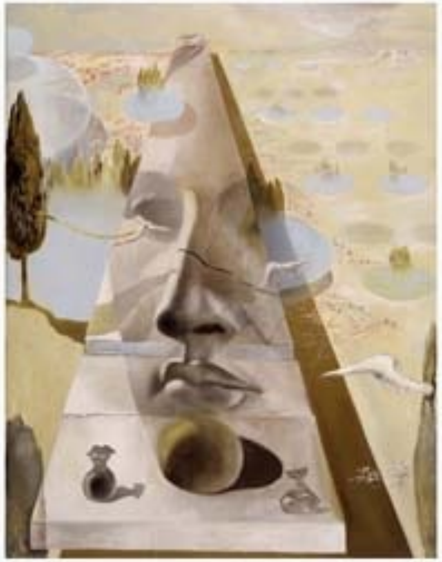
CS231A

Computer Vision:
From 3D Reconstruction to Recognition



Class Time

M-W; 11:30—13:00PM

CS231A

Instructors

Silvio Savarese

- ssilvio@stanford.edu
- Office: Gates Building, room: 256
- Office hour: tbd and by appointment

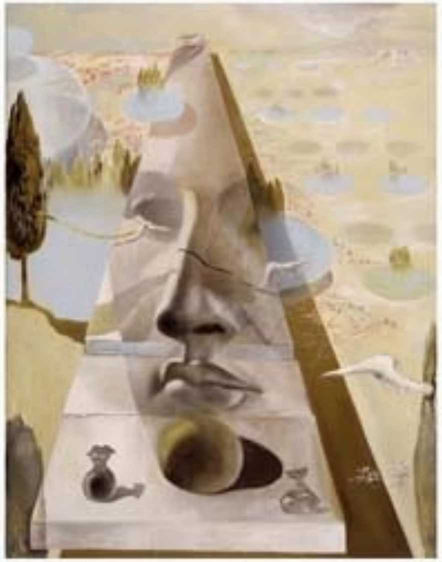
Jeannette Bohg

- bohg@stanford.edu
- Office: Gates Building, room: 244
- Office hour: Friday 9-10am or by appointment

CAs:

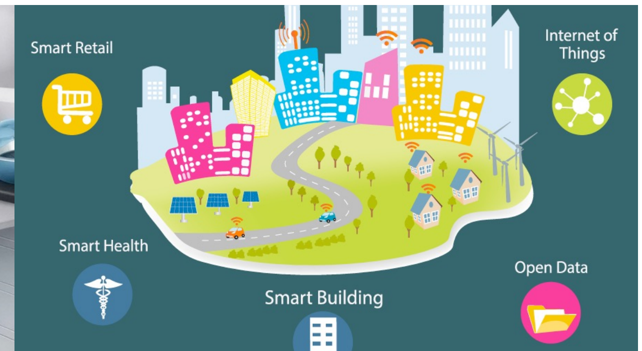
- Andrey Kurenkov, Krishnan Srinivasan, JunYoung Gwak, Yinan Zhang

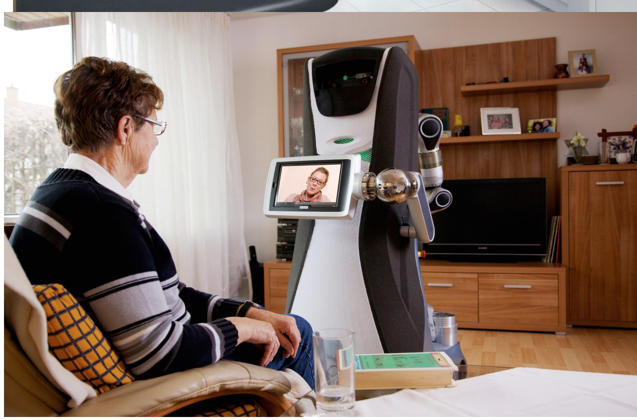
Lecture 1
Introduction



- An introduction to computer vision
- Course overview

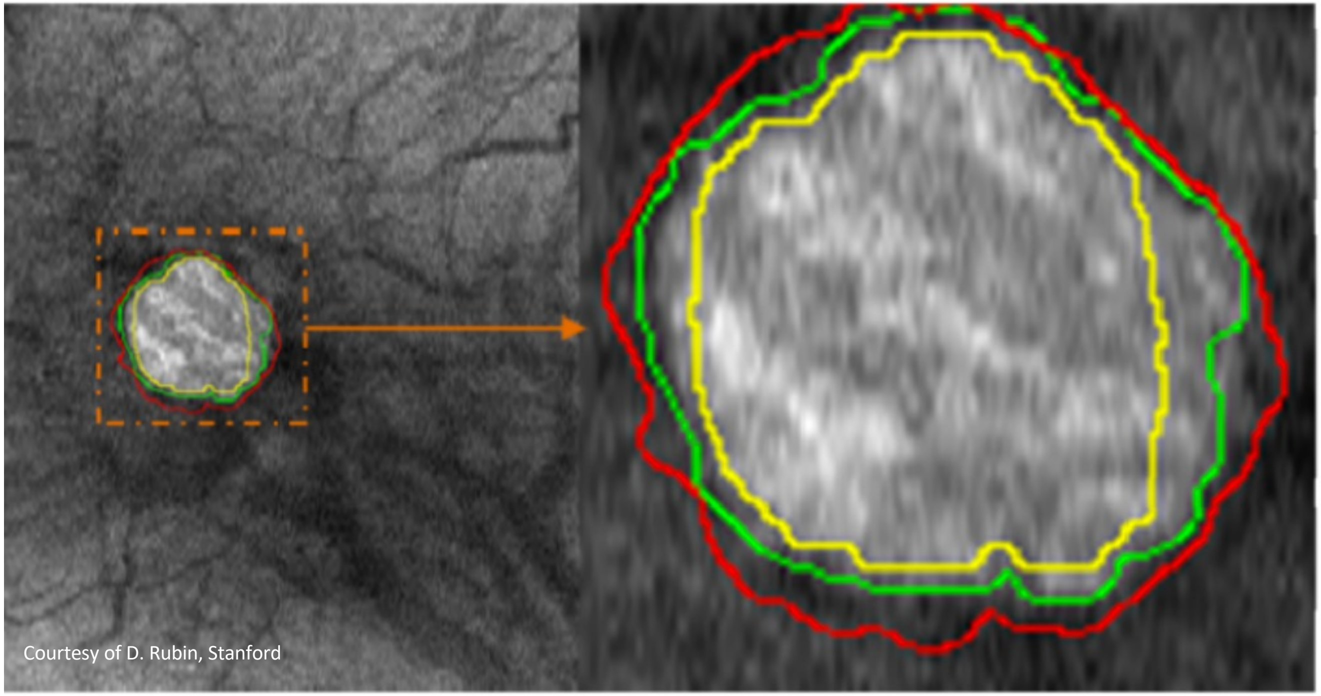
AI is a propelling force of today’s technology



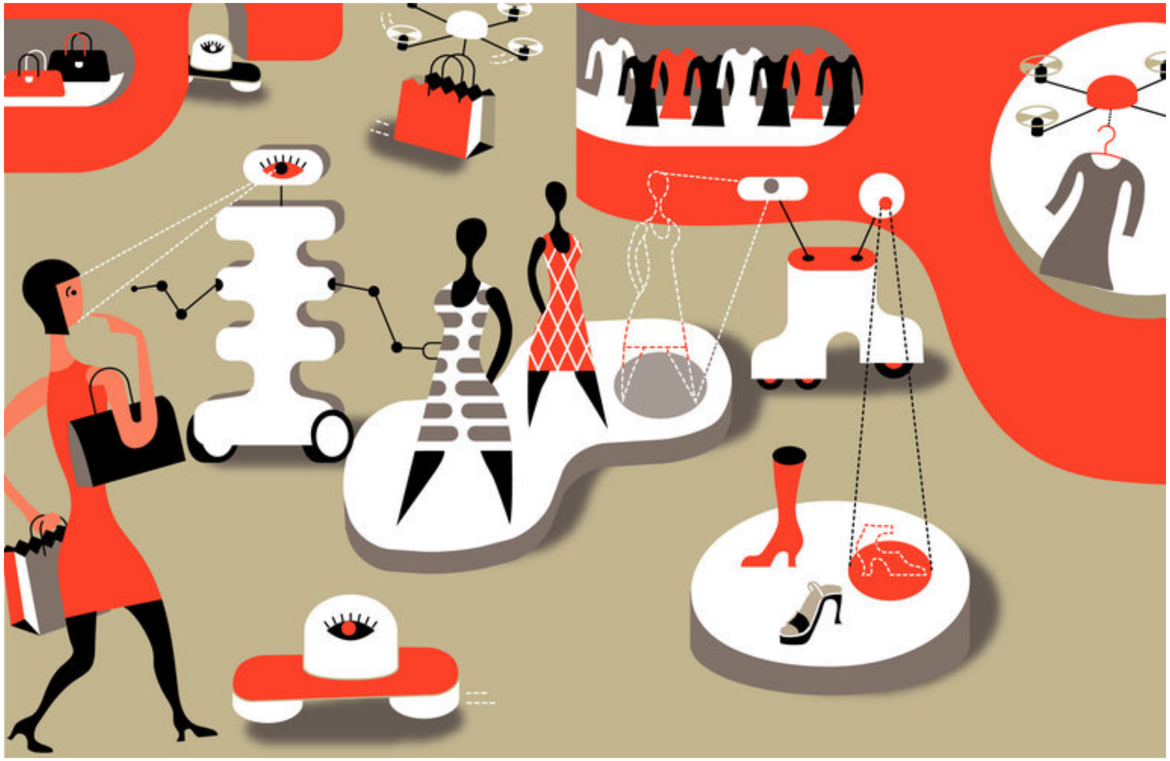


Smart Agriculture

Health care

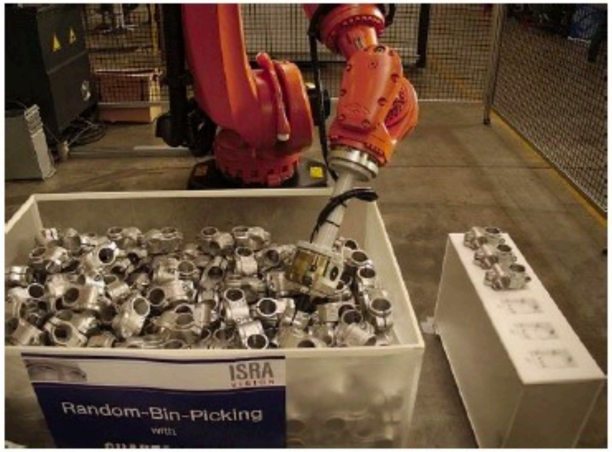


Retail



From Imagining the Retail Store of the Future - The New York Times, April 12, 2017

Manufacturing



Courtesy of ITRI, 2017

Transportation and Logistics

Construction Management

Why is this acceleration happening now?

Enabling factors

- Big data

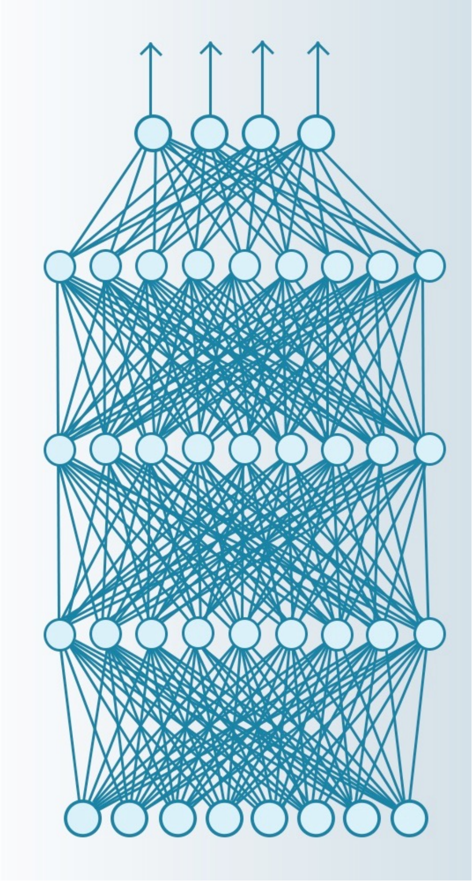
ImageNet, 2009
ShapeNet, 2015

Enabling factors

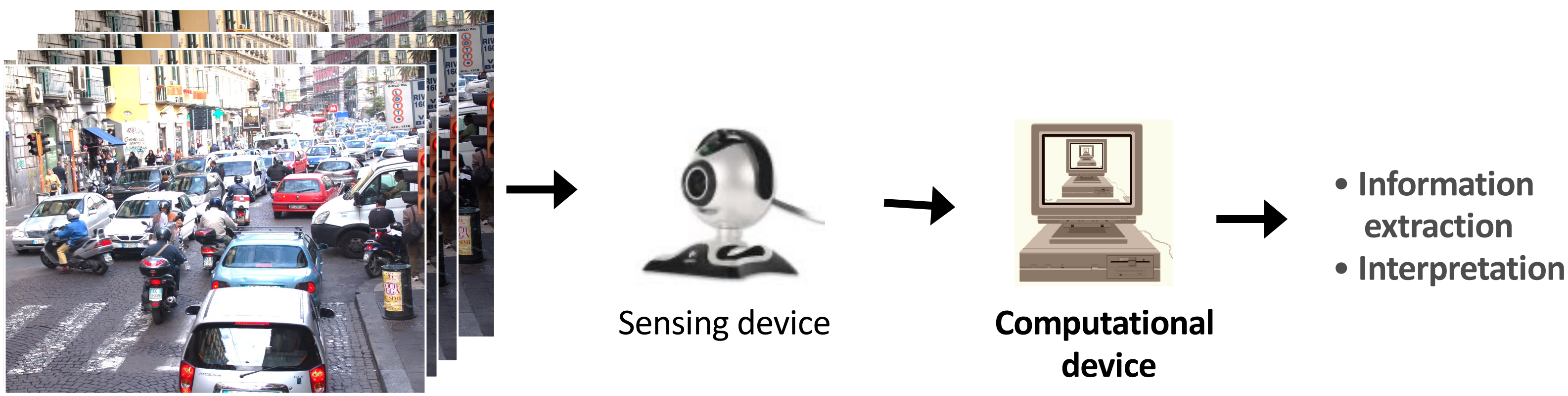
- Big data
- Faster hardware

Enabling factors

- Big data
- Faster hardware
- New algorithms
  - Representation learning
  - Neural networks
  - Inject learning to deterministic reasoning



Computer vision

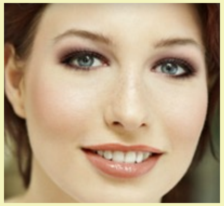


1. **Information extraction:** features, 3D structure, motion flows, etc…
2. **Interpretation:** recognize objects, scenes, actions, events

Major areas in Computer Vision

Space/Geometry

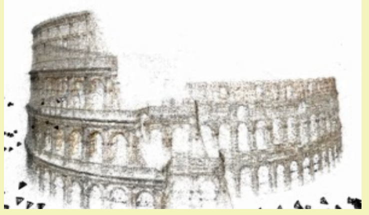
- Object shape recovery
- Depth estimation
- 3D scene reconstruction



Semantics/Learning

- Object detection and pose estimation
- Object tracking
- Scene understanding

Major areas in Computer Vision



Space/Geometry

- Object shape recovery
- Depth estimation
- 3D scene reconstruction

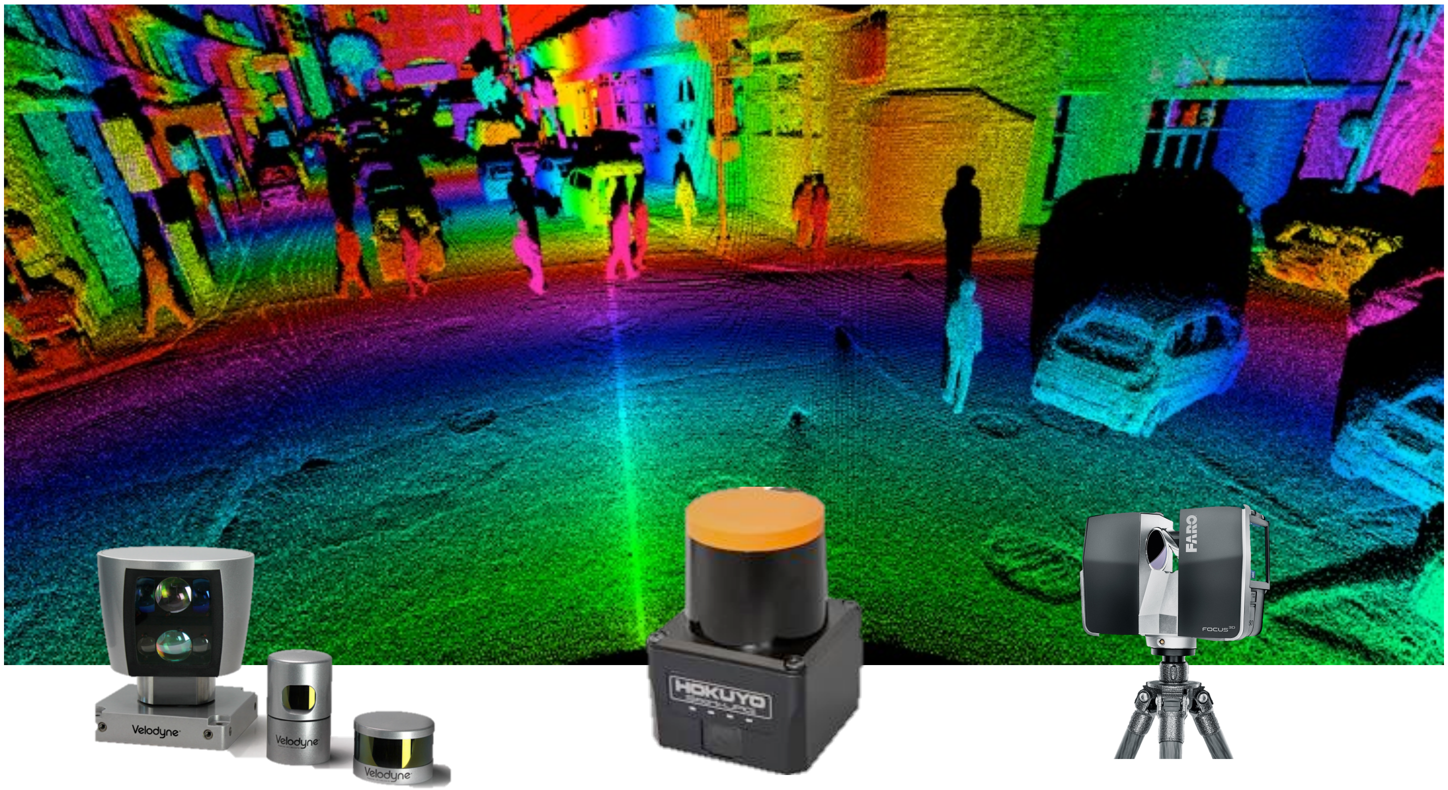
Semantics/Learning

- Object detection and pose estimation
- Object tracking
- Scene understanding

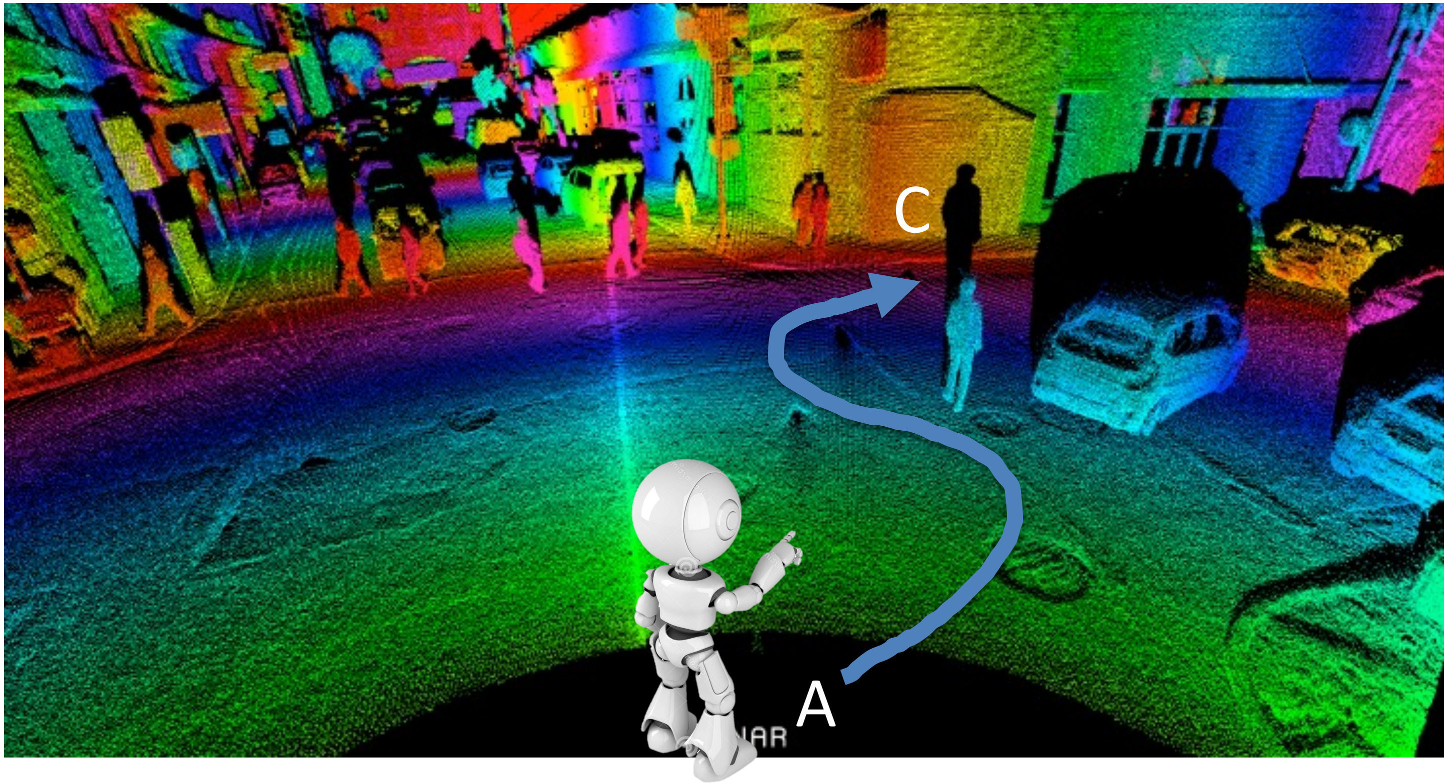
Recovering 3D models of the environments

Armeni et al. 2016

Recovering 3D models of the environments



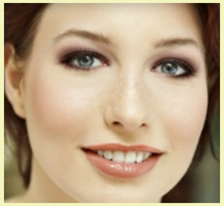
This is critical for autonomous driving or navigation!



Major areas in Computer Vision

Space/Geometry

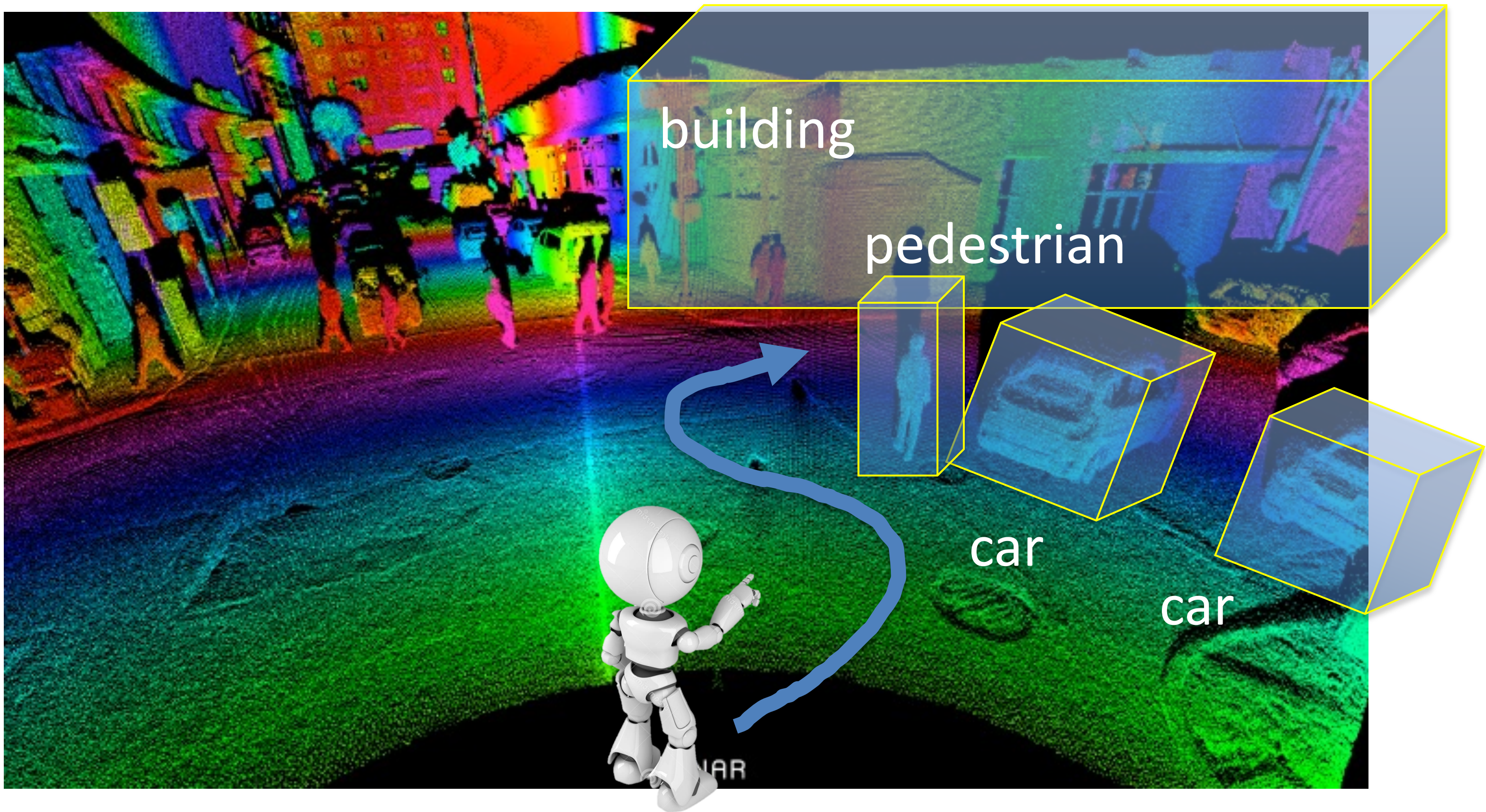
- Object shape recovery
- Depth estimation
- 3D scene reconstruction



Semantics/Learning

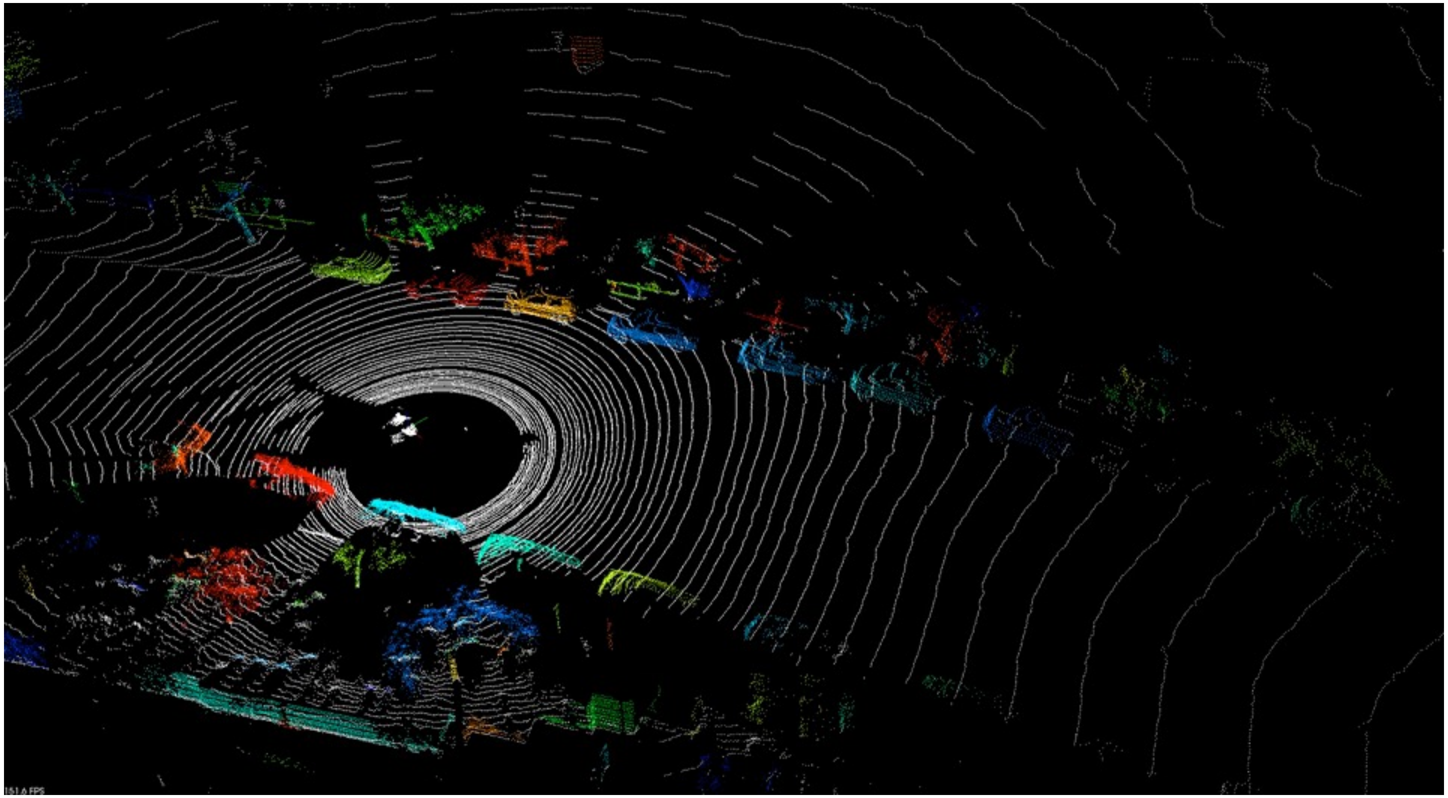
- Object detection and pose estimation
- Object tracking
- Scene understanding

Detecting and tracking objects in the environments

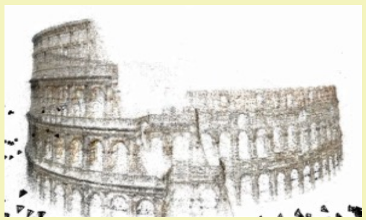


3D Scene Parsing

Held, Thrun, Savarese, 2016-206

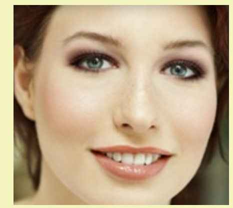


Major areas in Computer Vision



Space/Geometry

- Object shape recovery
- Depth estimation
- 3D scene reconstruction



Semantics/Learning

- Object detection and pose estimation
- Object tracking
- Scene understanding

CS 231A course overview

1. Space/Geometry

   Estimating spatial properties of objects and scene from images through geometrical methods

2. Semantics/Learning

CS 231A course overview

1. Space/Geometry

   Estimating spatial properties of objects and scene from images through geometrical methods

2. Semantics/Learning

   Estimating semantic and dynamic properties of scene elements from images through learning methods

CS 231A course overview

1. Space/Geometry

   Estimating spatial properties of objects and scene from images through geometrical methods

2. Semantics/Learning

   Estimating semantic and dynamic properties of scene elements from images through learning methods

Camera systems

Establish a mapping from 3D to 2D

How to calibrate a camera

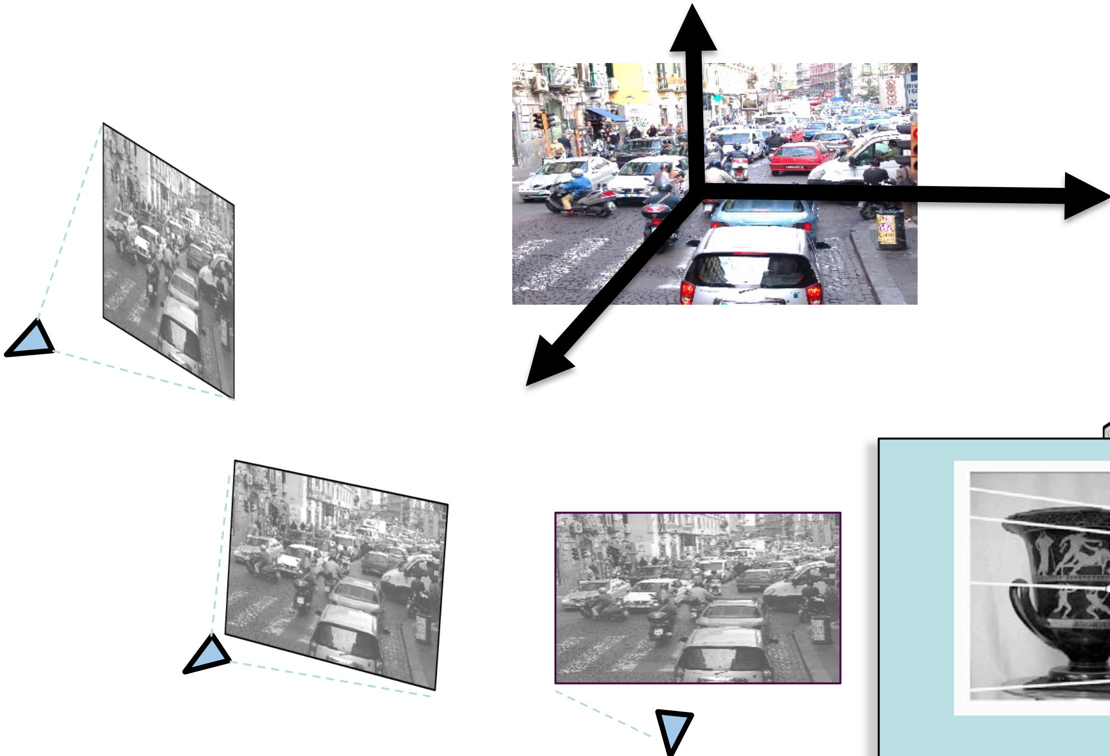
Estimate camera parameters such pose or focal length

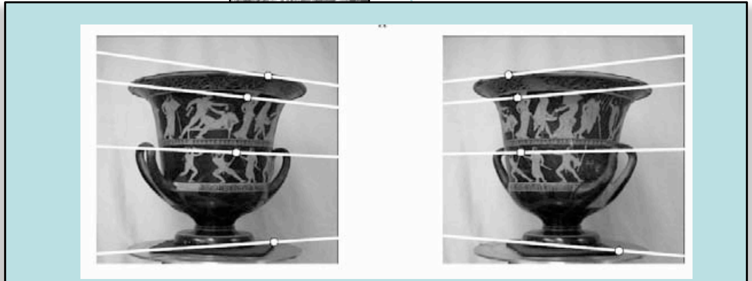
Single view metrology

Estimate 3D properties of the world from a single image

Multiple view geometry

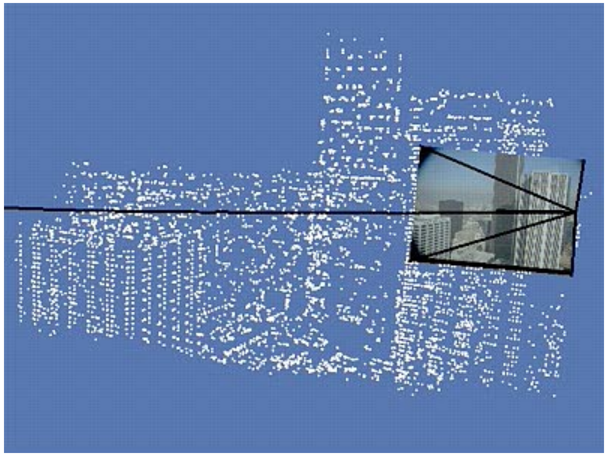
Estimate 3D properties of the world from multiple views





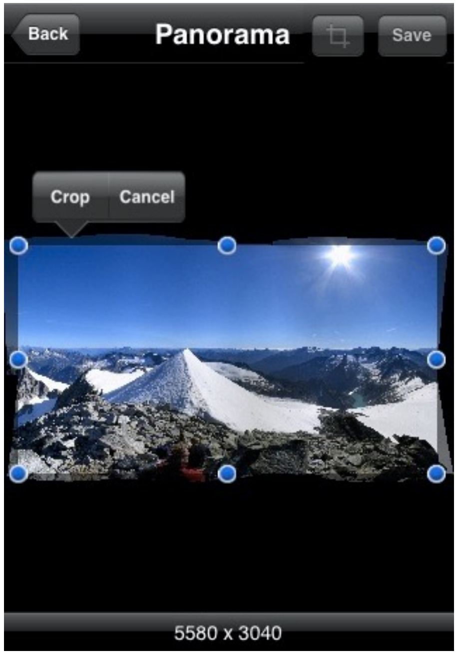
Epipolar geometry

Structure from motion



Courtesy of Oxford **Visual Geometry Group**

Panoramic Photography

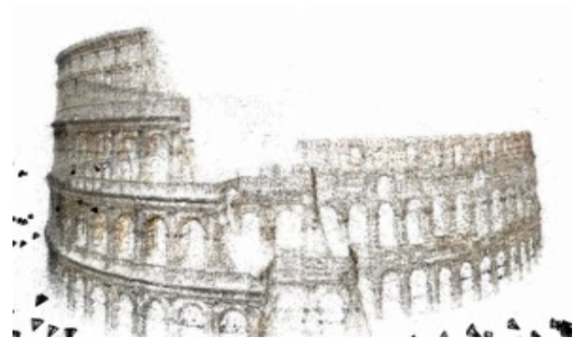


Cloudburst RESEARCH

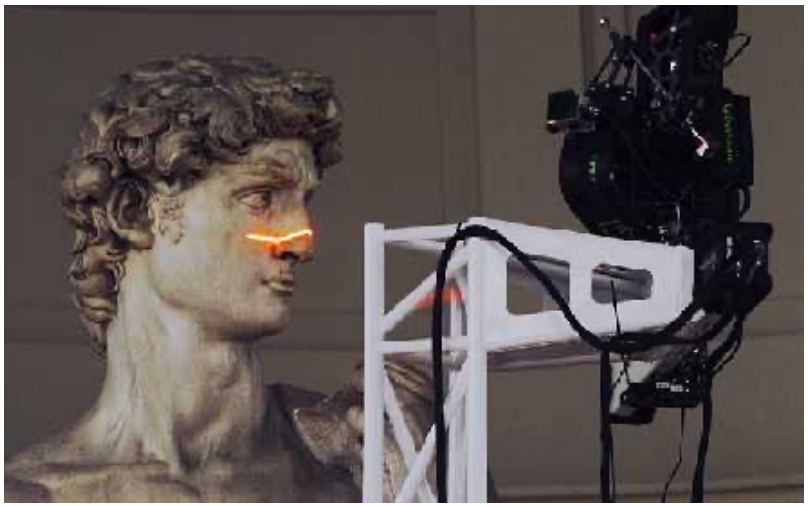
HumanEyes
Print Out of the Box

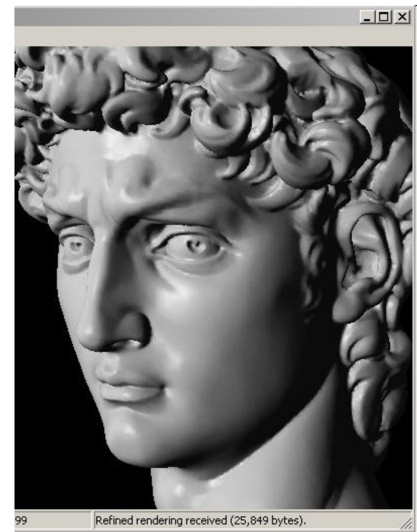
kolor

3D Modeling of landmarks



Accurate 3D Object Prototyping





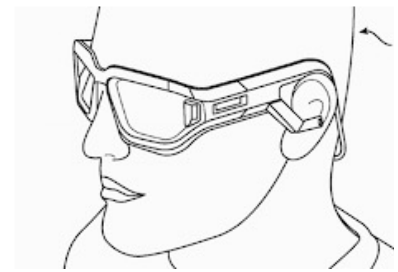
Scanning Michelangelo's *"The David"*

- The Digital Michelangelo Project
  - http://graphics.stanford.edu/projects/mich/
- 2 BILLION polygons, accuracy to .29mm

Augmented Reality

Mirriad
Advertising for the Skip Generation

AUGMENT



- Magic leap
- Daqri
- Meta
- Etc…

CS 231A course overview

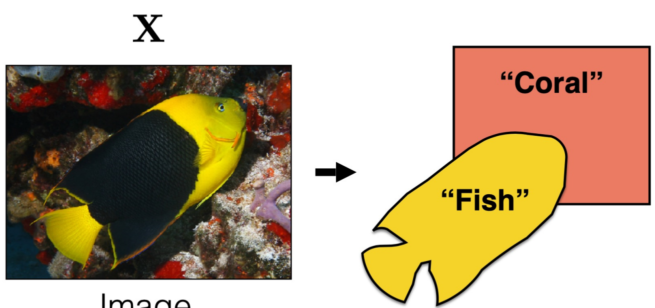
1. Space/Geometry

   Estimating spatial properties of objects and scene from images through geometrical methods

2. Semantics/Learning

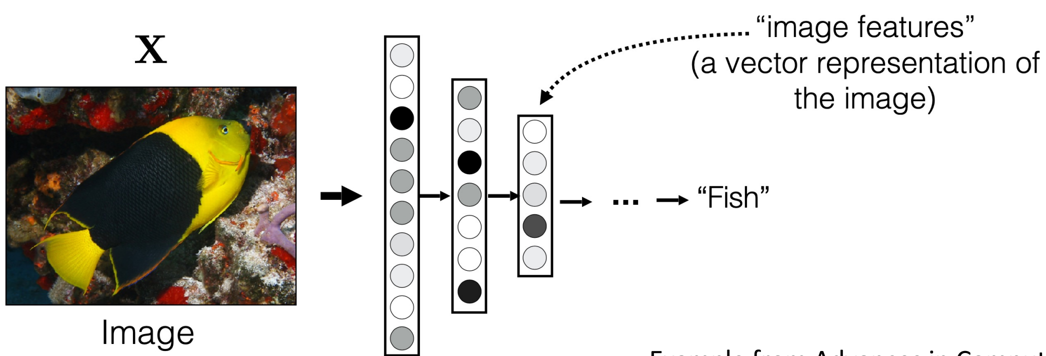
   Estimating semantic and dynamic properties of scene elements from images through learning methods

Representations and Representation Learning



Image

Compact mental representation

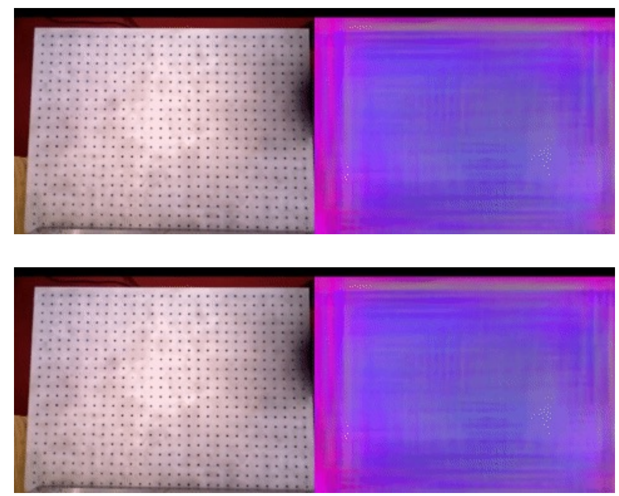


Image

Example from Advances in Computer Vision – MIT – 6.869/6.819

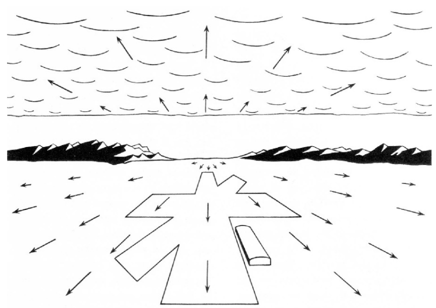
Monocular Depth Estimation and Feature Tracking

Monocular Depth Estimation



Feature Tracking

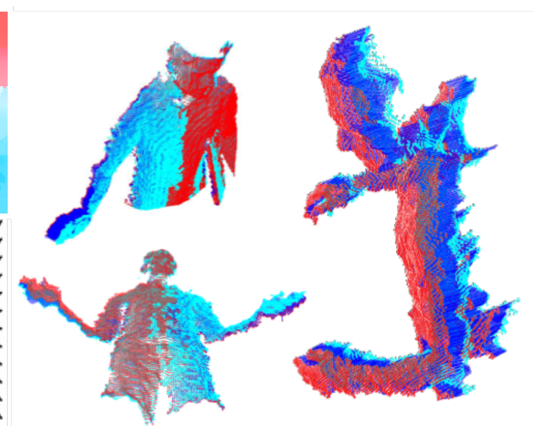
Optical and Scene Flow



J. J. Gibson, The Ecological Approach to Visual Perception

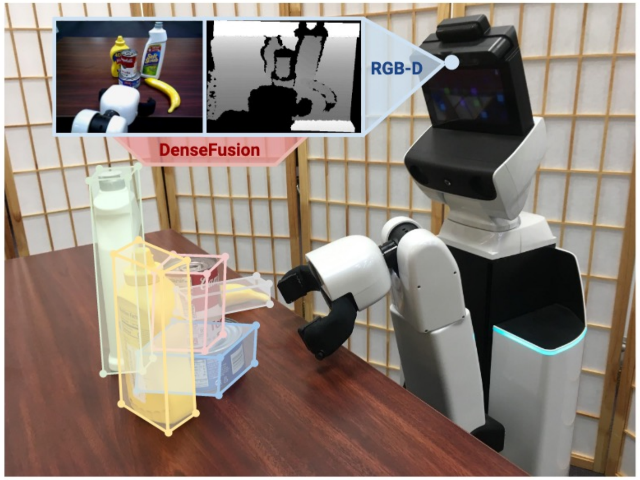
Lucas-Kanade Feature Tracking over multiple frames. Picture adopted from OpenCV Webpage.

A Database and Evaluation Methodology for Optical Flow. Baker et al. IJCV. 2011



A Primal-Dual Framework for Real-Time Dense RGB-D Scene Flow. Jaimez at al. ICRA, 2015.

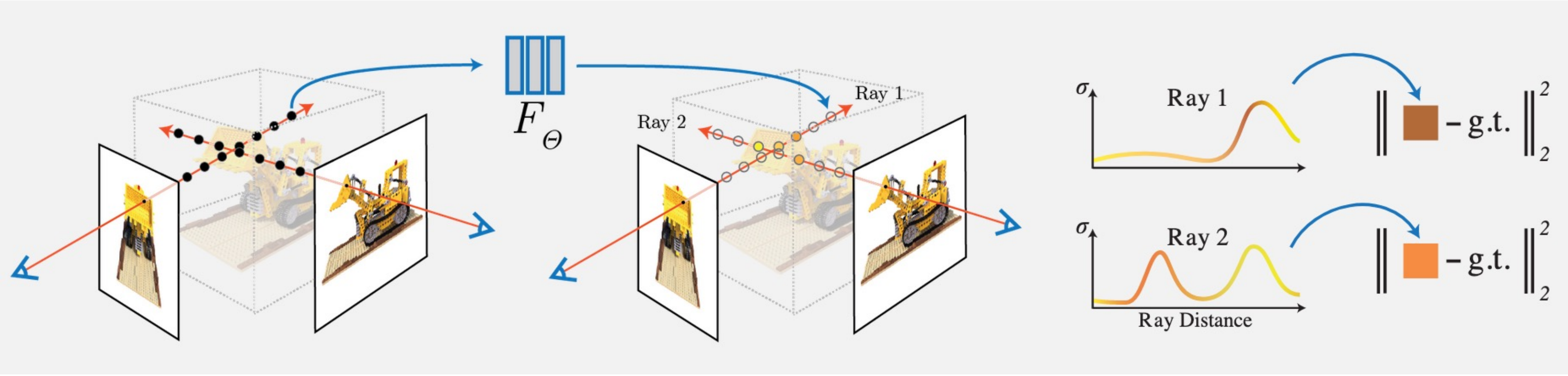
Optimal Estimation for Object Tracking



Wang et al. “Dense Fusion: 6D Object Pose Estimation by Iterative Dense Fusion”, CVPR 2019

Manuel Wühtrich et al. “Probabilistic Object Tracking using a Depth Camera”, IROS 2013

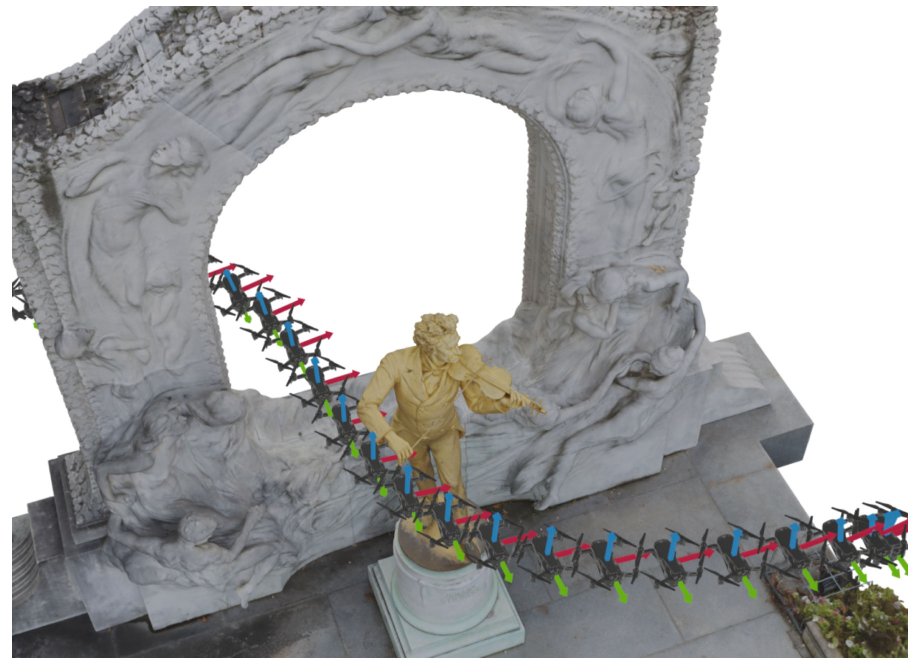
Neural Radiance Fields for View Synthesis



Autonomous navigation and safety

[Mobileye](): Vision systems in high-end BMW, GM, Volvo models. But also, Toyota, Google, Apple, Tesla, Nissan, Ford, etc.…

Source: A. Shashua, S. Seitz



Navigation in a Neural Radiance World using a Monocular camera only. Adamkiewicz, Chen et al. 2021

Personal robotics

jibo

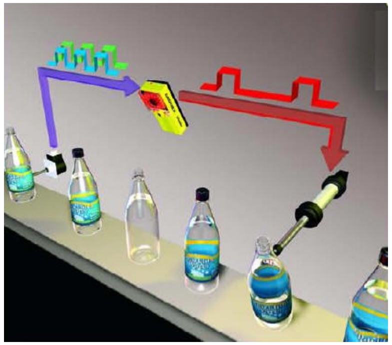
FIVE ELEMENTS ROBOTICS

a

ECOVACS

BLUE FROG robotics

More Applications



Factory inspection

Assistive technologies

Surveillance

Exploration and remote operations

Syllabus

Lecture	Topic
1	Introduction
2	Camera models
3	Camera calibration
4	Single view metrology
5	Epipolar geometry
6	Multi-view and stereo geometry
7	Volumetric Stereo
8	Structure from Motion
9	Fitting and Matching
10	Representations and Representation Learning
11	Depth Estimation, Feature Tracking
12	Optical and Scene Flow I
13	Optical and Scene Flow II
14	Optimal Estimation I
15	Optimal Estimation II
16	Neural Radiance Fields I
17	Neural Radiance Fields II
18	Guest Lecture

Project presentations

January

February

March

Geometry

Learning

Proposal due

Mid term

Final projects

Prerequisites

- This course requires knowledge of linear algebra, probability, statistics, machine learning and computer vision, as well as decent programming skills.
- Though not an absolute requirement, it is encouraged and preferred that you have at least taken either CS221 or CS229 or CS131A or have equivalent knowledge.
- We will leverage concepts from low-level image processing (CS131A) (e.g., linear filters, edge detectors, corner detectors, etc…) and machine learning (CS229) (e.g., **SVM, basic Bayesian inference, clustering**, neural networks, etc…) which we won't cover in this class.
- We will provide links to background material related to CS131A and CS229 (or discuss during TA sessions) so students can refresh or study those topics if needed.

Text books

Required:

- [FP] D. A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach* (2nd Edition). Prentice Hall, 2011.
- [HZ] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Academic Press, 2002.

Recommended:

- R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer, 2011.
- D. Hoiem and S. Savarese. *Representations and Techniques for 3D Object Recognition and Scene Interpretation*, Synthesis lecture on Artificial Intelligence and Machine Learning. Morgan Claypool Publishers, 2011
- Learning OpenCV, by Gary Bradski & Adrian Kaehler, O'Reilly Media, 2008.
- [PB] Probabilistic Robotics, by Thrun, Fox and Burgard, MIT Press, 2005. (PDF of relevant chapter will be provided)

Course assignments

- 1 warm up problem set (HW-0) released today
- 4 problem sets
- 1 mid-term exam
- 1 project

- Look up class schedule for release and due dates.
- Problems will be released through the schedule page and must be submitted through Gradescope (Use code 6P3K6D).

Midterm Exam

- The exam will be on 02/14.
- You will be updated with more details, e.g., platform, timing, material to be covered, review sessions etc., as we approach the midterm.

Course Projects

- Replicate an interesting paper
- Comparing different methods to a test bed
- A new approach to an existing problem
- Original research

- Write a 10-page paper summarizing your results
- Release the final code
- Give a final in-class presentation
- SCPD students can send videos instead.

- We will introduce project ideas in 1-2 weeks
- Important dates: look up class schedule

Course Projects

- Form your team:
  - 1-3 people
  - The larger is the team, the more work we expect from the team
  - Be nice to your partner: do you plan to drop the course?
- Evaluation
  - Quality of the project (including writing)
  - Final project in-class presentation (~ TBA minutes spotlight presentations)

Grading policy

- Homeworks: 37%
  - 1% for HW0
  - 9% for HW1, HW2, HW3, HW4 (each)
- Mid term exam: 20%
- Course project: 38%
  - Project proposal 1%
  - mid term progress report 5%
  - final report 25%
  - presentation 7%
- Attendance and class participation: 5%
  - Questions, answers, remarks, Ed posts, Quizzes, OH attendance …
  - Submitting questions/corrections for new course notes 9-15
  - Class participation are waived for SCPD students. For the project presentation, SCPD students can send videos instead.

Grading policy (HWs)

– 25% will be deducted per day late.
– Four 24-hours one-time late submission “bonuses” are available; that is, you can use this bonus to submit your HW late after at most 24 hours. This is one time deal: After you use all your bonuses, you must adhere to the standard late submission policy.
– Max 2 bonuses can be used per assignment.
– No exceptions will be made.

Grading policy (project)

- If 1 day late, 25% off the grade for the project
- If 2 days late, 50% off the grade for the project
- Zero credits if more than 2 days
- No "late submission bonus" is allowed when submitting your progress report or project report

Respond at **pollev.com/jeannetteboh707**

2022_CS231a_Jan3

0 done

0 underway

Start the presentation to see live content. For screen share software, share the entire screen. Get help at **pollev.com/app**

When poll is active, respond at **pollev.com/jeannetteboh707**

Are you remote or local during the majority of the winter quarter?

remote

local

What time zone are you in during the majority of the quarter?

- Pacific
- Eastern
- Mountain
- Central
- Europe +Africa
- Asia + Australia

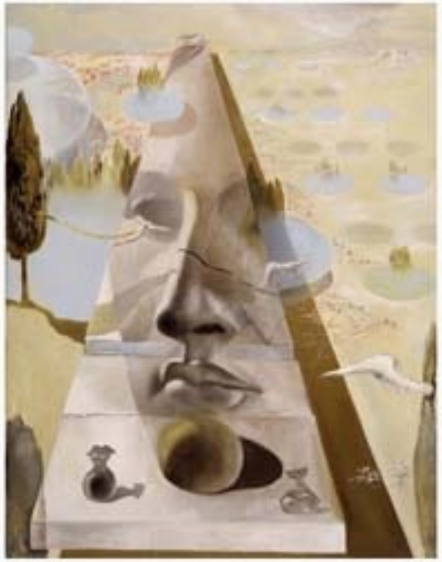
Are you an SCPD student?

yes

no

CS231

Introduction to Computer Vision



Next lecture: Camera systems