CSE 152: Computer Vision

Hao Su

Lecture 0: Introduction

Credit: Manmohan Chandraker

Defining computer vision

Wall-E: Fact and Fiction (Minh Do, Princeton University)

Defining computer vision

- Old: Computer programs that can
  - Process image information
  - Recognize instances of objects
  - Find distances of objects

- Modern: Understanding the world based on visual cues
  - Determining factors that govern image formation
  - Recognition across variations
  - Estimate semantic properties of a scene
  - Recognize complex actions
  - Predict long-term behaviors

Studying computer vision

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Studying computer vision

- Images are everywhere around us

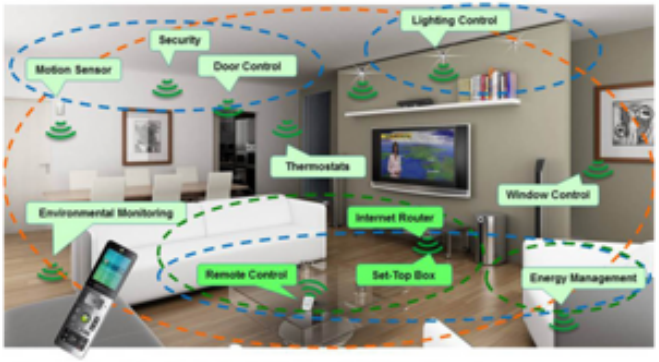
Source: Domo

Studying computer vision

- Images are everywhere around us
- Rapidly emerging technologies

Autonomous driving

Gaming

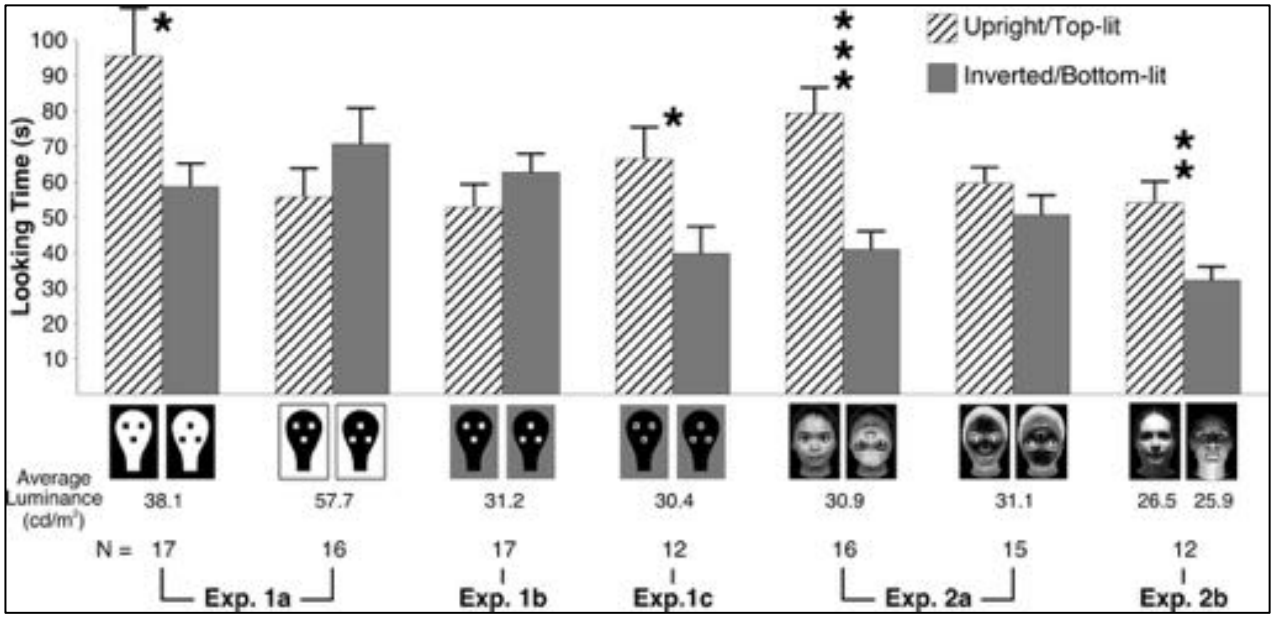


Smart homes

Factory automation

Studying computer vision

- Images are everywhere around us
- Rapidly emerging technologies
- Deep and attractive scientific problems
  - How do we recognize objects?
  - Why do newborn babies respond to face-like shapes?
  - Beautiful marriage of math, physics, biology, CS, engineering



[Farroni et al., 2005]

We Use Computer Vision

Computer vision in living rooms

Microsoft Kinect Xbox

Sportvision first down line

Vision to explore the world

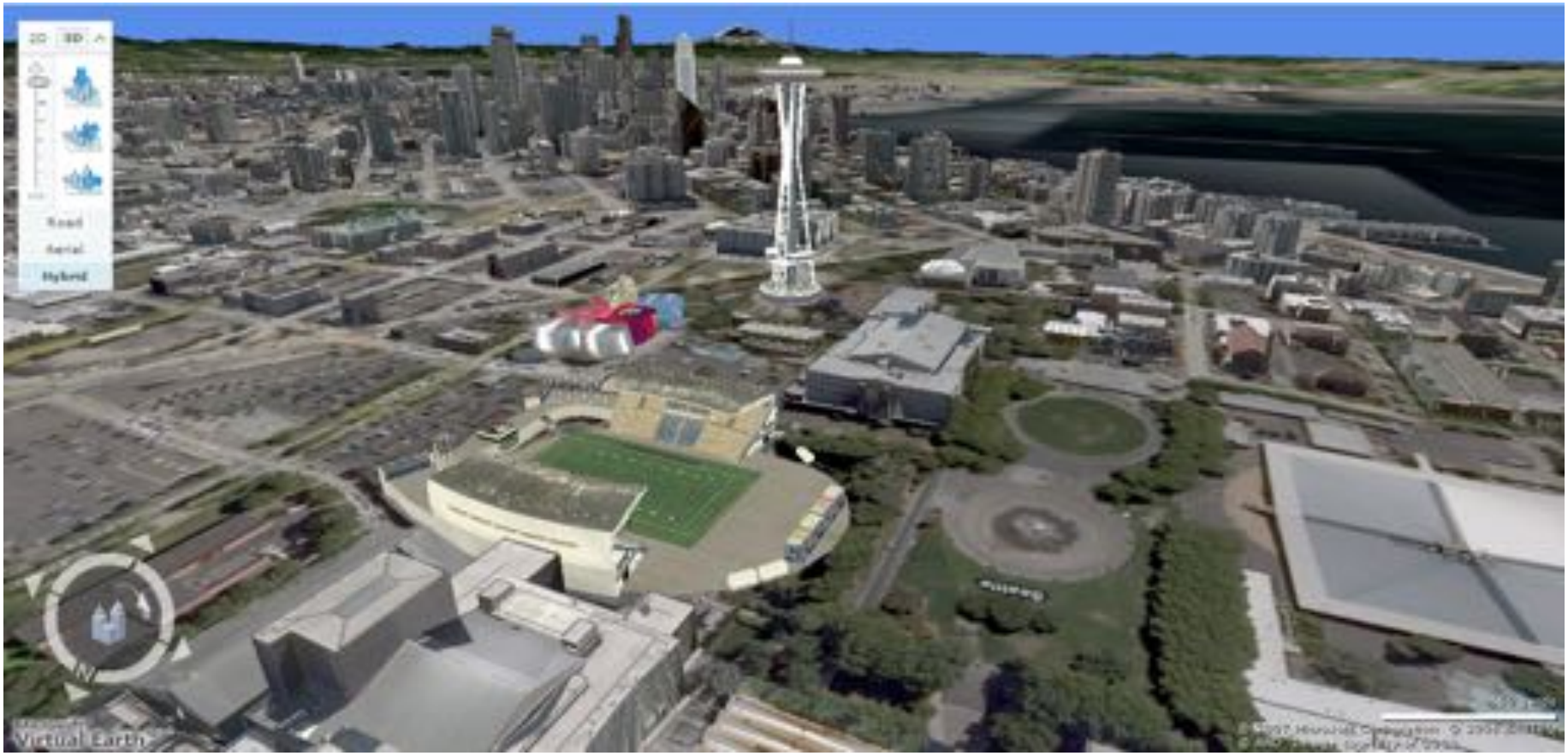


Image from Microsoft Virtual Earth

Vision to explore other worlds

Image from NASA's Mars Exploration Rover Spirit

- Panorama stitching
- Stereo imaging
- Navigation
- ….

Vision to explore all worlds

Including virtual ones!

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Organizing Computer Vision

Broad classes of vision applications

Sense

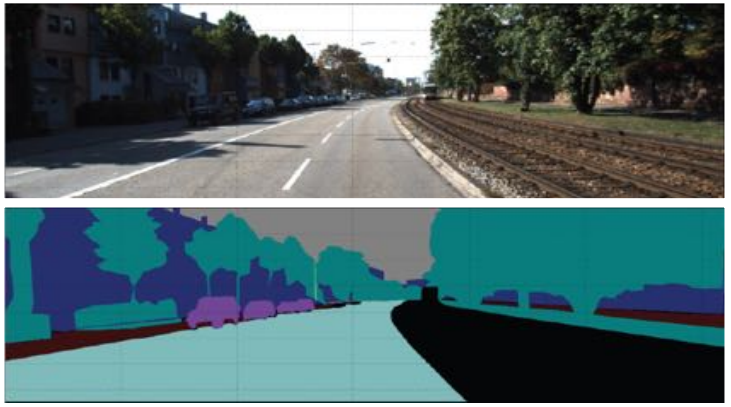
Understand

Interface

Reconstruct

Recognize

Reorganize



Broad classes of vision applications

Sense

Understand

Interface

Scenes

People

Broad classes of vision applications

Sense

Understand

Interface

Human-Human

Human-Machine

Machine-Machine

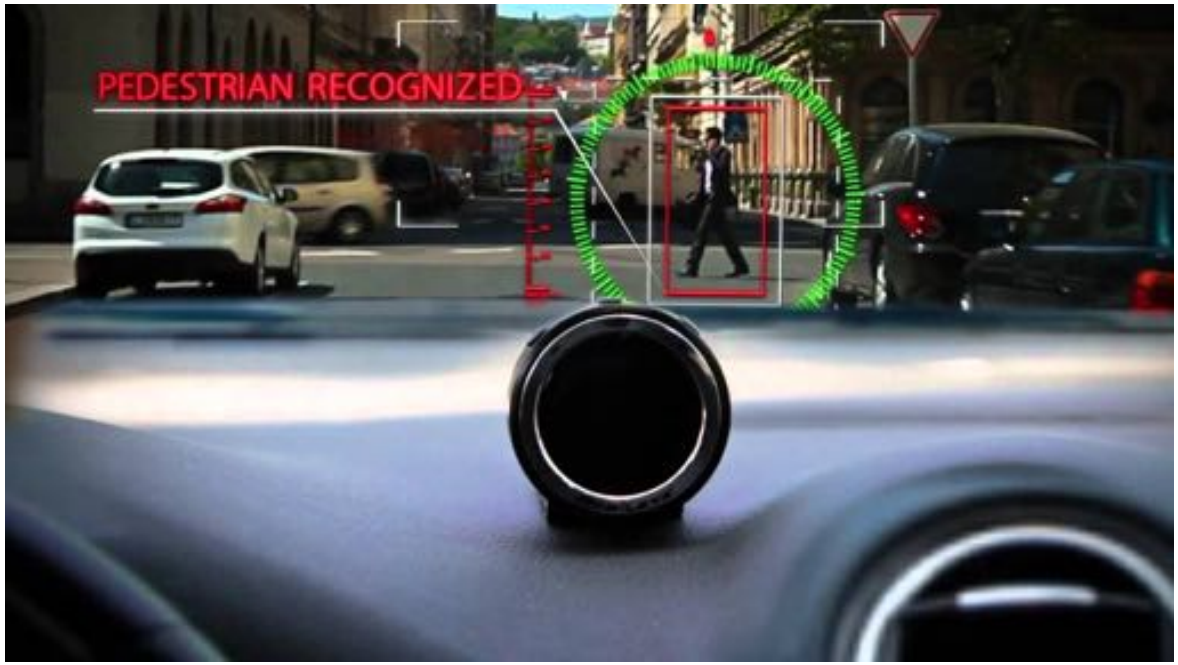
Significant progress in recent years

Sense

Understand

Interface

Advanced Driver Assistance Systems

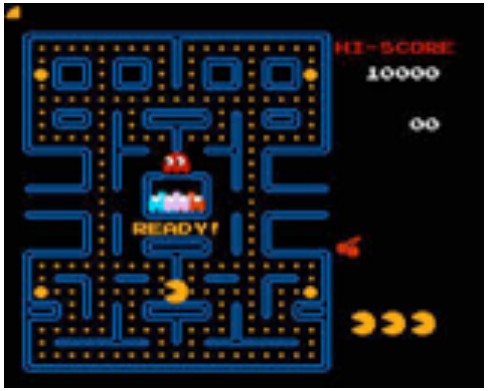


Deep learning is revolutionizing AI

Tic-tac-toe (1952)

Checkers (1994)

Chess (1997)



Atari (2015)

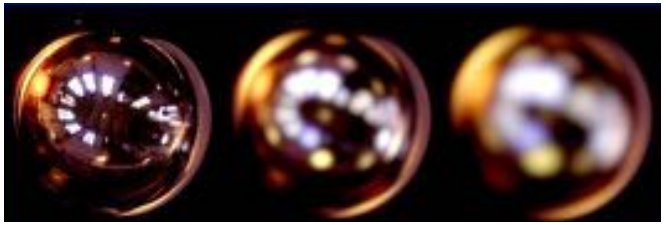
Go (2016)

Computer vision is also riding the wave

- Autonomous driving (Google, Tesla, Mobileye, ...)
- Augmented reality (HoloLens, Oculus, MagicLeap)
- Social networks (Google, Facebook, ...)
- Mobile applications
- Surveillance

Augmented Reality

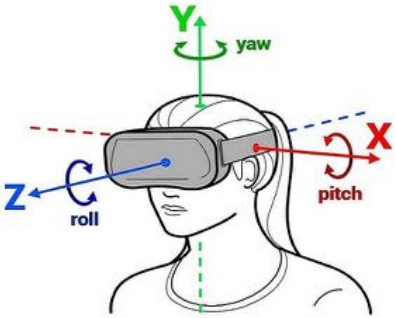
Vision in augmented reality devices



Material and lighting estimation

Gaze tracking

Depth estimation



Head pose estimation

Semantic segmentation

Object detection

Autonomous Driving

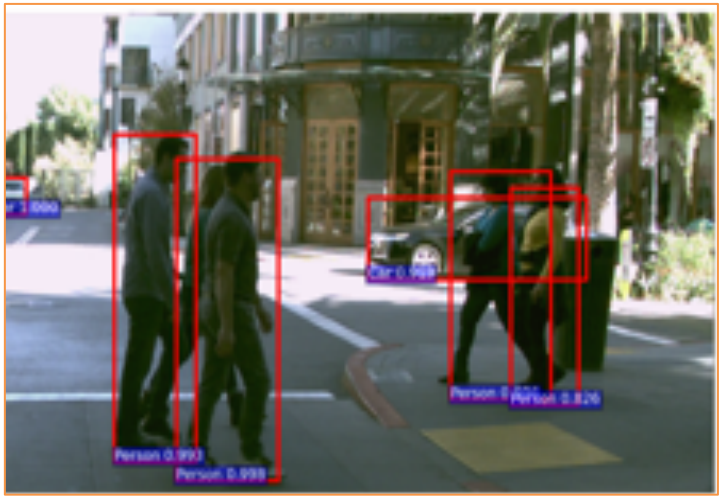
Autonomous navigation

Source: Wired

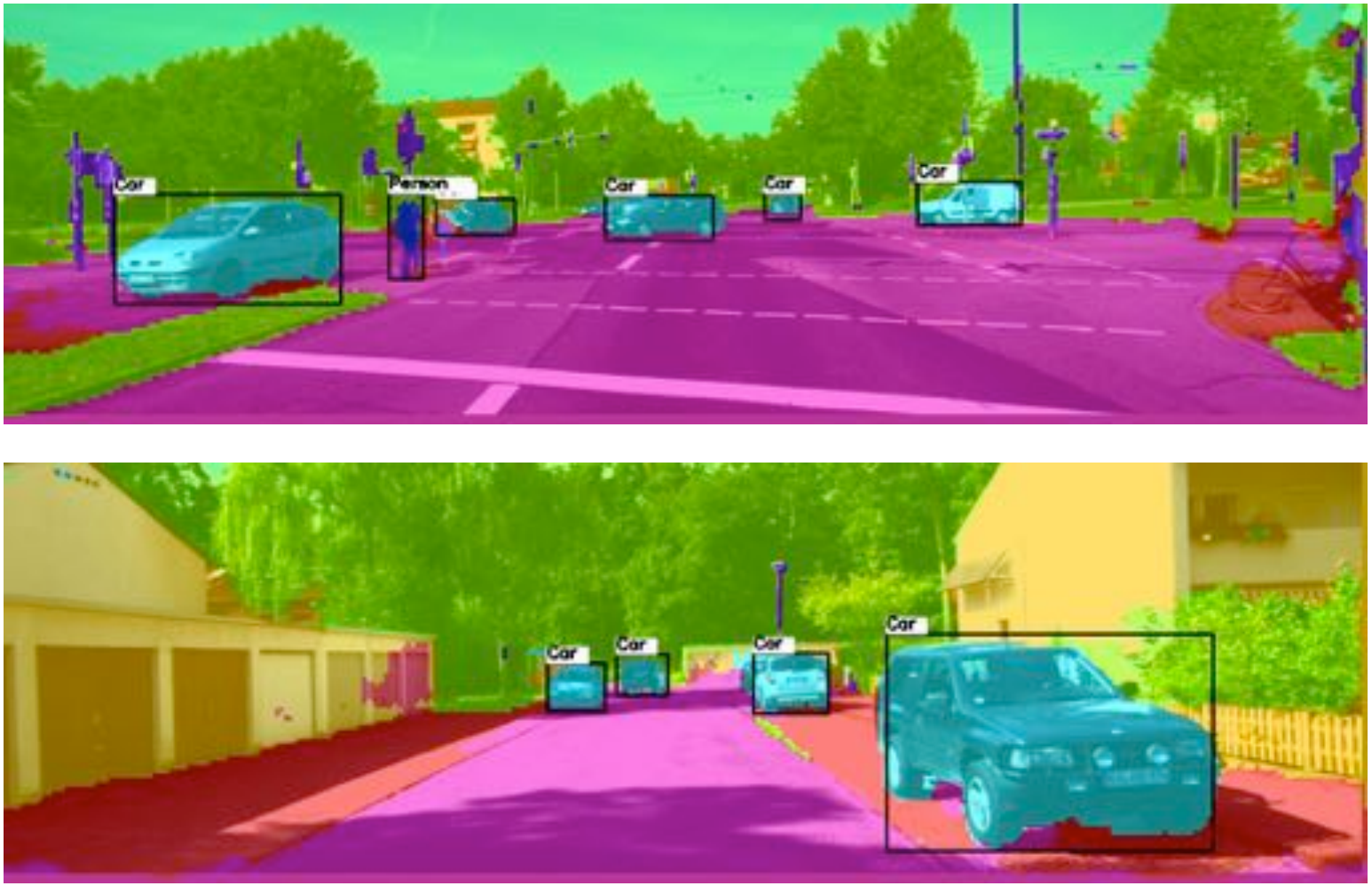
The hardness of the problem

- Finding locations
- Localize objects
- Estimate distances
- Understand relations
- Be aware of traffic rules
- Predict future behaviors
- Understand intentions
- Interdependent decisions

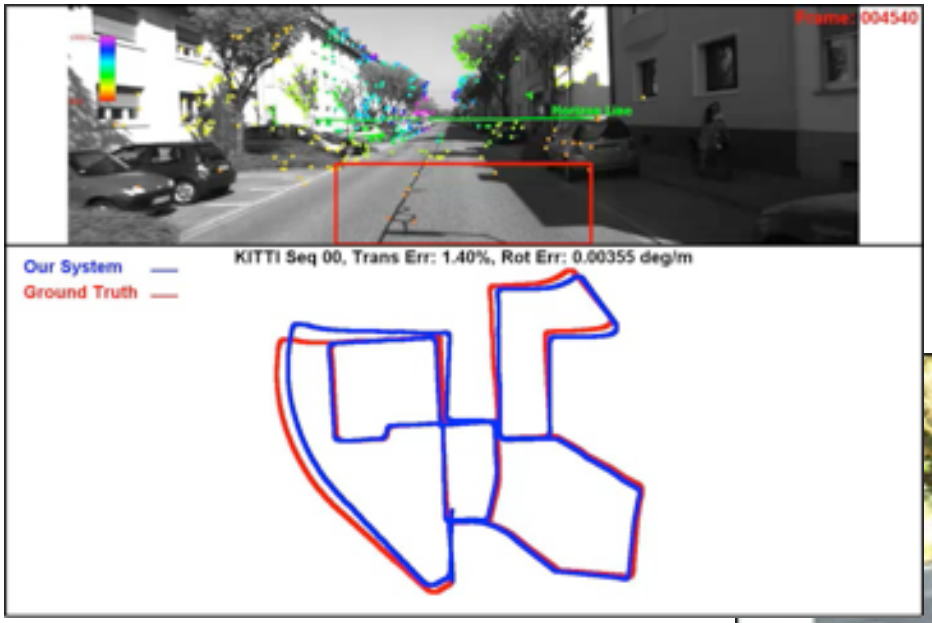
Object detection

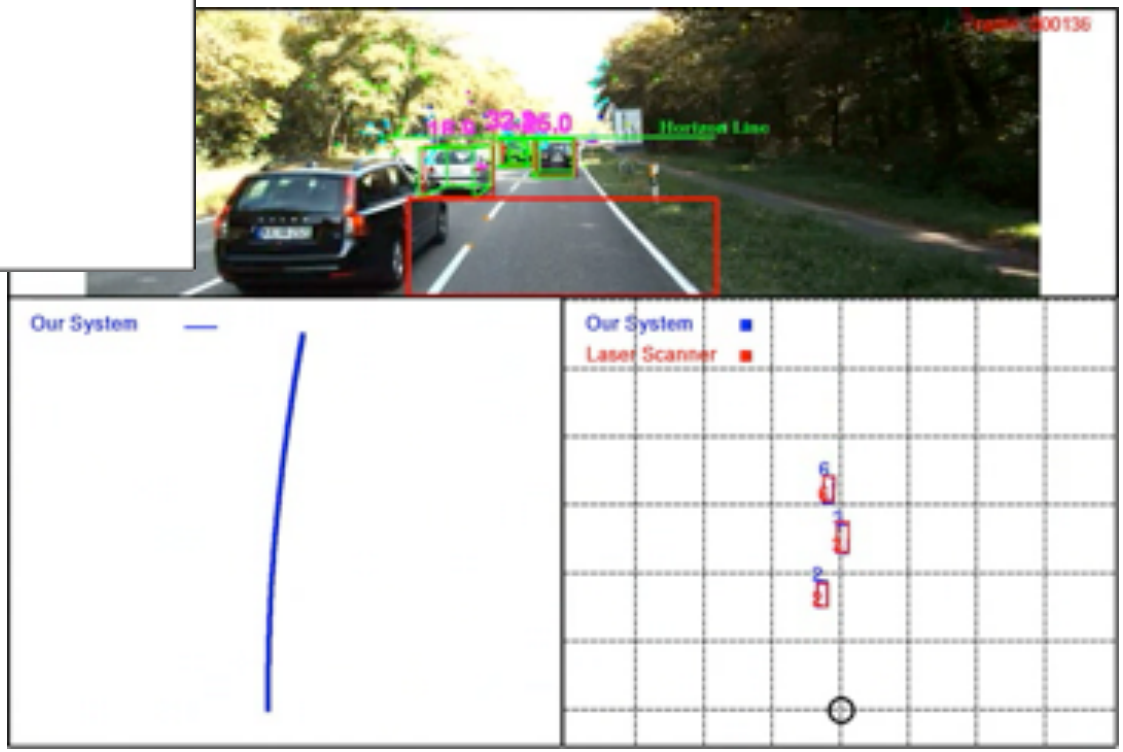


Semantic segmentation



Real-time navigation





Song and Chandraker, CVPR 2014

Future behavior prediction

Plenty of Other Applications

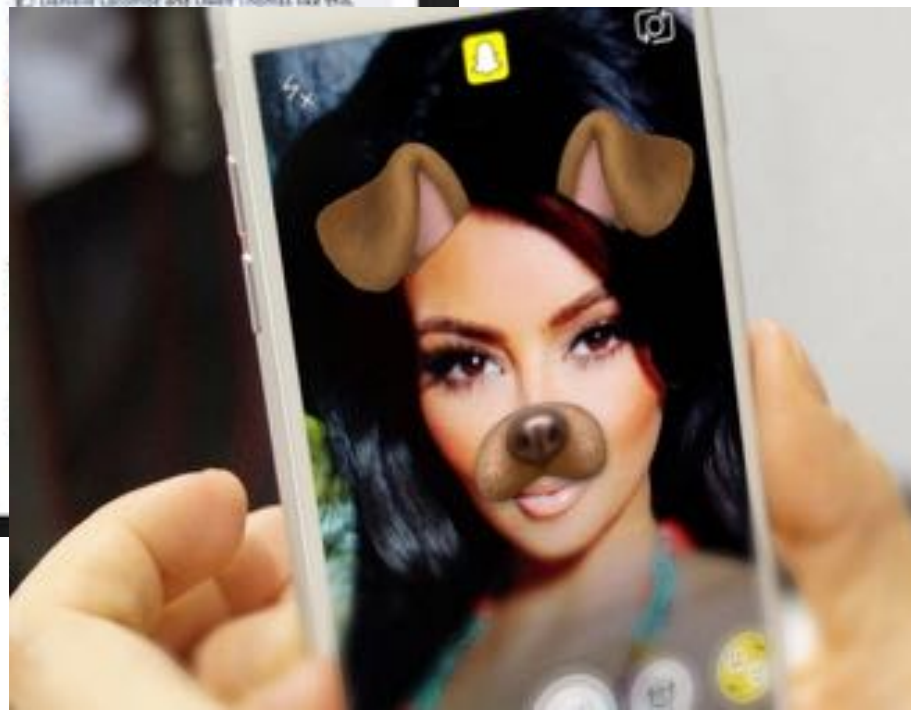
Mobile phones and tablets

Face recognition

Place recognition

Surveillance

Social media



A Few Challenges in Computer Vision

Why is computer vision difficult?

Viewpoint

Scale

Lighting

Deformation

Why is computer vision difficult?

Intra-class variation

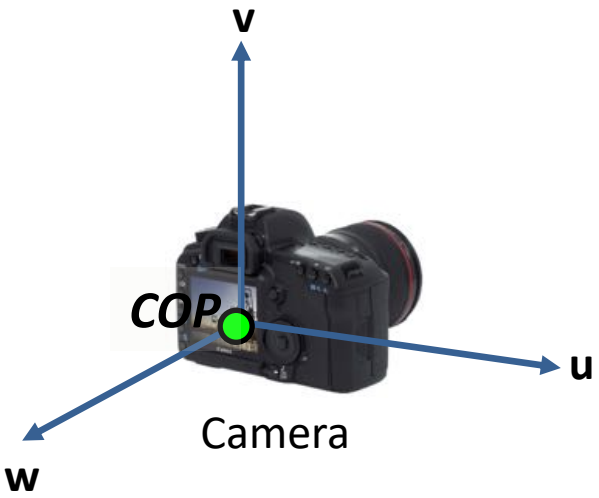
Motion (Source: S. Lazebnik)

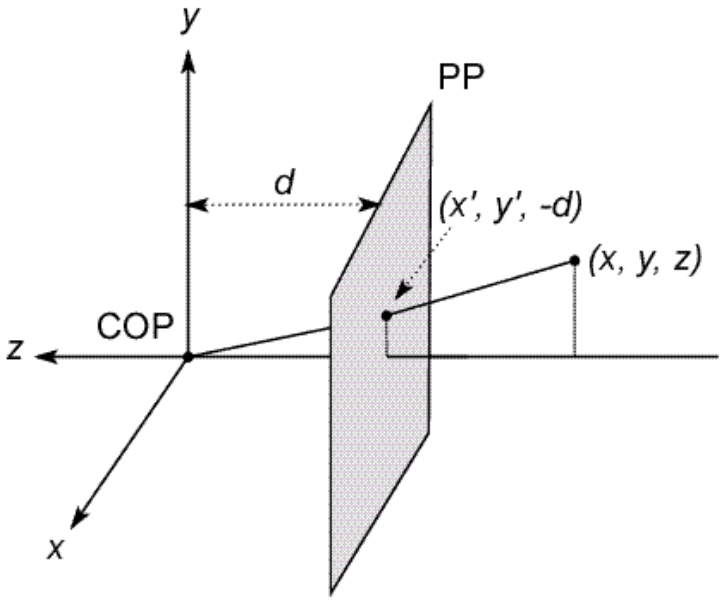
Background clutter

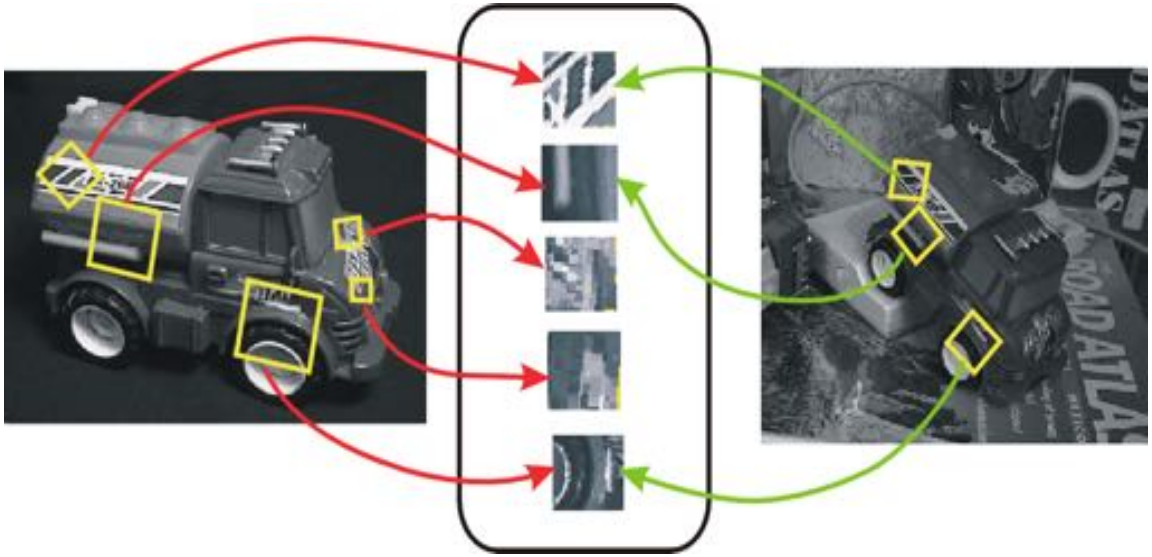
Occlusion

Building Models for Computer Vision

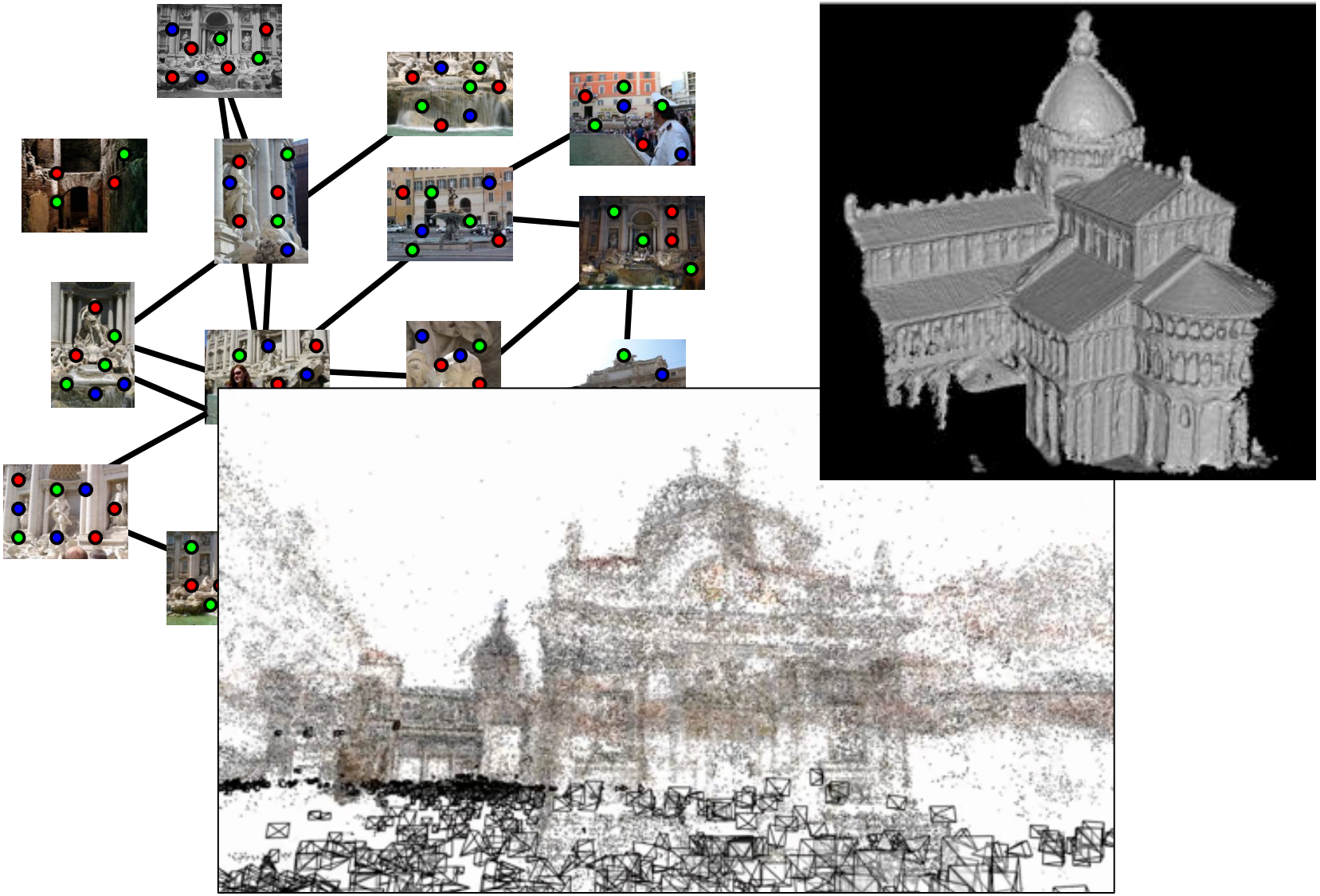
Representation of images



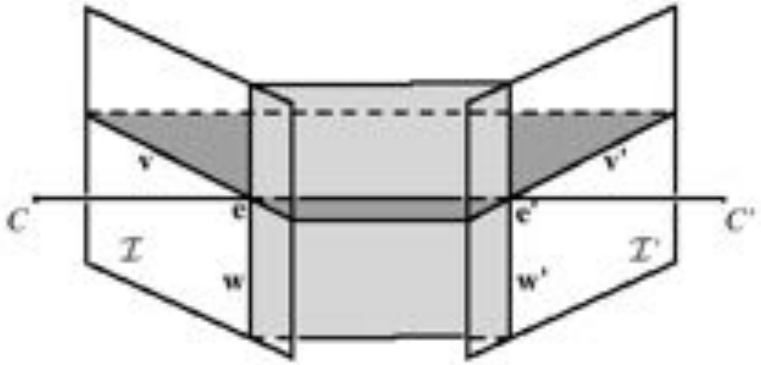


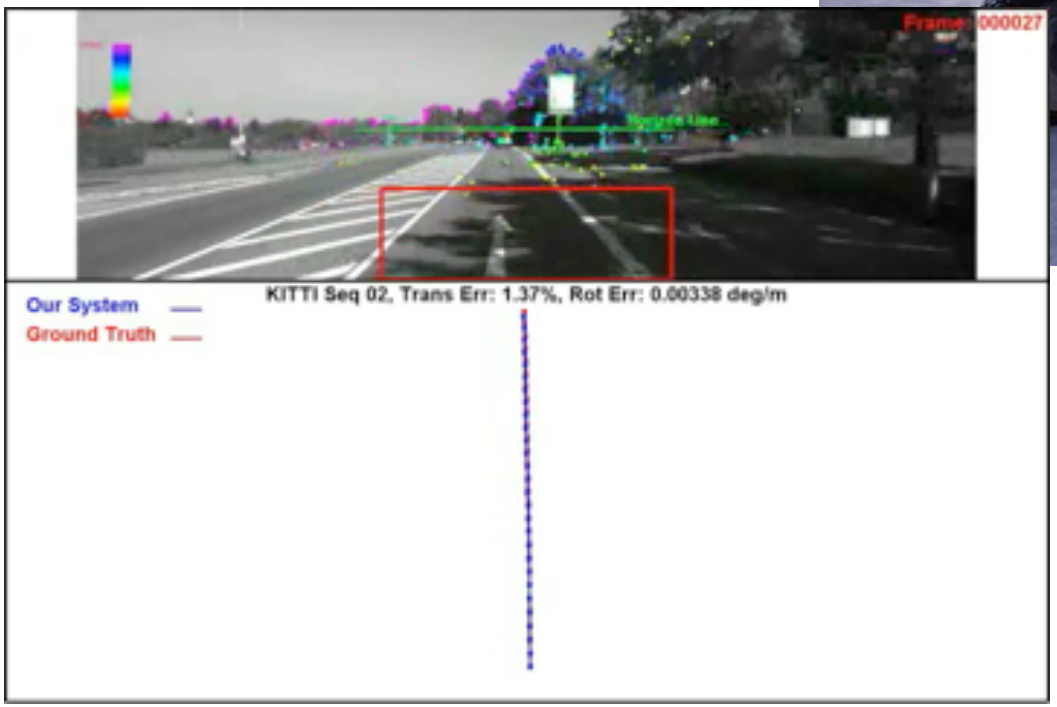


Estimate 3D structure from images

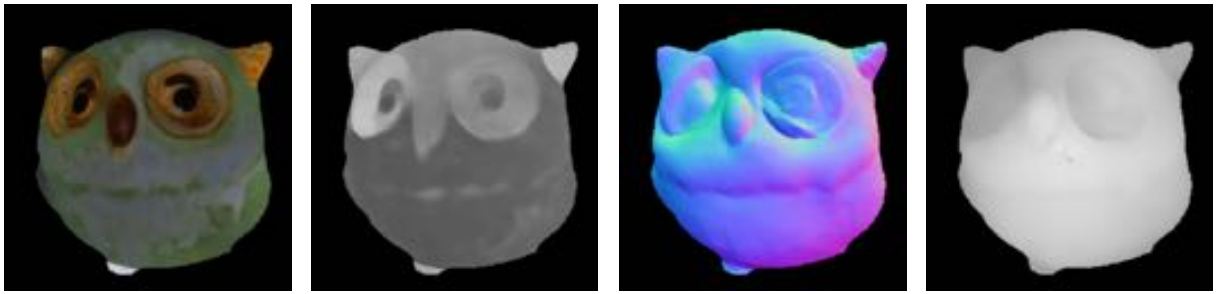


Estimate camera motion





Estimate lighting and material



Machine Learning

Recognition






[Raquel Urtasun]

Machine learning

- Typically in CS: write a program to execute a set of rules
- Computer vision: sometimes very hard to specify rules
- Machine learning: develop own program based on examples
- Training data: input-output pairs



[Raquel Urtasun]

So what does recognition involve?

Verification: is that a bus?

Detection: are there cars?

Identification: is that a picture of Mao?

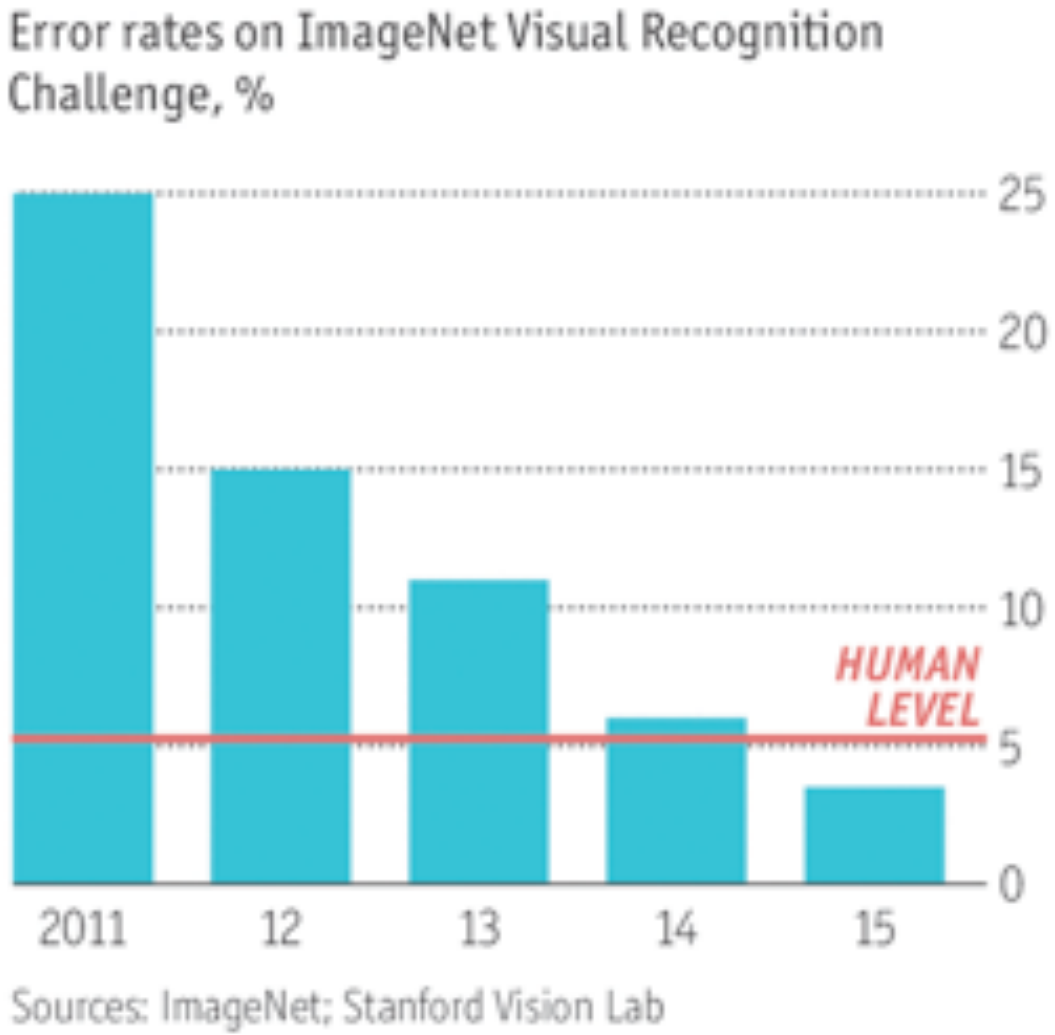
Object categorization

Scene categorization

Machine learning is a key player

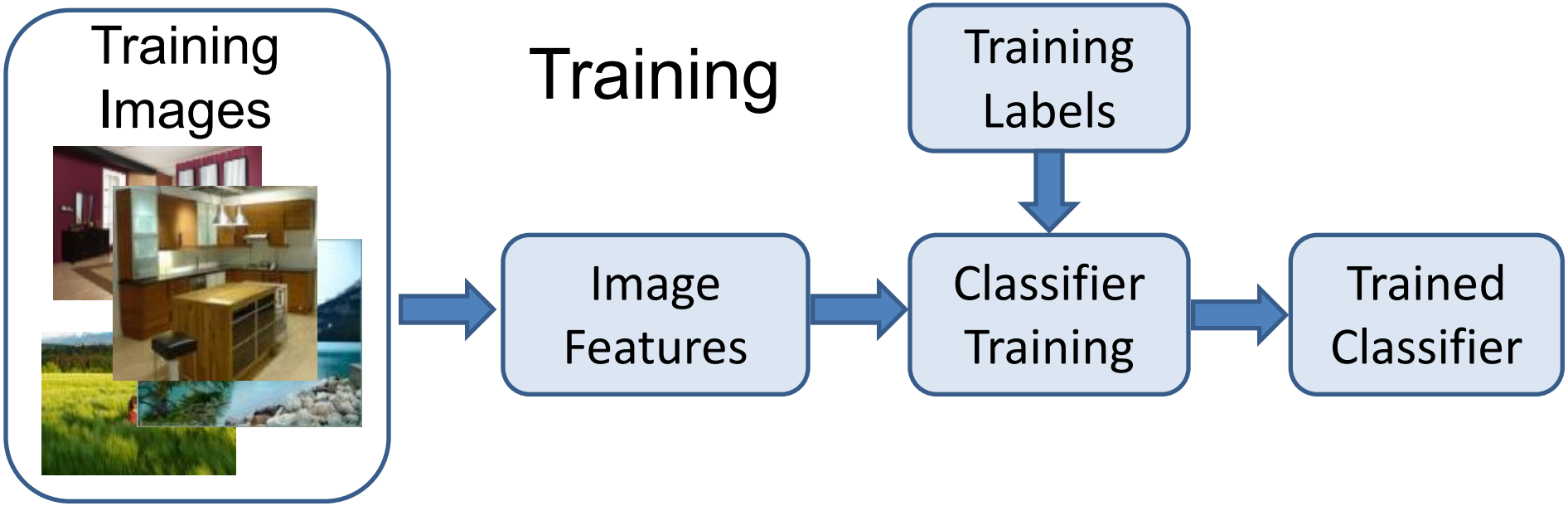
- What is it?
  - Object and scene recognition
- Who is it?
  - Identity recognition
- Where is it?
  - Object detection
- What are they doing?
  - Activities
- All of these are **classification** problems
  - Choose one class from a list of possible candidates

Recognition has progressed rapidly



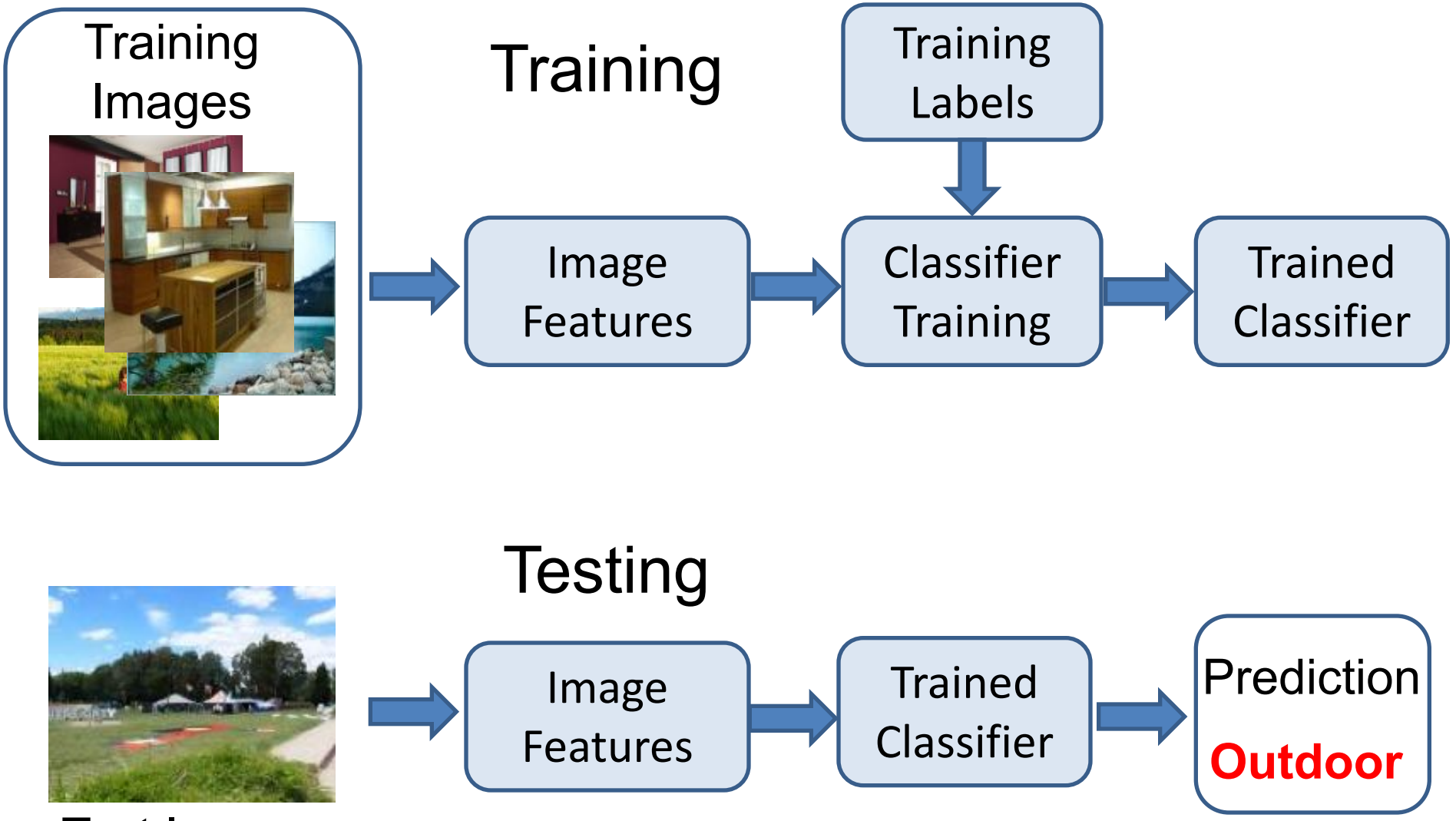
Neural Networks

Traditional Image Categorization: Training phase



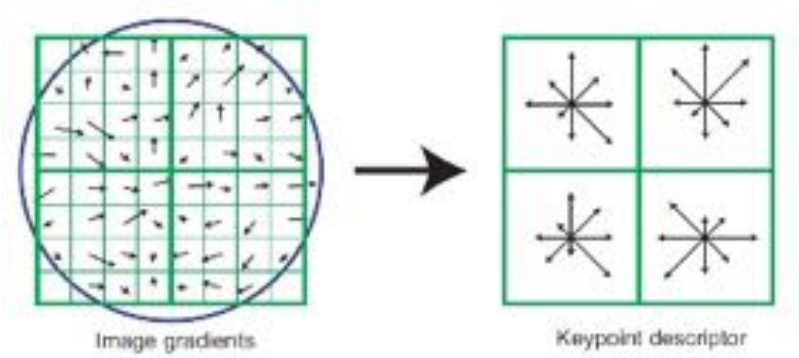
Slide credit: Jia-Bin Huang

Traditional Image Categorization: Testing phase

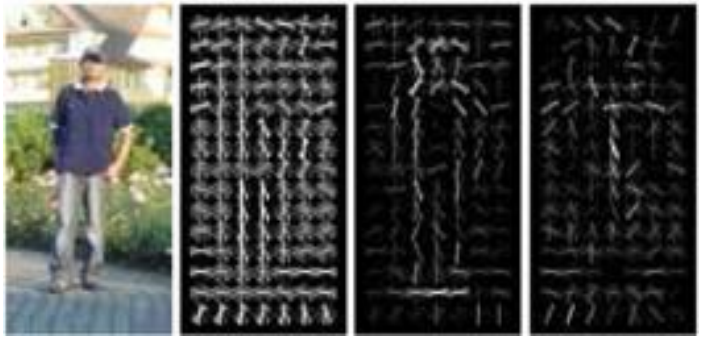


Slide credit: Jia-Bin Huang

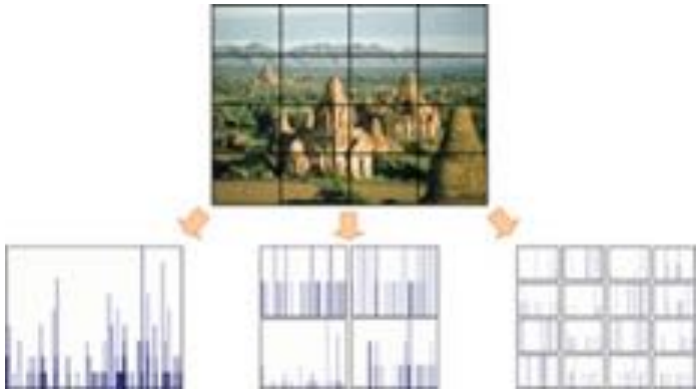
Features have been key



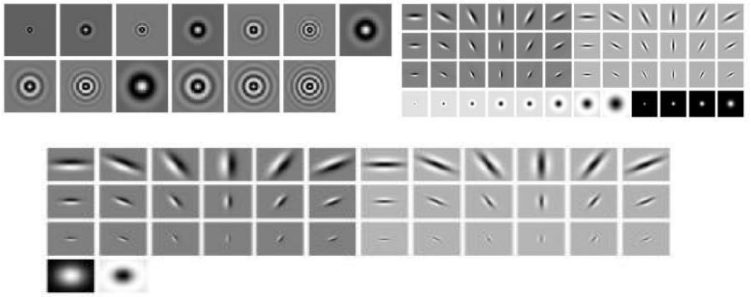
SIFT [Lowe IJCV 04]



HOG [Dalal and Triggs CVPR 05]



SPM [Lazebnik et al. CVPR 06]

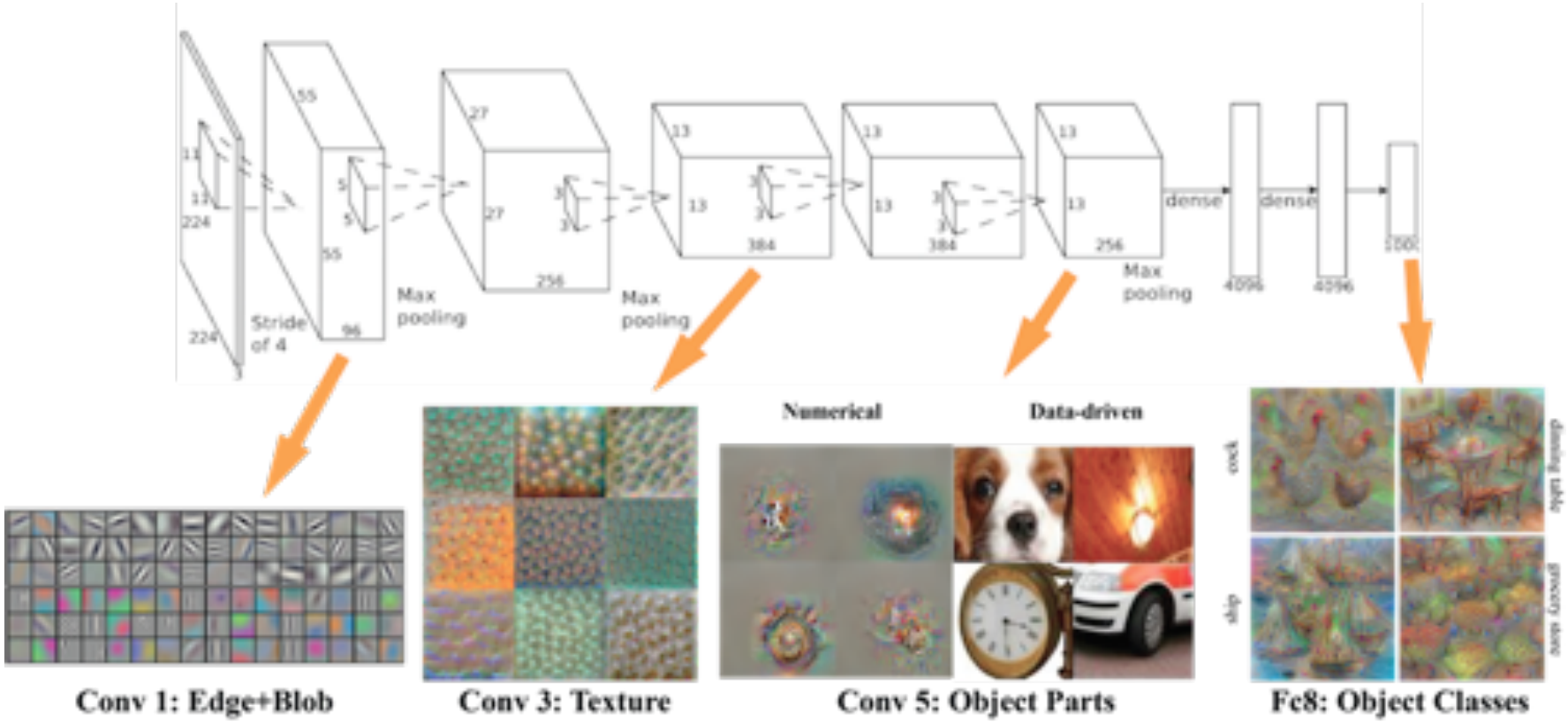


Textons

and many others:

SURF, MSER, LBP, GLOH, …..

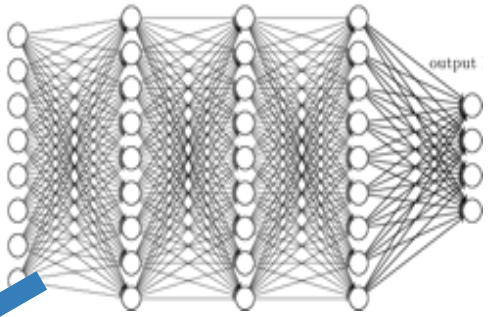
Deep learning has led to large gains



- Hierarchical and expressive feature representations
- Trained end-to-end, rather than hand-crafted for each task
- Remarkable in transferring knowledge across tasks

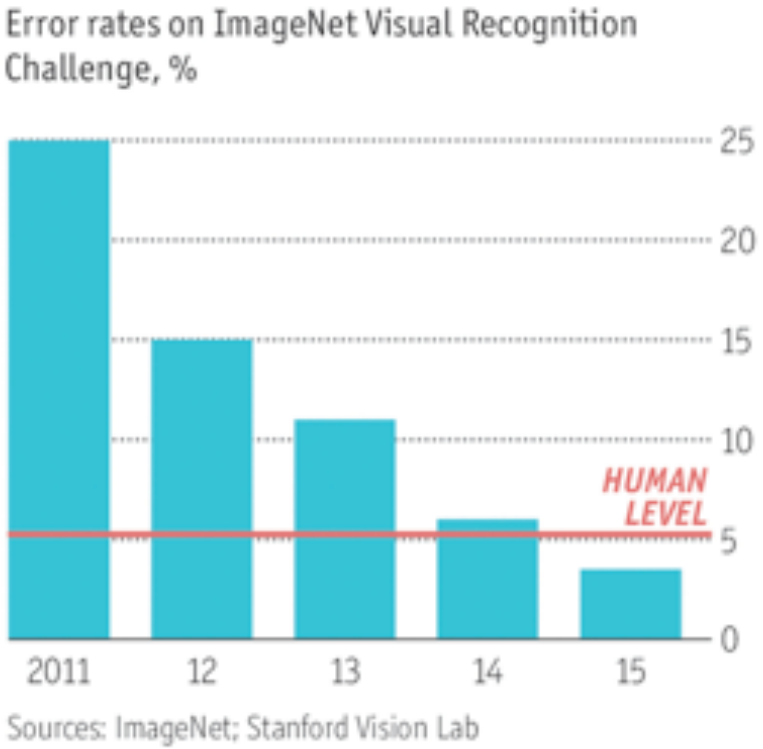
Significant recent impact on the field

Big labeled datasets



Deep learning

GPU technology

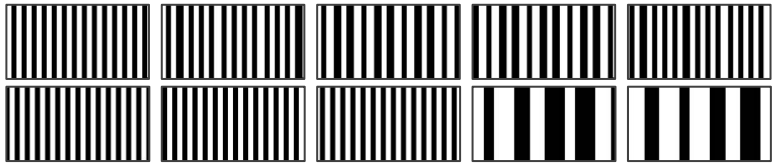


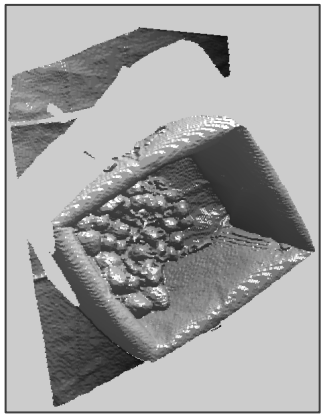
Deep learning has opened new areas

- Availability of large-scale image and video data
- Availability of computational power
  - Better and cheaper GPUs
  - Cloud computing resources
- Better understanding of how to train deep neural networks
- Advantages available for many areas of computer vision
  - Recognize objects across shape and appearance variations
  - Data-driven priors for 3D reconstruction
  - Predict long-term future behaviors in complex scenes
  - End-to-end training rather than expensive feature design.

New devices

- Time-of-flight sensors
- Structured light systems
- Light field cameras
- Coded apertures





Large-scale reconstructions

- Internet images pose challenges of scale and outliers
- Reconstructions with millions of images
- Choices to handle data
- Specific optimization approaches

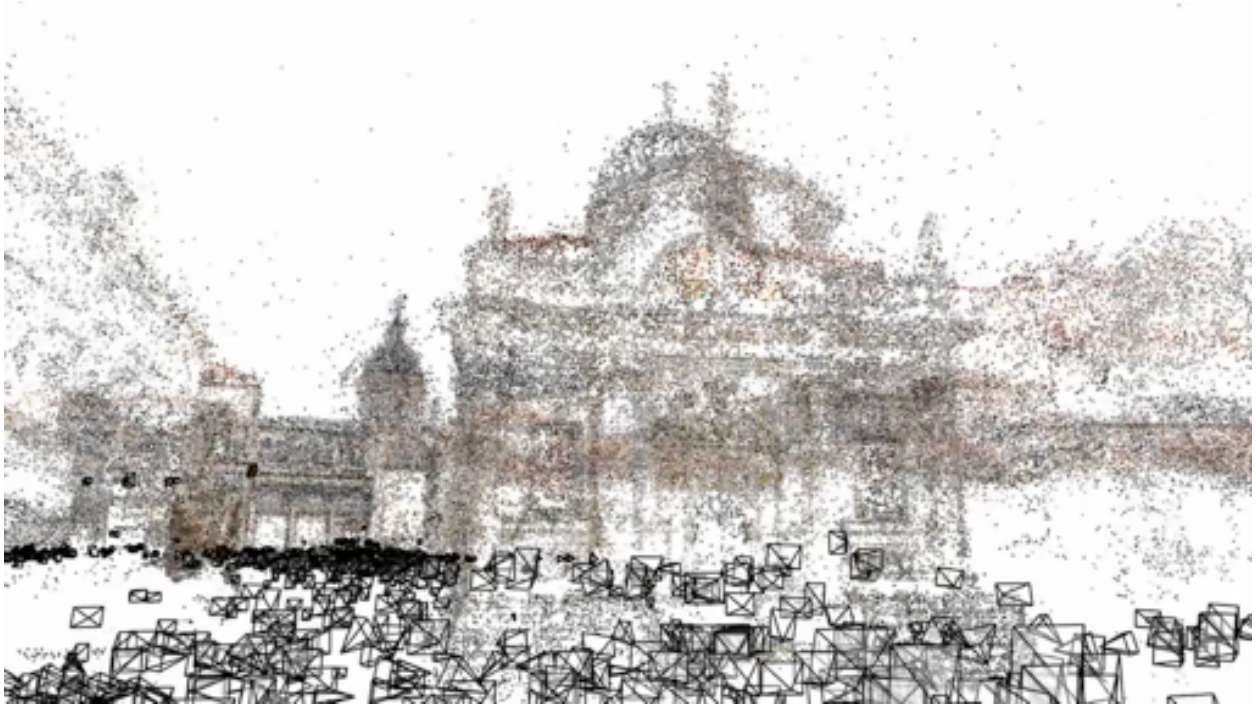
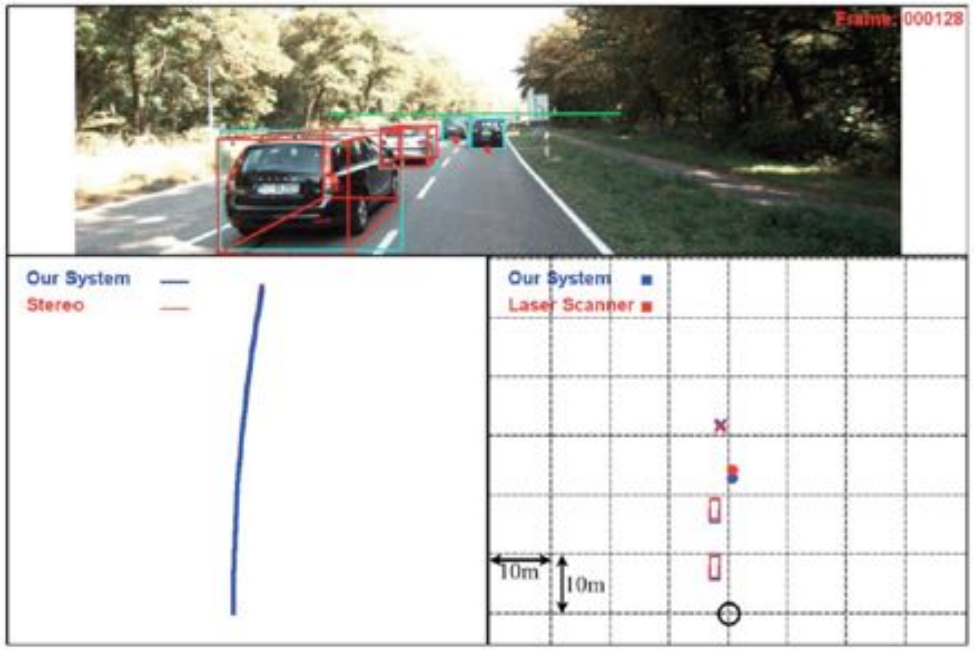


Figure from Agarwal et al.

Real-time 3D vision

- Mobile platforms, embedded systems (IoT devices)
- Stringent demands on computational resources
- Low power platforms (wattage) for automobile ECUs
- Carefully designed and multithreaded architectures



Song and Chandraker, CVPR 2014

Newcombe et al., CVPR 2015

A Few Topics That We Will Study

- Cameras and image formation
- Feature detection and matching
- Structure from Motion
- Multiview stereo
- Optical flow
- Image classification
- Object recognition
- Object detection
- Semantic segmentation
- Support Vector Machines
- Deep Neural Networks

Take-home message

- Computer vision is a key branch of AI
- Enables several modern applications around us
- A lot of highly visible and high-impact activity
- Huge industry interest
- This is a great time to study computer vision!

Course Details

Course details

- Homework assignments
  - Easy problems based directly on class discussions
  - Harder problems may require additional reading
  - Programming in Python might be required
  - Submit PDF to Gradescope before deadline
- Final exam
- Mid-term
- Participation
  - Ask questions, answer questions, engage in discussions

Course details

- Class webpage:
  - https://ucsd-cse-152.github.io/
- Instructor email:
  - haosu@eng.ucsd.edu
- Grading
  - 40% final exam
  - 30% homework assignments (3)
  - 25% mid-term
  - 05% participation
- Aim is to learn together, discuss and have fun!

Course details

- TAs:
  - Fangchen Liu: fliu@eng.ucsd.edu
  - Fanbo Xiang: fxiang@eng.ucsd.edu
  - Stephen Guerin: sguerin@eng.ucsd.edu
- Discussion section: Fr 4-4:50pm
- TA office hours to be posted on class webpage
- Piazza for questions and discussions:
  - https://piazza.com/class/k0tedcp5nj24g

Test of Background

Background

- Linear algebra
- Calculus
- Probability
- Python

- For each question
  - Write down the answer
  - Self-assess your confidence : scale of 1 (lowest) to 5 (highest)

Linear algebra

1(a) What is the rank of a matrix?

Consider the matrix $A = \begin{bmatrix} x & x^2 \\ y & y^2 \end{bmatrix}$

1(b) What is the rank of A when x = 1, y = 2?

1(c) What is the rank of A when x = 0, y = 1?

1(d) What is the null space of A when x = 2, y = 2?

1(e) Rate your confidence

1: Null space? Where Jedi master Yoda goes to sleep?

3: I kind of know, but not sure.

5: I can do this in my sleep!

Linear algebra

Consider the matrix

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 4 \end{bmatrix}$$

2(a) What is the transpose of A?

2(b) Define eigenvalues and eigenvectors of a matrix.

2(c) What are the eigenvalues of A?

2(d) Rate your confidence

1: Transpose? My phone autocorrects it to transport.

3: I kind of know, but not sure.

5: Transcended it!

Linear algebra

Given two vectors, a = [1, 2, 3] and b = [-1, 0, 1]

3(a) What is the dot product a.b?

3(b) What is the cross product a x b?

3(c) If R is a 3 x 3 rotation matrix, what is $R^T R$?

3(d) Rate your confidence

1: Rotation matrix? Gets my head spinning.

3: I kind of know, but not sure.

5: You spin me right round!

Probability

4(a) If P(A) = 0.5, P(B) = 0.4 and P(AB) = 0.2, what is P(A U B)?

4(b) In the above, what is P(A|B)?

4(c) State the Bayes rule.

4(d) Rate your confidence

1: Probability? It's a coin toss.

3: I kind of know, but not sure.

5: Keep the dice rolling!

Calculus

5(a) What is the derivative of $f(x) = x^2$?

5(b) What is the partial derivative of $f(x,y) = x^2y$ with respect to y?

5(c) What is the gradient of $f(x, y) = x^2y$?

5(d) State the chain rule of differentiation.

5(e) Rate your confidence

1: Gradient? Seems a steep climb.

3: I kind of know, but not sure.

5: Top of the hill!

Python

6(a) Have you used Python in the past?

6(b) Briefly describe a program or project you wrote in Python.

6(c) Write a snippet: use a loop to print numbers from 1 to 10.

6(d) Have you used NumPy in the past?

6(e) Rate your confidence

1: NumPy? Does it taste like ApplePie?

3: I kind of know, but not sure.

5: I breathe and eat code!