CSE 185 Introduction to Computer Vision

Lecture 1: Introduction

Slides credit: Yuri Boykov, Ming-Hsuan Yang, Boqing Gong, Richard Szeliski, Steve Seitz, Alyosha Efros, Fei-Fei Li, etc.

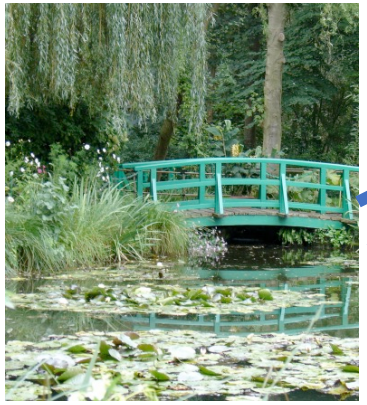
What is Computer Vison?

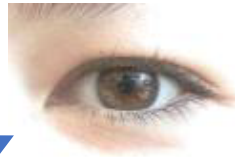
image (or video)

Sensing(device

Interpreting device

Interpretations





garden, spring, bridge, water, trees, 3D scene geometry, object shapes, etc.

Source: Fei-Fei Li

Every picture tells a story

- Goal of computer vision is to make computer interpret images

Goal of computer vision

- Bridge the gap between pixels and “meaning”

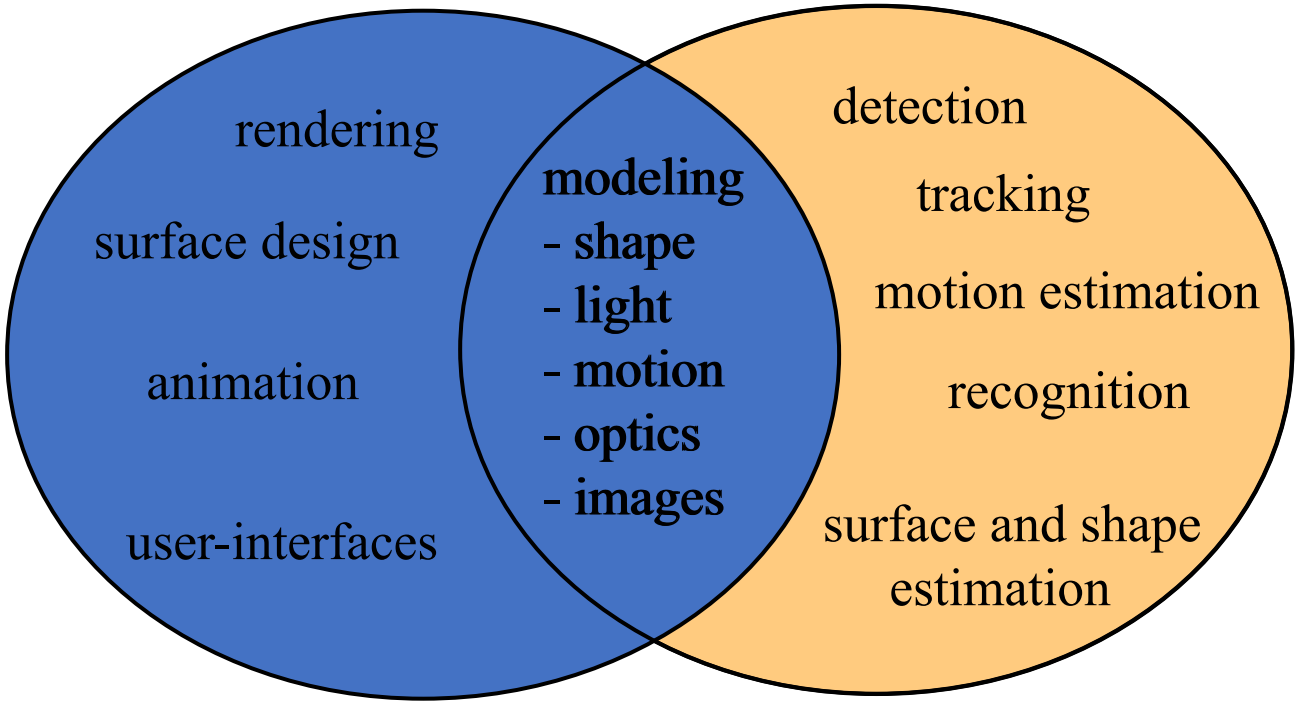
What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

What is it related to?

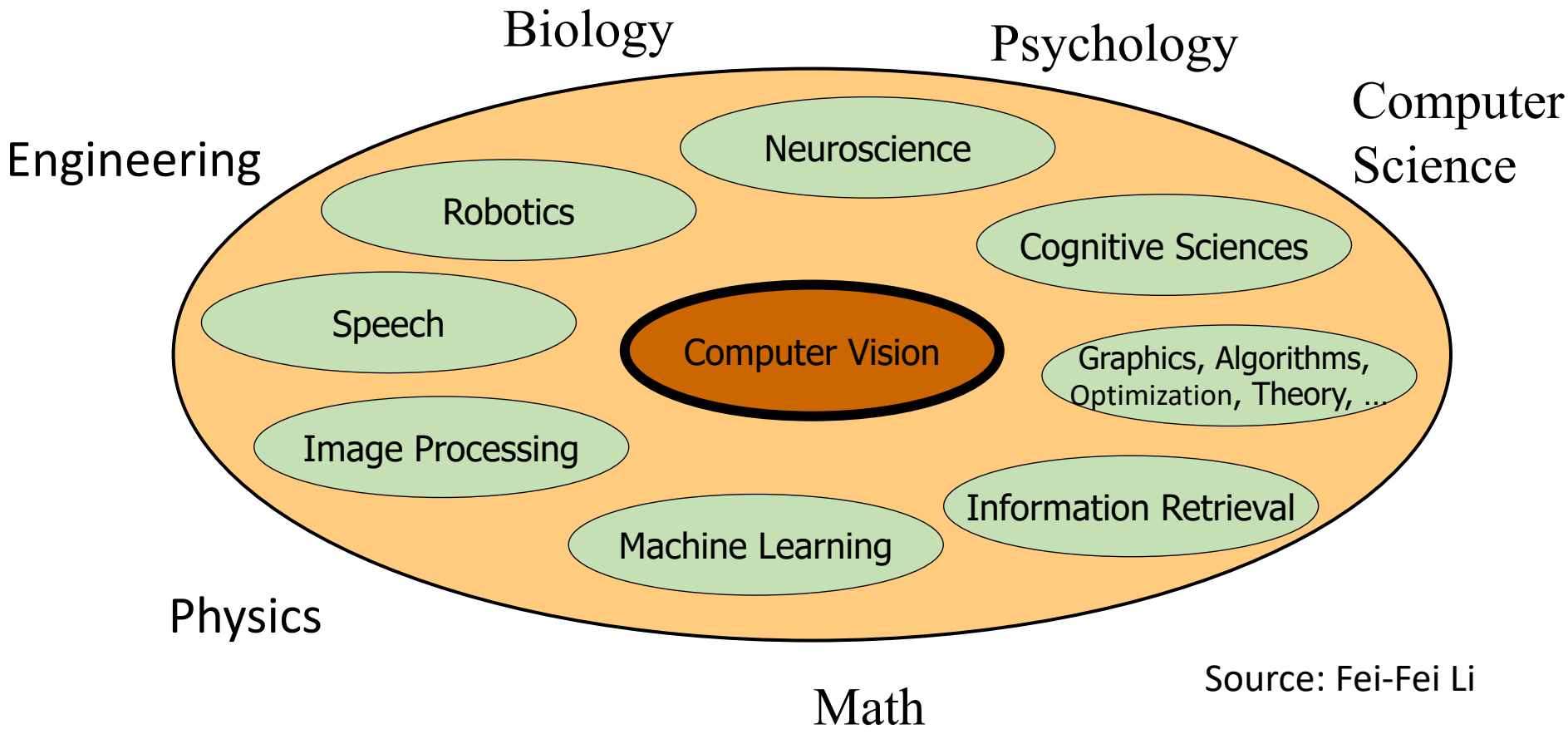


Computer Graphics
(Image Rendering)

Computer Vision
(Image Analysis)

Source: S. Seitz

Interdisciplinary area, math is important



Source: Fei-Fei Li

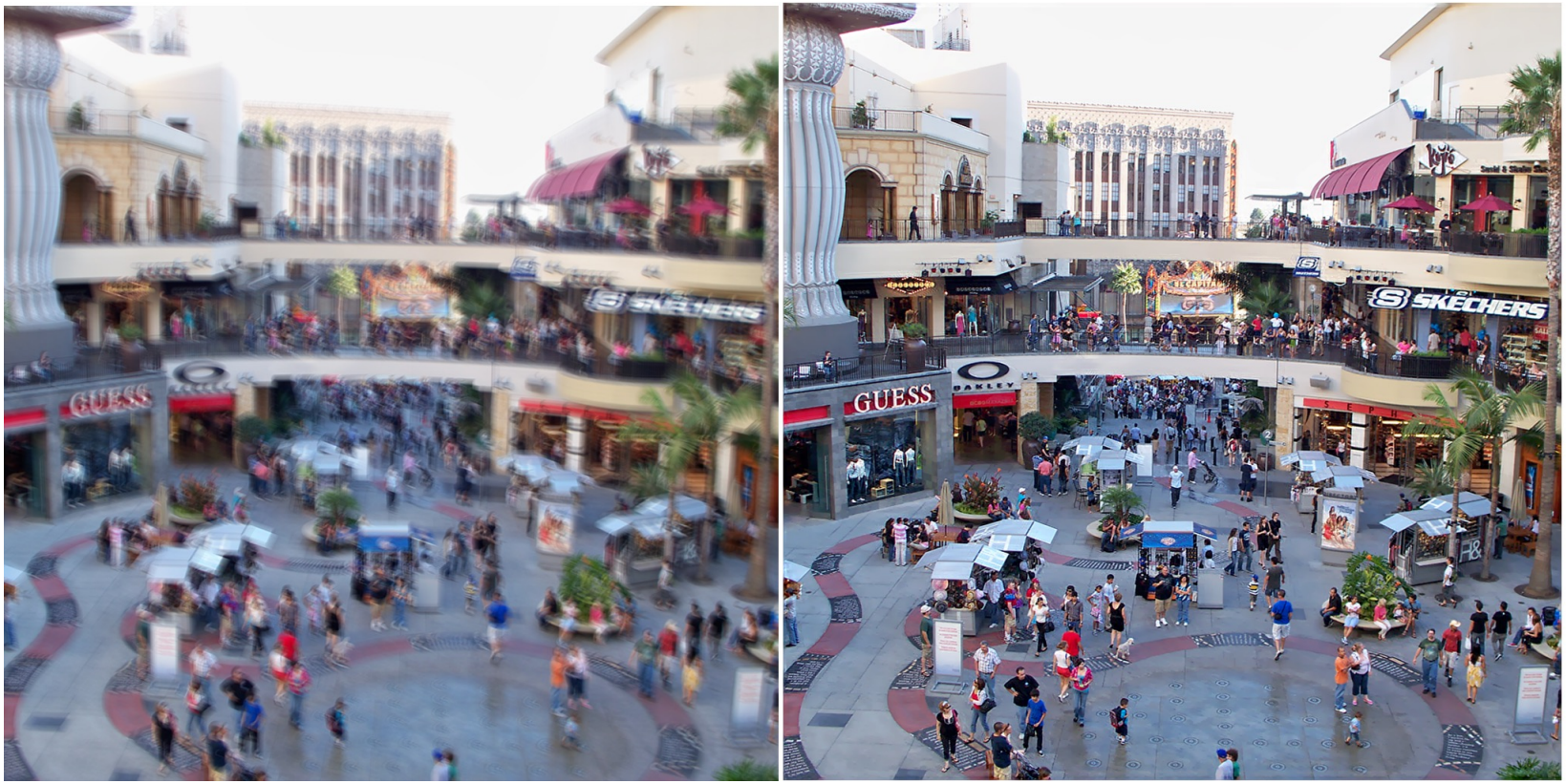
Applications of Computer Vision

Dehazing

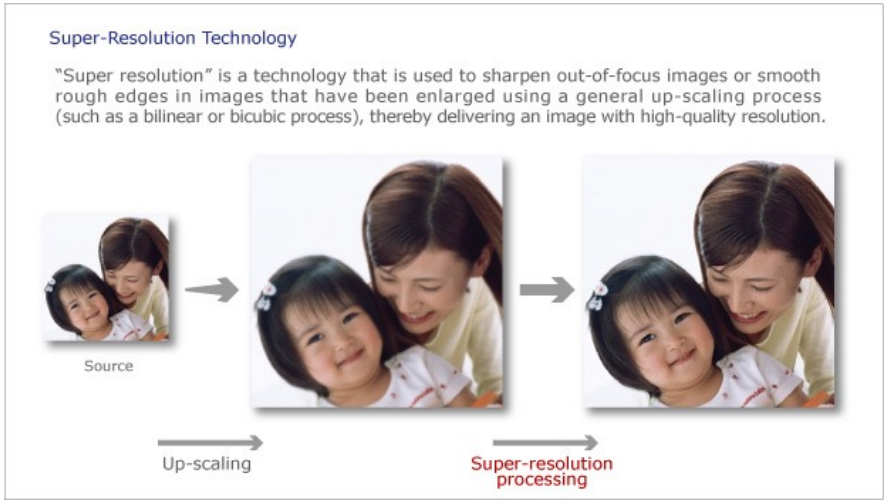
Jia et al. CVPR 08

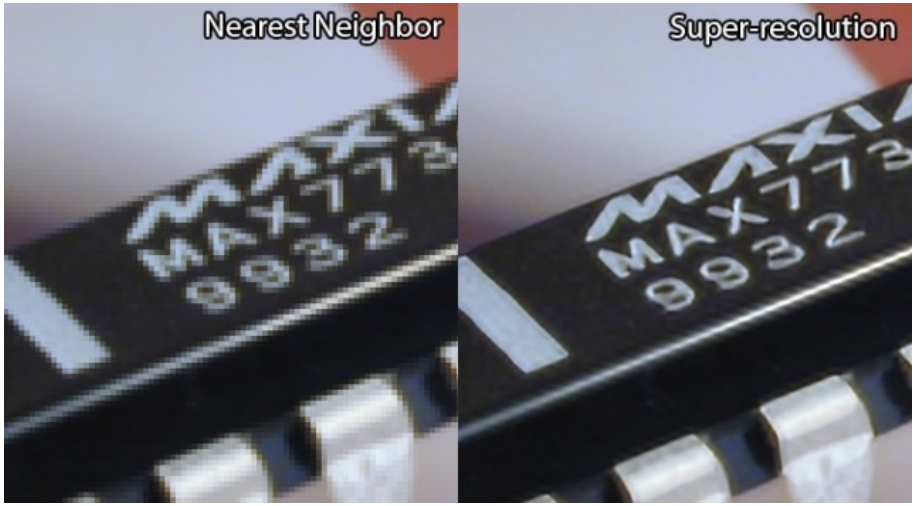
He et al. CVPR 09

Deblurring

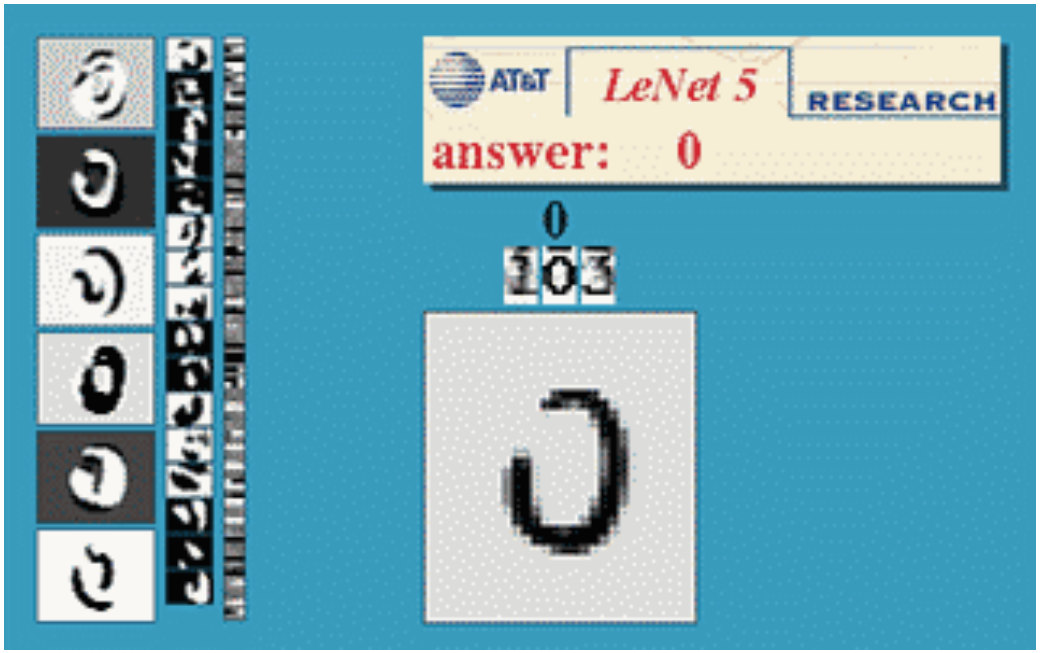


Super-resolution





Optical character recognition

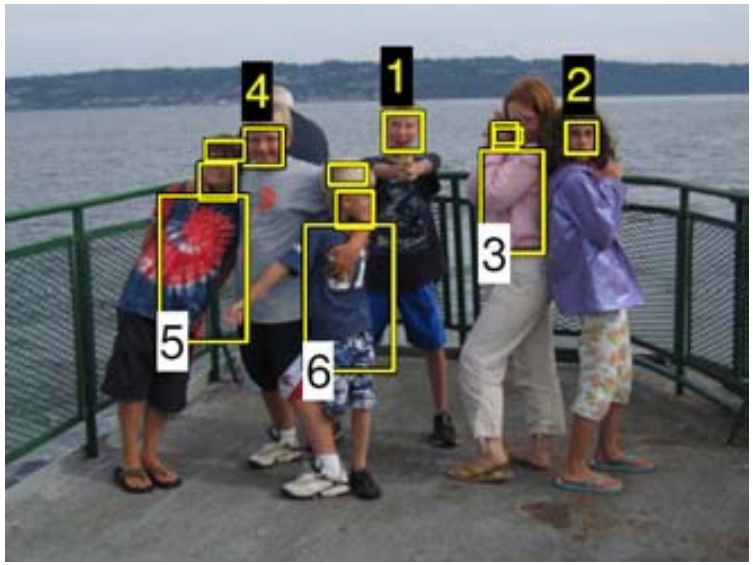


early 90's

must be older

Source: S. Seitz

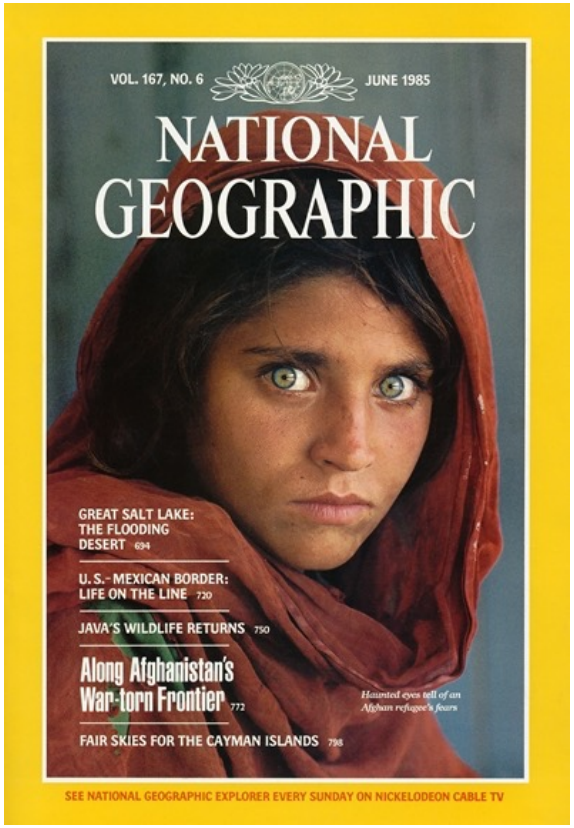
Face detection



face detection around 00’s

and now

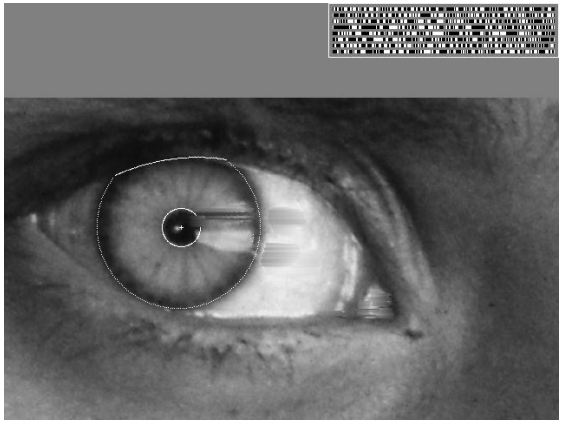
Face recognition

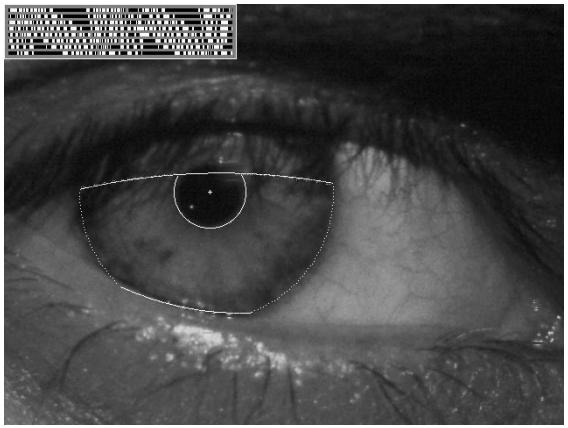


Who is she?

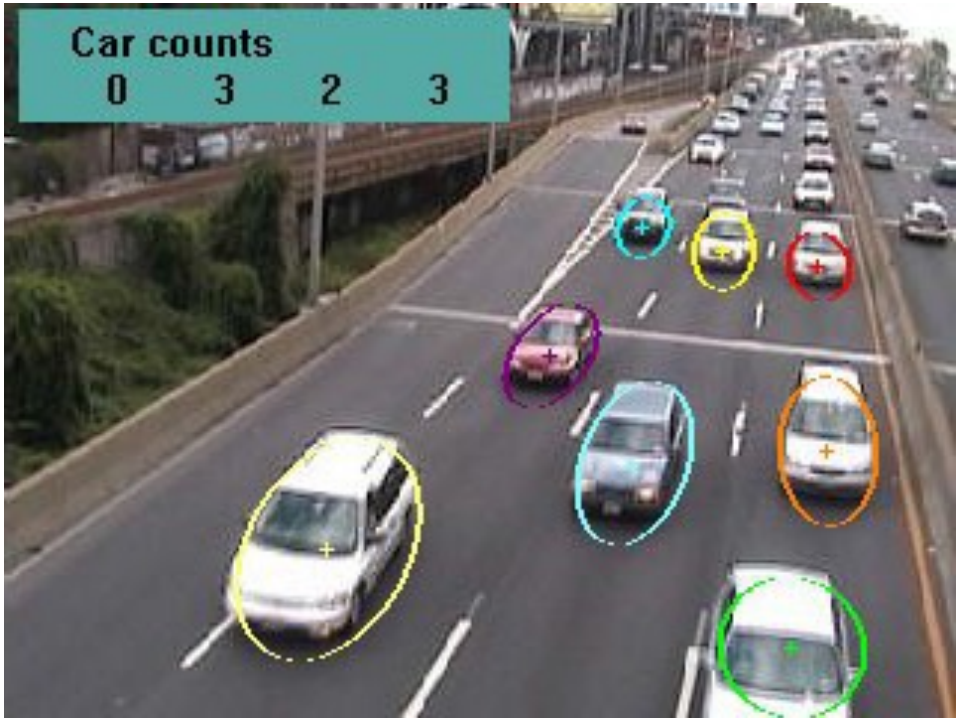
Face recognition

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story





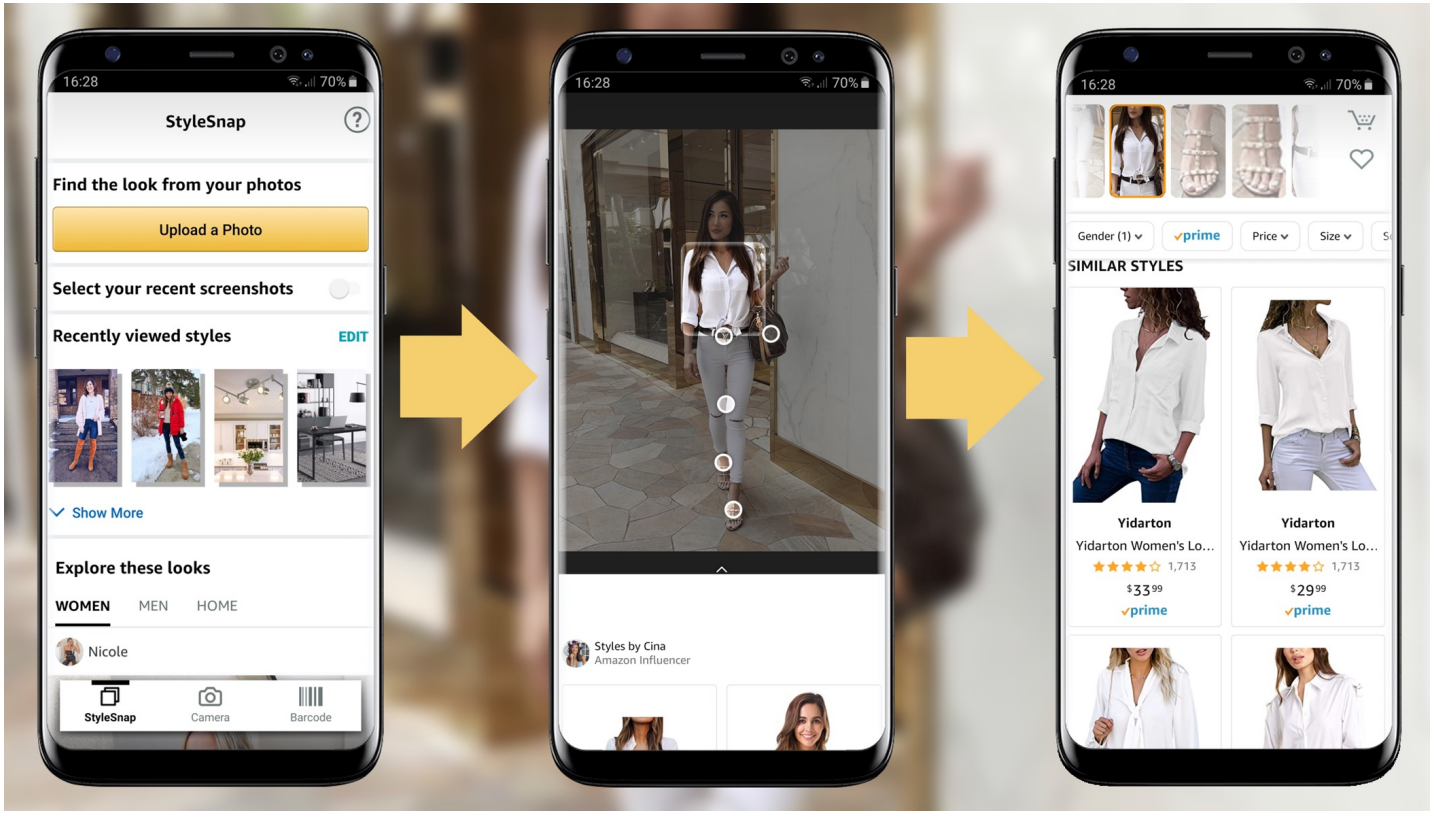
Object detection and tracking



Biometrics

[Amazon One Palm Payment is Coming to Whole Foods]

Visual search



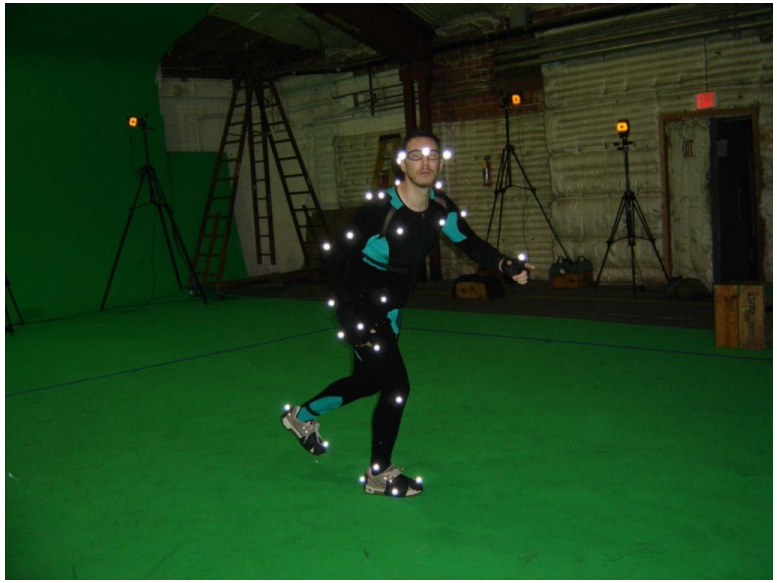
[Amazon Shop the Look: A Visual Search System for Fashion and Home]

Sports Analytics

https://www.sportlogiq.com/

Motion Capture

- Marker-based motion capture
  - http://www.youtube.com/watch?v=V0yT8mwg9nc



Autonomous Driving

Zoox https://www.youtube.com/watch?v=rfVjvpsCvFg

Avatar

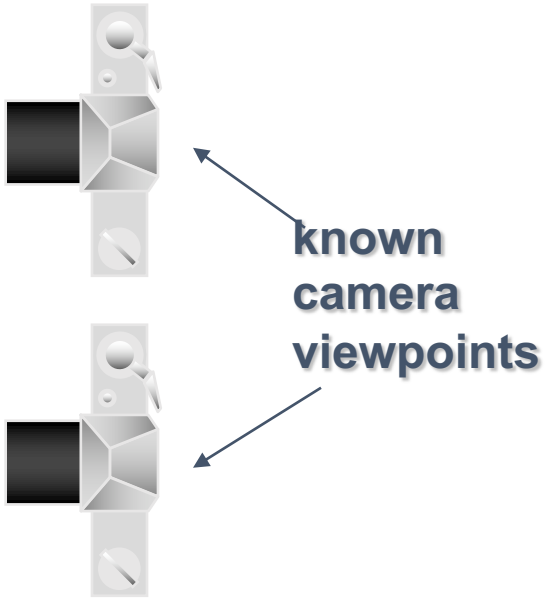
Meta Reality Labs Research - Codec Avatar

Object Detection and Segmentation

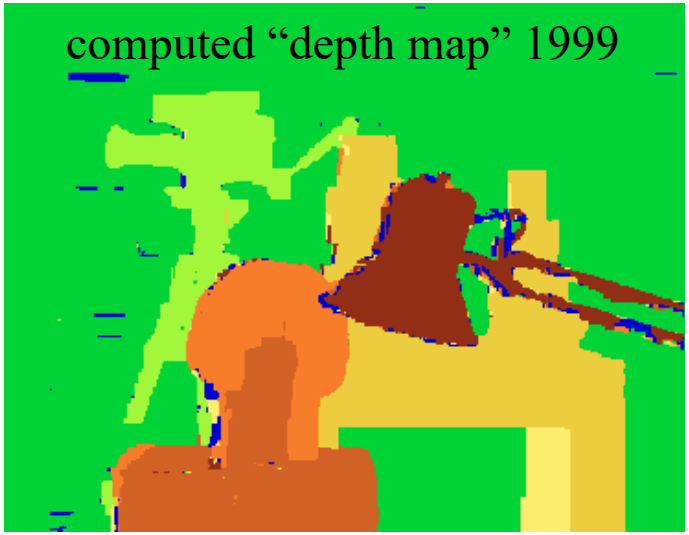
[Detectron2]

Stereo Reconstruction

- Shape from two (or more) images
- Biological motivation



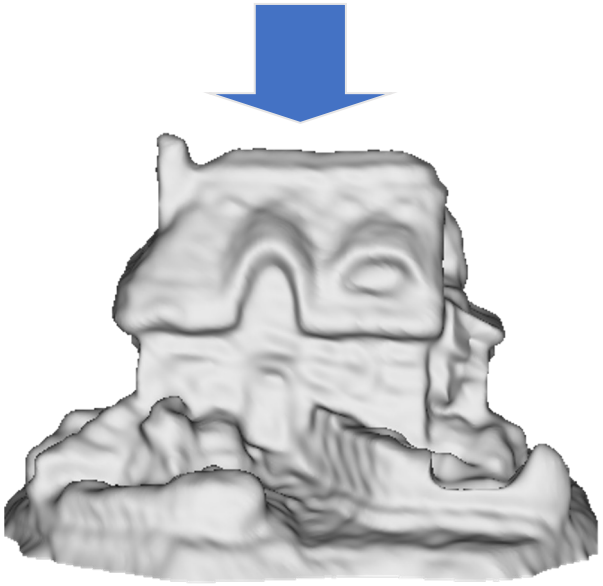
Stereo Reconstruction



Multiview Reconstruction

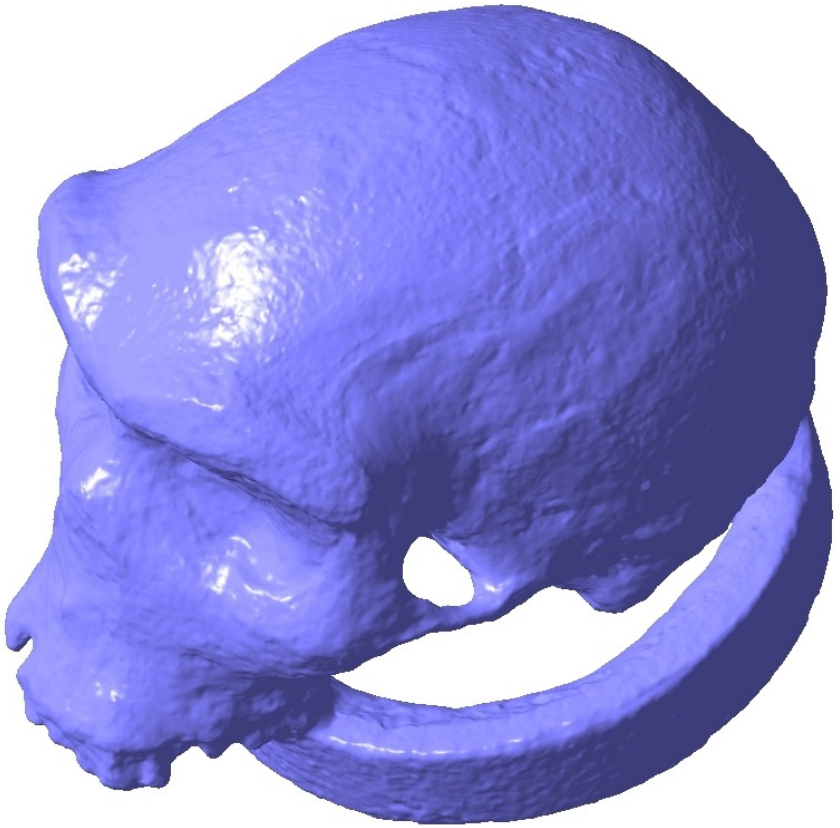


More then
2 images



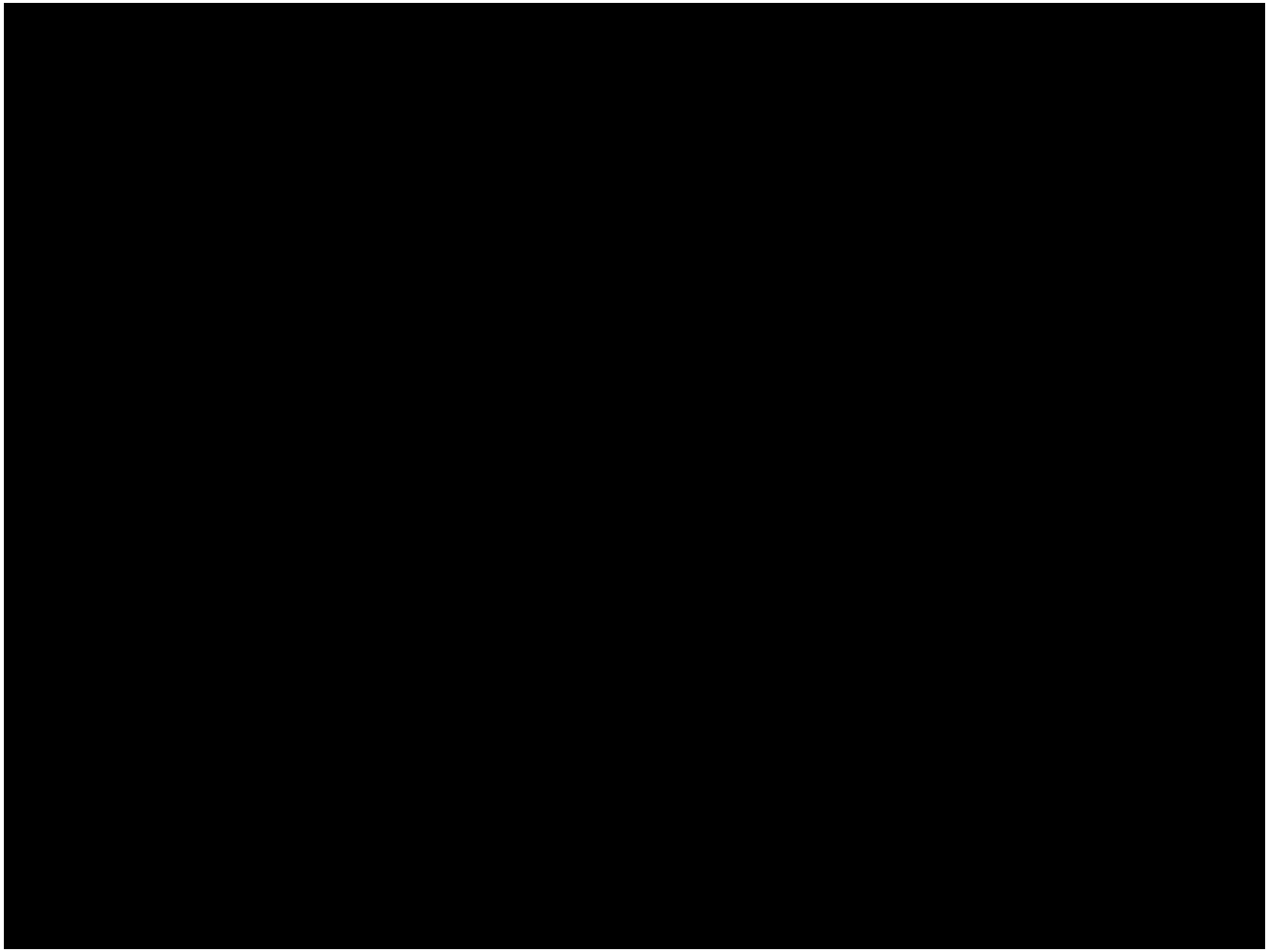
Multiview Reconstruction

multi-view reconstruction
from high resolution cameras



3D model
2006

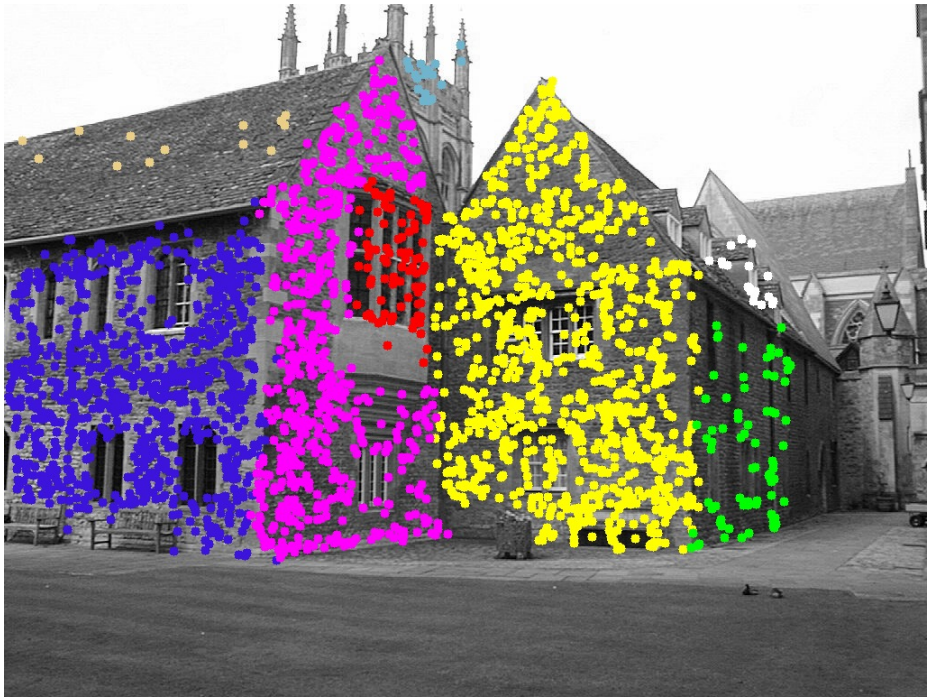
Structure from motion



(video from Carl Olsson)

Plane Reconstruction

Multiple views

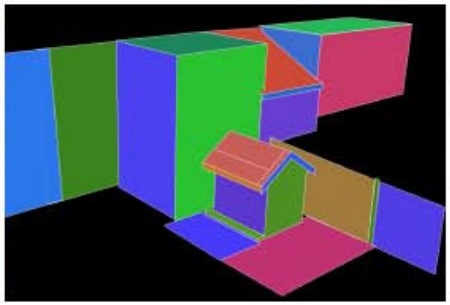


feature detection +
plane fitting

Plane Reconstruction

Input Photographs

2D Sketching Interface



Geometric Model

Texture-mapped model

3D model constructed form 9 images

3D Scene Reconstruction

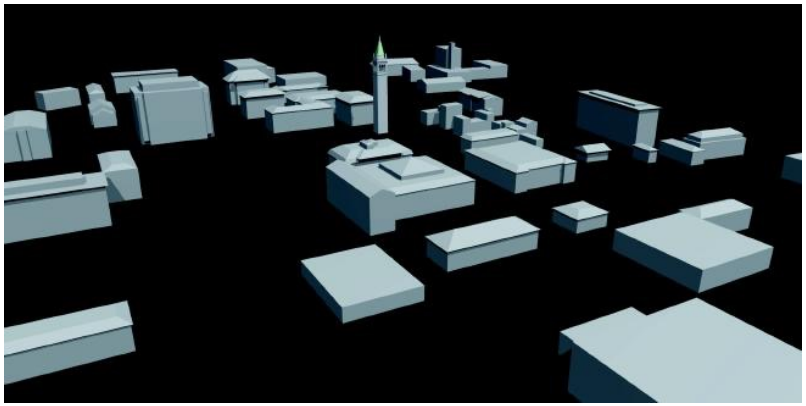


Image synthesis

- Generative Adversarial Network (GAN)

Image synthesis

- Generative Adversarial Network (GAN)

Image synthesis from text

- DALLE-2
- Imagen

"A modern, sleek Cadillac drives along the Gardiner expressway with downtown Toronto in the background, with a lens flare, 50mm photography."

"A photo of an astronaut riding a horse."

"A man walking through the bustling streets of Kowloon at night, lit by many bright neon shop signs, 50mm lens."

History of Computer Vision

- “In the 1960s, almost no one realized that machine vision was difficult.” – David Marr, 1982
- Marvin Minsky asked Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw” – Crevier, 1993
- 50+ years later, we are still working on this

1970s

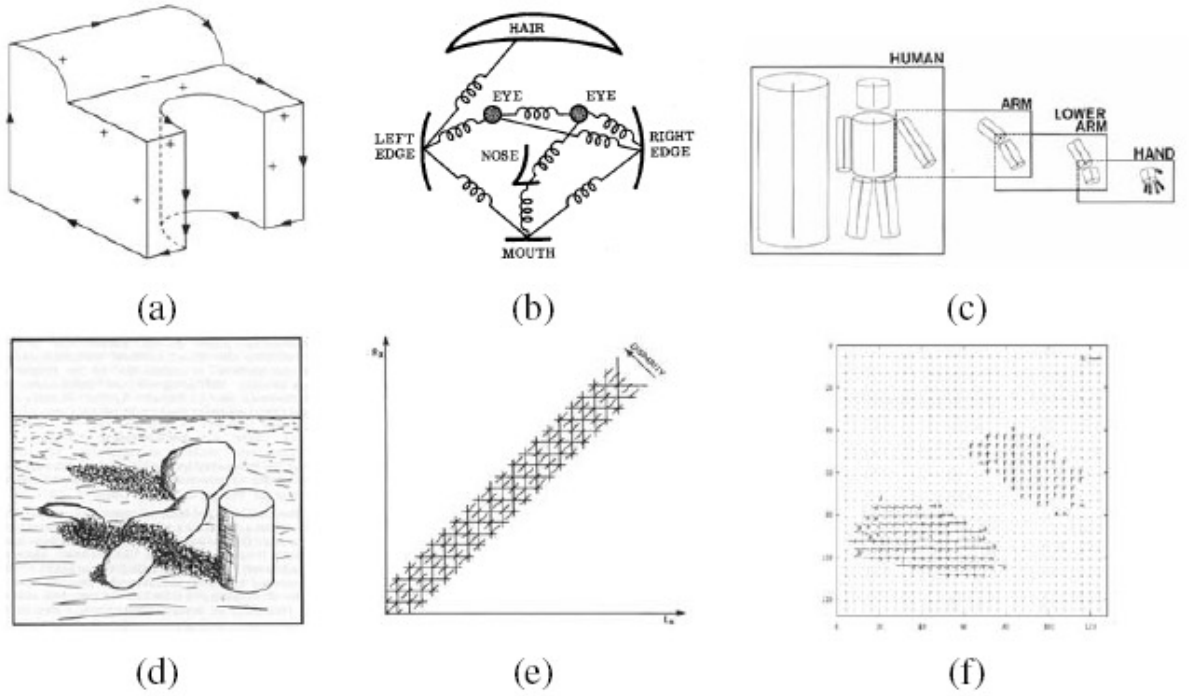


Figure 1.6: *Some early examples of computer vision algorithms (1970s): (a) line labeling (Nalwa 1993), (b) pictorial structures (Fischler and Elschlager 1973), (c) articulated body model (Marr 1982), (d) intrinsic images (Barrow and Tenenbaum 1981), (e) stereo correspondence (Marr 1982), (f) optical flow (Nagel and Enkelmann 1986).*

1980s

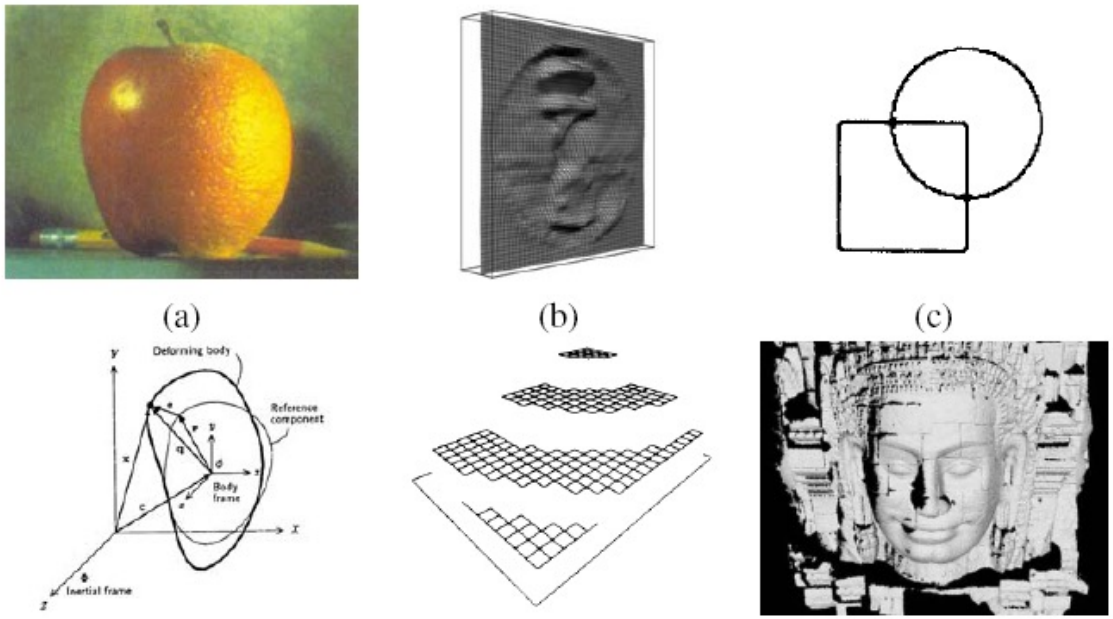


Figure 1.7: *Examples of computer vision algorithms from the 1980s: (a) pyramid blending (Burt and Adelson 1983b), (b) shape from shading (Freeman and Adelson 1991), (c) edge detection (Freeman and Adelson 1991), (d) physically-based models (Terzopoulos and Witkin 1988), (e) regularization-based surface reconstruction (Terzopoulos 1988), (f) range data acquisition and merging (Banno* et al. *2008).*

1990s

- ❑ Face detection
- ❑ Particle filter
- ❑ Normalized cut

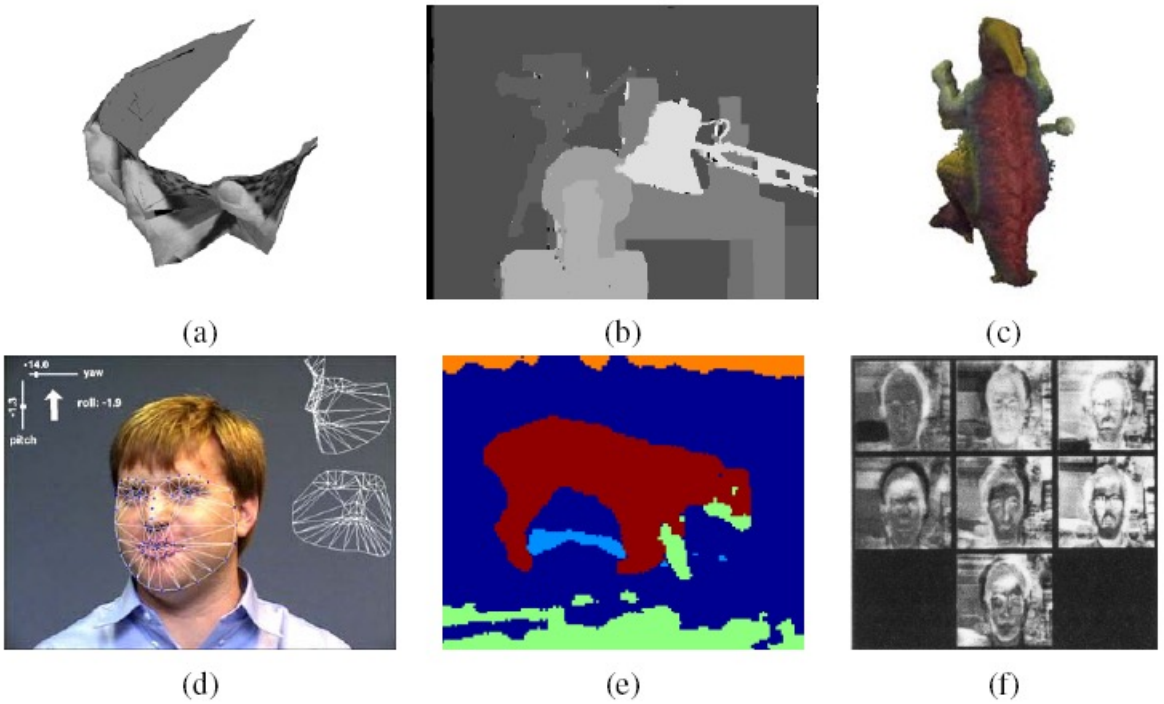
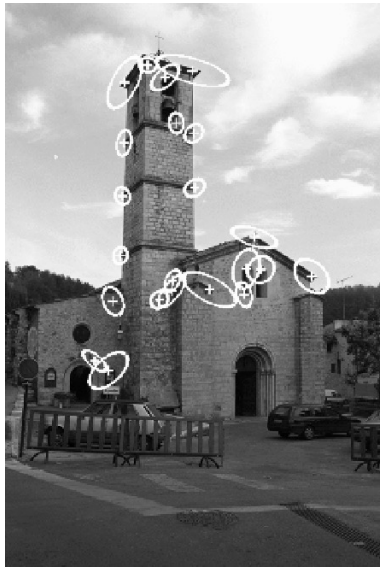
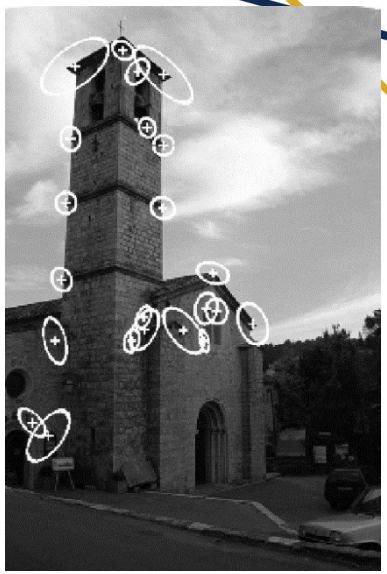


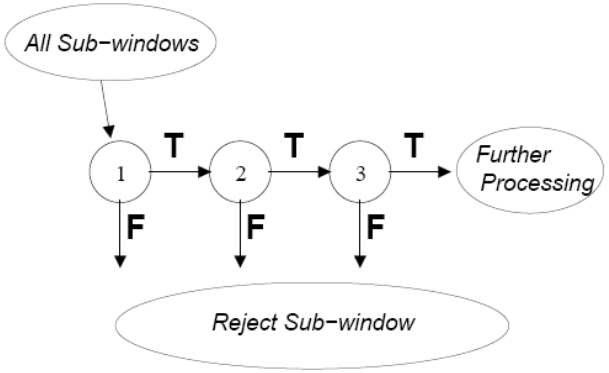
Figure 1.8: *Examples of computer vision algorithms from the 1990s: (a) factorization-based structure from motion (Tomasi and Kanade 1992), (b) dense stereo matching (Boykov* et al. *2001), (c) multi-view reconstruction (Seitz and Dyer 1999), (d) face tracking (Matthews and Baker 2004, Matthews* et al. *2007), (e) image segmentation (Fowlkes* et al. *2004), (f) face recognition (Turk and Pentland 1991a).*

2000s





- SIFT
  - Mosaicing, panorama
  - Object recognition
  - Photo tourism, photosynth
  - Human detection
- Adaboost-based face detector



2010s

- Real-world applications
  - Google: image search, glass, X, autonomous driving, product search
  - Adobe, Microsoft, Facebook, Apple, Toyota, Honda, Amazon
  - Applications for autonomous driving
  - Mobile phones
- Deep learning

2020s

- Deep learning dominates the landscape
- Vision and language
- Foundation models
- Large-scale applications

Topics for this course

- ❑ Image modalities and camera model
- ❑ Low-level vision: light, color, image filtering, edge, interest points, features, stereo, optical flow, 3D reconstruction
- ❑ Mid-level vision: clustering, grouping, segmentation
- ❑ High-level vision: object detection, object recognition, visual tracking
- ❑ Classical methods as well as modern methods (deep learning)

Course Materials

- Course webpage
  - http://www.mengtang.org/cse185
  - Syllabus
  - Lecture notes
  - Assignments
  - Labs
- Textbook
  - Computer Vision: Algorithms and Applications, Richard Szeliski (available online)
- Reference for background study:
  - Computer Vision: Models, Learning, and Inference, Simon Prince
  - Introductory Techniques for 3-D Computer Vision, Emanuele Trucco and A. Verri
  - Programming Computer vision with Python, Jan Erik Solem
  - Learning OpenCV: Computer Vision with OpenCV Library, Gary Bradski and Adrian Kaehler

Grading

- Exams (40%)
  - Midterm:20%
  - Final: 20%
- Labs (20%)
  - 11 labs
  - 2% for attending one lab
  - Nothing to submit for labs
- Assignment (40%)
  - 4 assignments
  - 5%, 10%, 10%, 15%