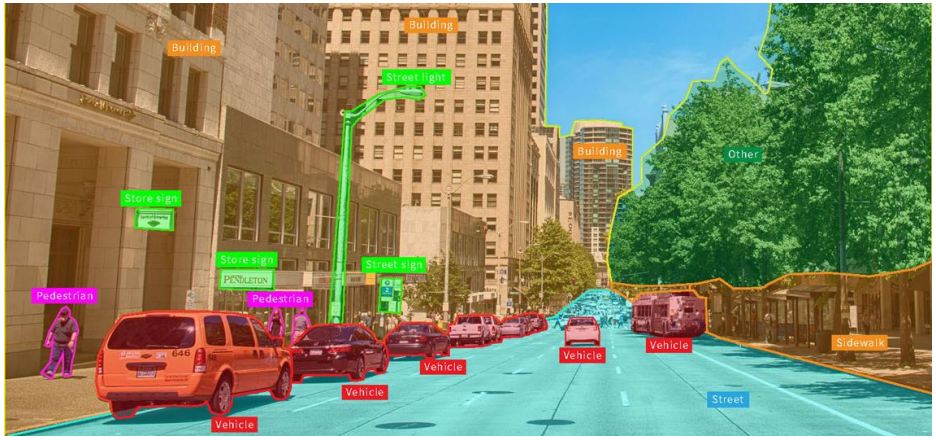
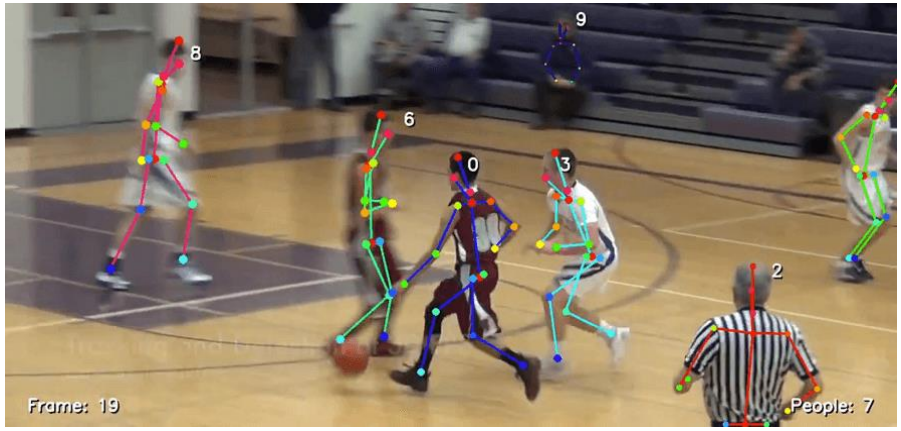
Computer Vision

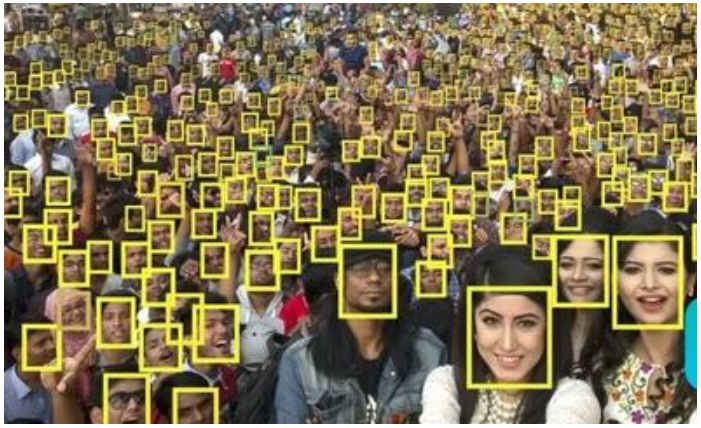
Lec 1 - Introduction

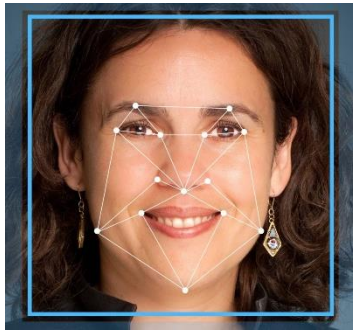
University of Haifa

Simon Korman


Building
Building
Street light
Building
Other
Store sign
Store sign
Street sign
Pedestrian
Pedestrian
Vehicle
Vehicle
Vehicle
Vehicle
Vehicle
Sidewalk
Street


9
8
6
0
3
2
Frame: 19
People: 7





Computer Vision

- Automatic understanding of images and video

1. **measurement**: Computing properties of the 3D world from visual data

2. **perception and interpretation**: Algorithms and representations to allow a machine to recognize objects, people, scenes and activities

3. **search and organization**: Algorithms to mine, search and interact with visual data

4. **generation**: Learning to generate conditional visual content

vision for measurement

Real-time stereo

Wang et al.

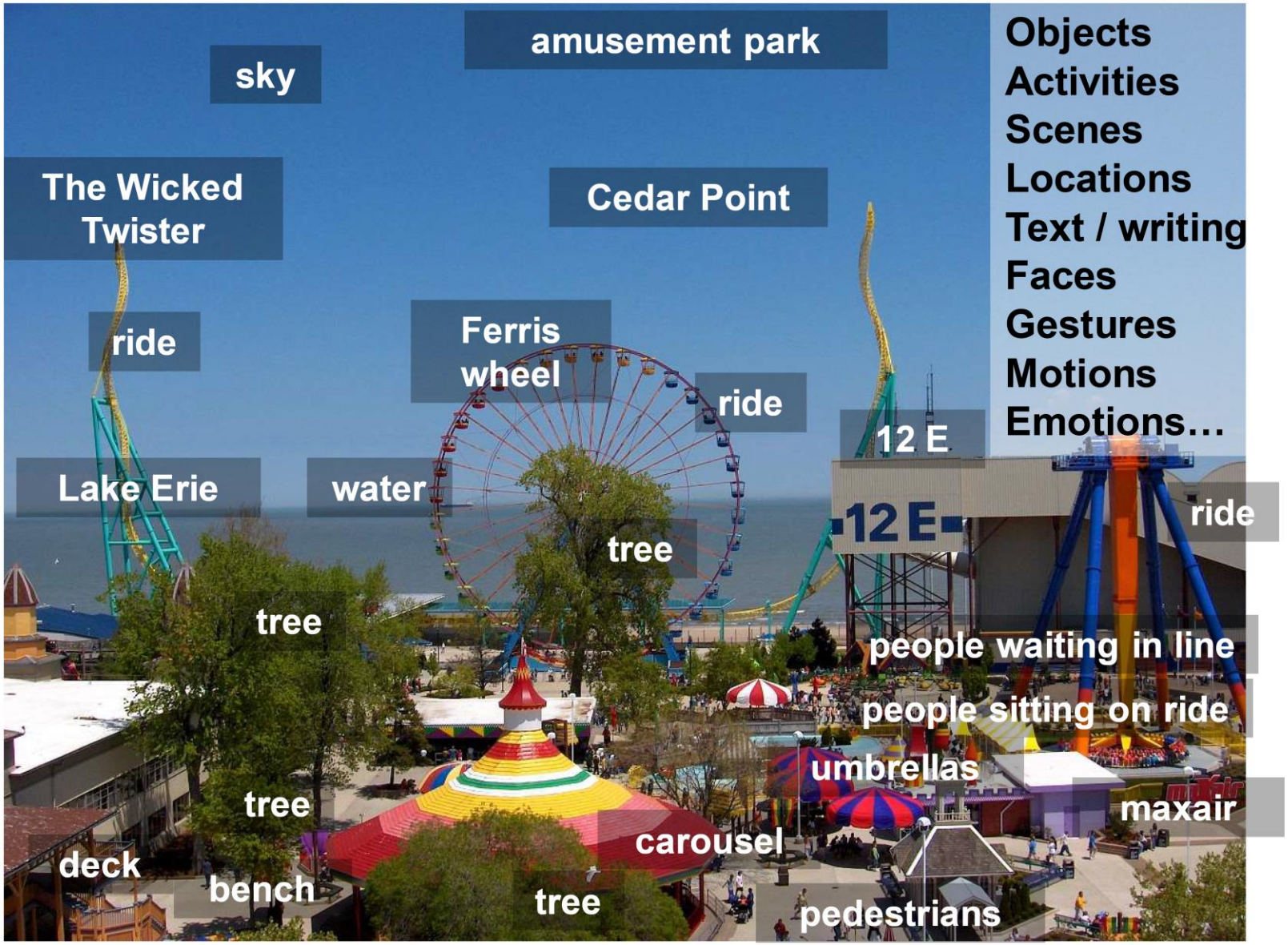
Structure from motion

Snavely et al.

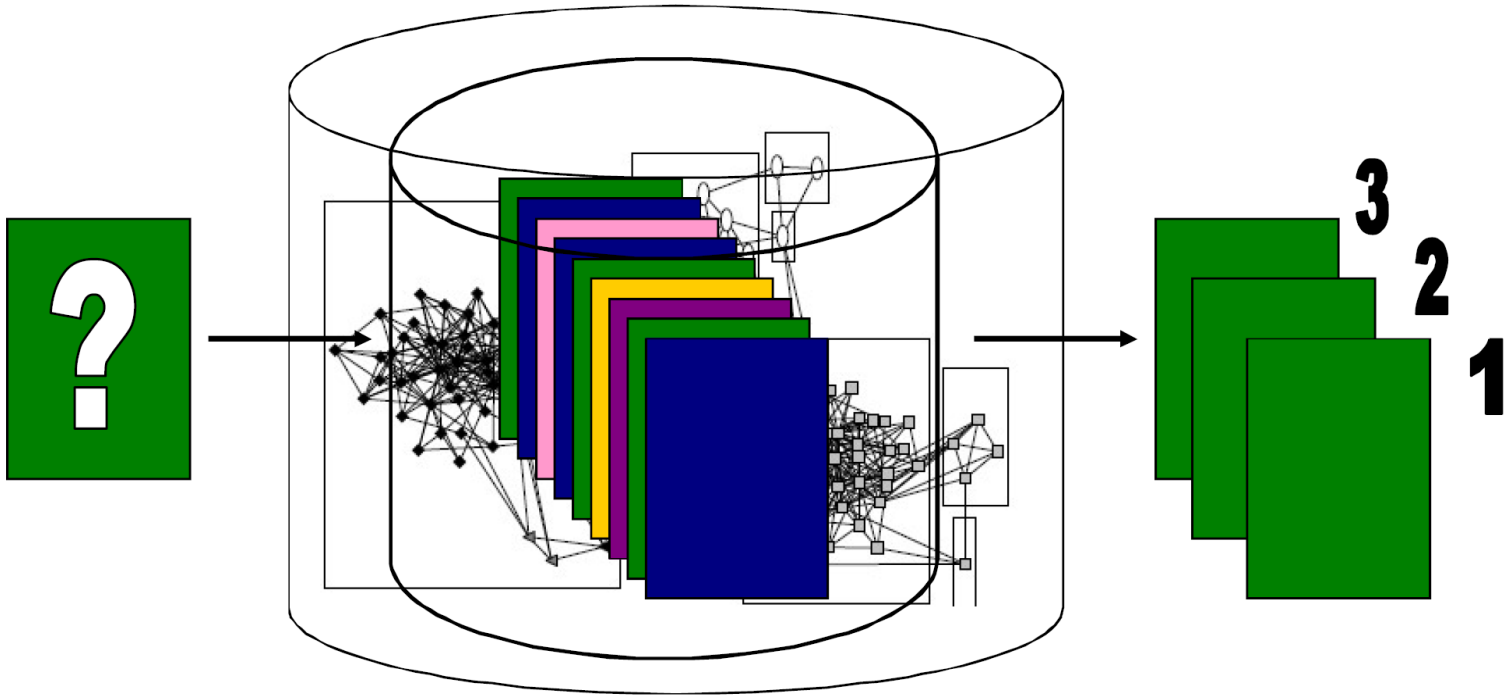
Tracking

Demirdjian et al.

vision for perception, interpretation



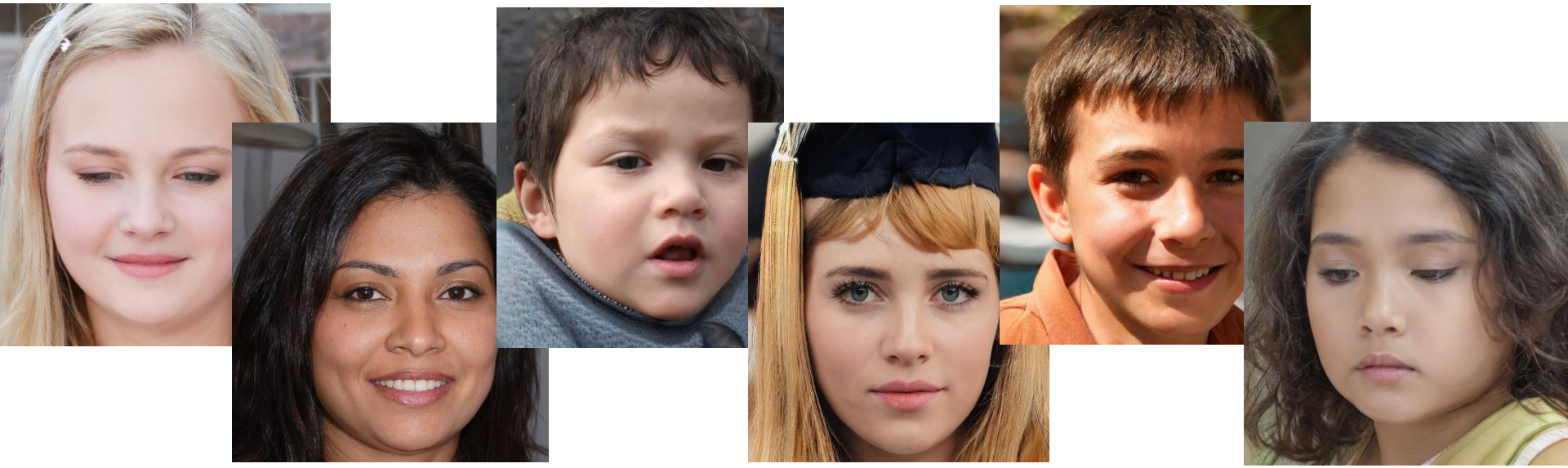
visual search, organization



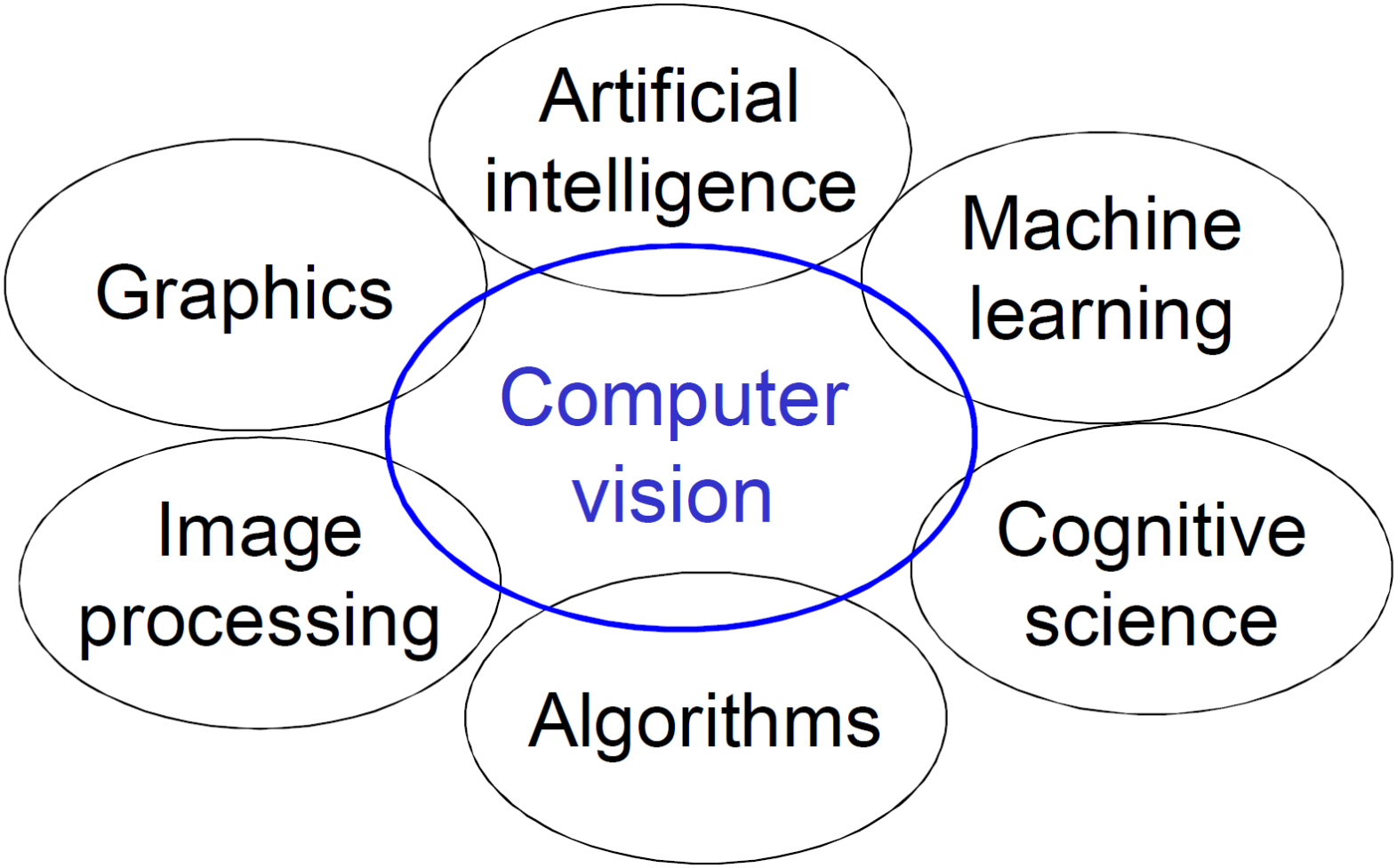
generation

A hamburger in the shape of a Rubik’s cube

A velociraptor working at a hotdog stand



Related Disciplines



Vision and Graphics

Inverse problems:
analysis and synthesis.

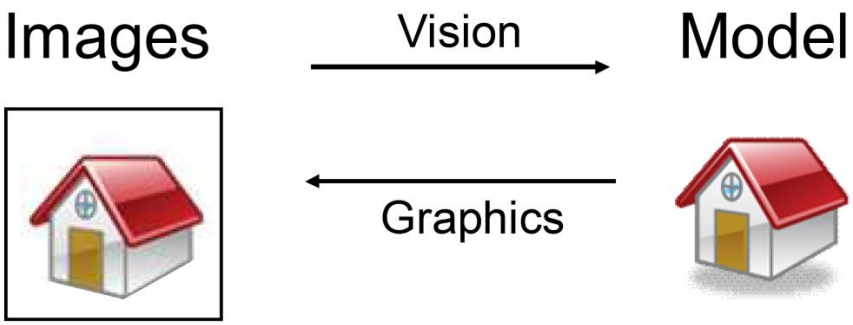
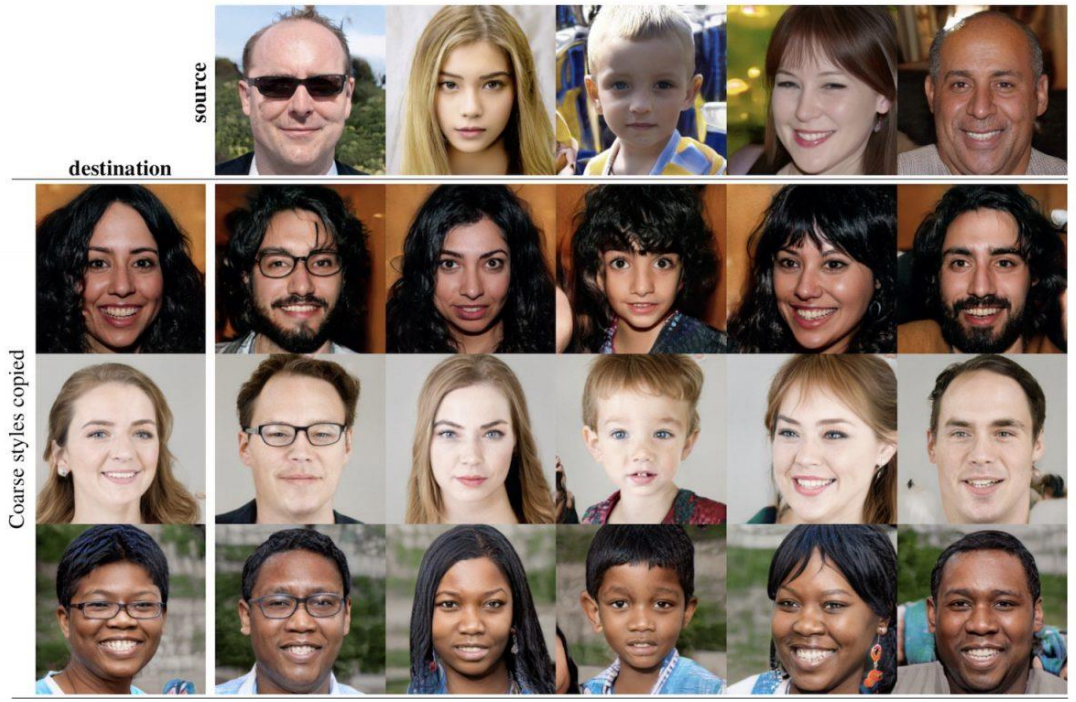
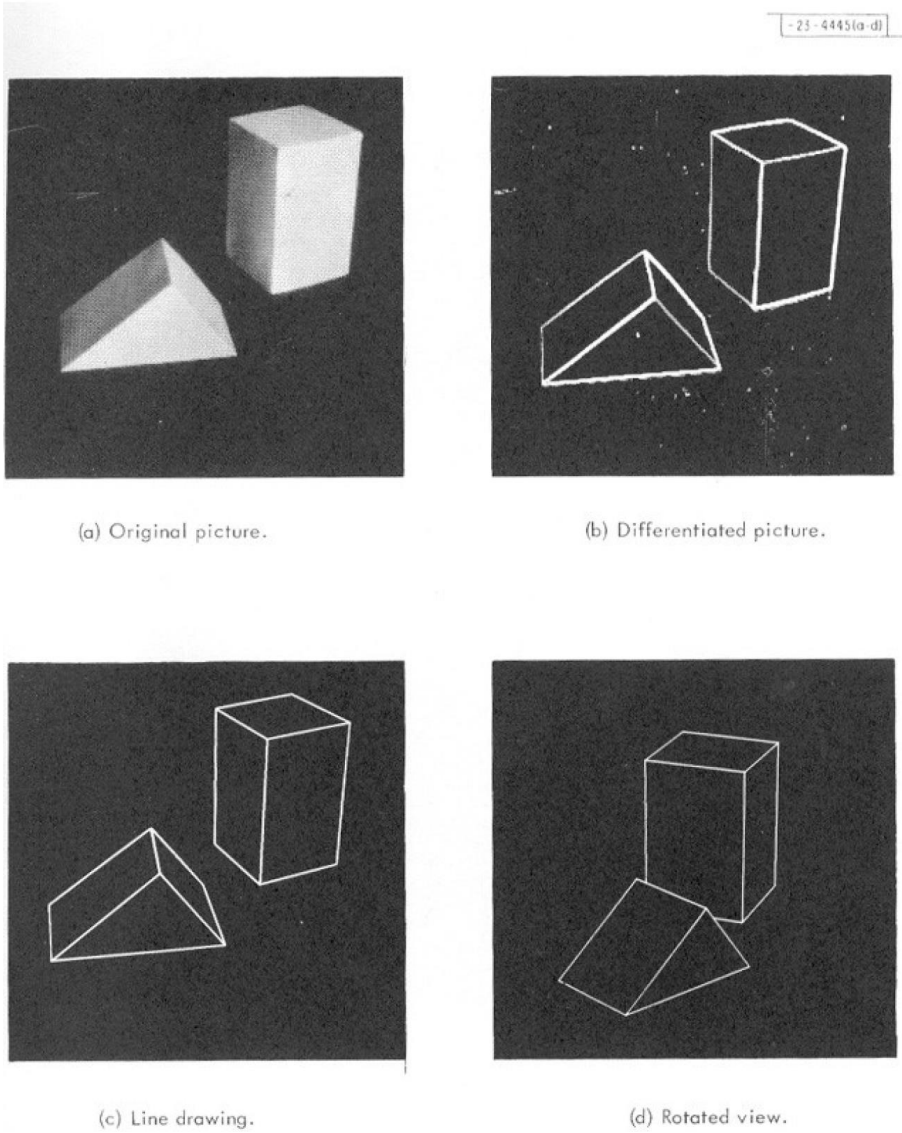


Image Generation



Visual Data in 1963



(a) Original picture. (b) Differentiated picture.

(c) Line drawing. (d) Rotated view.

L. G. Roberts, *Machine Perception of Three Dimensional Solids*, Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

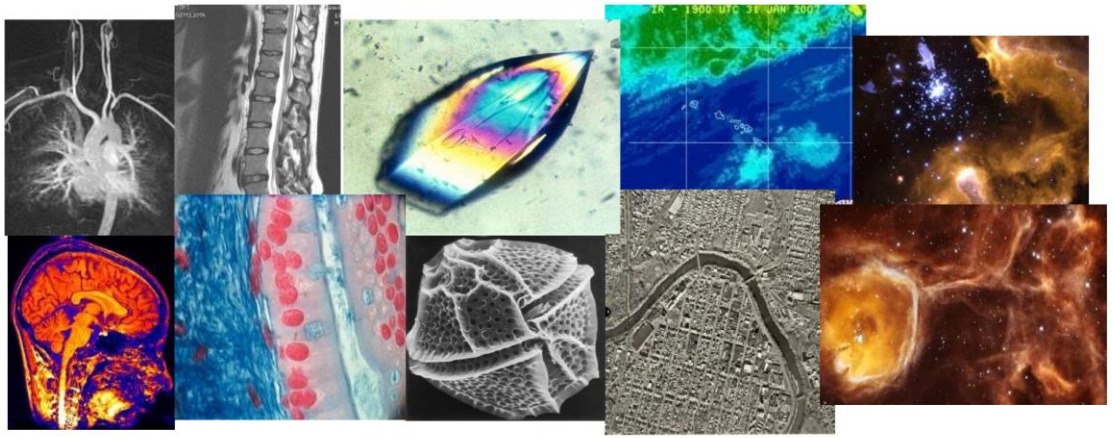
Visual Data Nowadays

Personal photo albums

Movies, news, sports

Google Image Search Picasa™ flickr webshots picsearch™ YouTube Broadcast Yourself™

Surveillance and security

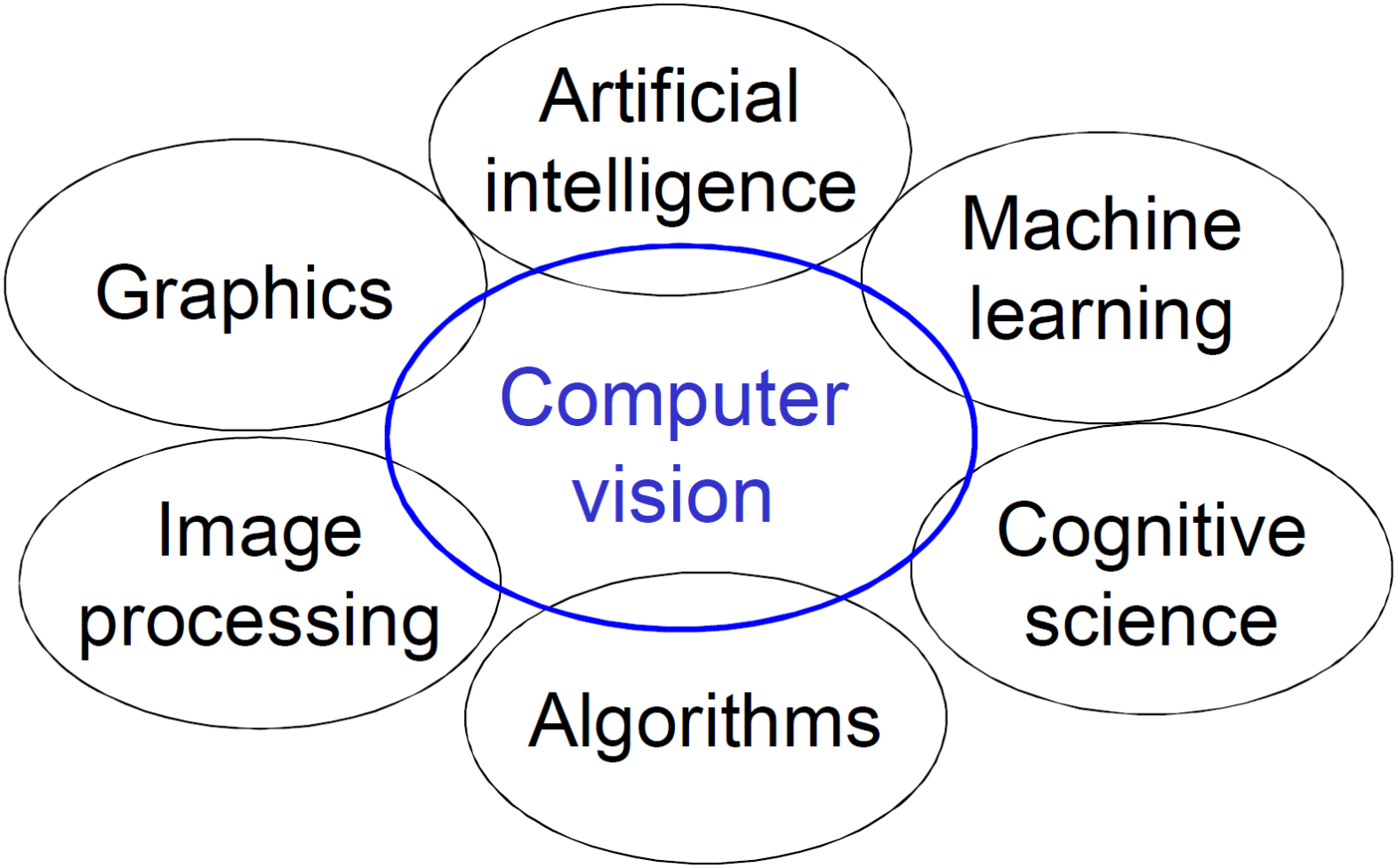


Medical and scientific images

+ many dedicated large-scale datasets

Slide credit; L. Lazebnik

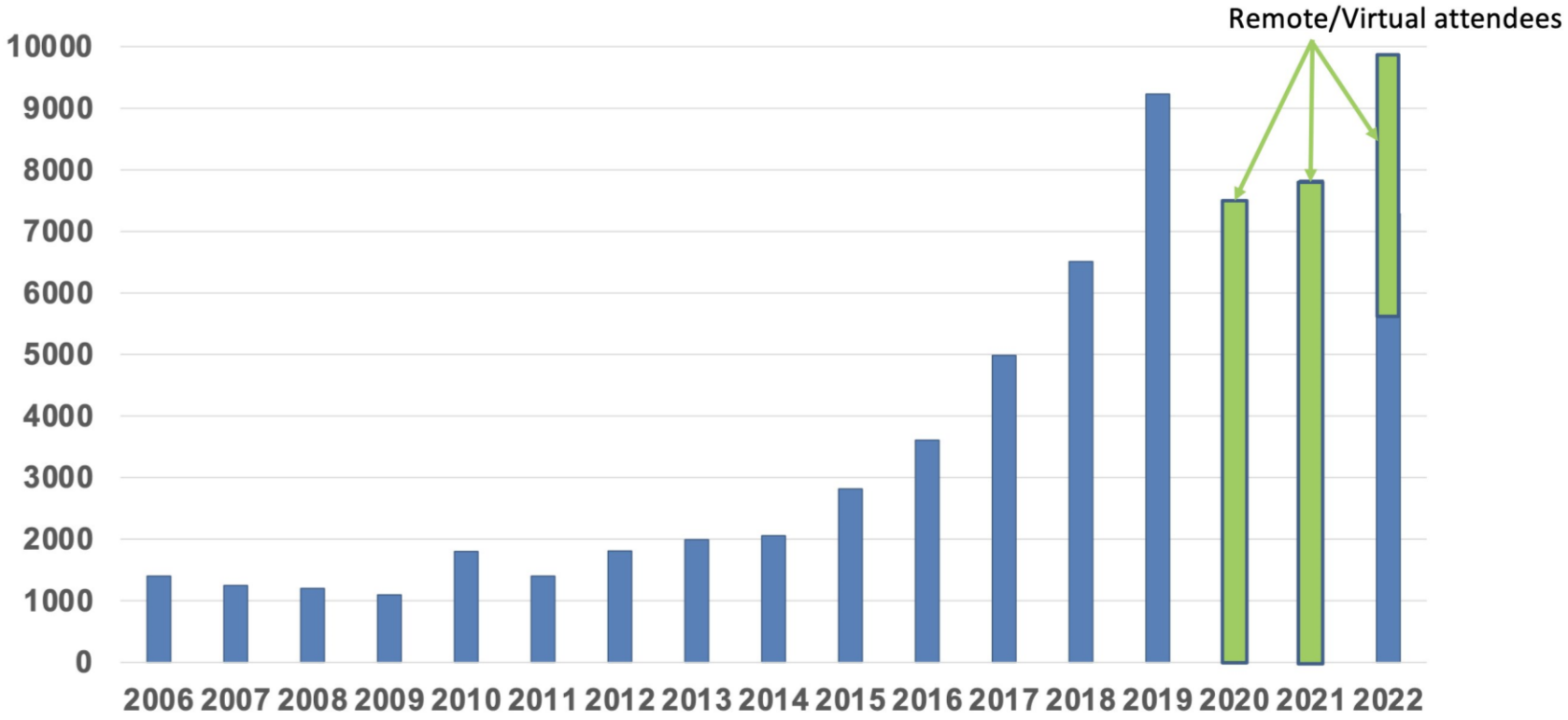
Related Disciplines



Why Vision?

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

CVPR Attendance Trend (as of June 20, 2022)



StartupHub.ai

ISRAEL'S **COMPUTER VISION** STARTUPS

Faces and Digital Cameras

Camera waits for everyone to smile to take a photo [Canon]

Setting camera focus via face detection

night mode [Apple]

Linking to Info with a Mobile Device

Faces and Digital Cameras

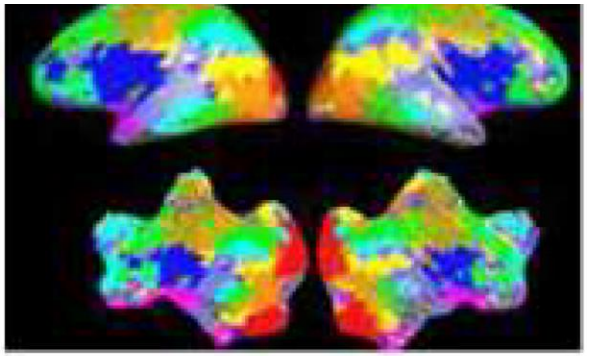
Human joystick, NewsBreaker Live

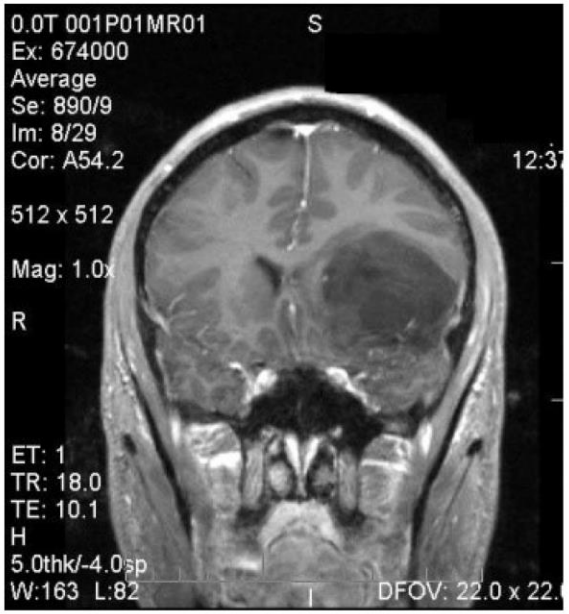
Assistive technology systems
Camera Mouse, Boston College

Microsoft Kinect

vision for medical and neuro images

Image guided surgery
MIT AI Vision Group





Special Visual Effects

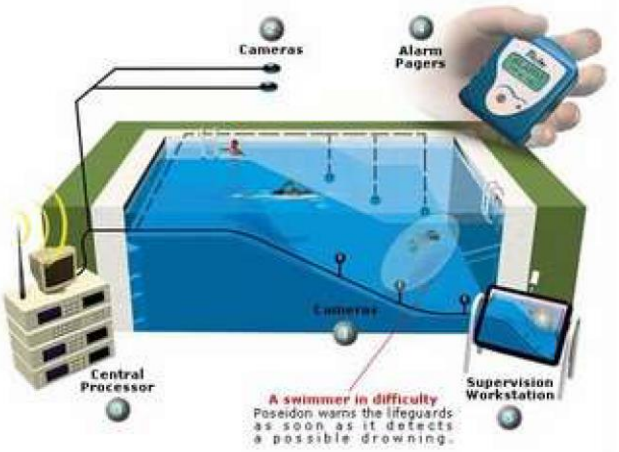
The Matrix

What Dreams May Come

Mocap for *Pirates of the Carribean*,
Industrial Light and Magic
Source: S. Seitz

Safety and Security

Navigation,
driver safety

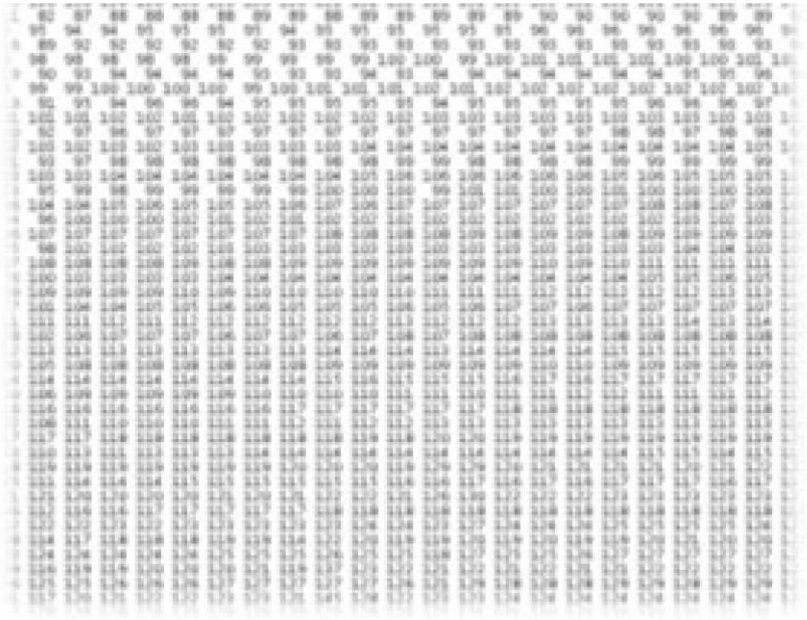


Monitoring pool
(Poseidon)

Pedestrian detection
MERL, Viola et al.

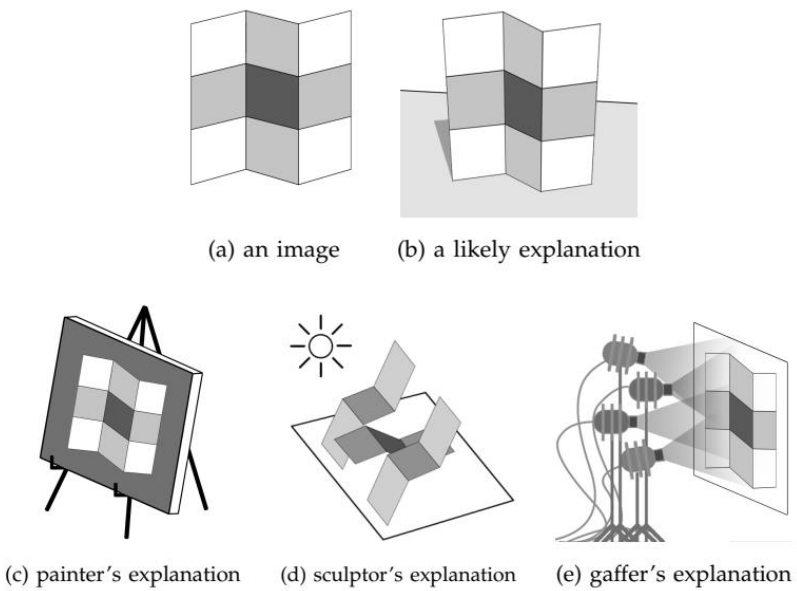
Surveillance

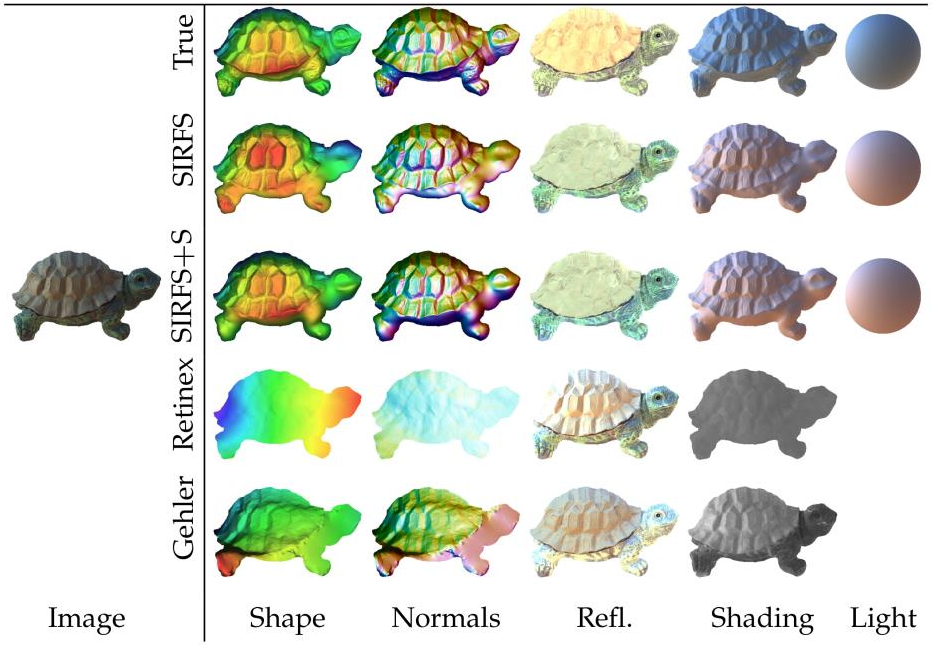
What the computer gets



Why is vision difficult?

- Ill-posed problem: real world is much more complex than what we can measure in images
  - e.g. 3D→2D
- Impossible to "invert" the image formation process

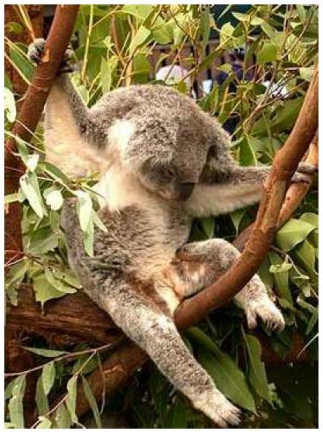




"SIRFs", Barron and Malik, PAMI 2020

Challenges: Many nuisance parameters

Illumination



Object pose

Clutter

Occlusions

Intra-class appearance

Viewpoint

Challenges: Intra-class variation

slide credit: Fei-Fei, Fergus & Torralba

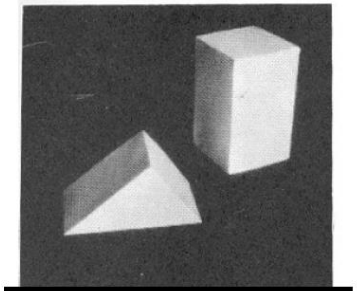
Challenges: Importance of context

Challenges: Importance of context

Challenges: Complexity

- Millions of pixels in an image
- 30,000 human recognizable object categories
- 30+ degrees of freedom in the pose of articulated objects like humans
- Billions of images online
- 500 hours of video uploaded to YouTube every minute
- About half of the cerebral cortex in primates is devoted to processing visual information

Progress charted by datasets



Roberts 1963

COIL

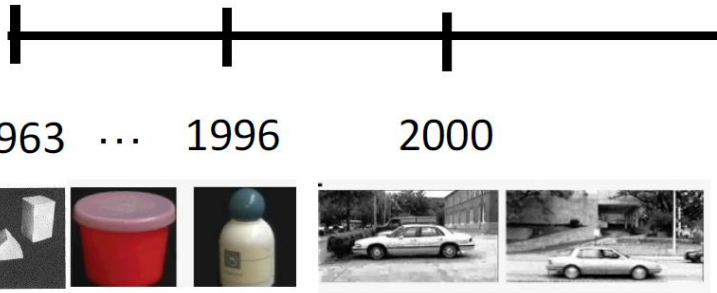
1963 ⋯ 1996

Progress charted by datasets

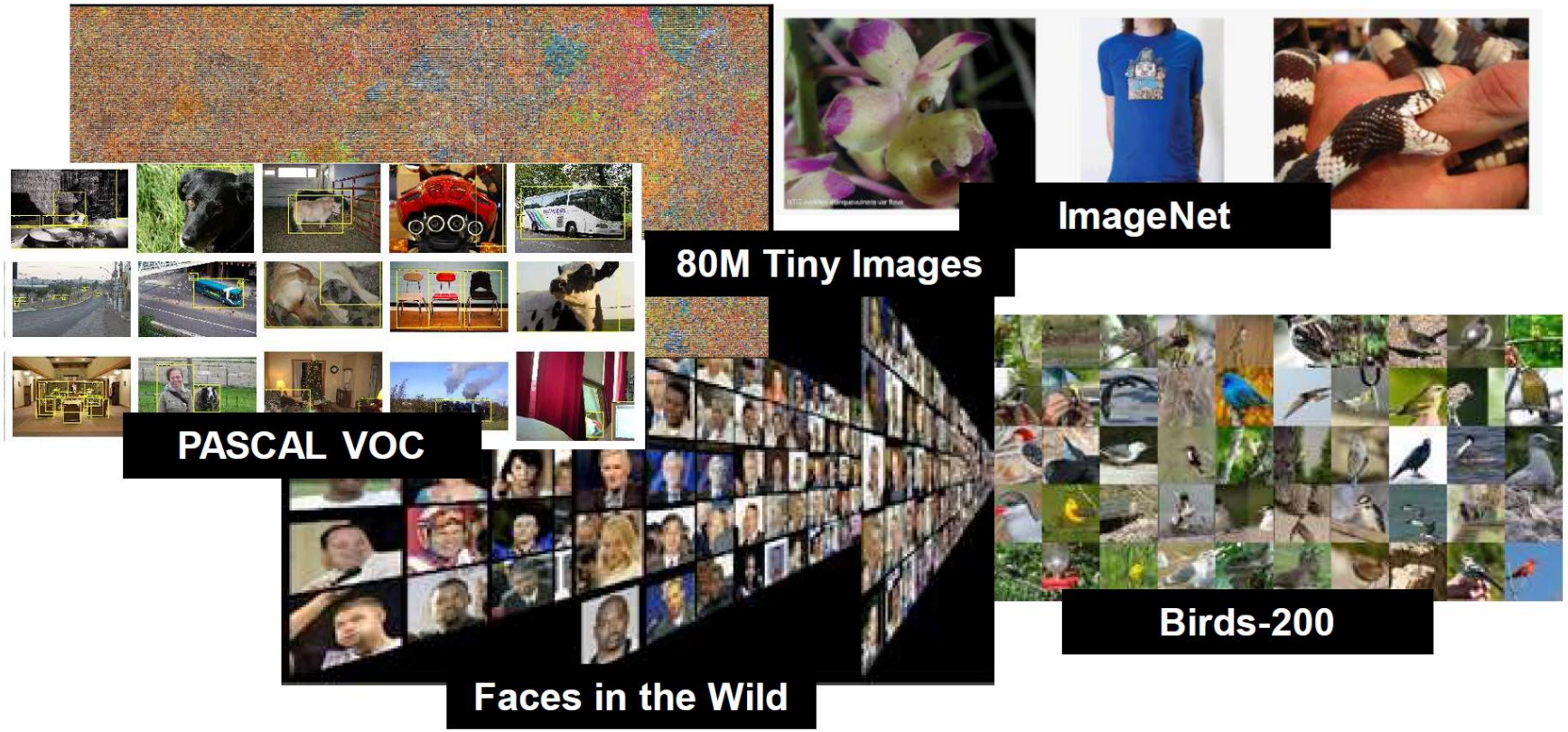
MSRC 21 Objects

Caltech-101

Caltech-256



Progress charted by datasets



Expanding horizons: large-scale recognition

clarifai

ABOUT TECHNOLOGY API NEWS BLOG CAREERS CONTACT

Paste a url here...

USE THE URL CHOOSE A FILE INSTEAD

*By using the demo you agree to our terms of service

Predicted Tags

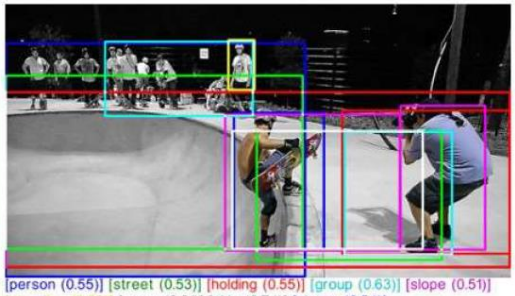
mammal livestock cattle

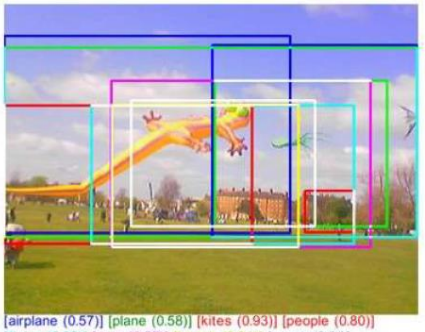
pasture agriculture bovine

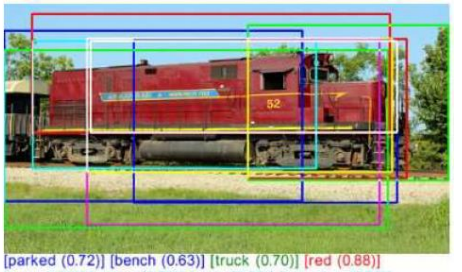
farm nobody meadow grass

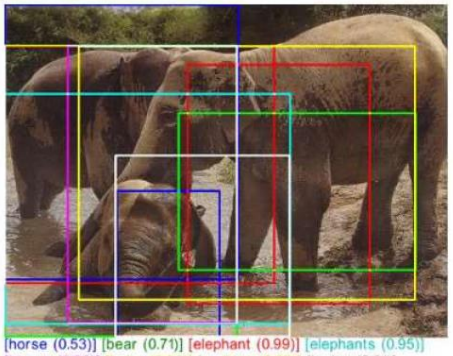
Similar Images

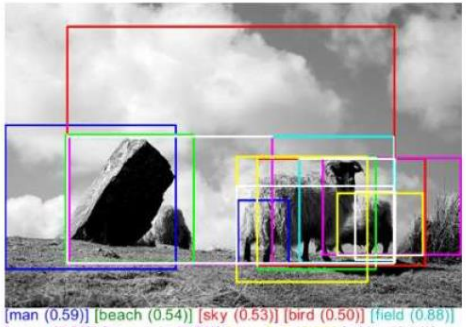
Expanding horizons: Captioning

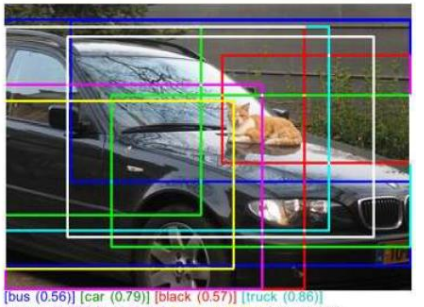




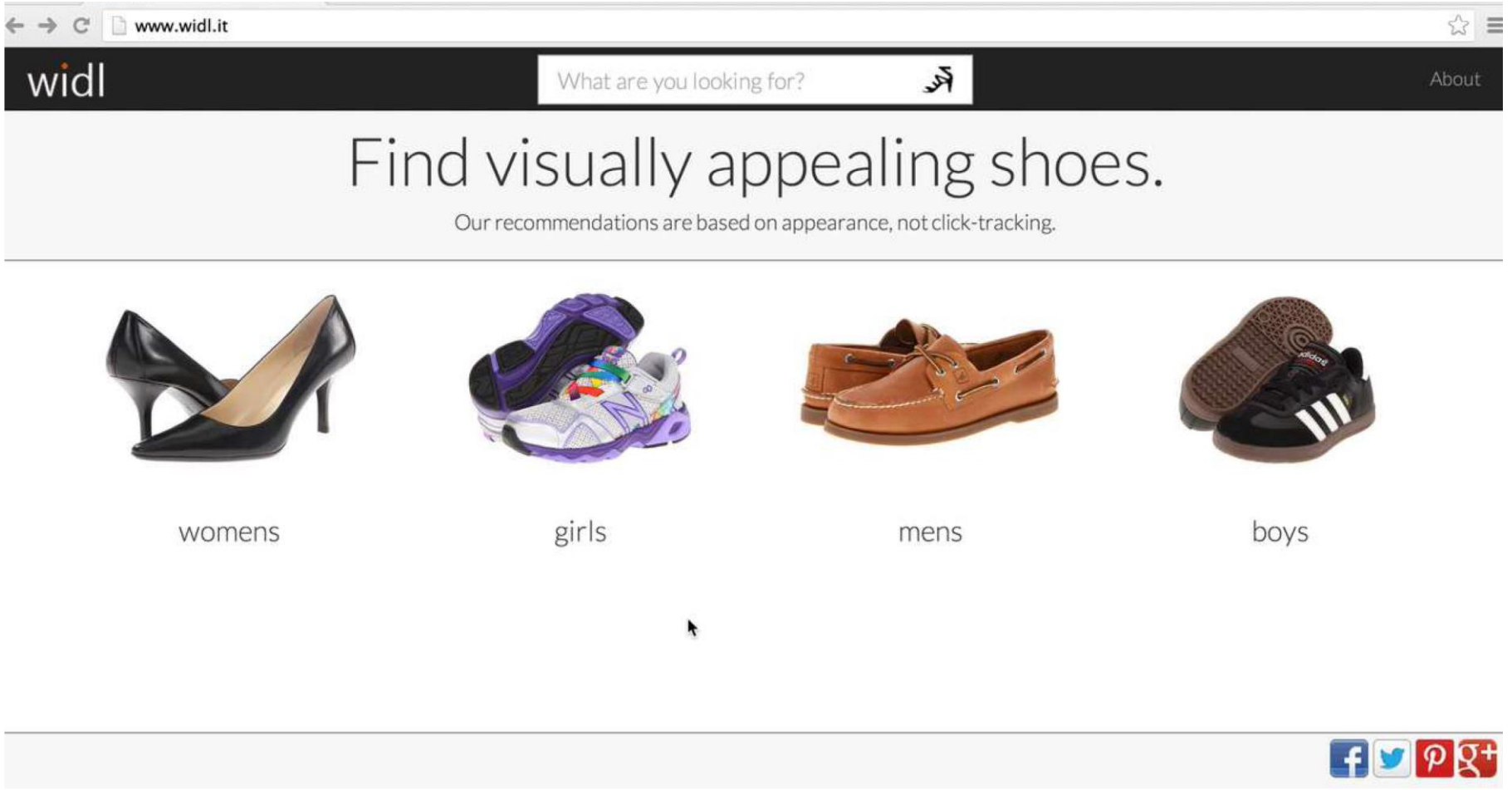




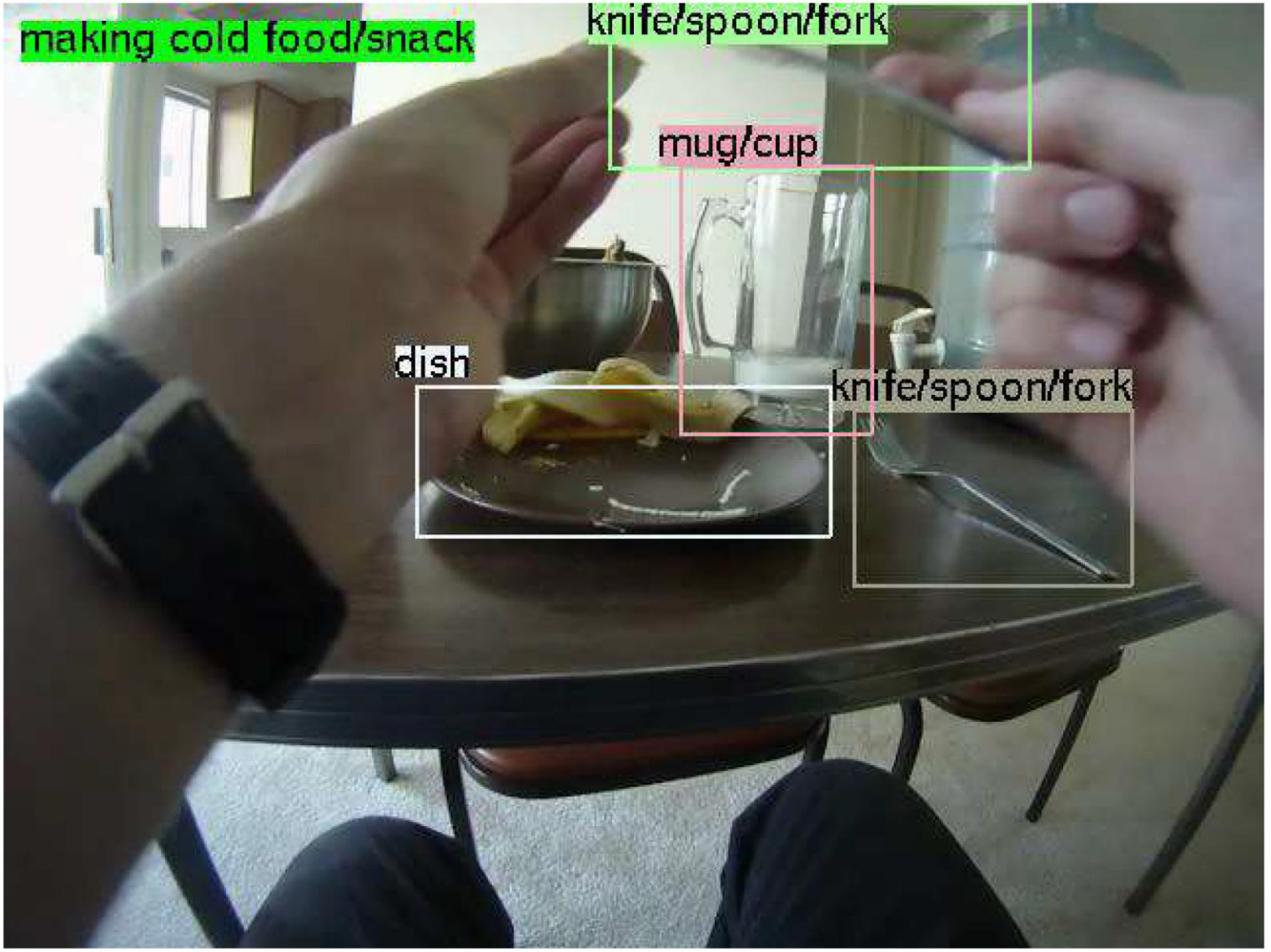




Expanding horizons: interactive visual search



Expanding horizons: first-person vision



Expanding horizons: enhancement / generation

Goals of the course

- Introduction to primary topics
- Hands-on experience with algorithms
- Views of vision as a research area

Topics mostly covered

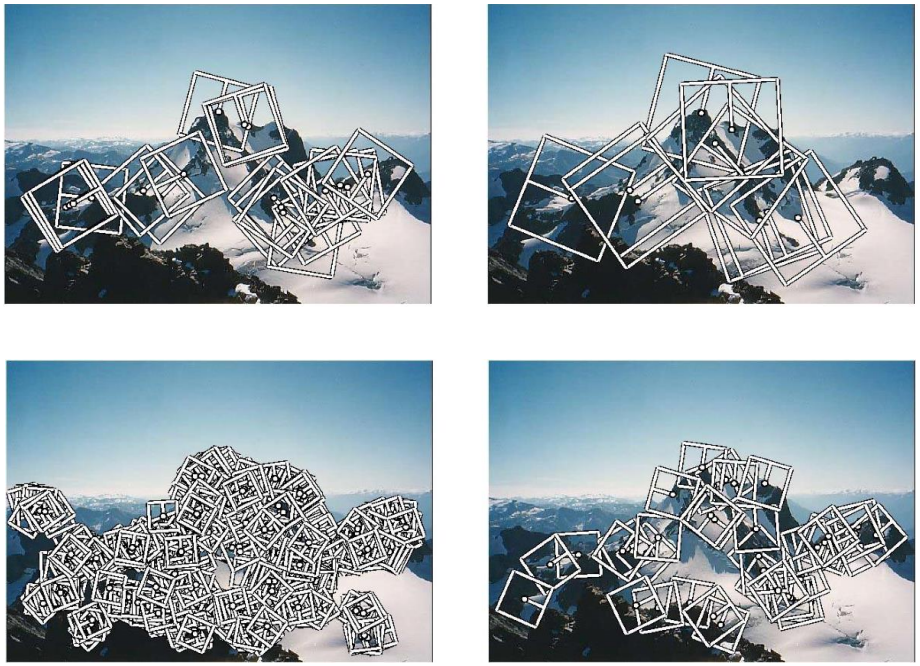
Image processing:

- Basics of filtering.
- Image pyramids.
- Gradients and lines.
- Hough transforms.
- Corner detection.

Topics to be covered

Feature detection and correspondences:

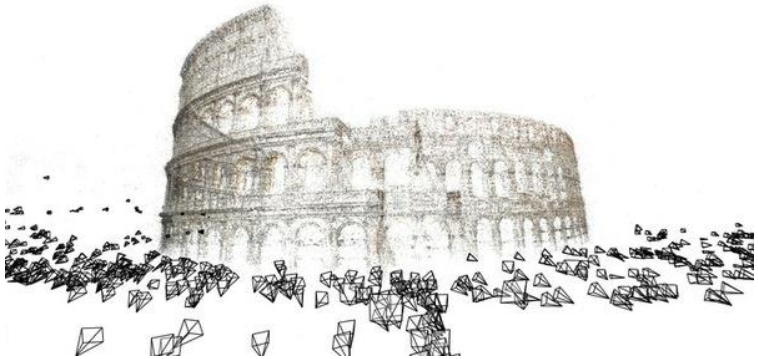
- SIFT ...
- Feature descriptors.
- RANSAC.



Topics to be covered

3D vision:

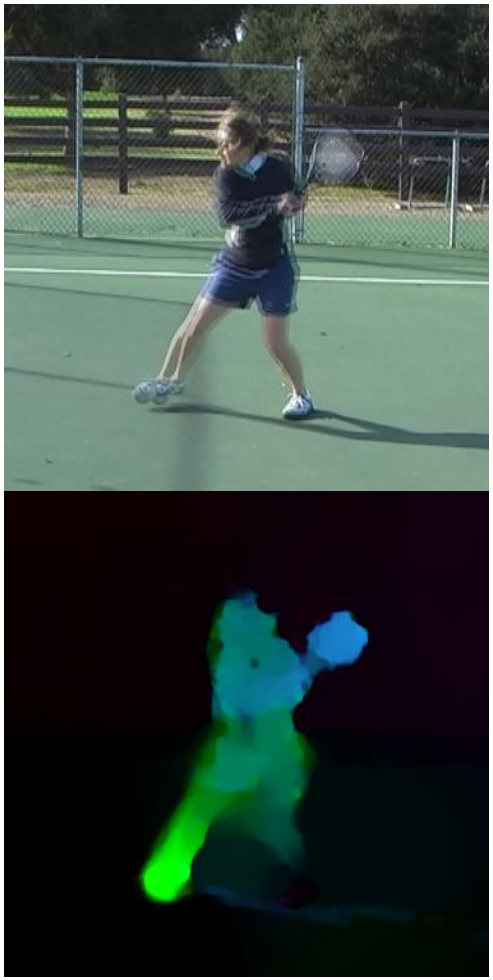
- Homographies and image alignment.
- Camera models.
- Fundamental matrix.
- Epipolar geometry.
- Structure from motion.



Topics to be covered

Dealing with motion:

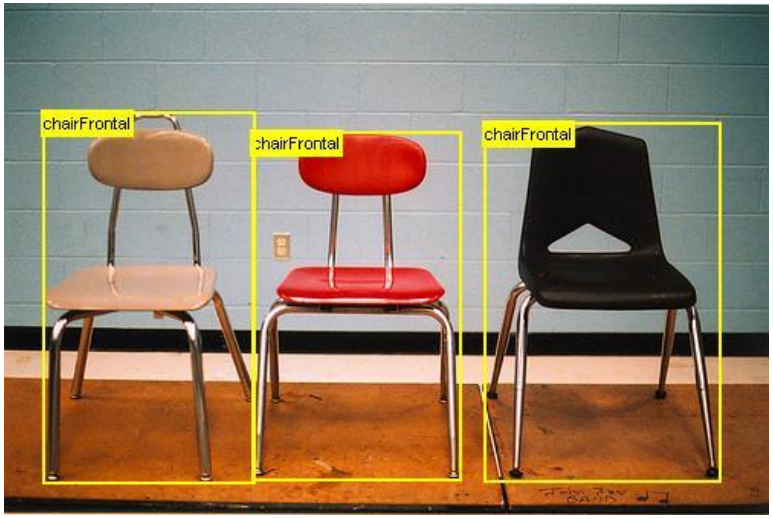
- Image registration.
- Photometric stereo.
- Optical flow.



Topics to be covered

Recognition - Objects and learning:

- Classical approaches
- Convolutional neural networks
- Current advances



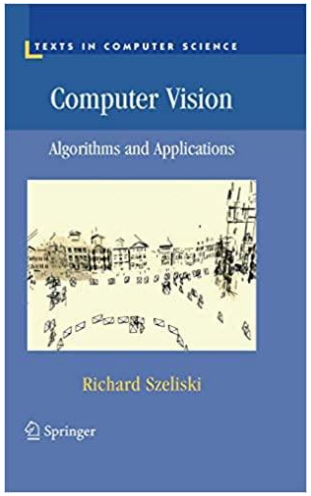
Topics to be covered

Generative models

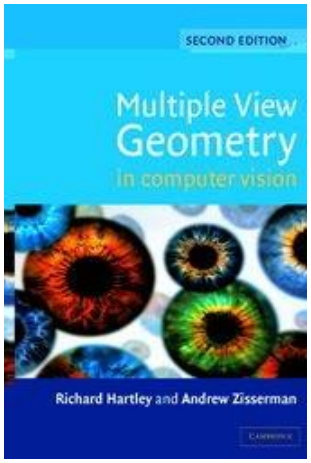
- Novel view synthesis
- Image generation

Textbooks

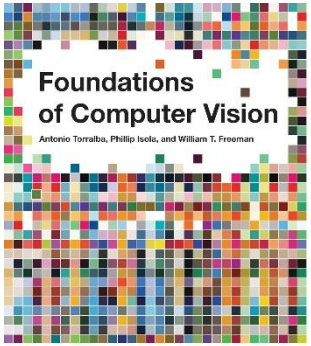
- Computer Vision: Algorithms and Applications
  - by Szeliski
  - https://szeliski.org/Book/



- Multiple View Geometry in Computer Vision
  - by Hartley and Zisserman
  - can be found online



- Foundations of Computer Vision
  - by Torralba, Isola and Freeman



Topics - outline

- **Intro and 2D vision** (~3 lectures)
  - introduction
  - feature detection, description and matching
  - image transformations
- **3D vision** (~3 lectures)
  - image transformations
  - camera models
  - 2-view geometry
  - stereo and reconstruction
- **Learning based vision** (~6 lectures)
  - neural networks
  - convolutional nets
  - training principles
  - applications: detection, recognition and generation

Grading

- Will be based on 4 assignments
  - 2D vision
  - 3D vision
  - DL basics
  - DL applications
- Majority will be programming problems
  - implementation
  - explanation
  - discussion of results
- Will use Python -> Python+Pytorch
- Must be done in pairs (or individually upon approval)
- Deadlines are firm

Misc

- **Slide credit**: Svetlana Lazebnik, Sanja Fidler, Kristen Grauman, Ioannis Gkioulekas, Kris Kitani, James Hays, Fredo Durand, Rick Szeliski, Andrew Zisserman, and more