



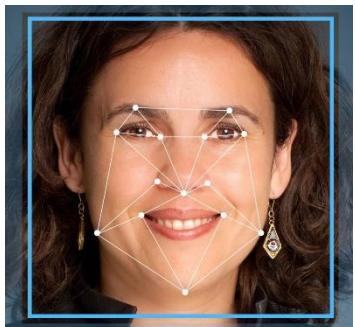
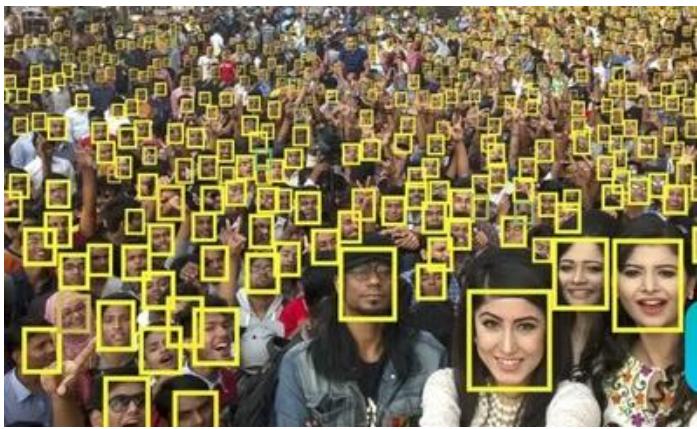
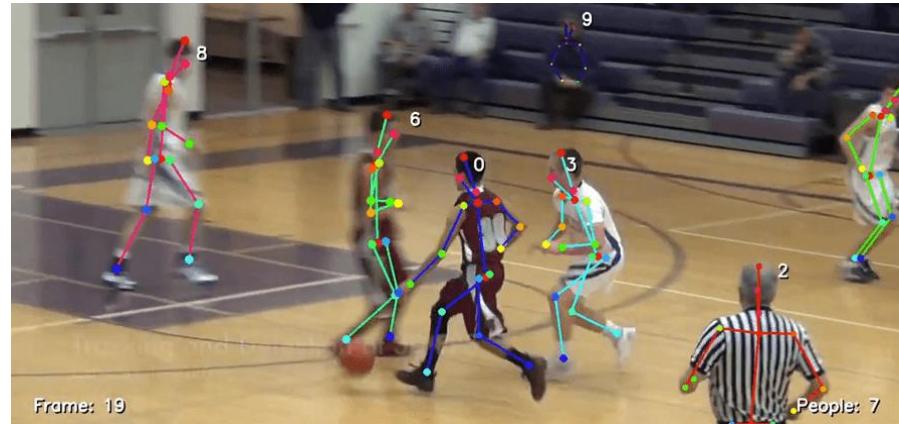
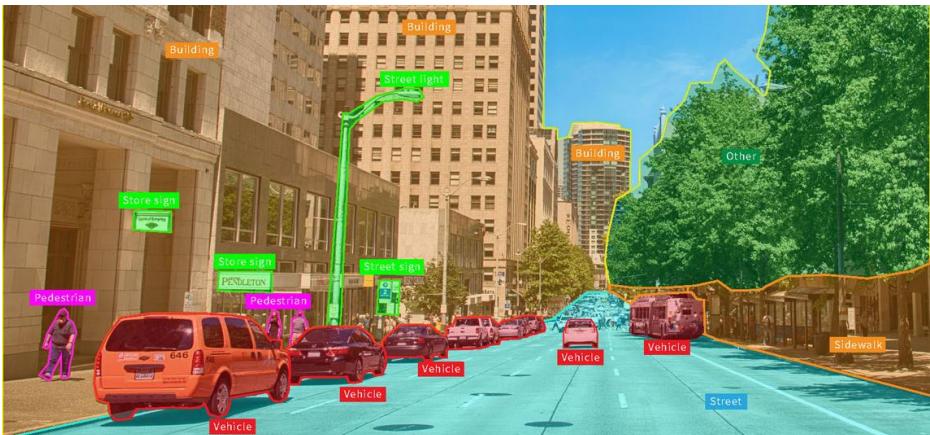
A large grid of small images covering various topics, including landscapes, vehicles, buildings, interior rooms, electronic devices, sports equipment, food items, insects, animals, and people.

Computer Vision

Lec 1 - Introduction

University of Haifa

Simon Korman

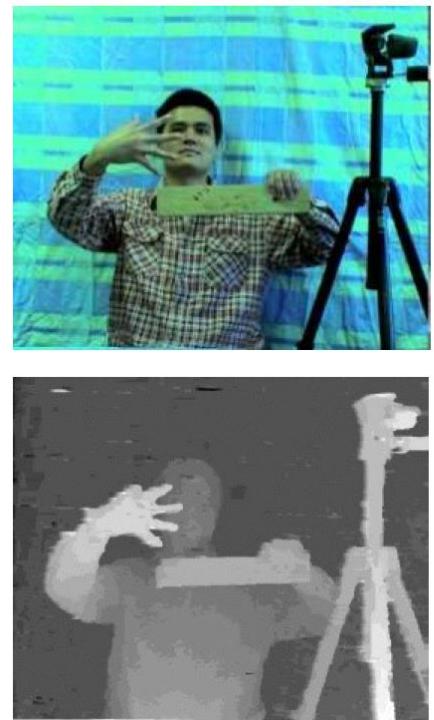


Computer Vision

- 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



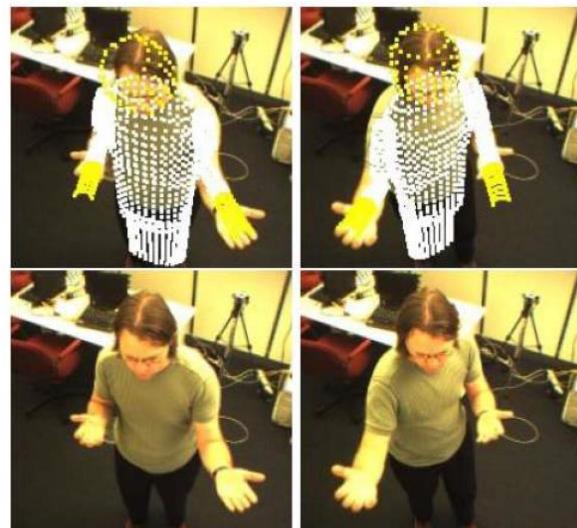
Wang et al.

Structure from motion



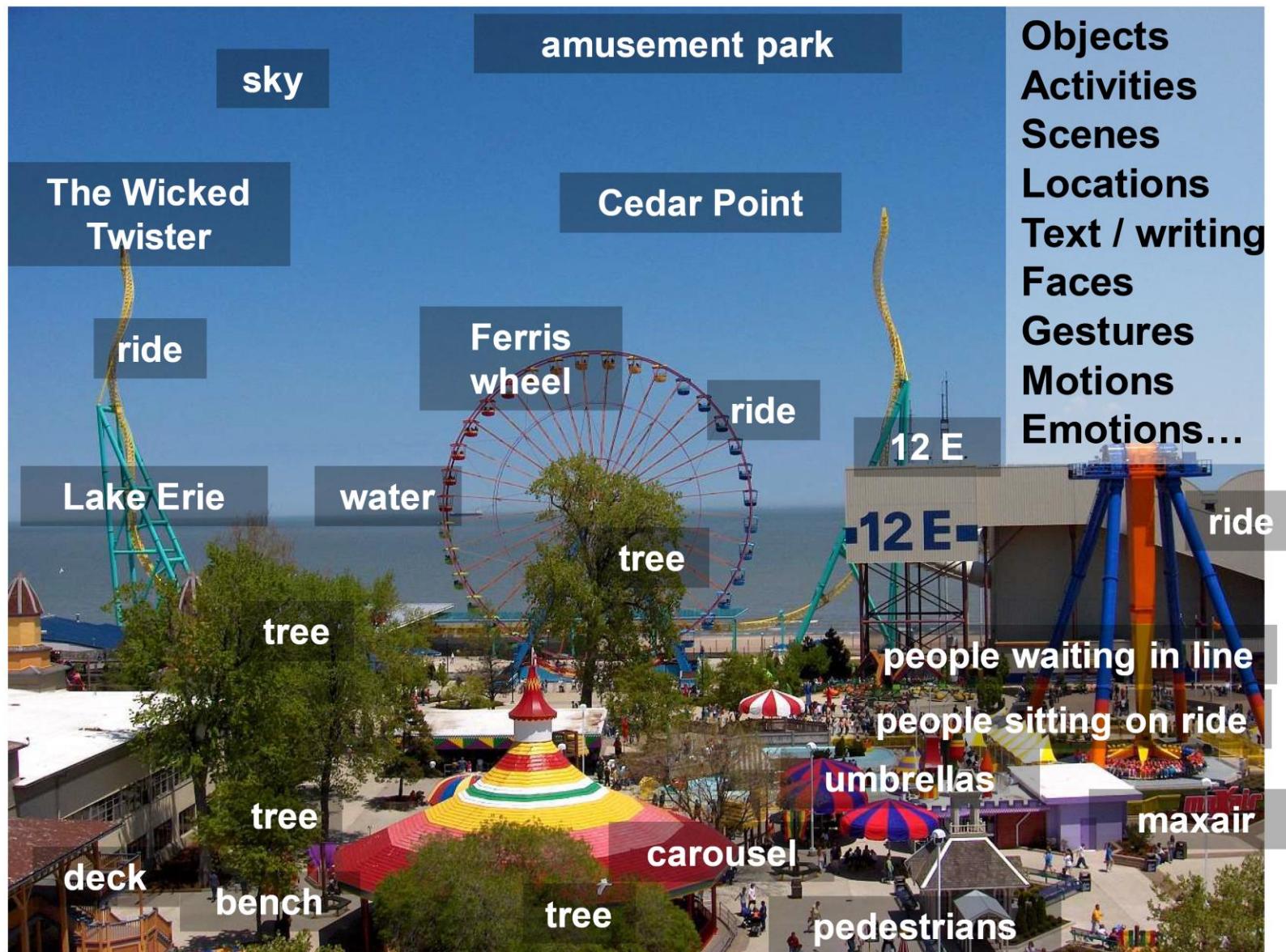
Snavely et al.

Tracking

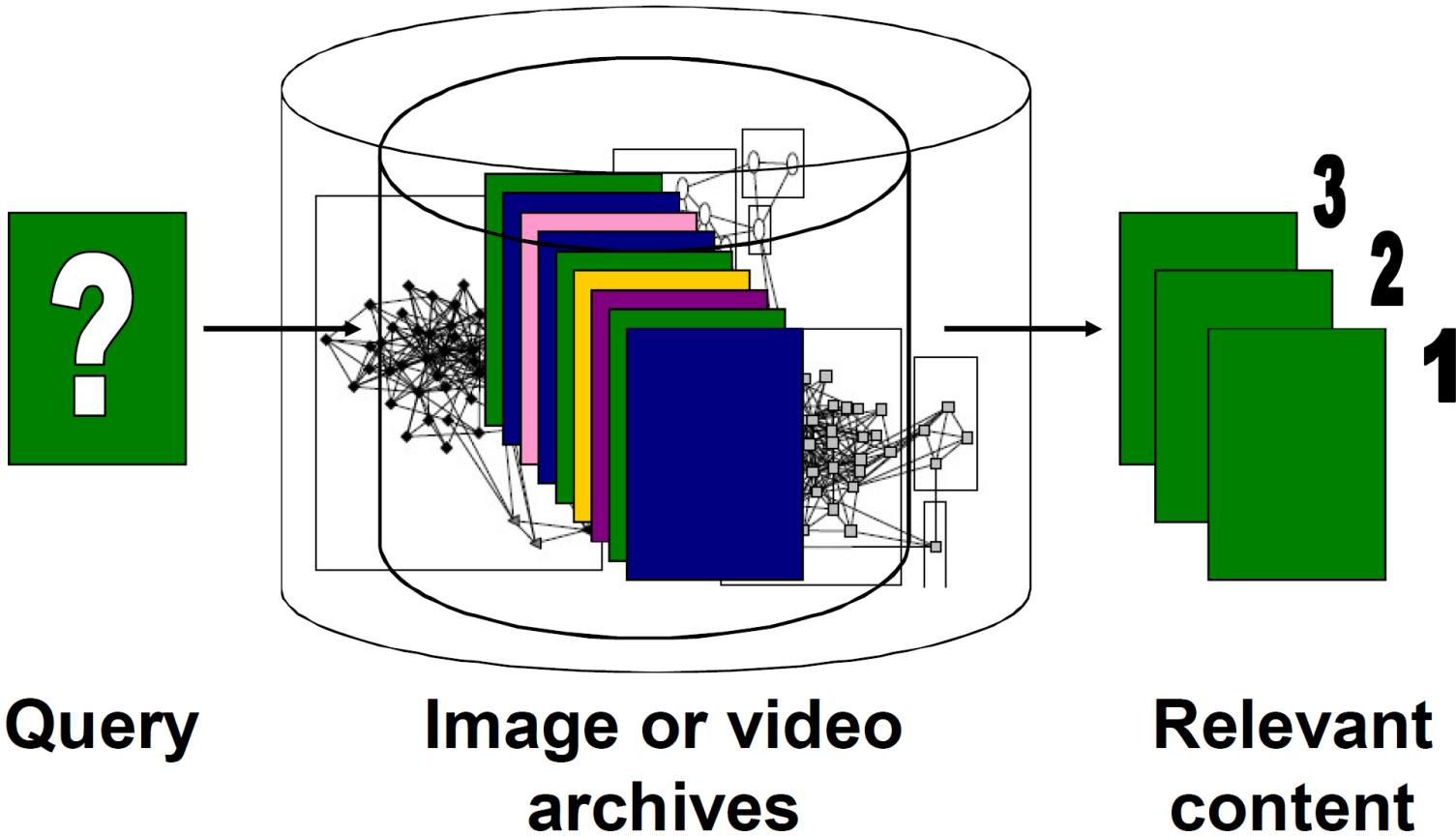


Demirdjian et al.

vision for perception, interpretation



visual search, organization



generation

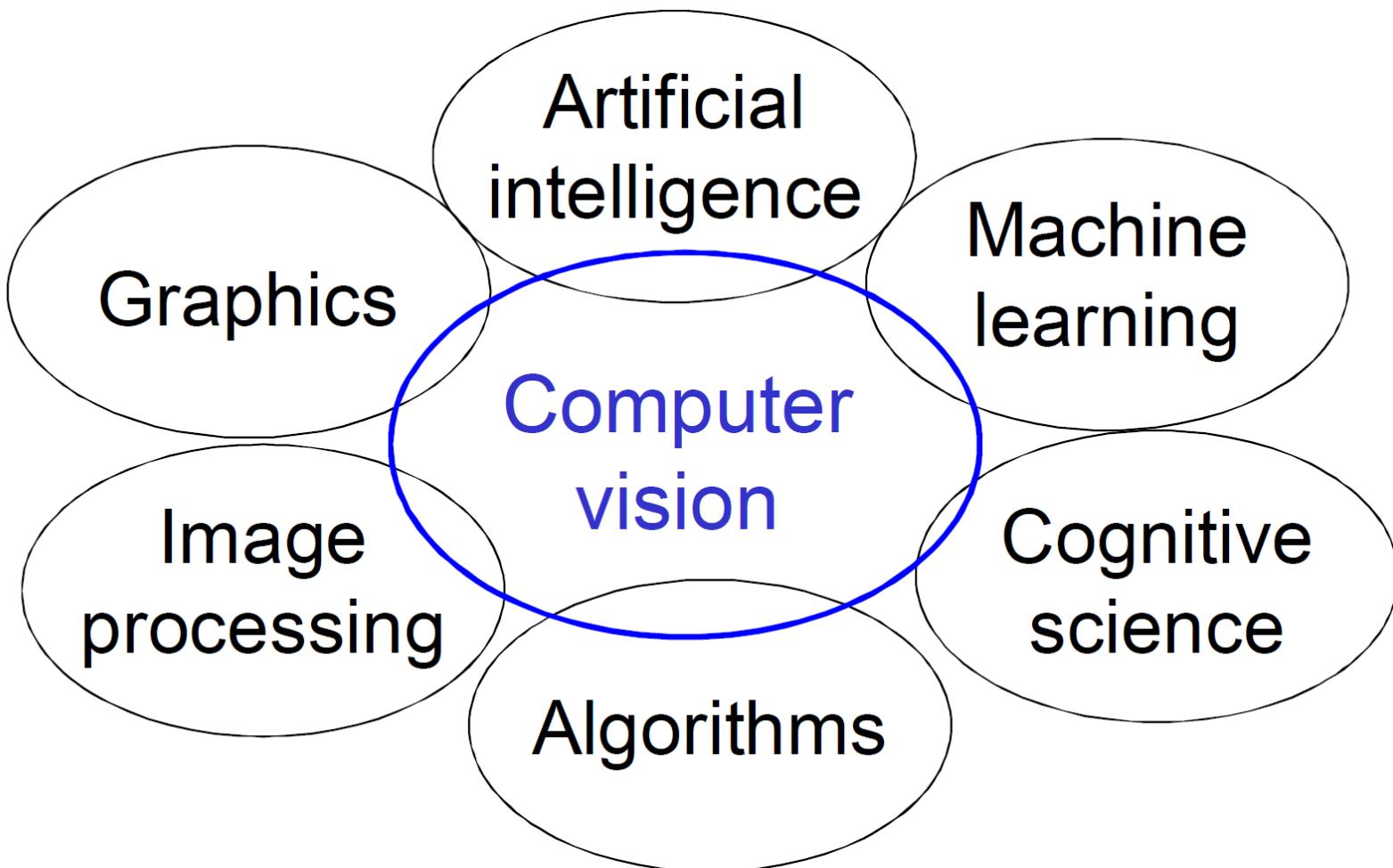
A hamburger in
the shape of a
Rubik's cube



A velociraptor
working at a
hotdog stand



Related Disciplines



Vision and Graphics

Inverse problems:
analysis and synthesis.

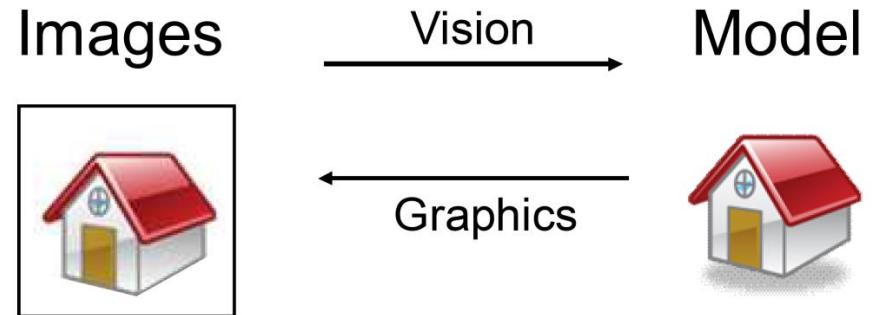
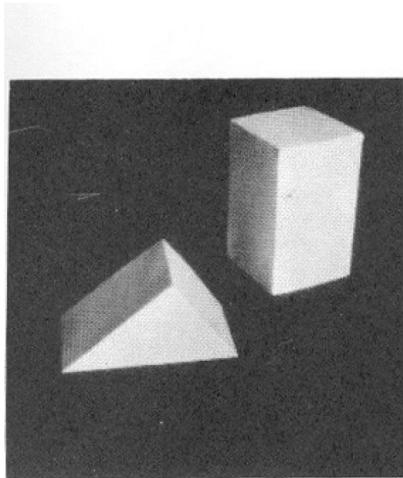


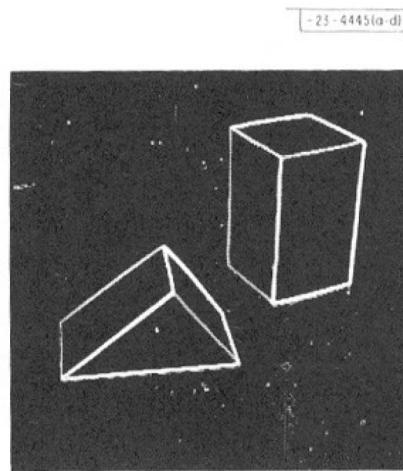
Image Generation



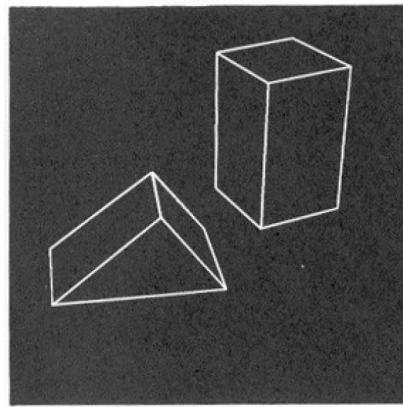
Visual Data in 1963



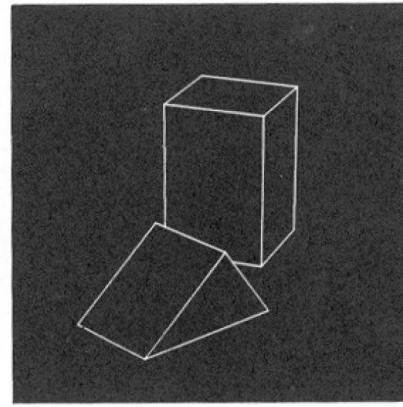
(a) Original picture.



(b) Differentiated picture.



(c) Line drawing.



(d) Rotated view.

- 23 - 4445(a-d)

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

Visual Data Nowadays



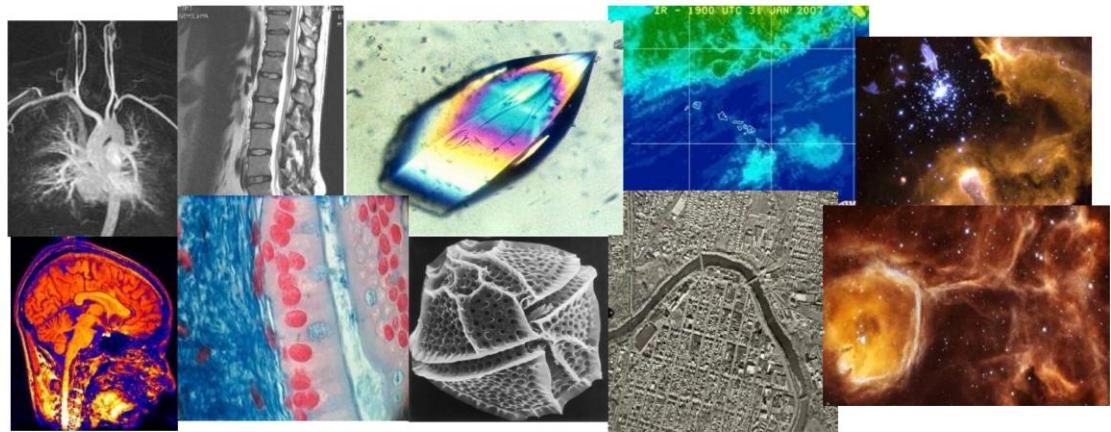
Personal photo albums



Movies, news, sports



Surveillance and security

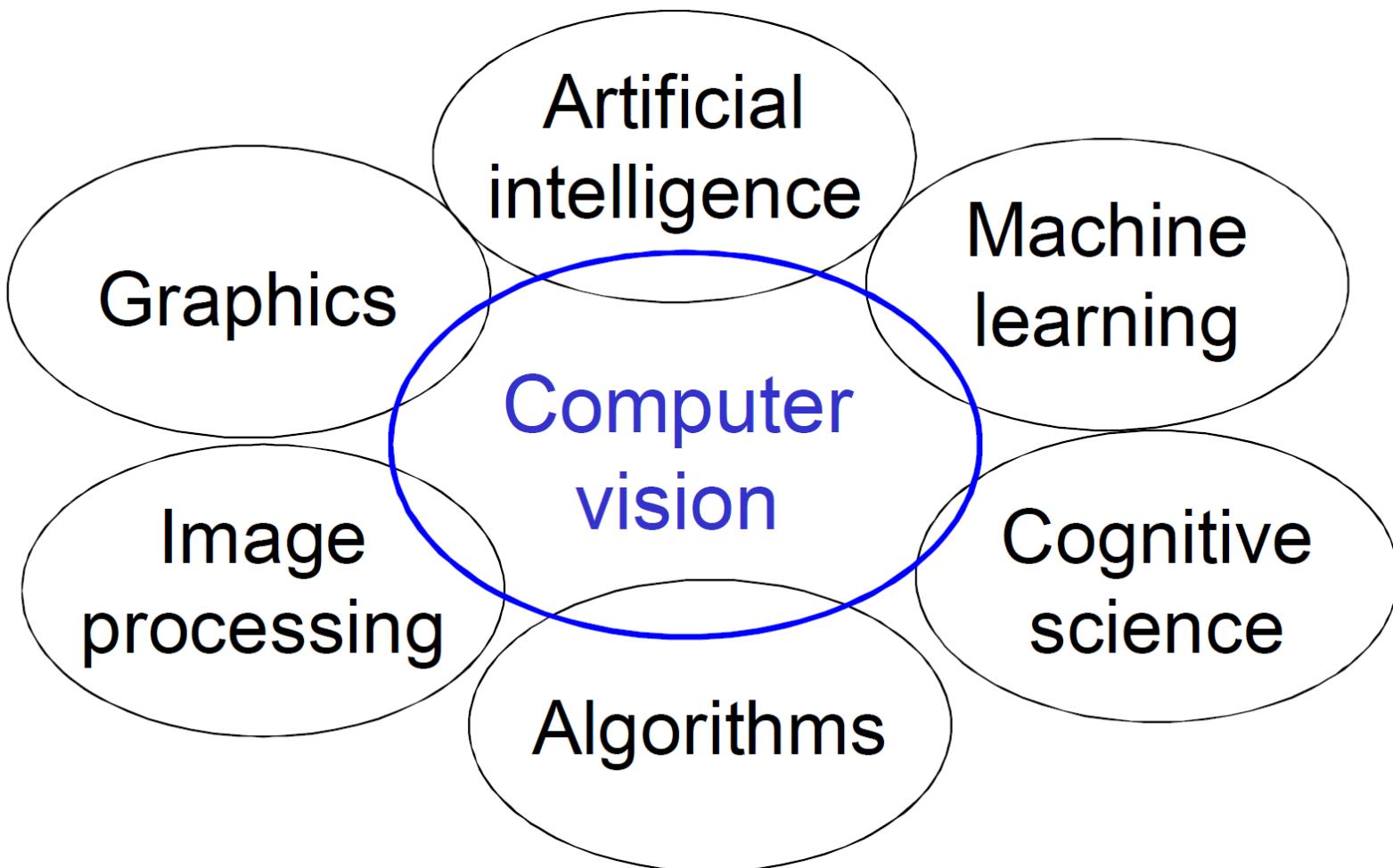


Medical and scientific images

+ many dedicated large-scale datasets

Slide credit: L. Lazebnik

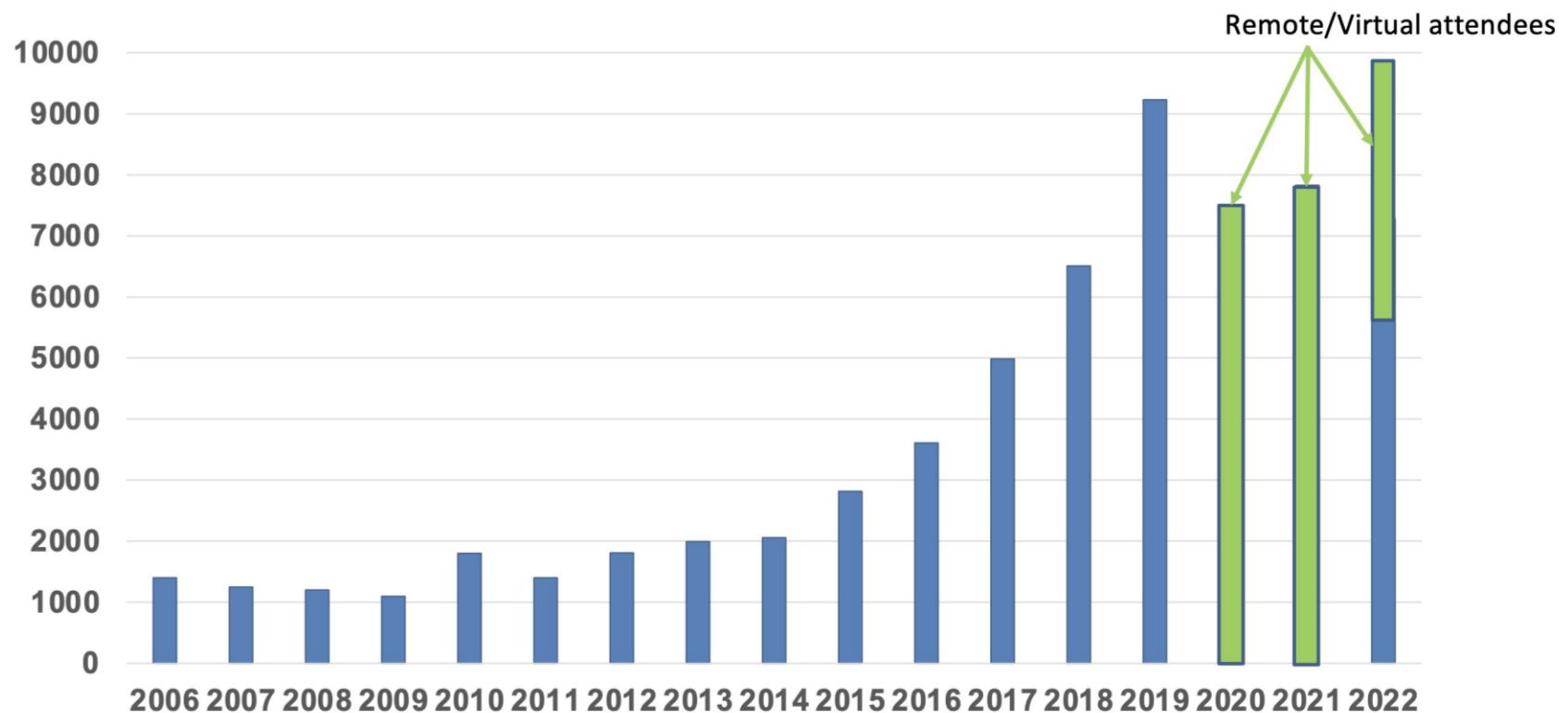
Related Disciplines



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)



ISRAEL'S COMPUTER VISION STARTUPS



COMPUTER VISION TECHNOLOGY

VIDEO INTELLIGENCE



HEALTHCARE

MEDICAL IMAGING

OPTICAL & SENSOR

FACIAL RECOGNITION

DATA CREATION

DEVELOPMENT

OPTICAL

AUGMENTED REALITY

PLATFORM

EYE TRACKING

AUTOMOTIVE

AUTONOMOUS

IN-CAR MONITORING

TRAFFIC & MOBILITY

DEVELOPMENT

IMPAIRMENT AID

VR, SURGERY & MONITOR

EV

CONTINUOUS MONITORING

IMPACT MITIGATION

TELEOPERATION

INSPECTION

DASHCAM & ROUTING

CONSTRUCTION

MARKETING

WATER VISION

EDUCATION, RAIL & TRAVEL

AGRICULTURE

CROP MANAGEMENT

HARVESTING

NEW MEDIA

TELEOPERATION

INSPECTION

DASHCAM & ROUTING

CONSTRUCTION

MARKETING

WATER VISION

EDUCATION, RAIL & TRAVEL

INDUSTRIAL

ROBOTICS & UTILITIES

DRONES

SURVEILLANCE & AUTONOMY

ENTERPRISE

SECURITY, DEV. & SUPPORT

MARKETING

VIDEO, CONTENT & SECURITY

WATER VISION

EDUCATION, RAIL & TRAVEL

RETAIL

MONITORING & ANALYTICS

SPORTS

ANALYTICS & CONTENT

ENTERPRISE

SECURITY, DEV. & SUPPORT

MARKETING

VIDEO, CONTENT & SECURITY

WATER VISION

EDUCATION, RAIL & TRAVEL

SECTORS

SMART CITY

VISUAL SEARCH

CONSUMER ROBOTICS & TECH

FITNESS

REAL ESTATE

FASHION

WATER VISION

EDUCATION

Faces and Digital Cameras



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

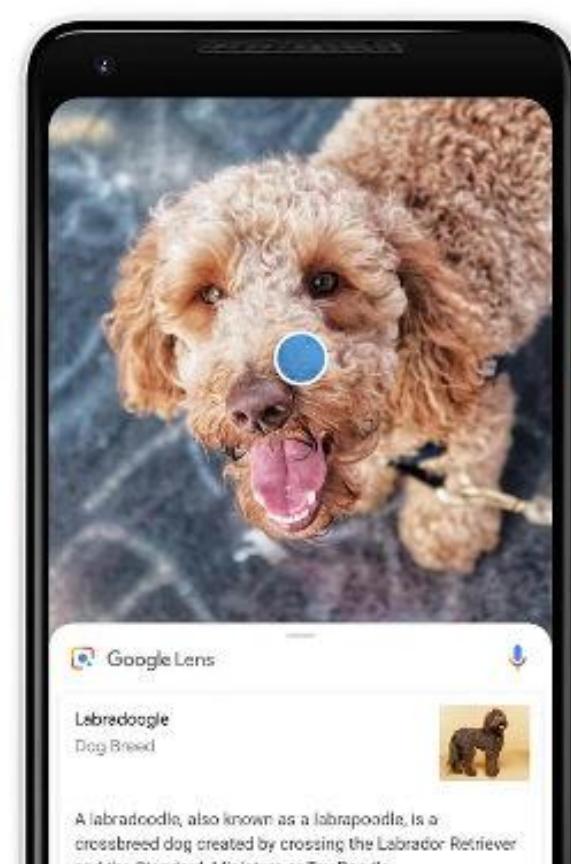
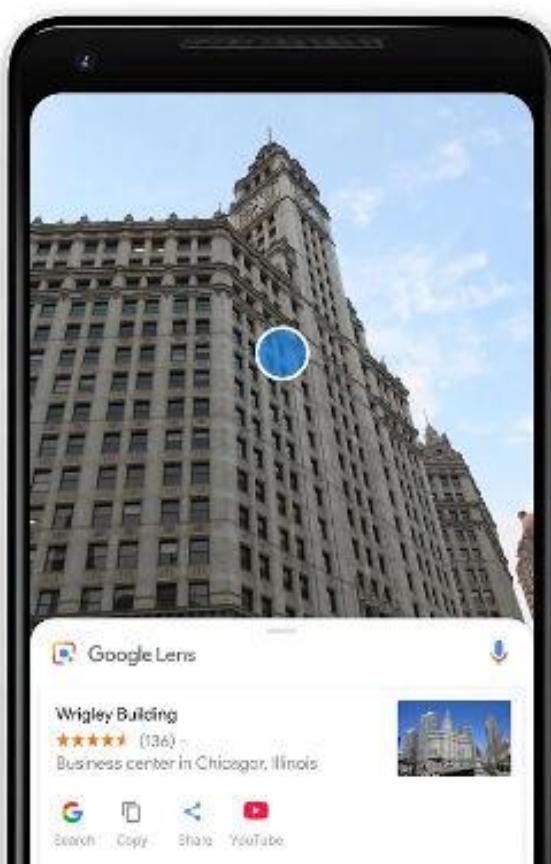


Setting camera focus via face detection

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

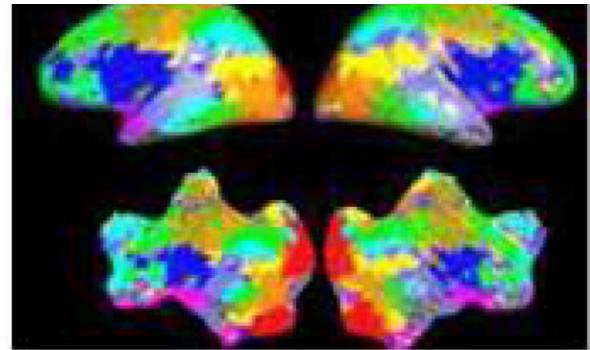


Microsoft Kinect

vision for medical and neuro images



Image guided surgery
MIT AI Vision Group



Special Visual Effects



The Matrix



What Dreams May Come

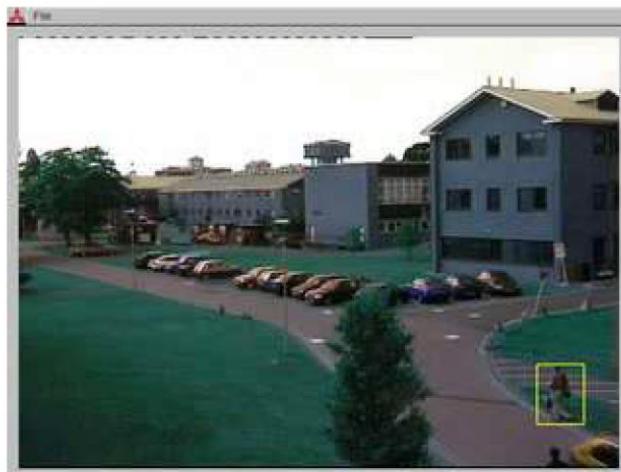


Mocap for *Pirates of the Caribbean*,
Industrial Light and Magic
Source: S. Seitz

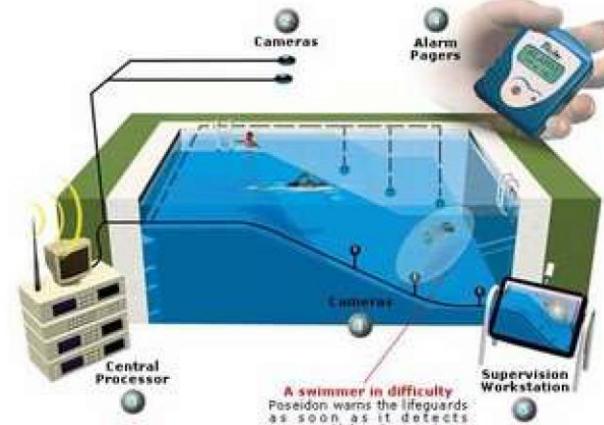
Safety and Security



Navigation,
driver safety



Pedestrian detection
MERL, Viola et al.

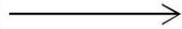


Monitoring pool
(Poseidon)



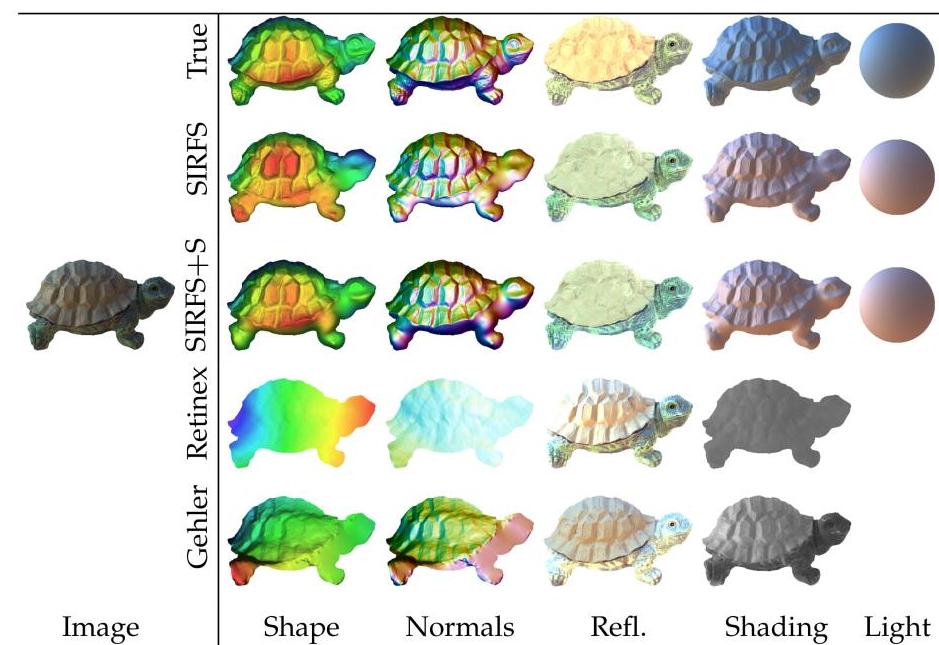
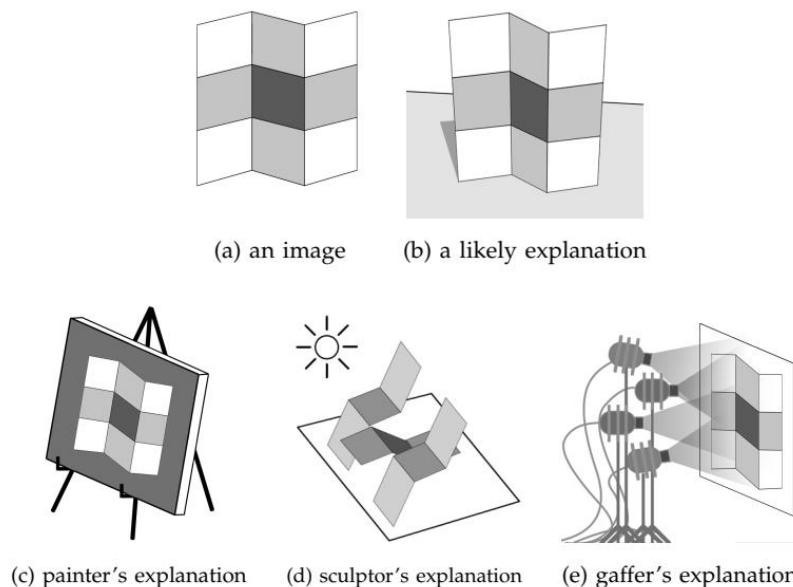
Surveillance

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



Challenges: Many nuisance parameters



Illumination



Object pose



Clutter



Occlusions



Intra-class
appearance



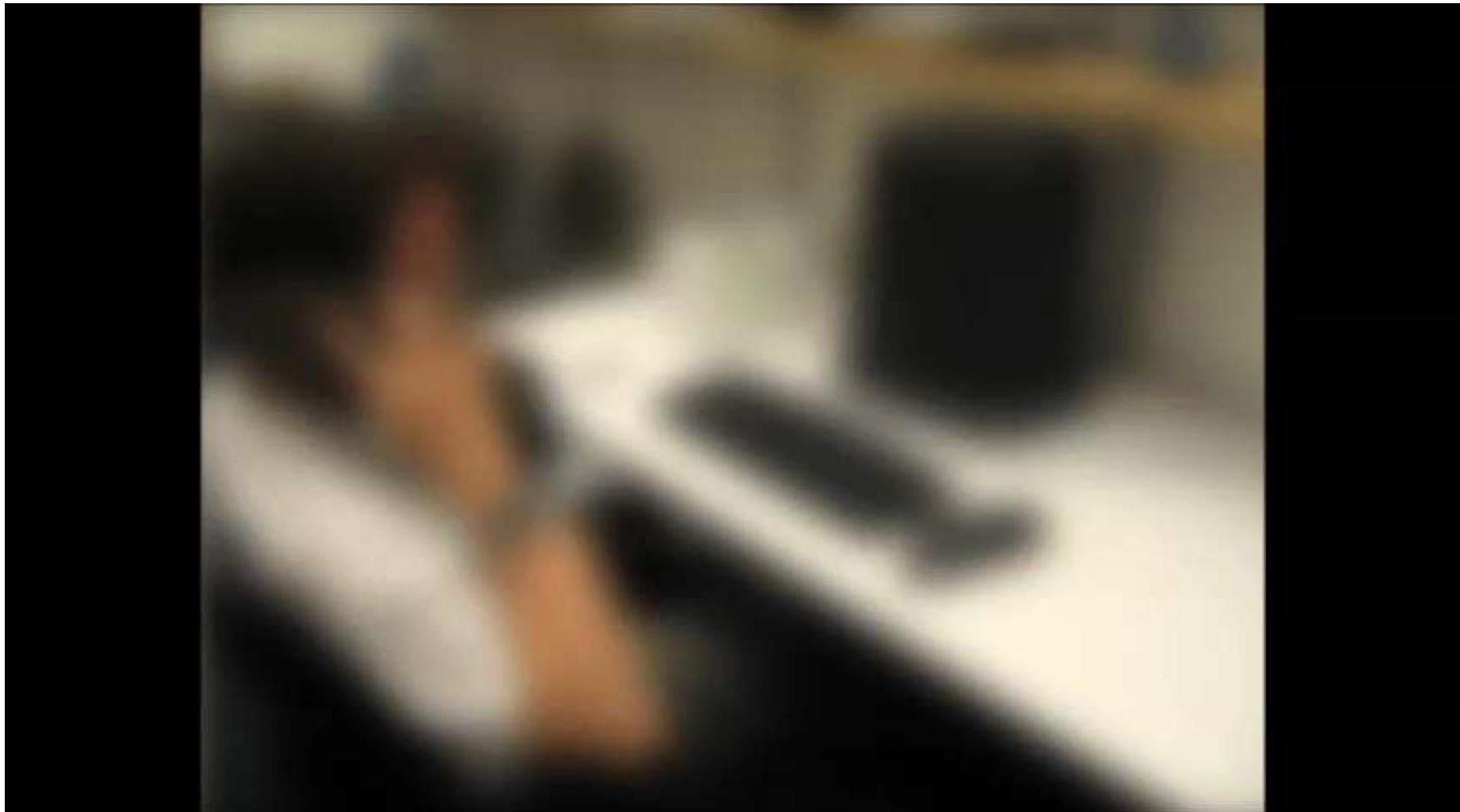
Viewpoint

Challenges: Intra-class variation



slide credit: Fei-Fei, Fergus & Torralba

Challenges: Importance of context



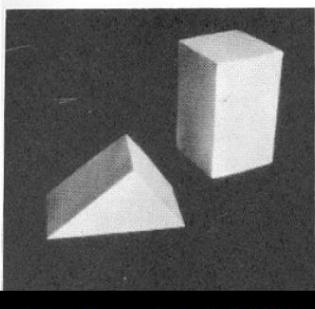
Challenges: Importance of context



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

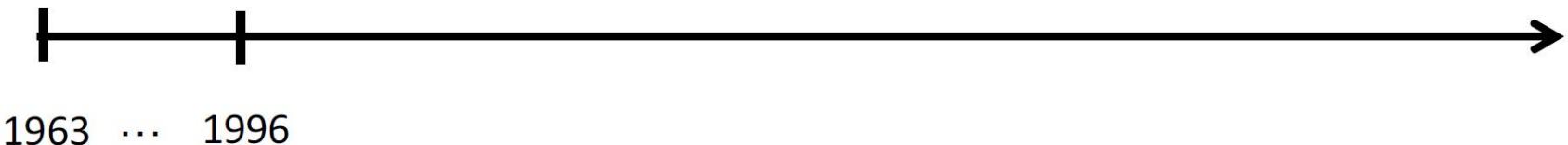
Progress charted by datasets



Roberts 1963



COIL



Progress charted by datasets



MSRC 21 Objects



Caltech-101



Caltech-256



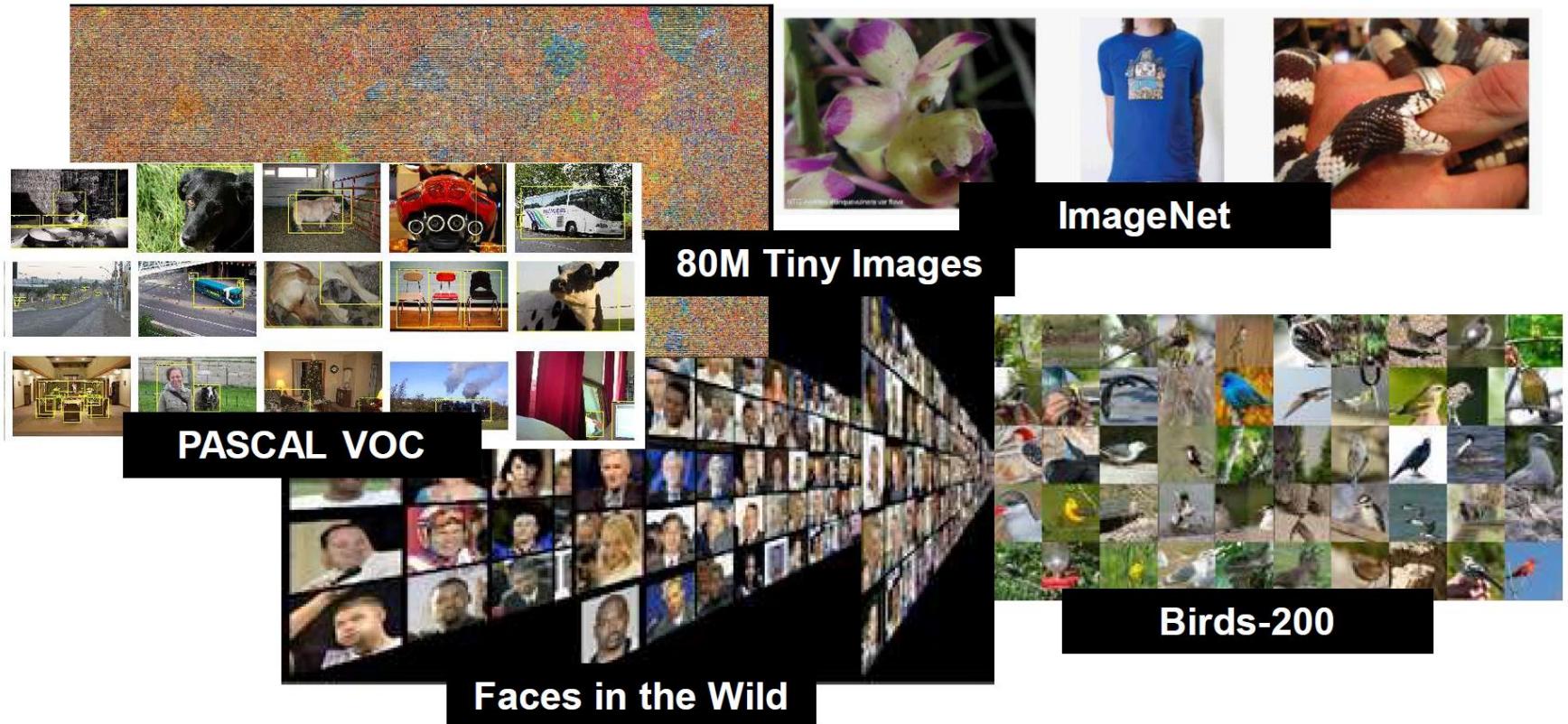
1963 ... 1996

2000

2005



Progress charted by datasets



Expanding horizons: large-scale recognition

[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



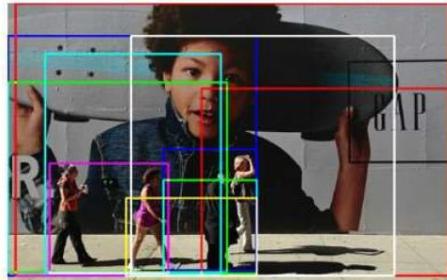
Predicted Tags

mammal livestock cattle
pasture agriculture bovine
farm nobody meadow grass

Similar Images

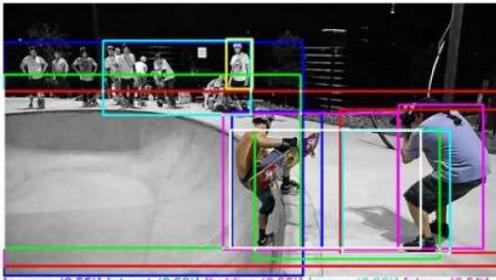


Expanding horizons: Captioning



[men (0.59)] [group (0.66)] [woman (0.64)]
 [people (0.89)] [holding (0.60)] [playing (0.61)] [tennis (0.69)]
 [court (0.51)] [standing (0.59)] [skis (0.58)] [street (0.52)]
 [man (0.77)] [skateboard (0.67)]

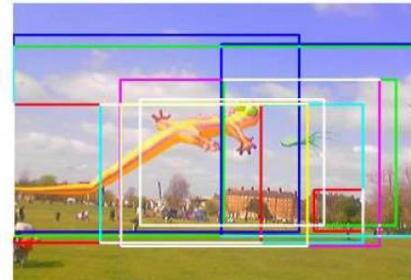
a group of people standing next to each other
 people stand outside a large ad for gap featuring a young boy



[person (0.55)] [street (0.53)] [holding (0.55)] [group (0.63)] [slope (0.51)]
 [standing (0.62)] [snow (0.91)] [skis (0.74)] [player (0.54)]
 [people (0.85)] [men (0.57)] [skiing (0.51)]

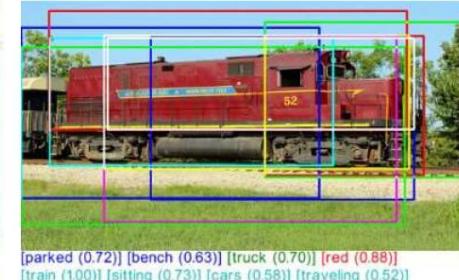
[skateboard (0.89)] [riding (0.75)] [tennis (0.74)] [trick (0.53)] [skate (0.52)]
 [woman (0.52)] [man (0.86)] [down (0.61)]

a group of people riding skis down a snow covered slope
 a guy on a skate board on the side of a ramp



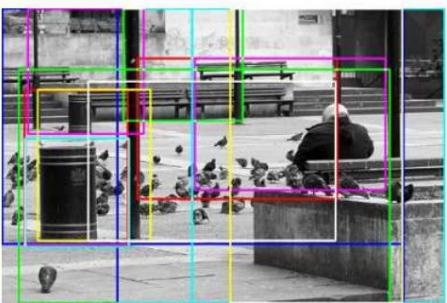
[airplane (0.57)] [plane (0.58)] [kites (0.93)] [people (0.80)]
 [flying (0.93)] [man (0.57)] [beach (0.84)] [wave (0.61)]
 [sky (0.61)] [kite (0.74)] [field (0.75)]

a couple of people flying kites in a field
 people in a field flying different styles of kites



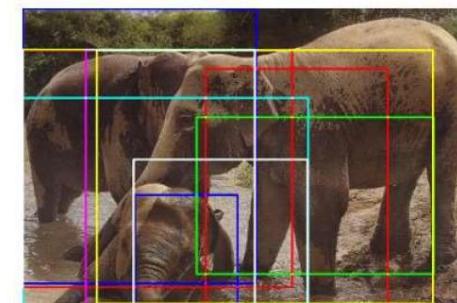
[parked (0.72)] [bench (0.63)] [truck (0.70)] [red (0.88)]
 [train (1.00)] [sitting (0.73)] [cars (0.58)] [traveling (0.52)]
 [grass (0.65)] [track (0.69)] [car (0.59)] [yellow (0.57)]
 [field (0.80)] [engine (0.56)] [down (0.54)] [tracks (0.94)]

a train traveling down train tracks near a field
 a red train is coming down the tracks



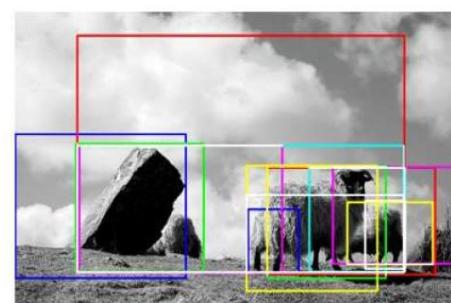
[umbrella (0.59)] [woman (0.52)]
 [fire (0.96)] [hydrant (0.96)] [street (0.79)] [old (0.50)]
 [bench (0.81)] [building (0.75)] [standing (0.57)] [baseball (0.55)]
 [white (0.82)] [sitting (0.65)] [people (0.79)] [photo (0.53)]
 [black (0.84)] [kitchen (0.54)] [man (0.72)] [water (0.56)]

a black and white photo of a fire hydrant
 a courtyard full of poles pigeons and garbage cans also has benches on either side of it one of which shows the back of a large person facin in the direction of the pigeons



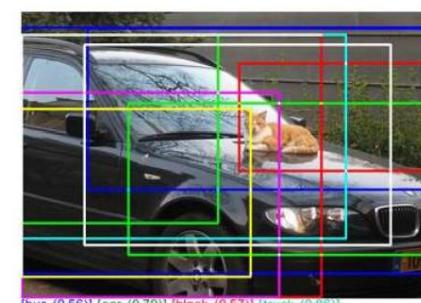
[horse (0.53)] [bear (0.71)] [elephant (0.99)] [elephants (0.95)]
 [brown (0.68)] [baby (0.62)] [walking (0.57)] [laying (0.61)]
 [man (0.57)] [standing (0.79)] [field (0.65)]
 [water (0.83)] [large (0.71)] [dirt (0.65)] [river (0.58)]

a baby elephant standing next to each other on a field
 elephants are playing together in a shallow watering hole



[man (0.59)] [beach (0.54)] [sky (0.53)] [bird (0.50)] [field (0.88)]
 [snow (0.86)] [mountain (0.59)] [standing (0.81)] [white (0.64)]
 [people (0.51)] [dog (0.60)] [cows (0.55)]
 [sheep (0.97)] [black (0.84)] [grass (0.64)] [horse (0.60)]
 [elephants (0.57)] [bear (0.81)]

a black bear standing on top of a grass covered field
 a couple of sheep standing up on a small hill



[bus (0.56)] [car (0.79)] [black (0.57)] [truck (0.86)]
 [street (0.57)] [bed (0.51)] [parked (0.55)] [dog (0.65)]
 [sitting (0.55)] [man (0.53)] [cat (0.72)]

a dog sitting on top of a car
 a cat is lying on the hood of a black car

Expanding horizons: interactive visual search

www.widl.it

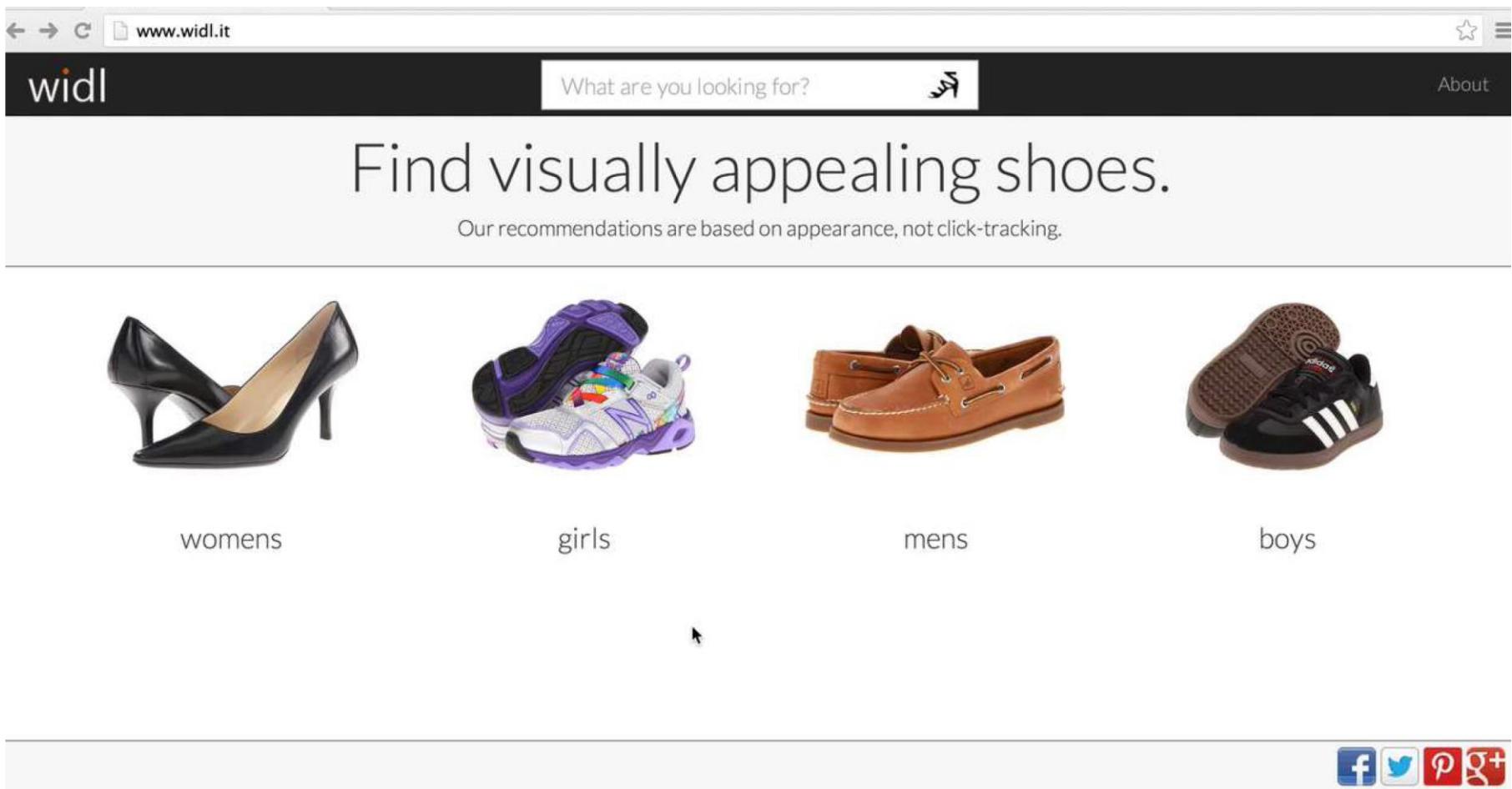
widl

What are you looking for? 

About

Find visually appealing shoes.

Our recommendations are based on appearance, not click-tracking.



womens

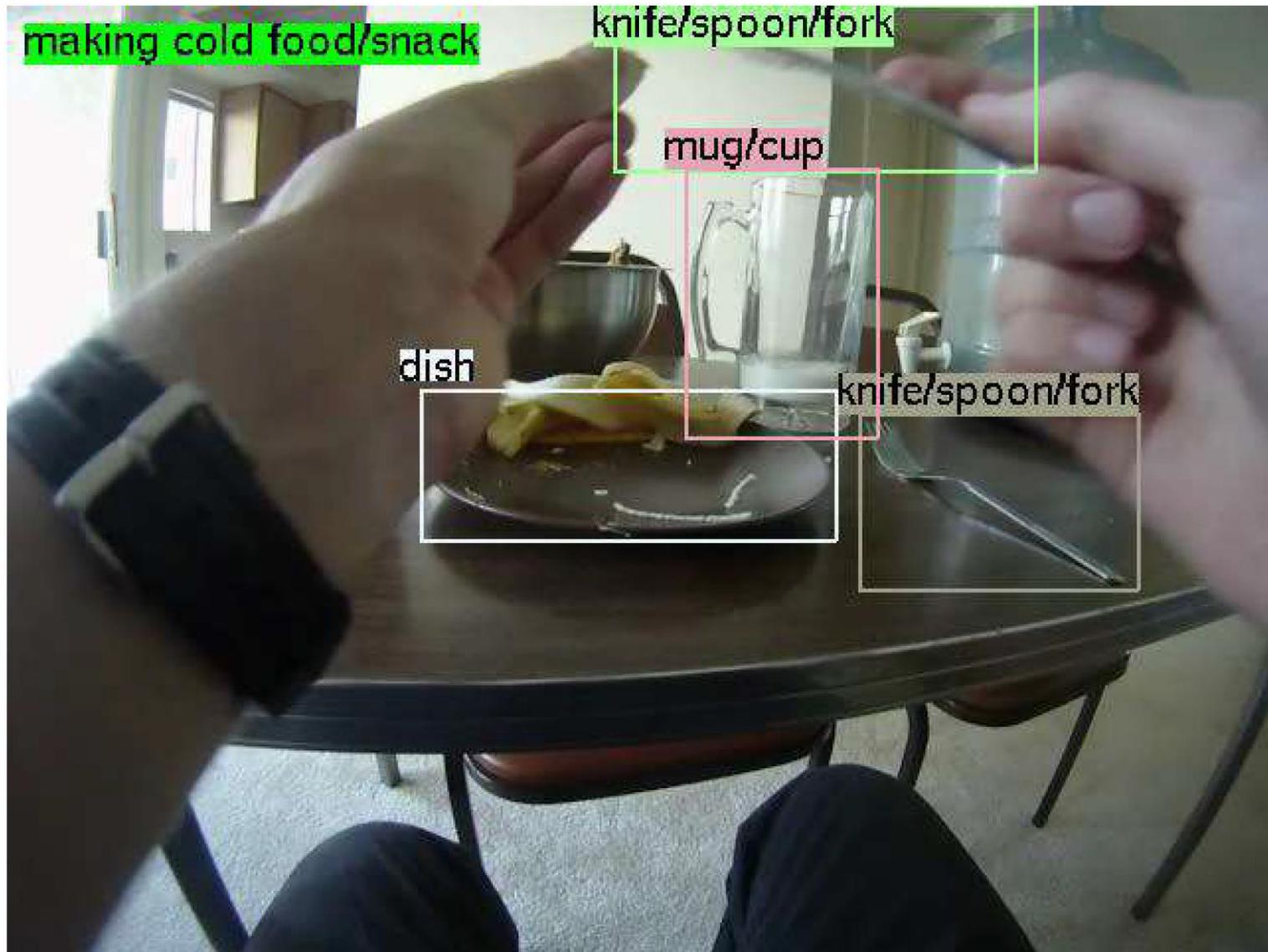
girls

mens

boys

Expanding horizons: first-person vision



Expanding horizons: enhancement / generation

The slide features a black header bar at the top and bottom. In the center, there is a white background area. On the left side, the MyHeritage logo is displayed above the text "Deep Nostalgia™". To the right of this text is a large, stylized orange scroll or flame graphic. A woman in a historical-style dress is shown from the waist up, holding the scroll. In the top right corner, there is a small rectangular video frame showing a man's face.

MyHeritage
**Deep
Nostalgia™**

Tal Erlichman
Director of Product Management

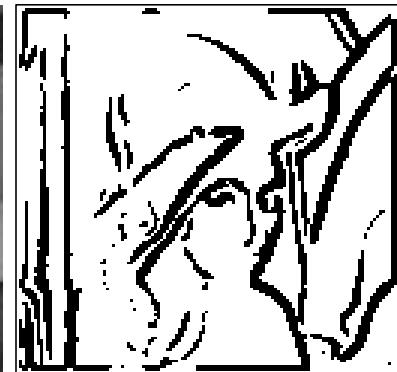
Goals of the course

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

Topics mostly covered

Image processing:

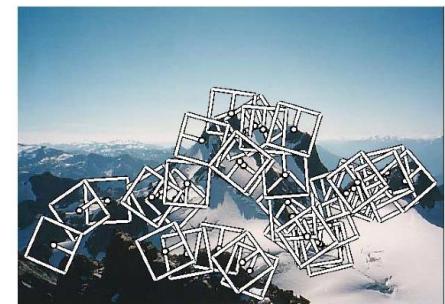
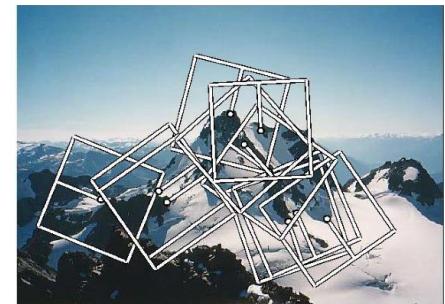
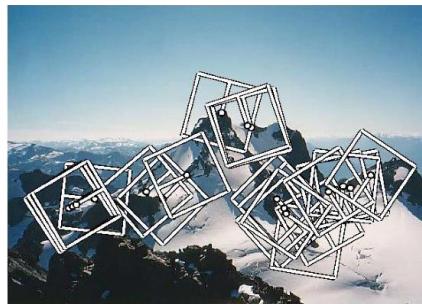
- Basics of filtering.
- Image pyramids.
- Gradients and lines.
- Hough transforms.
- Corner detection.



Topics to be covered

Feature detection and correspondences:

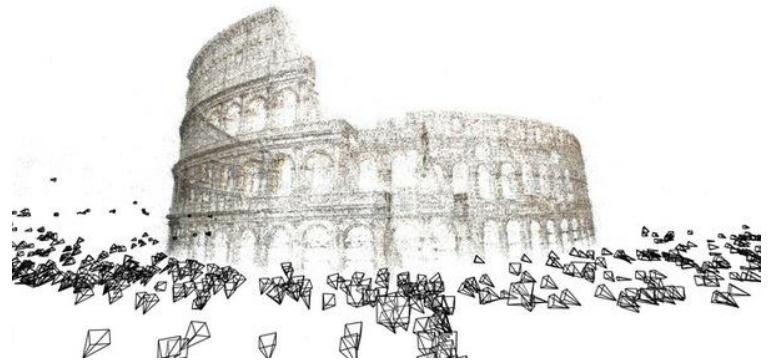
- SIFT ...
- Feature descriptors.
- RANSAC.



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:

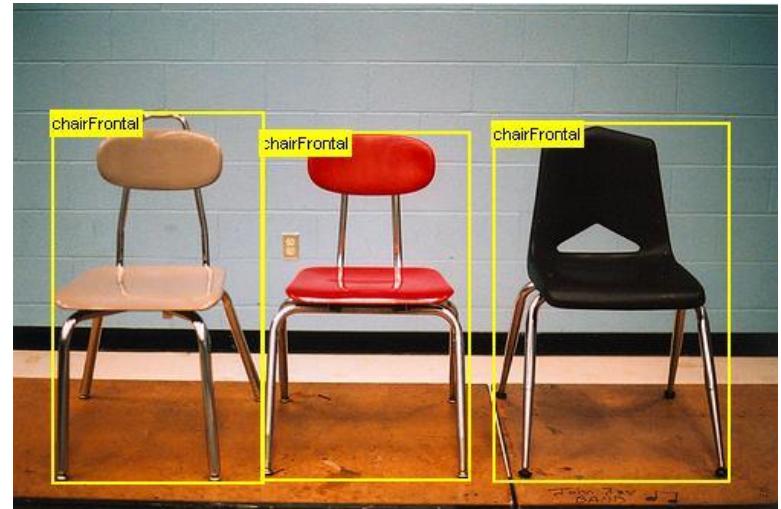
- Image registration.
- Photometric stereo.
- Optical flow.



Topics to be covered

Recognition - Objects and learning:

- Classical approaches
- Convolutional neural networks
- Current advances



Topics to be covered

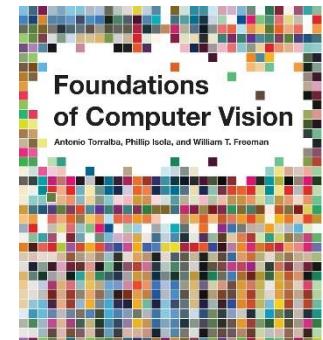
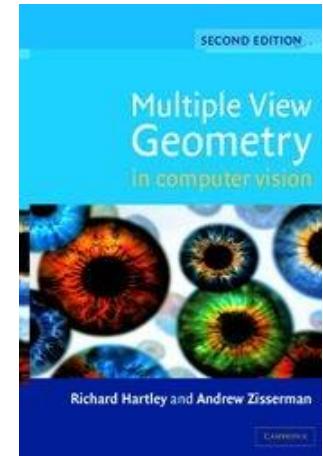
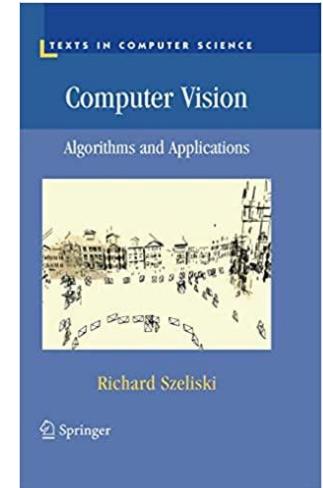
Generative models

- 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