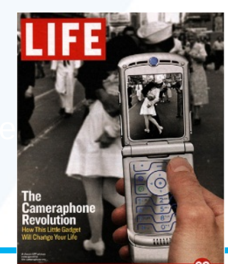
Ph.D. thesis, MIT Department of Electrical Engineering, 1963

Progress in Computer Vision

- First generation: **Military/Early Research**
  - Few systems, each custom-built, cost $Ms
  - “Users” have PhDs
  - 1 hour per frame
- Second generation: **Industrial/Medical**
  - Numerous systems, 1000s of each, cost $10Ks
  - “Users” have college degree
  - Special hardware
- Third generation: **Consumer**
  - 100000(00) systems, cost $100s
  - “Users” have little or no training
  - Emphasis on software







Computer Vision, Also Known As ...

- Computational Vision
  - Includes modeling of biological vision
- Image Understanding
  - Automated scene analysis (e.g., satellite images, robot navigation)
- Machine Vision
  - Industrial, factory-floor systems for inspection, measurements, part placement, etc.

Connections to other disciplines

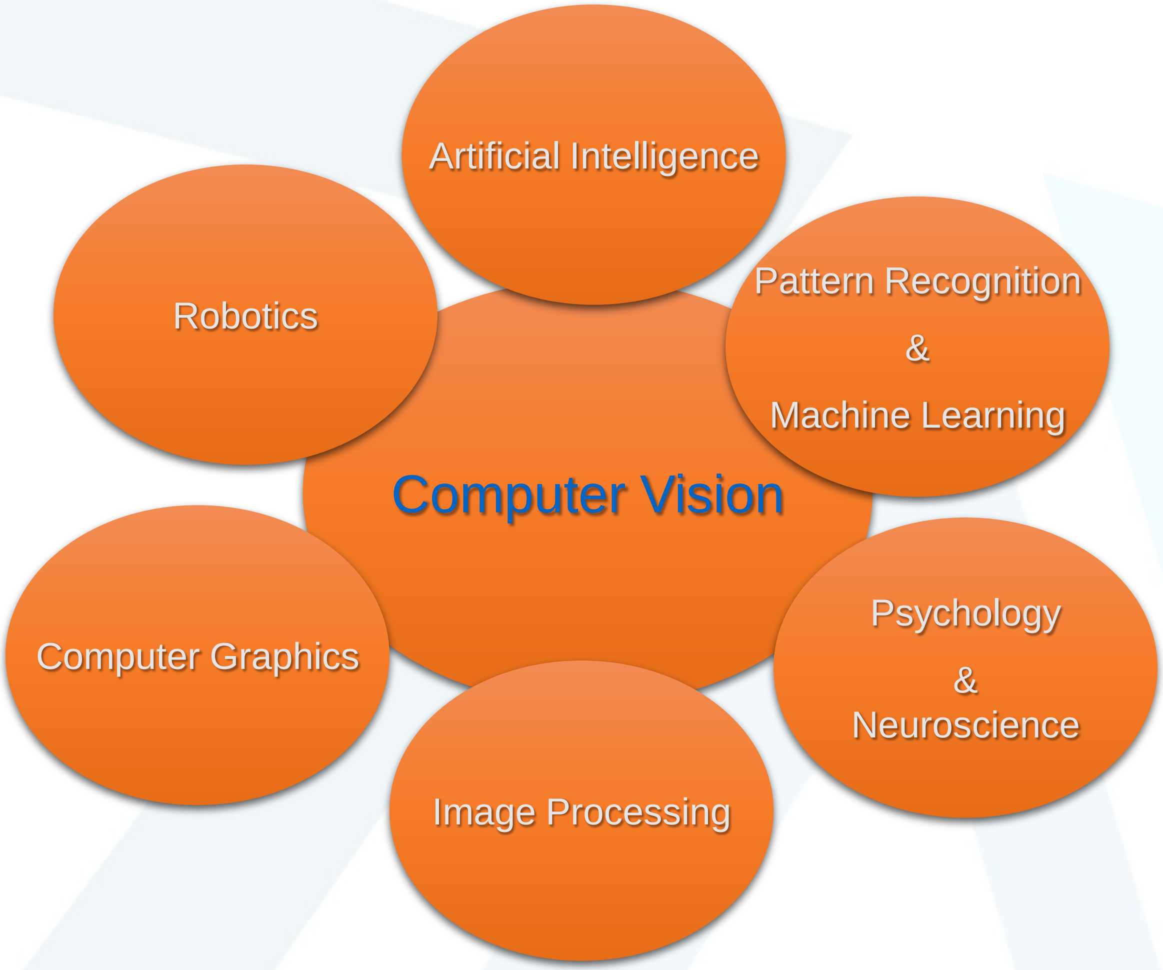
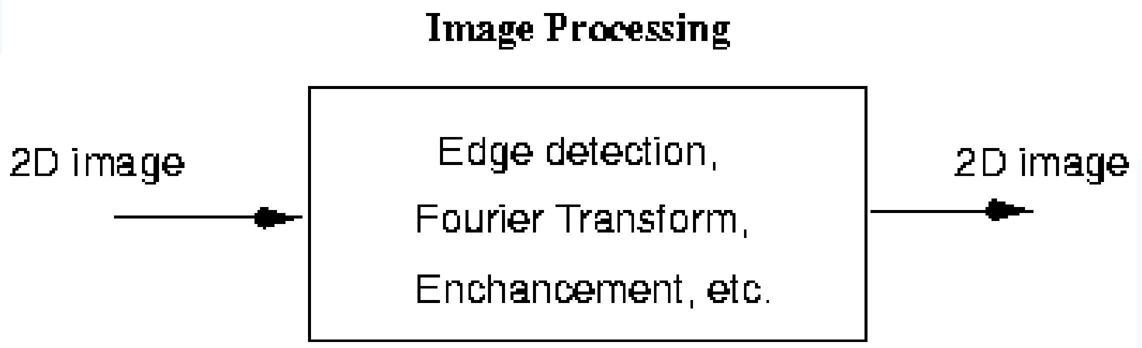


Image Processing



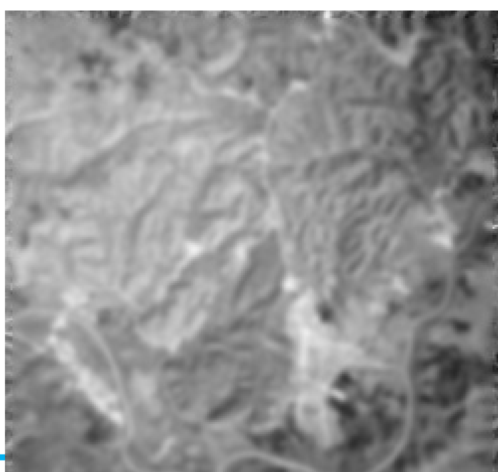
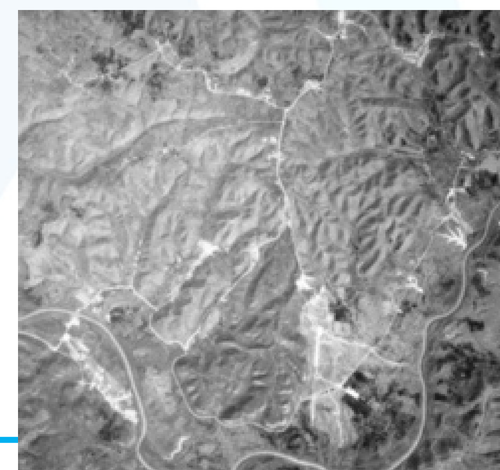
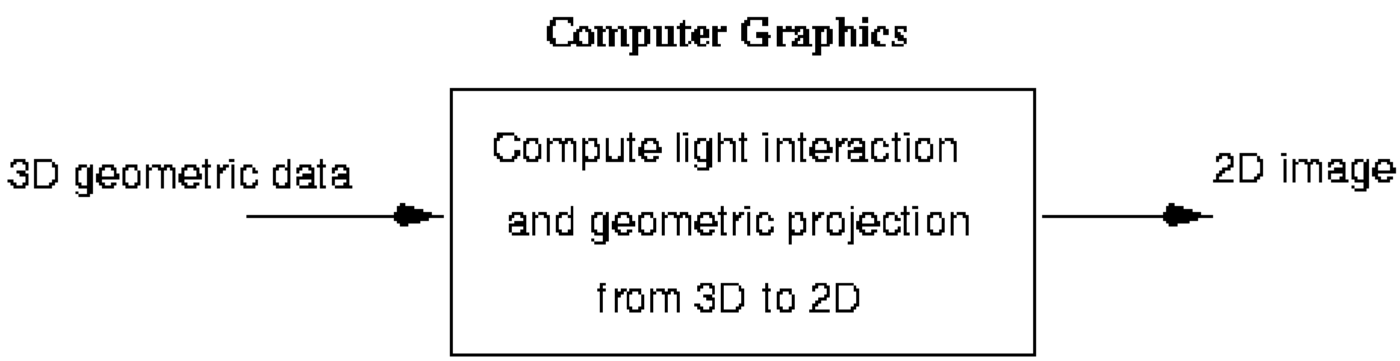


Image Enhancement

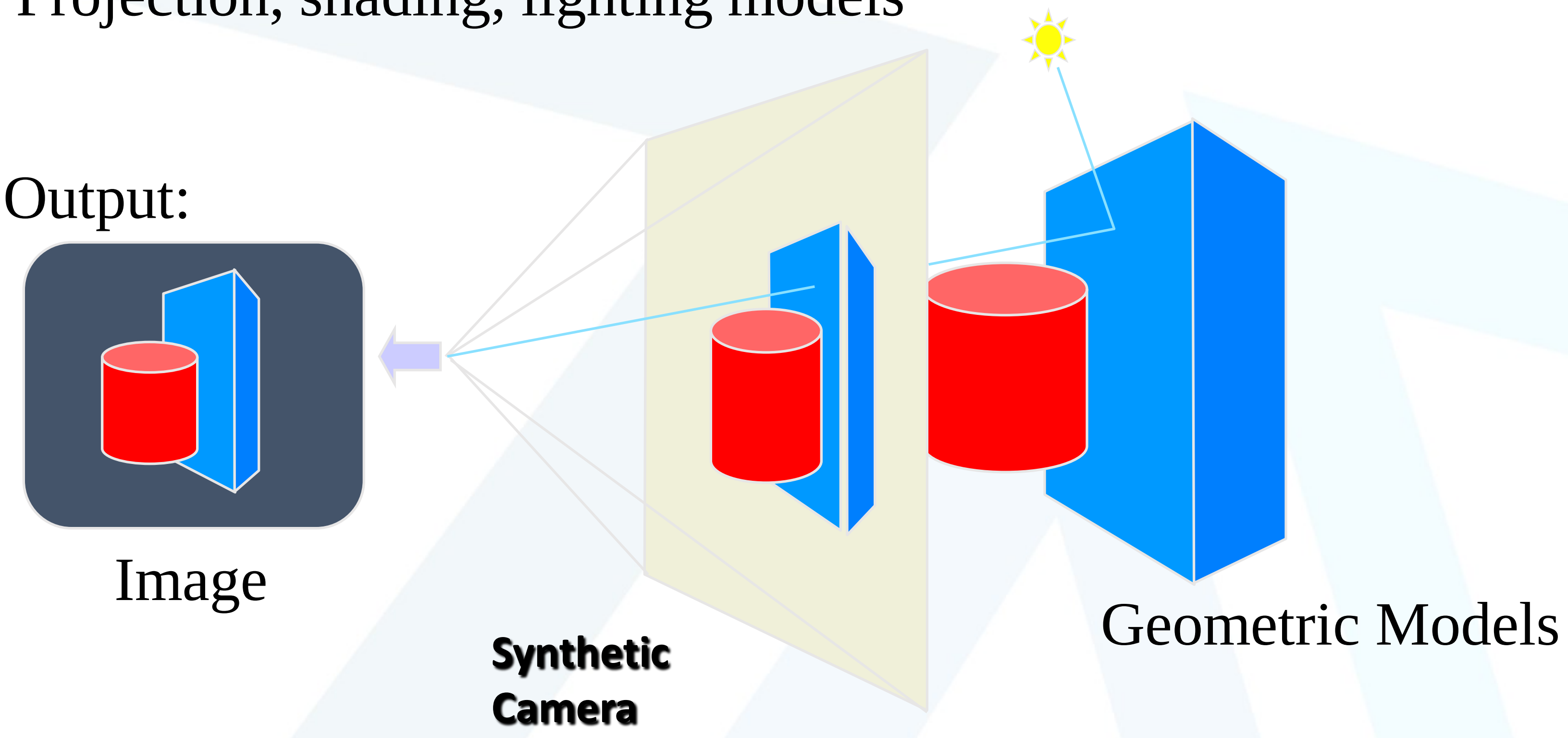


Computer Graphics

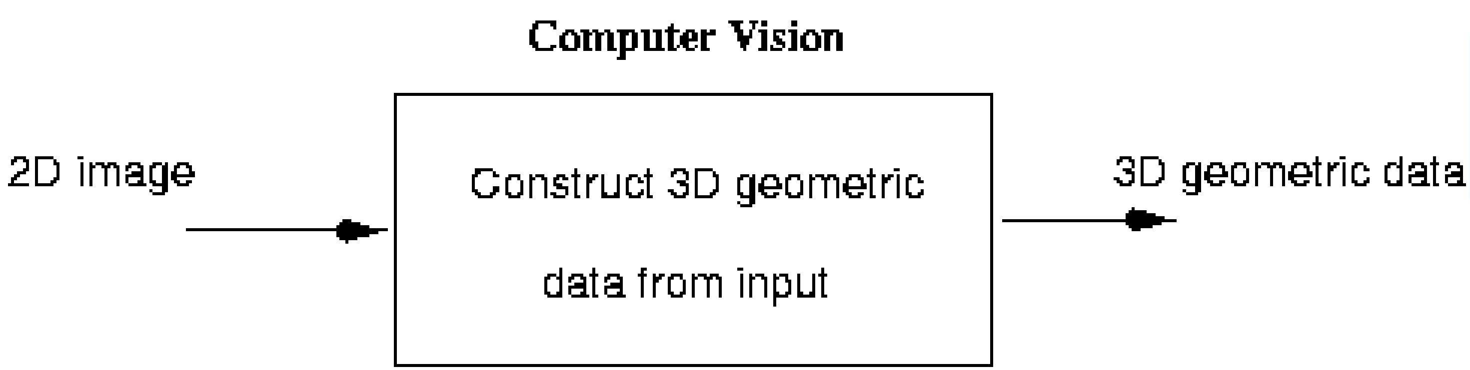


Computer Graphics

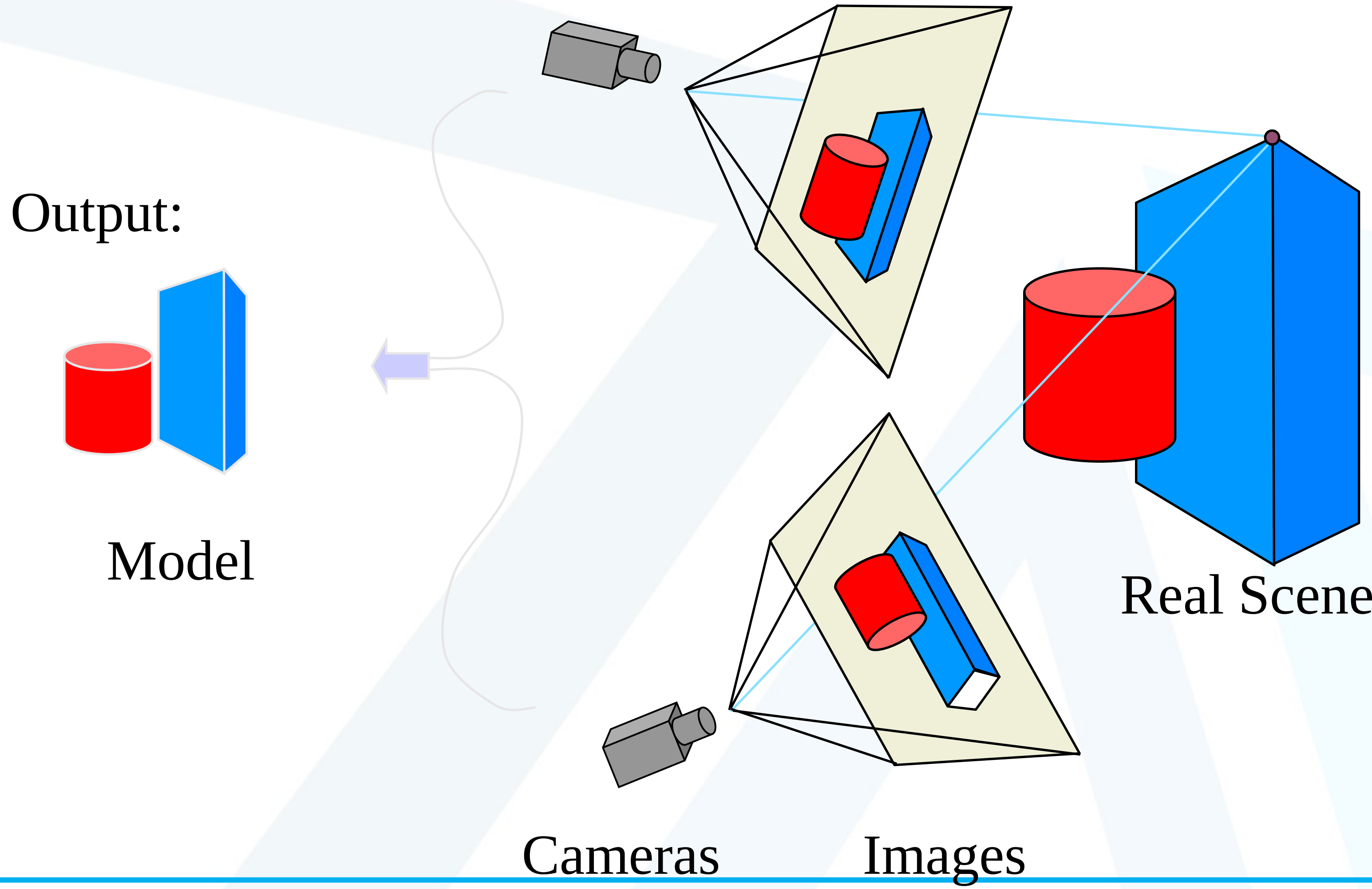
Projection, shading, lighting models



Computer Vision



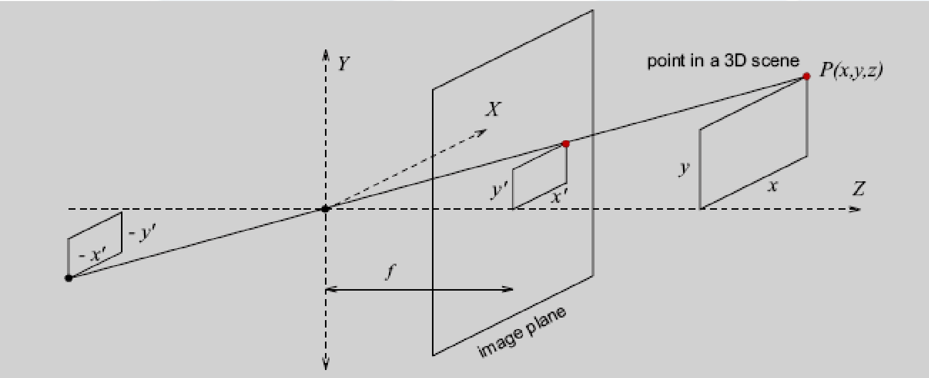
Computer Vision



Why is Computer Vision Difficult?

(1) It is a **many-to-one** mapping

- A variety of surfaces having different *material* and *geometrical* properties, possibly under different *lighting* conditions, could lead to similar images.
- Inverse mapping has non-unique solution; a lot of information is *lost* in the transformation from the 3D world to the 2D image.



Why is Computer Vision Difficult? (cont'd)

(2) It is **computationally intensive**

- A typical video is 30 frames / sec

(3) We **do not understand** the recognition problem!

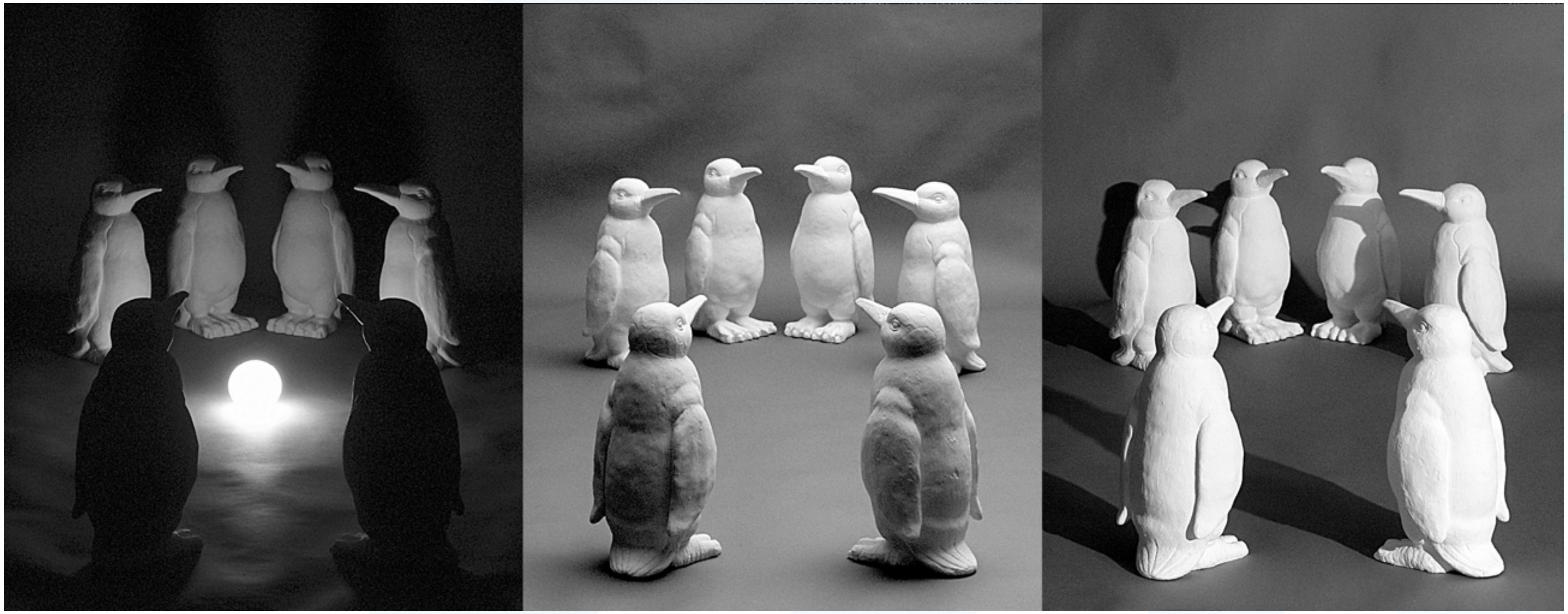
Main Challenges

- Viewpoint variations
- Illumination changes
- Scale changes
- Deformation
- Occlusions
- Background clutter
- Motion
- Intra/Inter-class variations

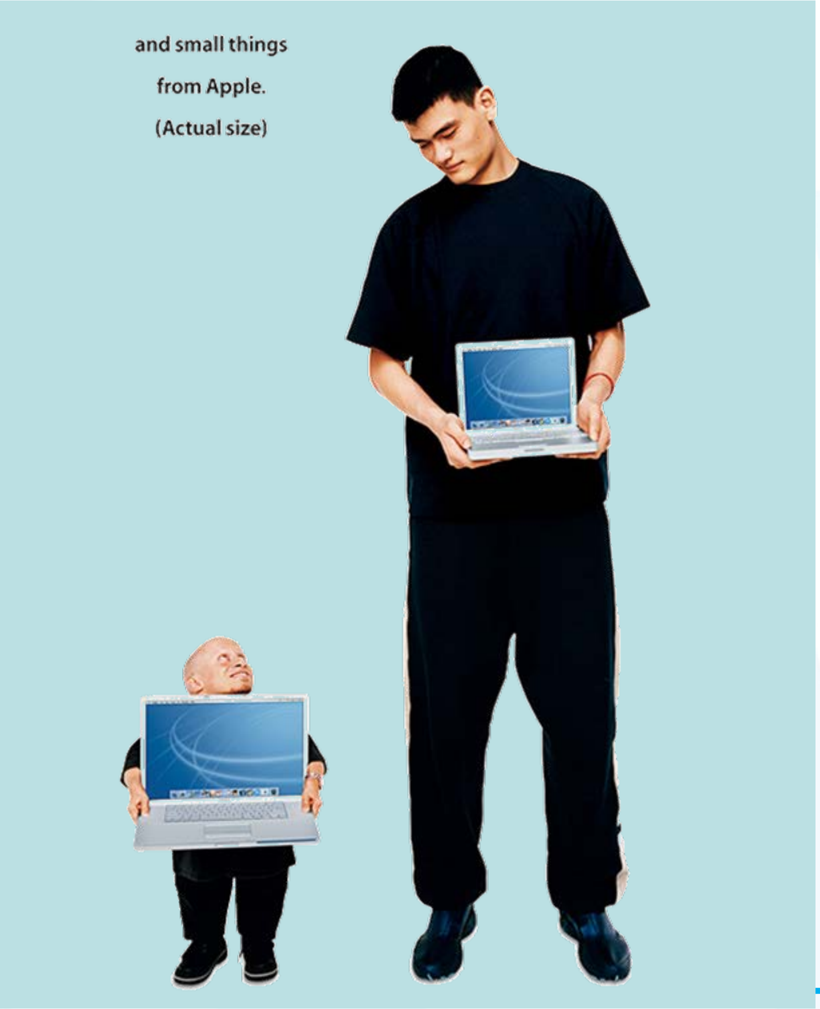
Viewpoint variations

Michelangelo 1475-1564

Illumination changes



Scale changes



Deformations

Occlusions

Background clutter

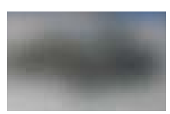
Motion blurring

Object intra-class variation






Local ambiguity

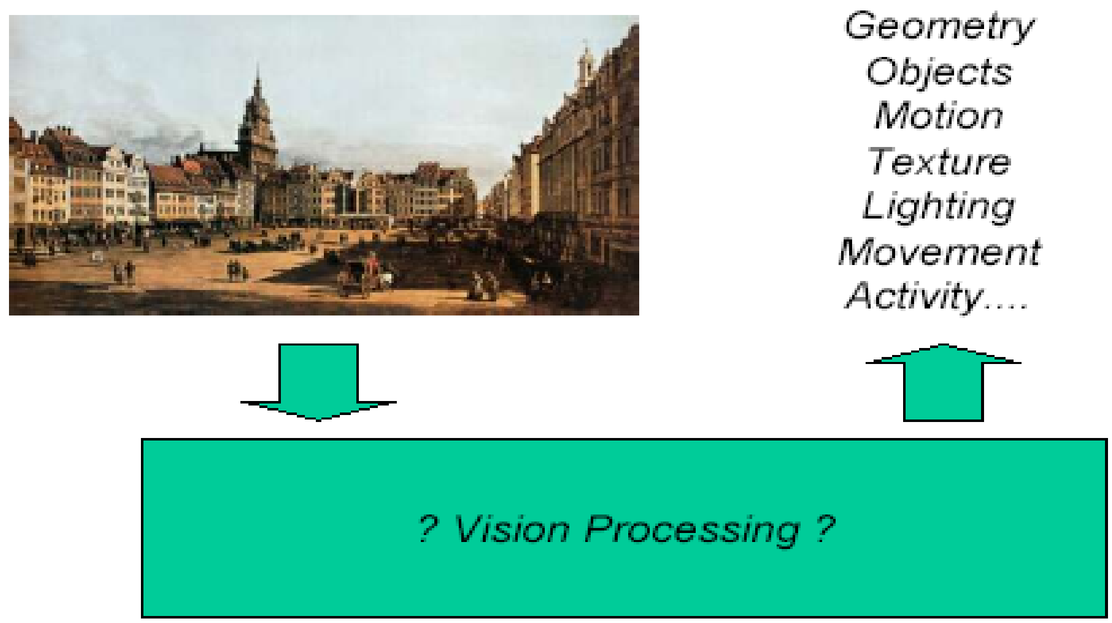


Three Processing Levels

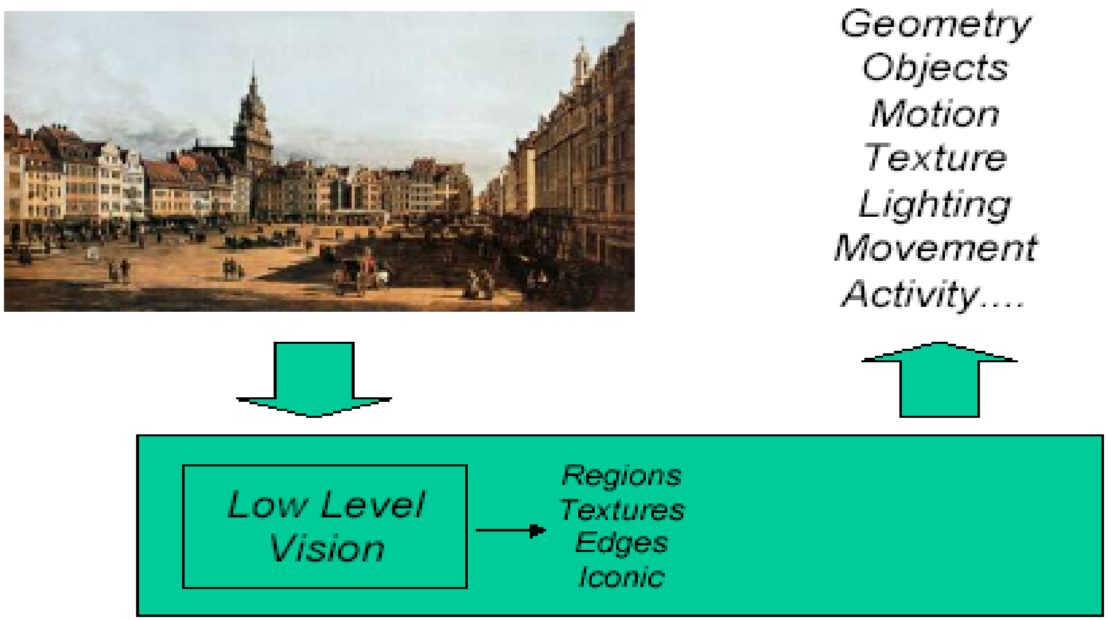
(1) **Low Level**

(2) **Mid Level**

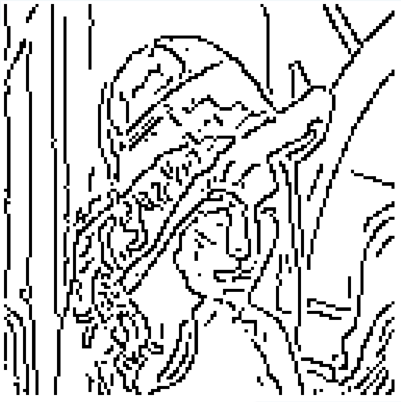
(3) **High Level**



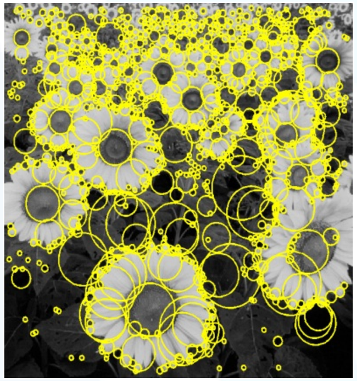
Low Level Vision



Low Level Vision - Examples



Edge detection



Corner and blob detection

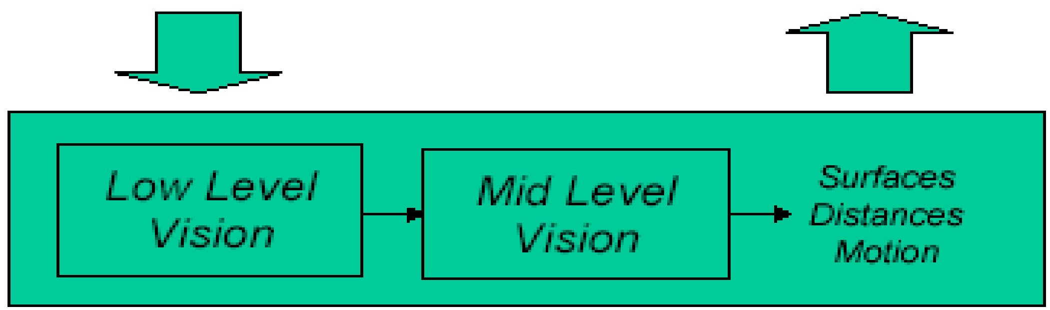
Low Level Vision - Examples

- Region segmentation



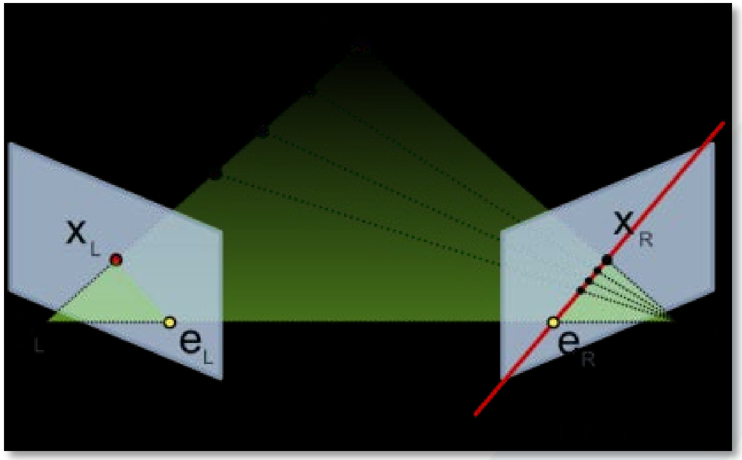
Mid Level Vision

Geometry
Objects
Motion
Texture
Lighting
Movement
Activity….



Mid Level Vision - Examples

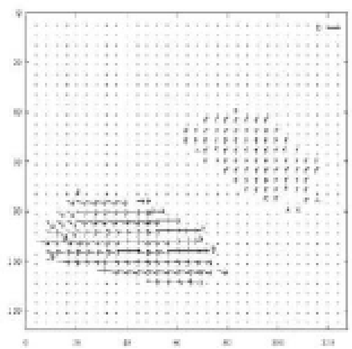
- 3D Reconstruction

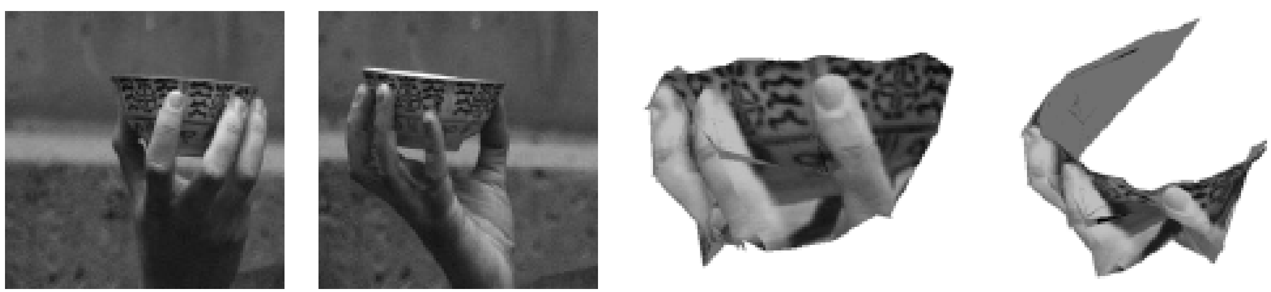


Mid Level Vision - Examples

- Structure (i.e., 3D) from motion

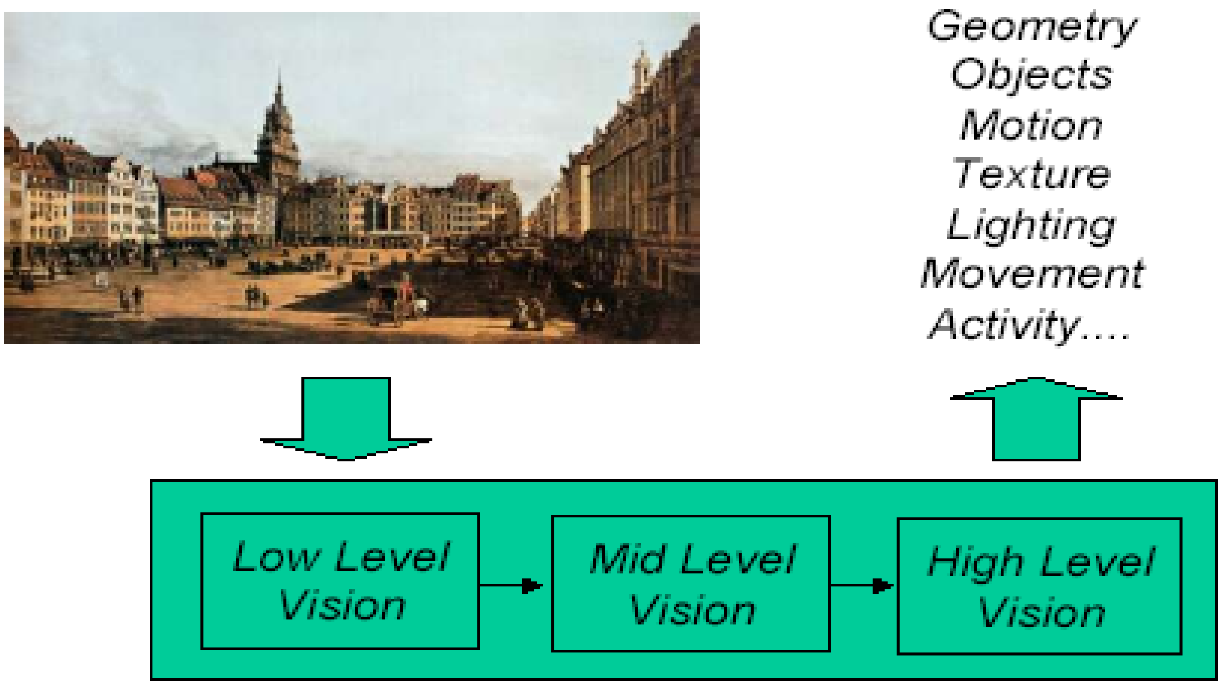
Optical flow





3D teacup model reconstructed from a 240-frame video sequence

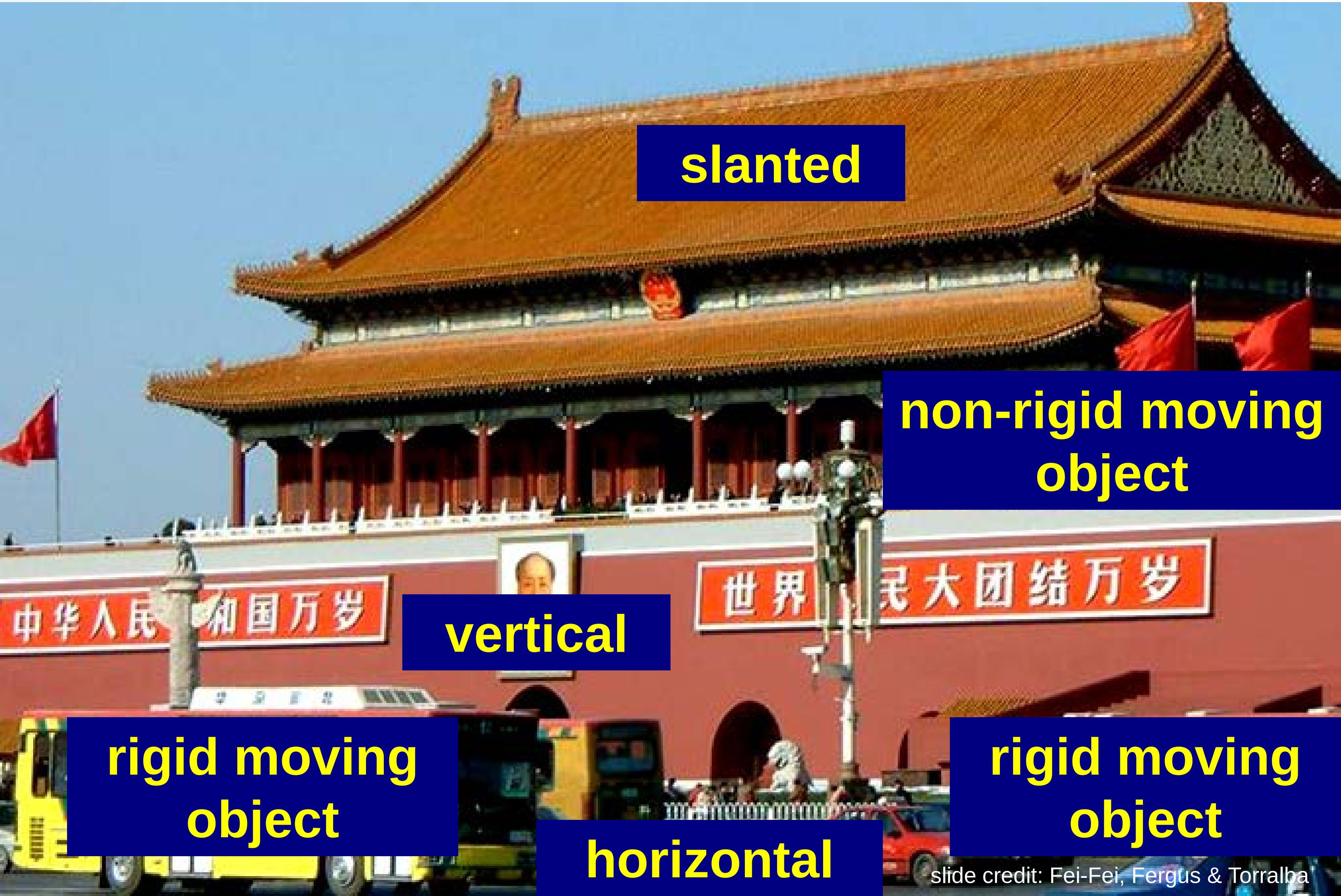
High Level Vision



Scene Interpretation

Object categorization

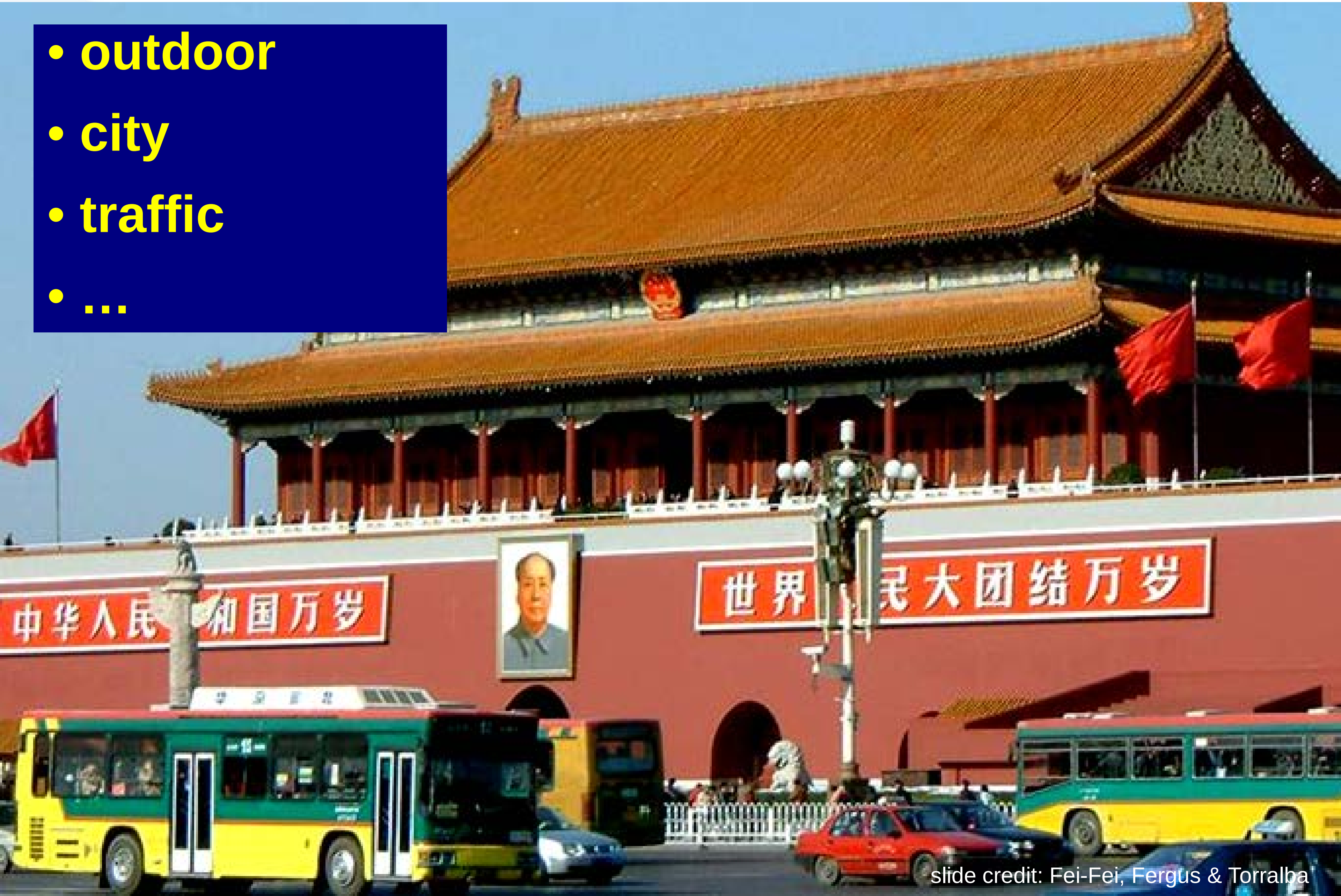
Qualitative geometric information



slide credit: Fei-Fei, Fergus & Torralba

Scene and context categorization

- **outdoor**
- **city**
- **traffic**
- **…**



slide credit: Fei-Fei, Fergus & Torralba

Visual Cues

- People use information from various visual cues for recognition (e.g., color, shape, texture etc.)

(1) How important is each visual cue?

(2) How do we combine information from various visual cues?

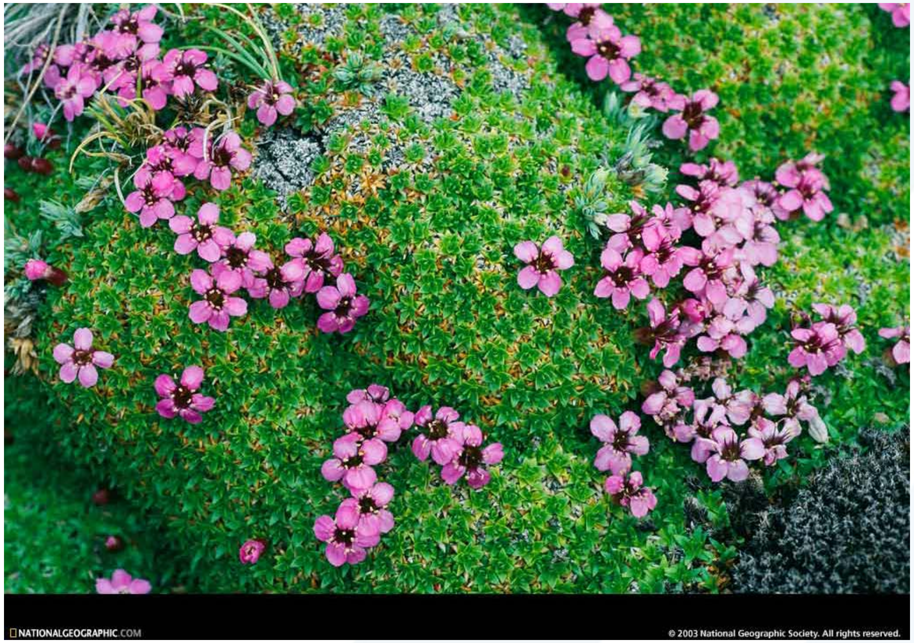
Color Cues

Texture Cues

Shape Cues

Grouping Cues

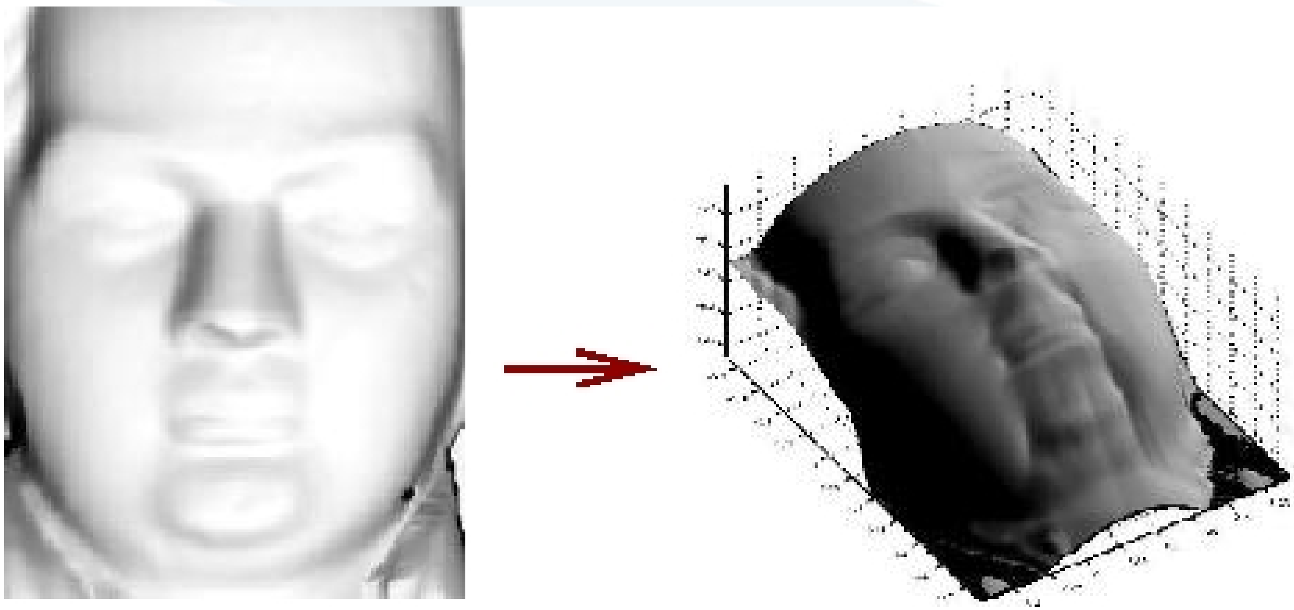
Similarity (color, texture, proximity)



Depth Cues


NATIONALGEOGRAPHIC.COM
© 2003 National Geographic Society. All rights reserved.

Shading Cues



a) Image

b) 3D surface reconstructed from the single image a)

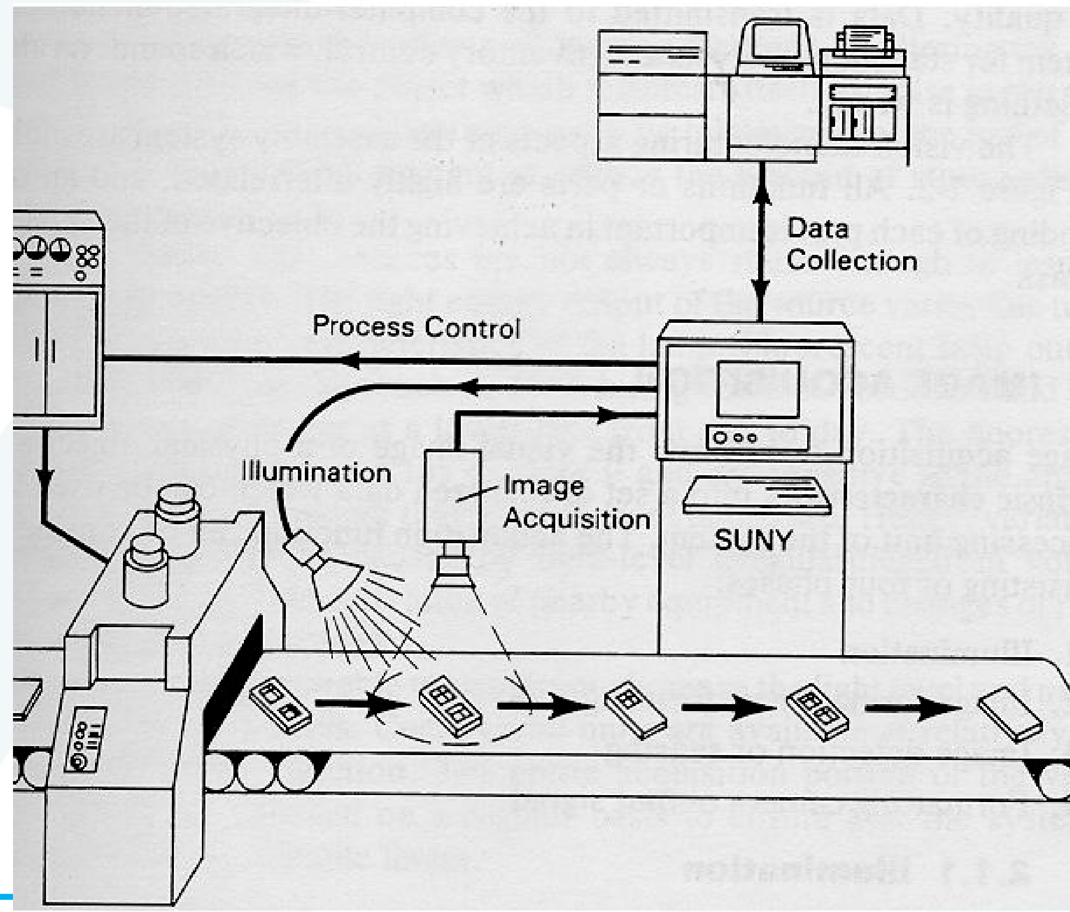
Computer Vision Applications

- Industrial inspection/quality control
- Surveillance and security
- Face recognition
- Gesture recognition
- Space applications
- Medical image analysis
- Autonomous vehicles
- Virtual reality and much more ……

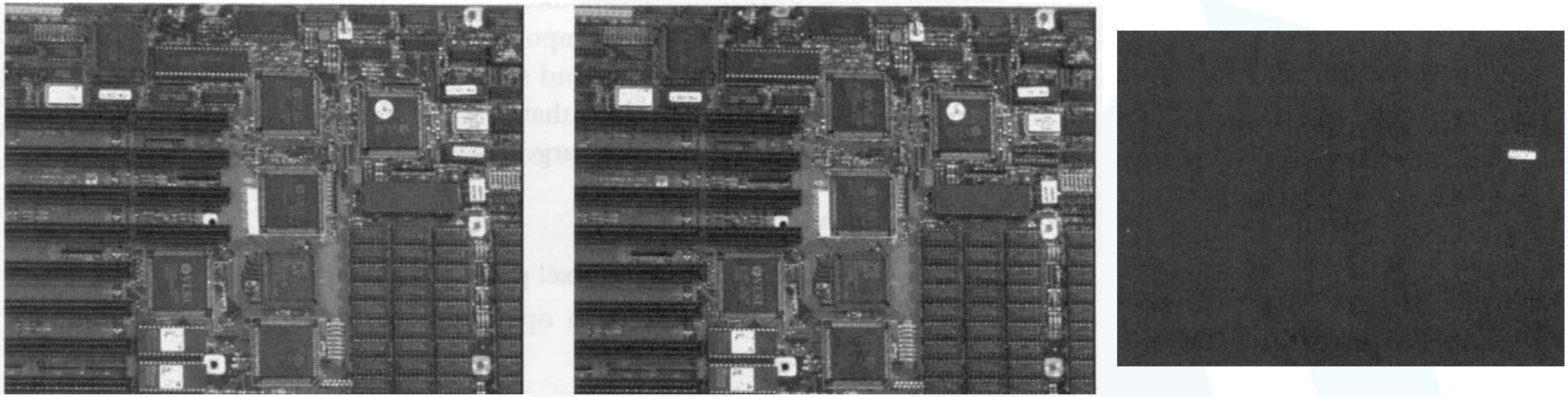
Industrial Computer Vision (Machine Vision)

Industrial computer vision systems work **really well.**

- Make strong assumptions about *lighting conditions*
- Make strong assumptions about the *position* of objects
- Make strong assumptions about the *type* of objects



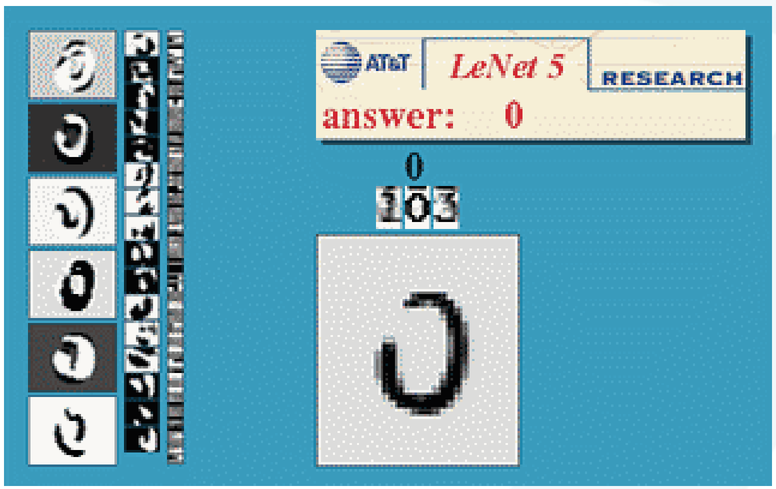
Visual Inspection



COGNEX

Optical character recognition (OCR)

- Technology to convert scanned docs to text

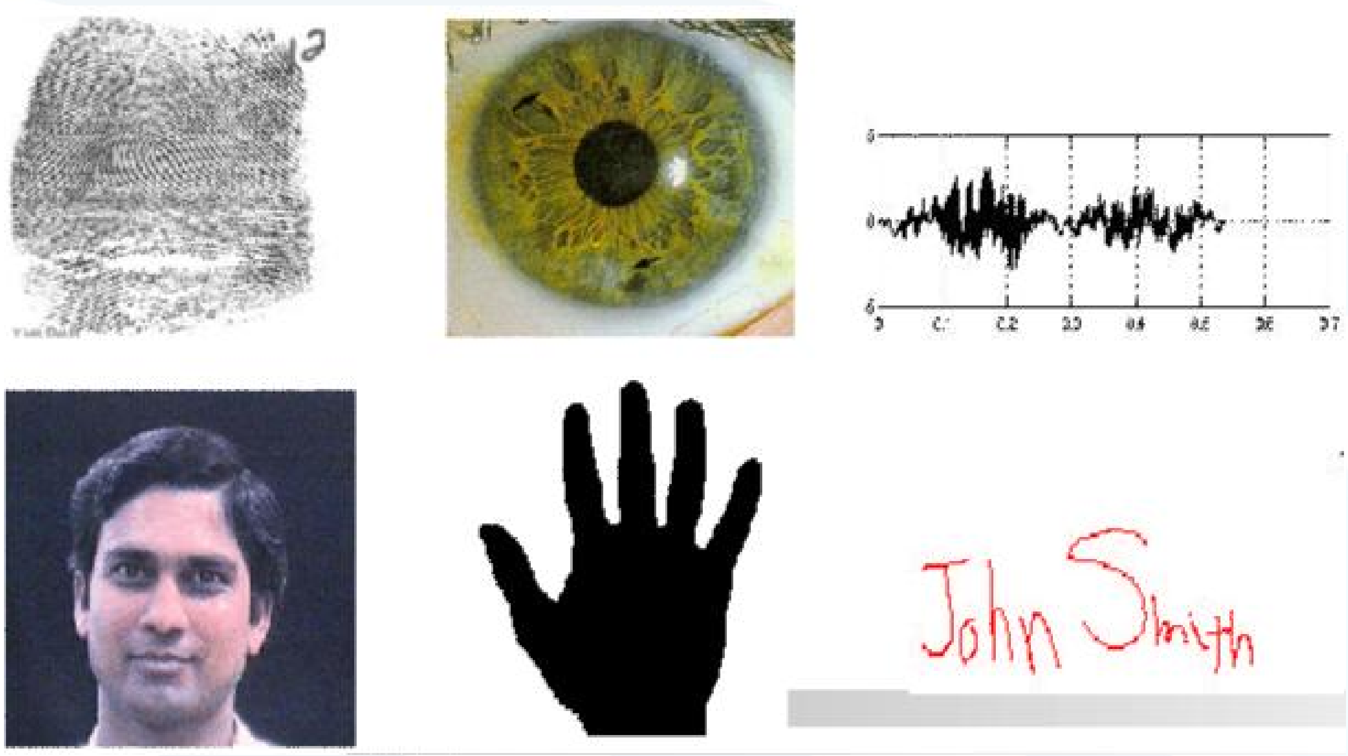


Digit recognition, AT&T labs
http://yann.lecun.com/exdb/lenet/

Automatic check processing

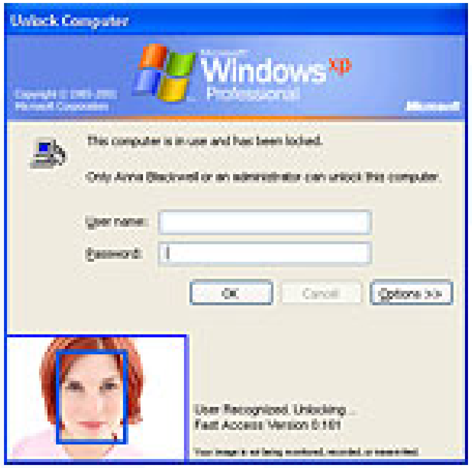
License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Biometrics



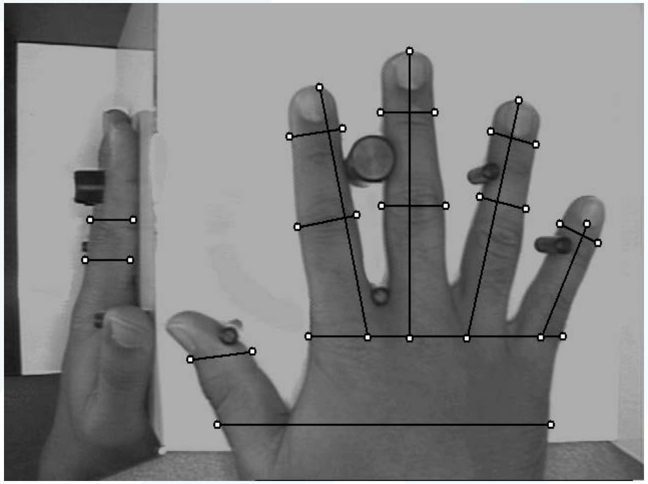
Login without a password…

Fingerprint scanners on many new laptops, other devices

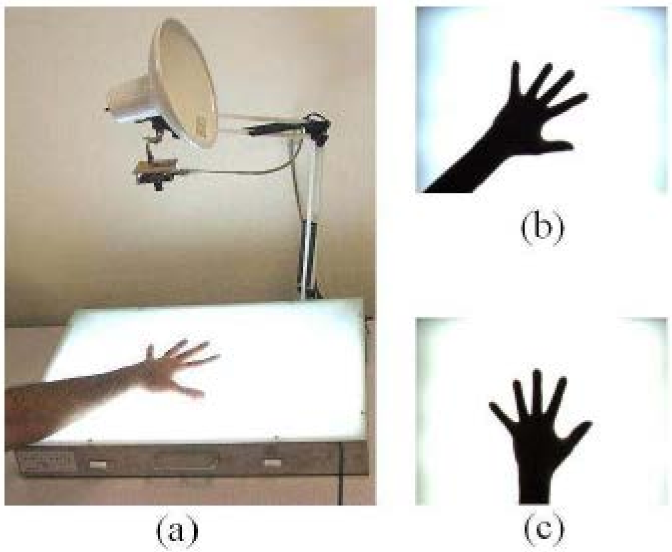


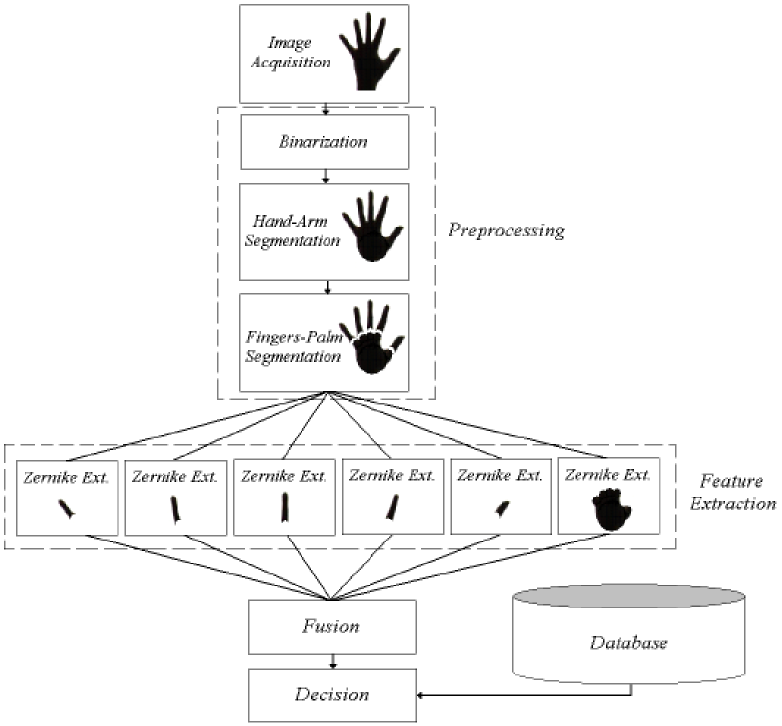
Face recognition systems now beginning to appear more widely http://www.sensiblevision.com/

Hand-based Biometrics



Hand-based Biometrics at

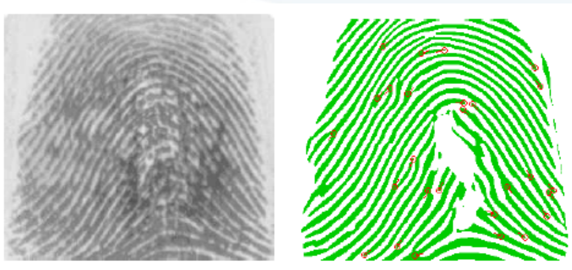




G. Amayeh, G. Bebis, A. Erol, and M. Nicolescu, "Hand-Based Verification and Identification Using Palm-Finger Segmentation and Fusion", **Computer Vision and Image Understanding (CVIU)** vol 113, pp. 477-501, 2009.

Fingerprint Biometrics

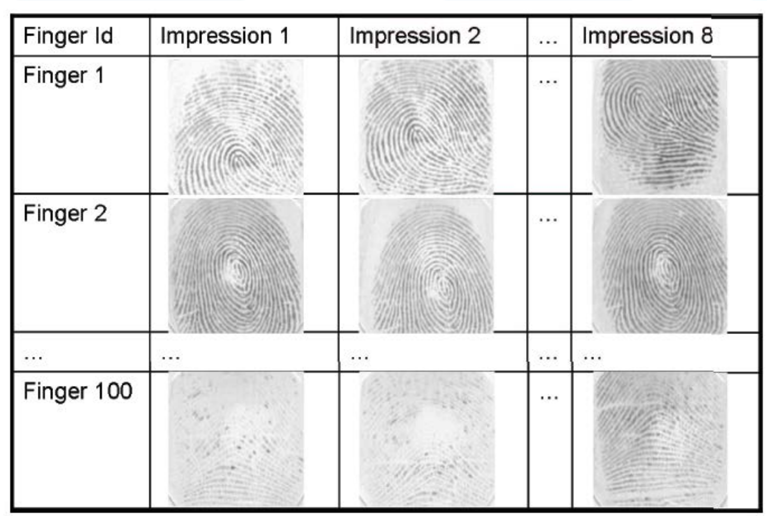
minutiae



small overlapping area

input

matching

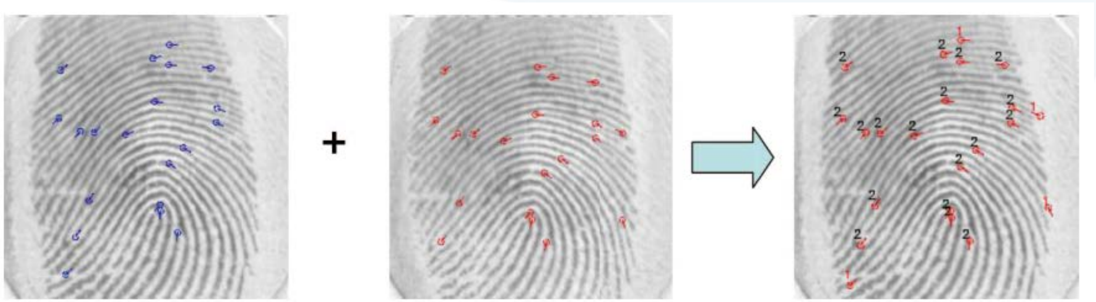


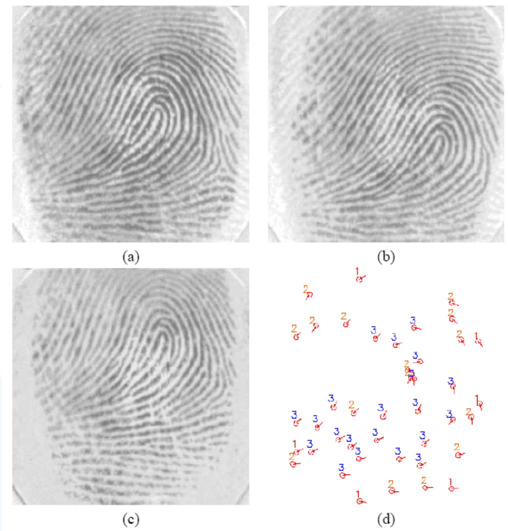
Finger Id	Impression 1	Impression 2	...	Impression 8
Finger 1			...	
Finger 2			...	
...	...	...	...	...
Finger 100			...	

ID

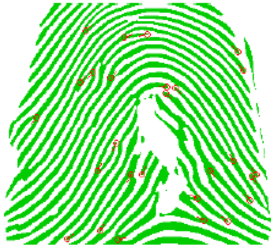
Fingerprint Biometrics at

Super-Template Synthesis

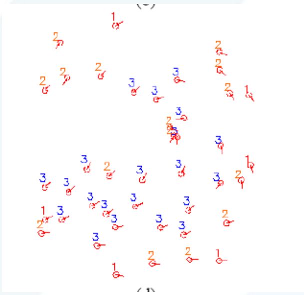




super-template



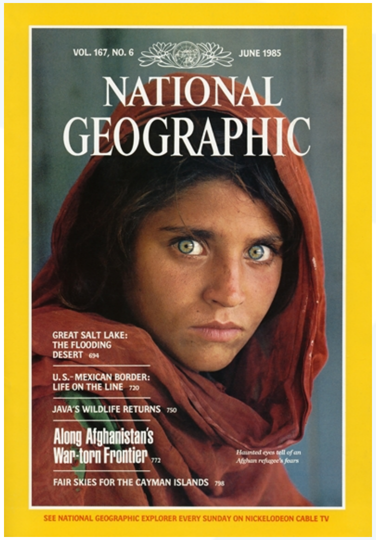
matching



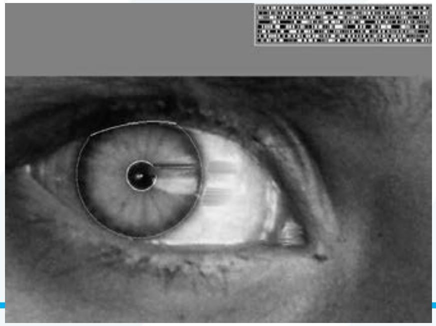
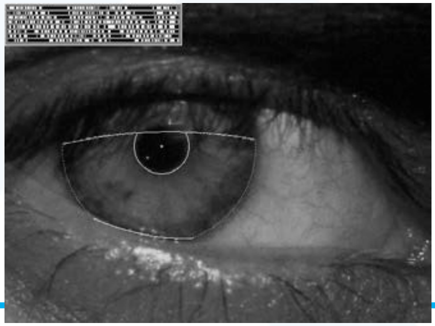
ID

T. Uz, G. Bebis, A. Erol, and S. Prabhakar, "Minutiae-Based Template Synthesis and Matching for Fingerprint Authentication", **Computer Vision and Image Understanding (CVIU)**, vol 113, pp. 979-992, 2009.

Iris Biometrics

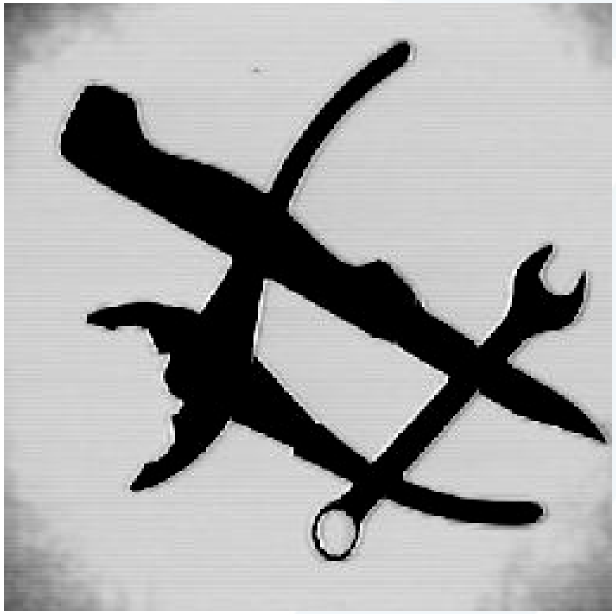


How the Afghan Girl was Identified by Her Iris Patterns

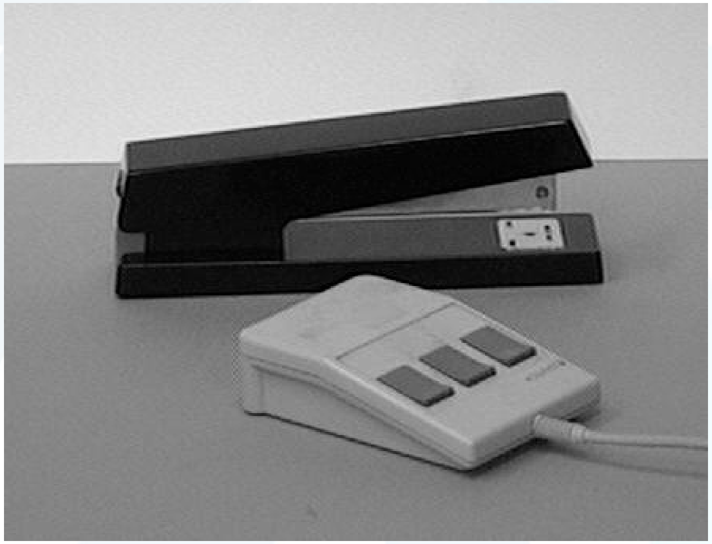



Object Recognition

2D

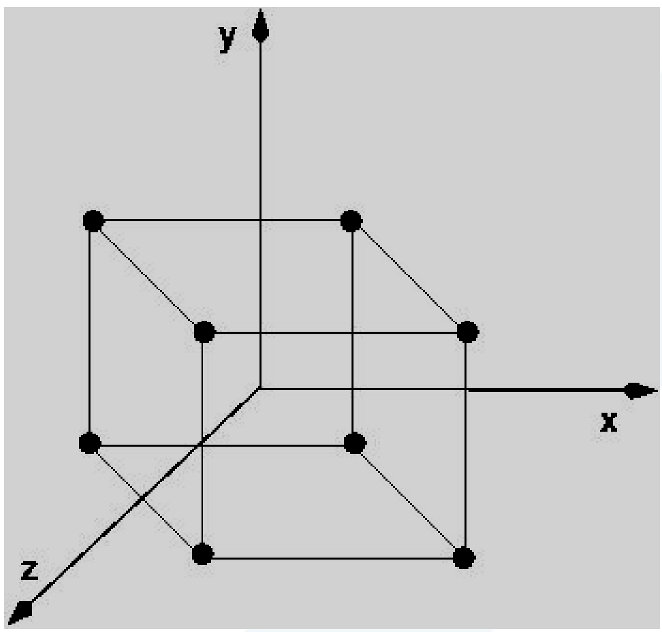


3D

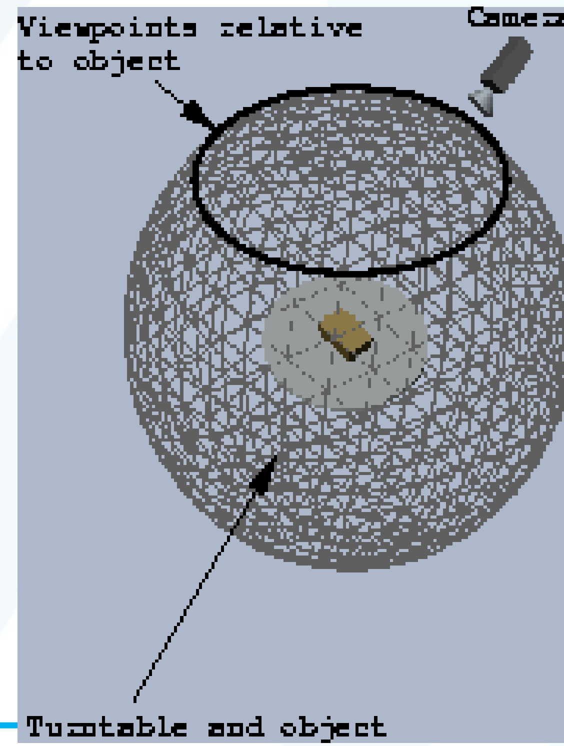


Object Recognition (cont'd)

(1) Object-centered

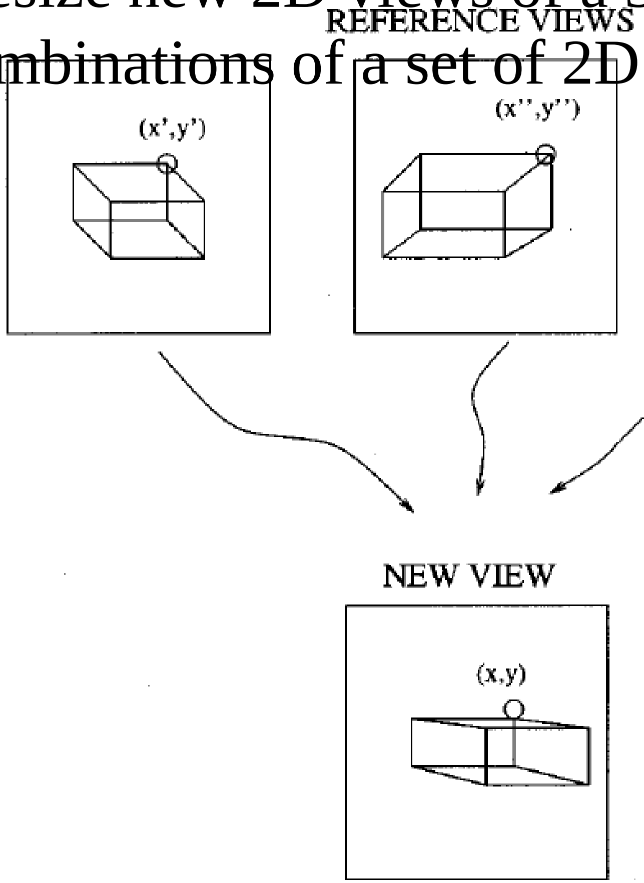


(2) Viewer-centered



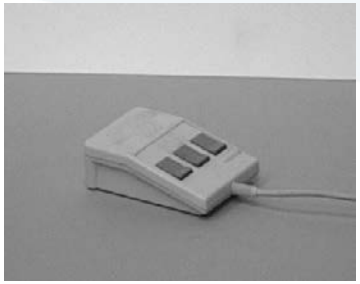
Object Recognition at

Synthesize new 2D views of a 3D object using linear combinations of a set of 2D reference views

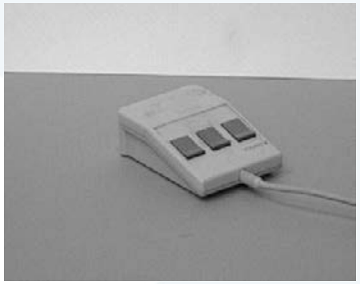


Object Recognition at

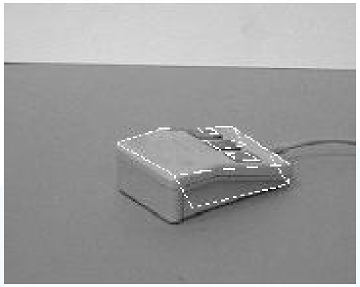
reference view 1



reference view 2



novel view recognized

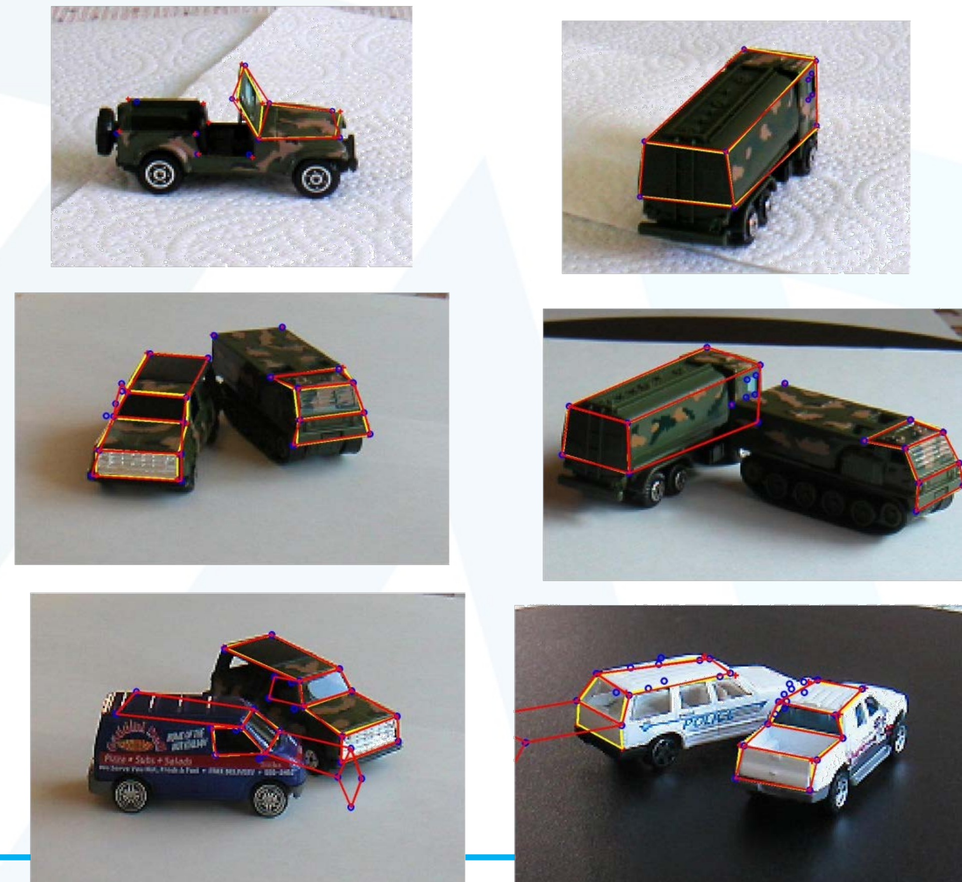


- No 3D models required.
- “Predict” novel 2D views from known 2D views

W. Li, G. Bebis, and N. Bourbakis, "3D Object Recognition Using 2D Views", **IEEE Transactions on Image Processing**, vol. 17, no. 11, pp. 2236-2255, 2008.

Object Recognition at

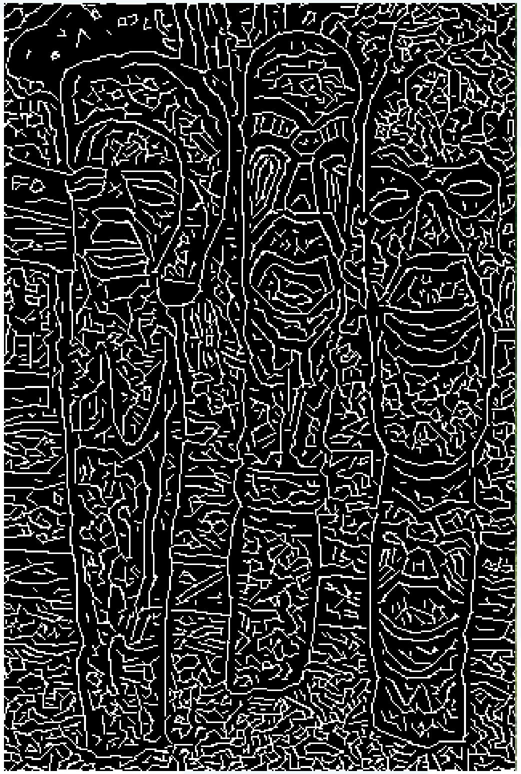
Reference Views

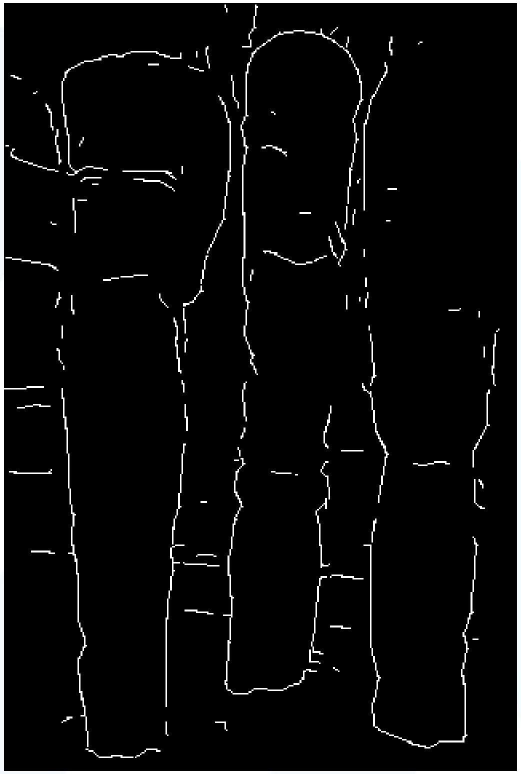


Segmentation

Segmentation at

Iterative Tensor Voting





L. Loss, G. Bebis, M. Nicolescu, and A. Skurikhin, "An Iterative Multi-Scale Tensor Voting Scheme for Perceptual Grouping of Natural Shapes in Cluttered Backgrounds", **Computer Vision and Image Understanding (CVIU),** vol. 113, no. 1, pp. 126-149, January 2009.

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… "

Image Retrieval

- Color, texture

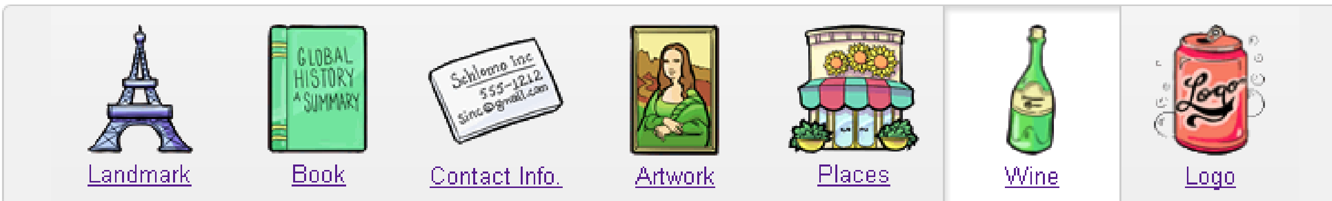
$T = 33.6\text{s}$, found 2 of 2

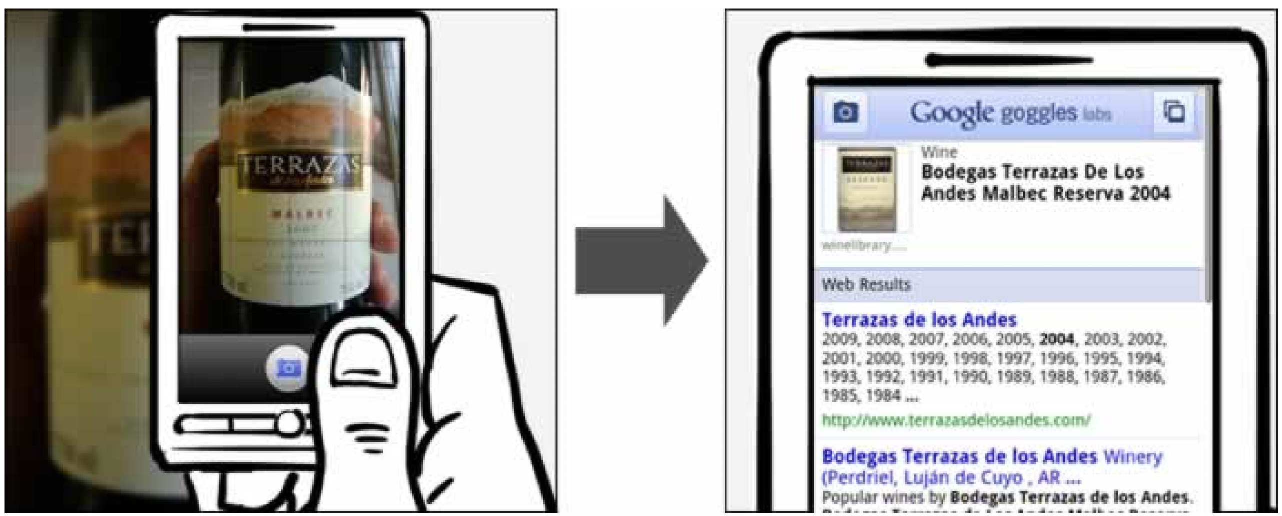
http://corbis.demo.ltutech.com/en/demos/corbis/

Mobile Visual Search:

Google Goggles in Action

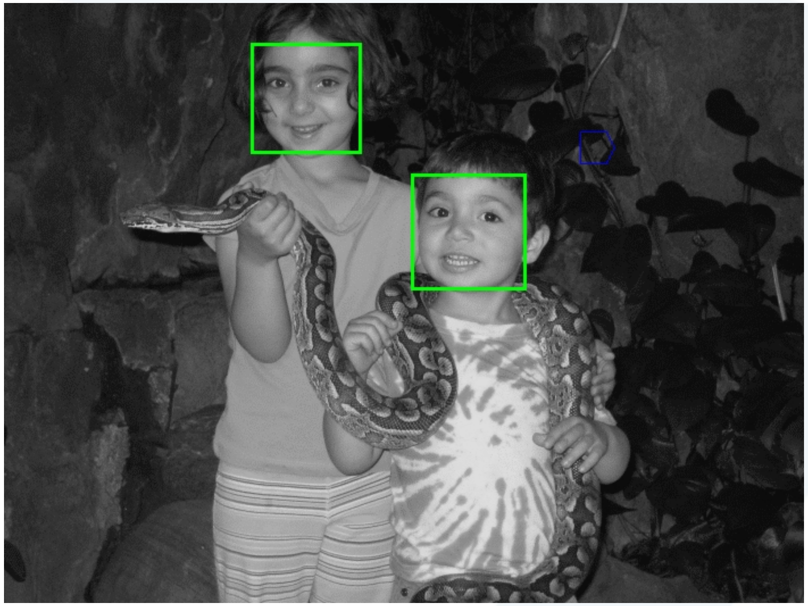
Click the icons below to see the different ways Google Goggles can be used.

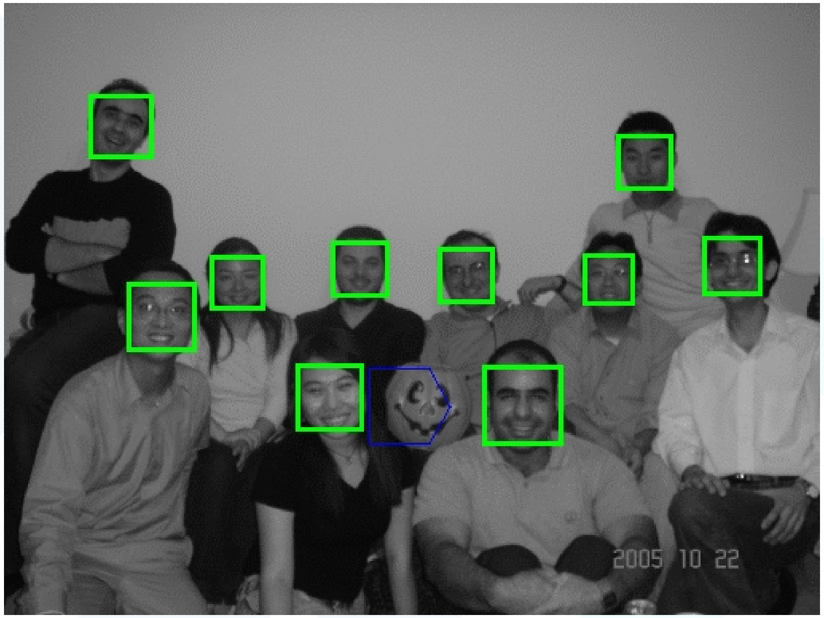




http://www.google.com/mobile/goggles/

Face Detection



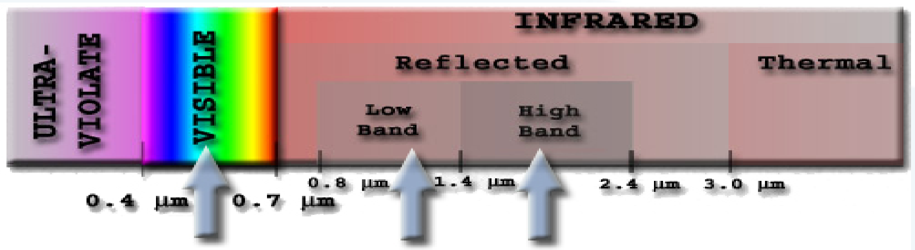


http://www.facedetection.com/

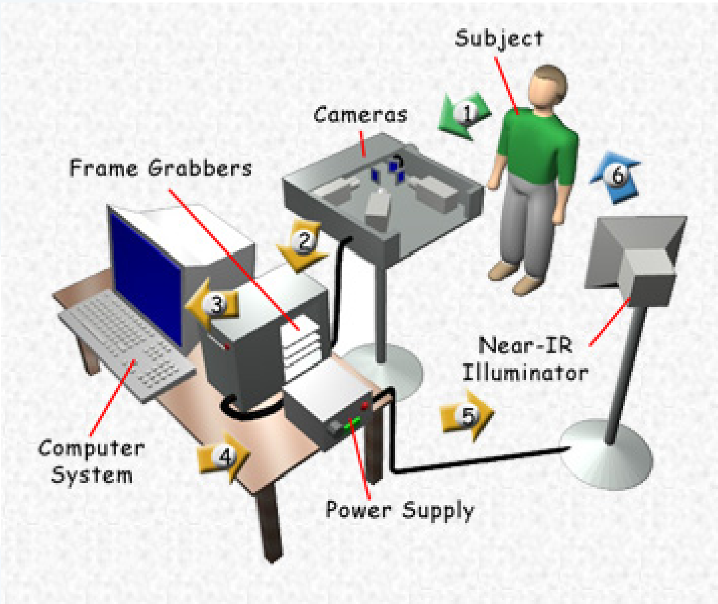
Face Detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

Face Detection at



Human skin exhibits an abrupt change in reflectance around 1.4 mm.



J. Dowdall, I. Pavlidis, and G. Bebis, "[Face Detection in the Near-IR Spectrum]", **Image and Vision Computing,** vol 21, no. 7, pp. 565-578, 2003.

Face Recognition

http://www.face-rec.org/

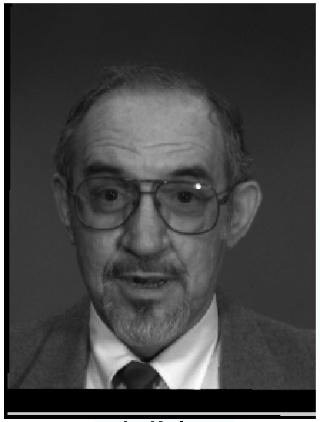
appearance changes

Face Recognition: Apple iPhoto

http://www.apple.com/ilife/iphoto/

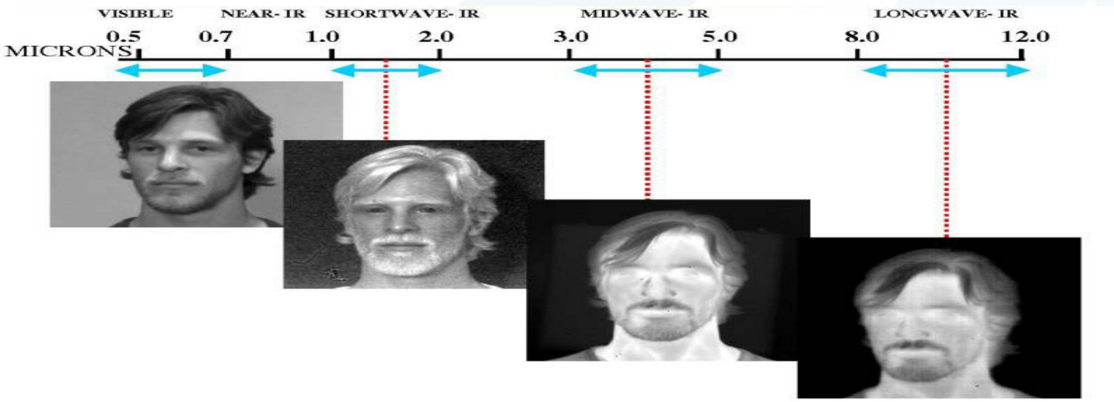
Face Recognition at

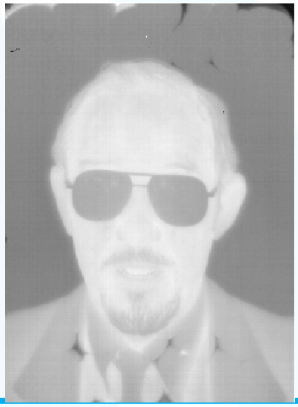
- Visible spectrum
  - High resolution, less sensitive to the presence of eyeglasses.
  - Particularly sensitive to changes in illumination direction and facial expression.



visible

- Thermal IR spectrum
  - Not sensitive to illumination changes.
  - Low resolution, sensitive to air currents, face heat patterns, aging, and the presence of eyeglasses (i.e., IR is opaque to glass).

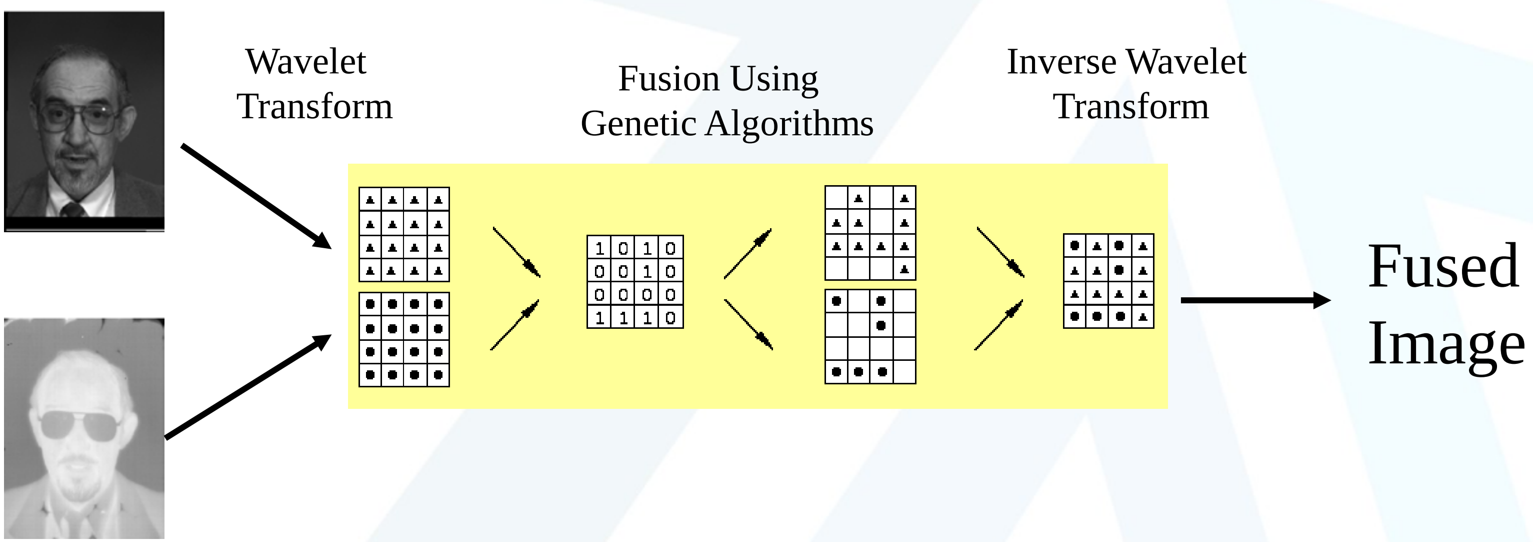




LWIR

Face Recognition at

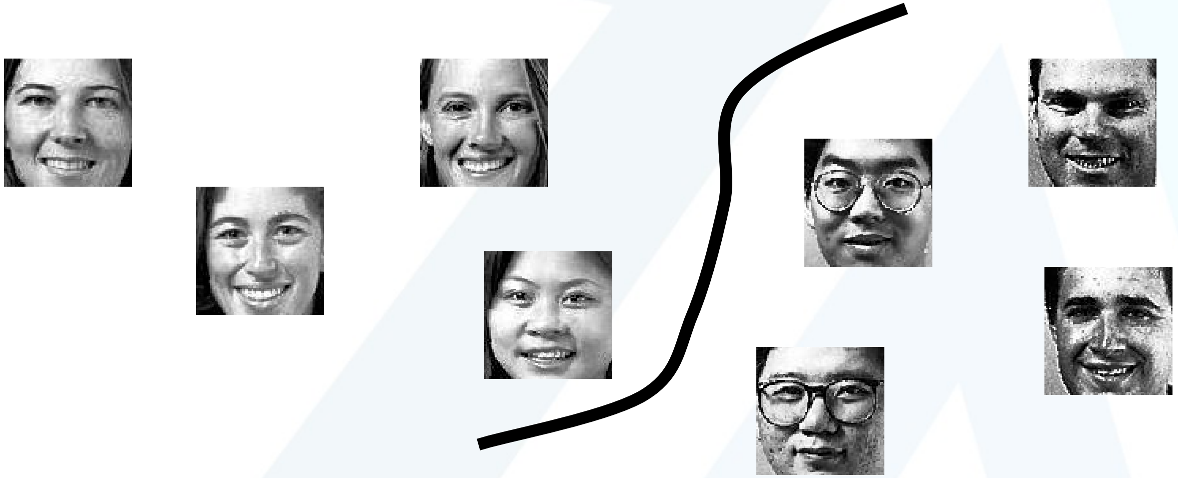
Fuse visible with thermal infrared imagery



G. Bebis, A. Gyaourova, S. Singh, and I. Pavlidis, "Face Recognition by Fusing Thermal Infrared and Visible Imagery", **Image and Vision Computing**, vol. 24, no. 7, pp. 727-742, 2006.

Gender Classification at

Discover gender-specific features using Genetic Algorithms (GAs)



Z. Sun, G. Bebis, X. Yuan, and S. Louis, "Genetic Feature Subset Selection for Gender Classification: A Comparison Study", **IEEE Workshop on Applications of Computer Vision**, pp. 165-170, 2002.

Gender Classification at

Original images

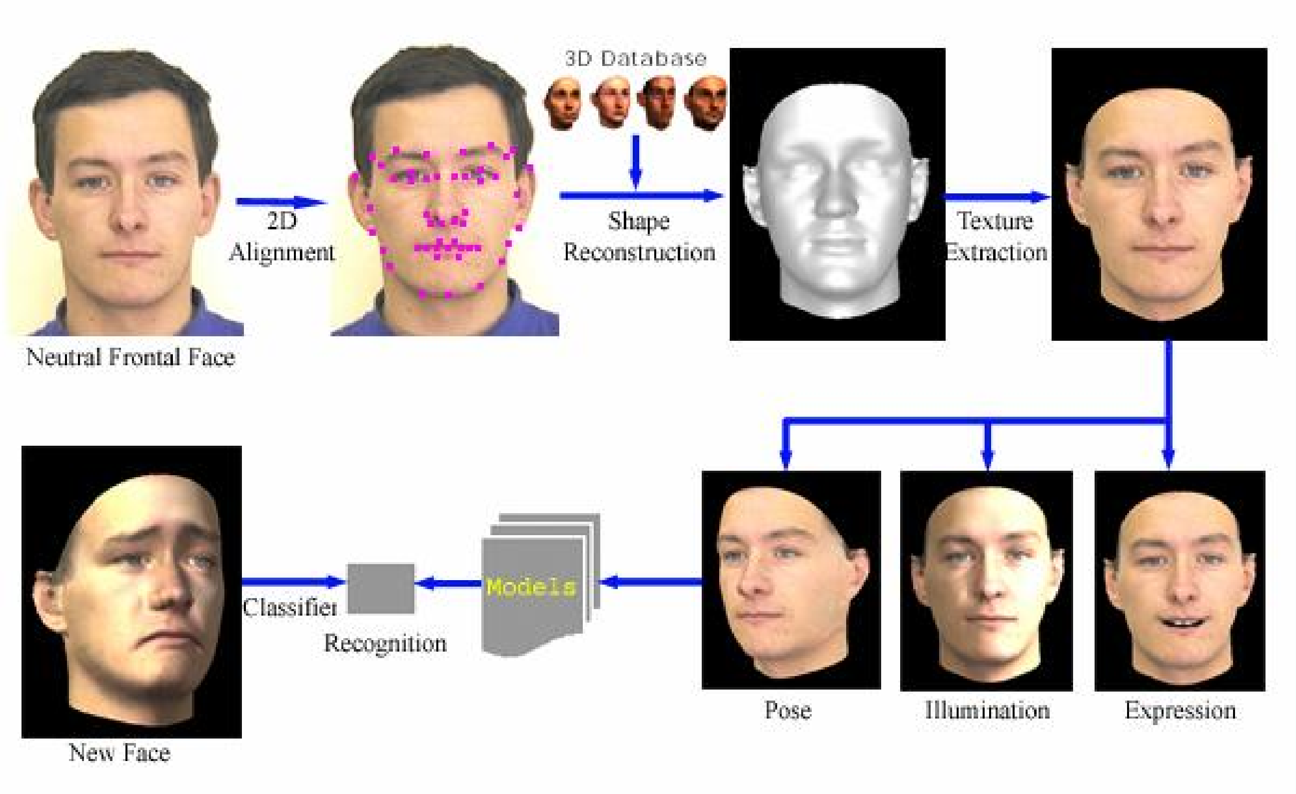
Reconstructed using traditional features

Reconstructed using GA-based features

Preserve gender-related information but not identity specific features!

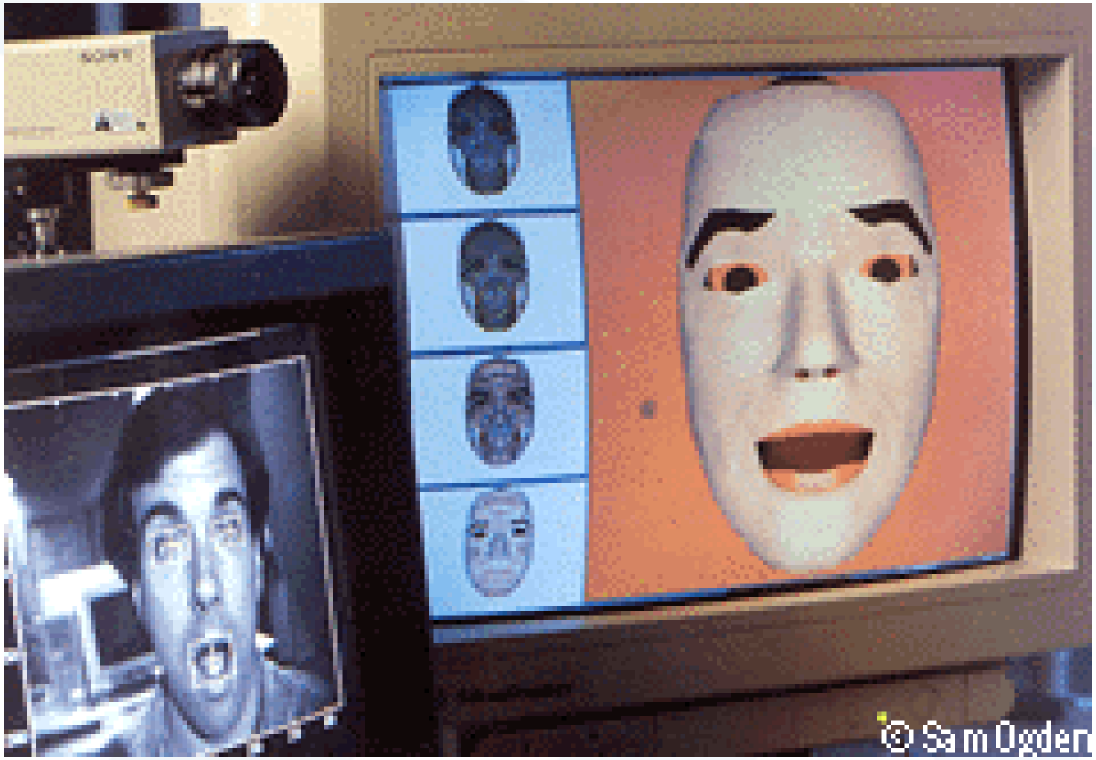
Z. Sun, G. Bebis, and R. Miller, "Object Detection Using Feature Subset Selection", **Pattern Recognition**, vol. 37, pp. 2165-2176, 2004.

3D Face Recognition



Demo: http://www.youtube.com/watch?v=VuGvlMB13pw

Facial Expression Recognition



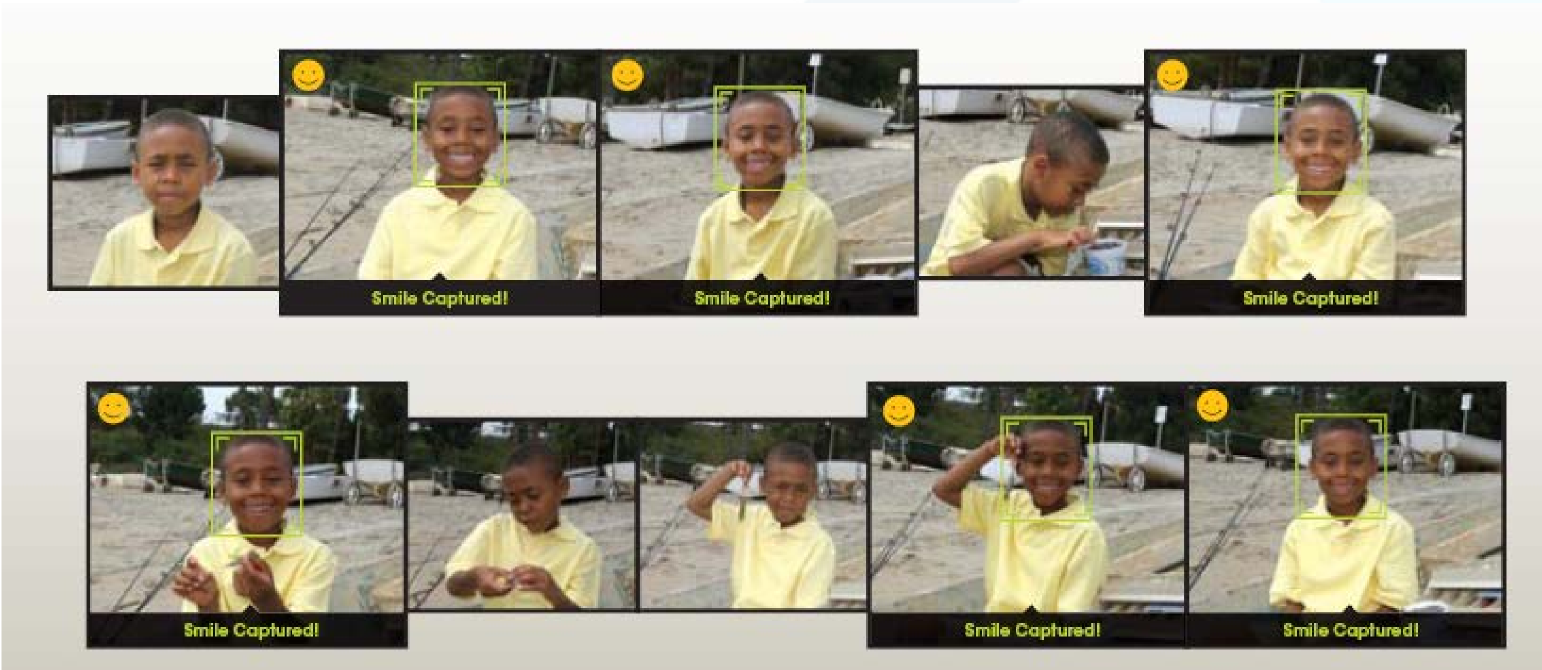
http://www.youtube.com/watch?v=M1WgnisIyPQ&feature=related

Smile detection?

جَامعة
المَنارة
MANARA UNIVERSITY

The Smile Shutter flow

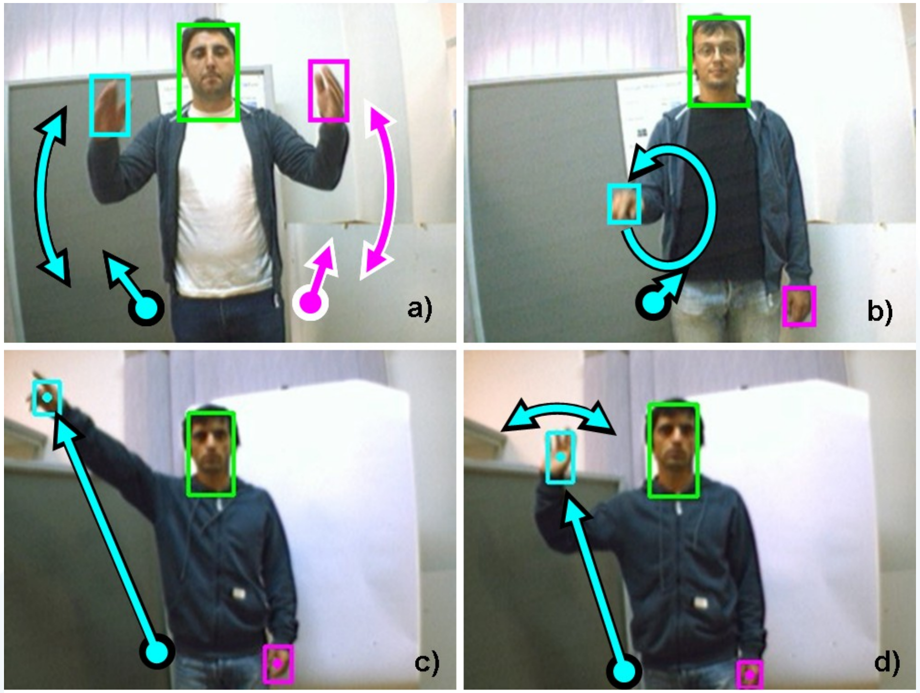
Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



Sony Cyber-shot® T70 Digital Still Camera

Hand Gesture Recognition

- Smart Human-Computer User Interfaces
- Sign Language Recognition



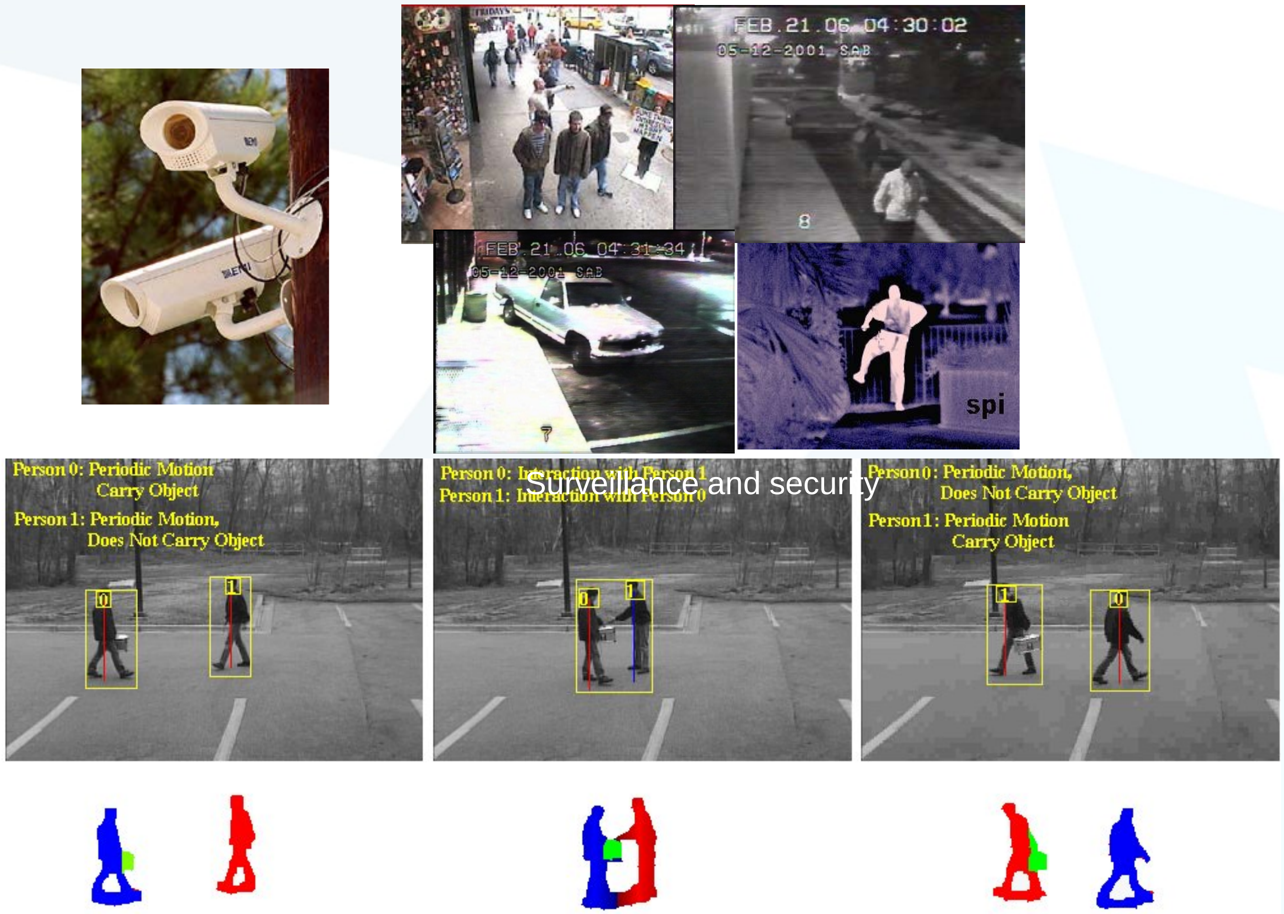
Vision-based Interaction and Games

Kinect

Assistive technologies

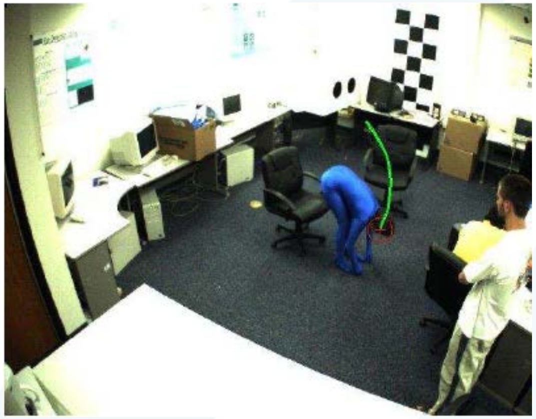
Nintendo Wii has camera-based IR tracking built in. See Lee's work at CMU on clever tricks on using it to create a multi-touch display!

Visual Surveillance and Human Activity Recognition



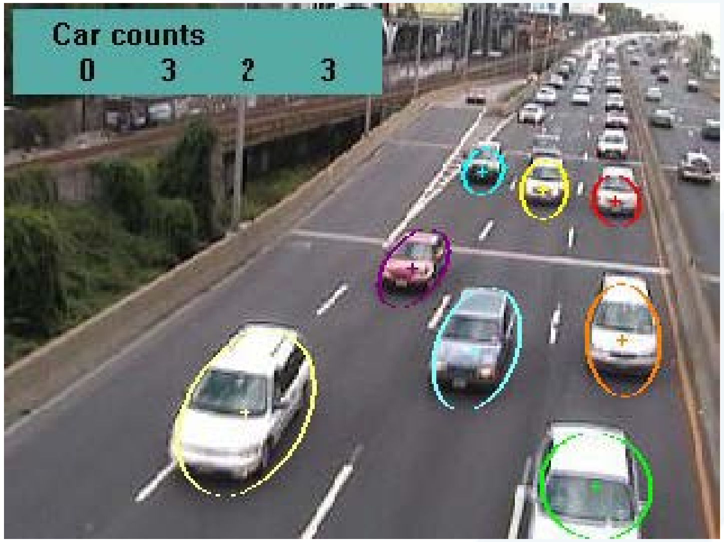
Human Activity Recognition at

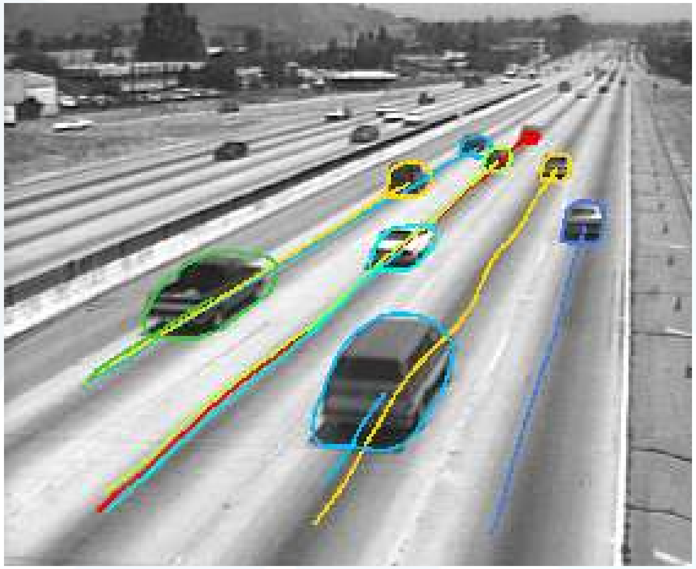
- Recognize simple human actions using 3D head trajectories



J. Usabiaga, G. Bebis, A. Erol, Mircea Nicolescu, and Monica Nicolescu, "Recognizing Simple Human Actions Using 3D Head Trajectories", **Computational Intelligence** (special issue on Ambient Intelligence), vol. 23, no. 4, pp. 484-496, 2007.

Traffic Monitoring





http://www.honeywellvideo.com/

Vehicle Detection and Tracking at

’s low light camera

Ford’s Concept Car

Vehicle Detection and Tracking

- Can process 10 fps on average; 6% errorrs (FP + FN)

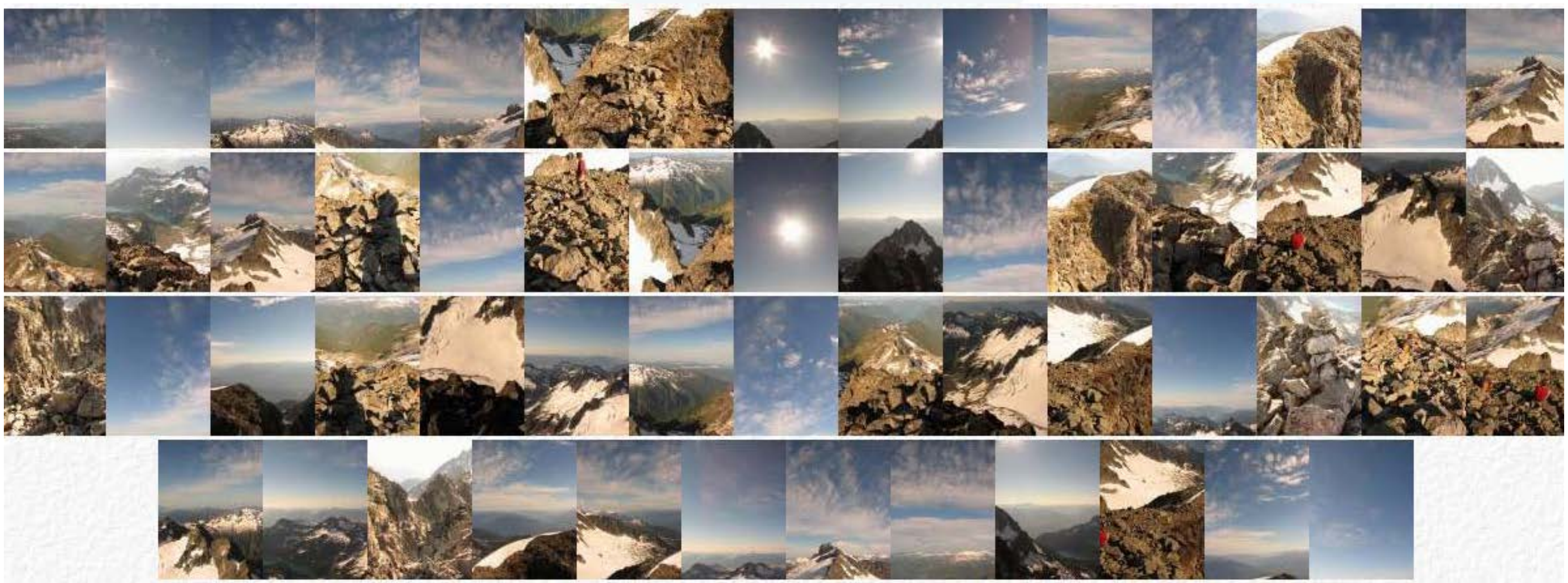
Z. Sun, G. Bebis, and R. Miller, "Monocular Pre-crash Vehicle Detection: Features and Classifiers", **IEEE Transactions on Image Processing** , vol. 15, no. 7, pp. 2019-2034, July 2006.

Smart cars:

Mobileye

- Vision systems currently in high-end BMW, GM, Volvo models.

Automatic Panorama Stitching



Automatic Panorama Stitching (cont'd)

3D urban modeling: Photosynth

http://photosynth.net/

Photosynth allows you to take a bunch of photos of the same scene or object and automatically stitch them all together into one big interactive 3D viewing experience

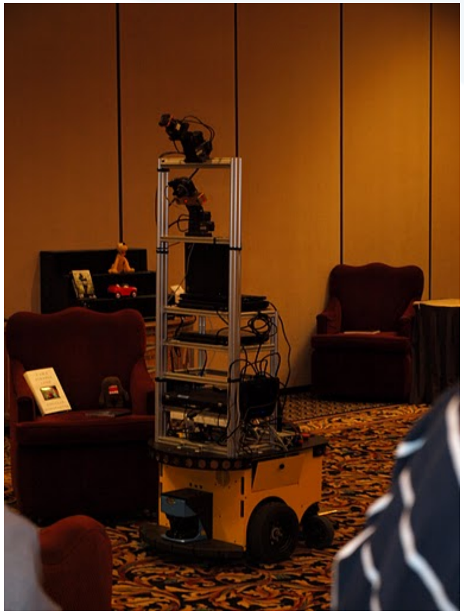
Automatic 3D reconstruction from internet photo collections

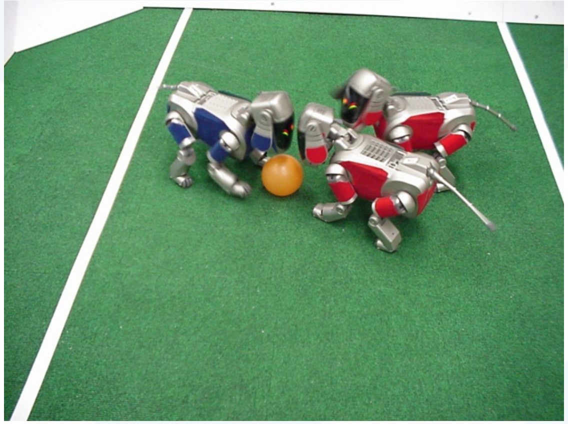
“Statue of Liberty” “Half Dome, Yosemite” “Colosseum, Rome”

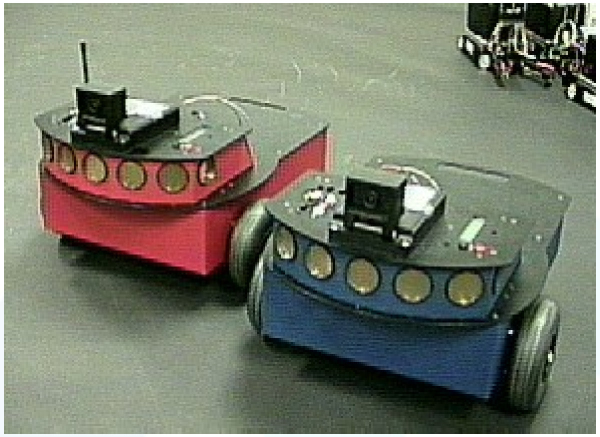
Flickr photos

3D model

Robotics







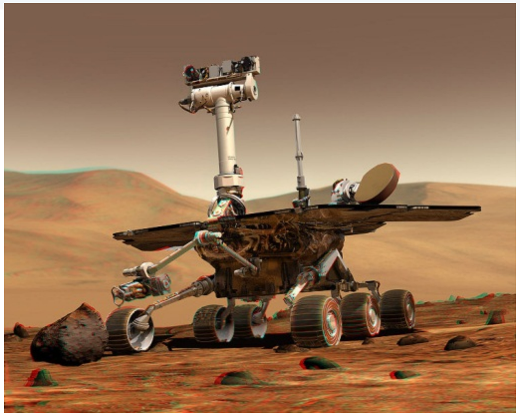
http://www.robocup.org/

Semantic Robot Vision Challenge

http://www.semantic-robot-vision-challenge.org/

http://www.youtube.com/watch?v=GItjILILB50

Vision in space



NASA'S Mars Exploration Rover Spirit

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

Movie Special Effects

- Insert synthetic objects in real image sequences;
- Change artificially the position or the orientation of a camera;
- Freeze a moving 3D scene.

Medical Imaging

Skin/Breast Cancer Detection

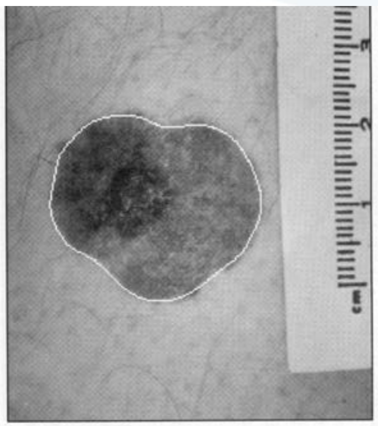
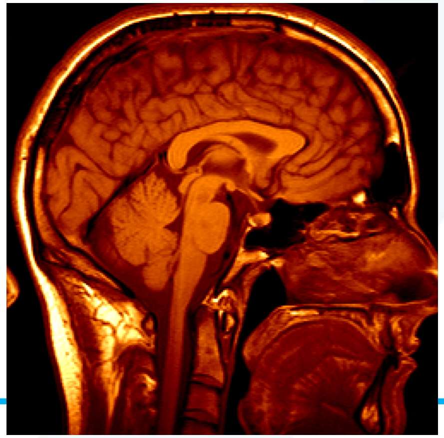


Image guided surgery
[Grimson et al., MIT]

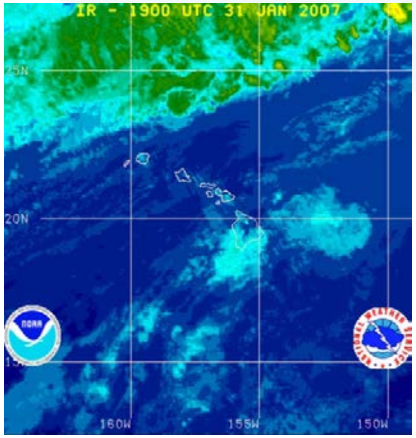
Enable surgeons to visualize internal structures through an automated overlay of 3D reconstructions of internal anatomy on top of live video views of a patient.

3D imaging
MRI, CT



Other Scientific Applications

Weather



Aerial/Satellite

Astronomy

Computer Vision Jobs !!

- Academia
  - MIT, UC-Berkeley, CMU, UIUC, USC …… UNR!
- National Labs and Government
  - Los Alamos National Lab
  - Lawrence Livermore National Lab
  - Navy, Air-force, Army
- Industry
  - Microsoft, Intel, IBM, Xerox, Compaq, Siemens, HP, TI, Motorola, Phillips, Honeywell, Ford

See: http://www.cs.ubc.ca/spider/lowe/vision.html

What skills do you need to succeed in this field?

- Strong programming skills (i.e., C, C++, Matlab,Python)
- Good knowledge of Data Structures and Algorithms
- Good skills in analyzing algorithm performance (i.e., time and memory requirements).
- Good background in mathematics, especially in:
  - Linear Algebra
  - Probabilities and Statistics
  - Numerical Analysis