Computer Vision Systems Programming VO

Computer Vision: Past, Present, Future

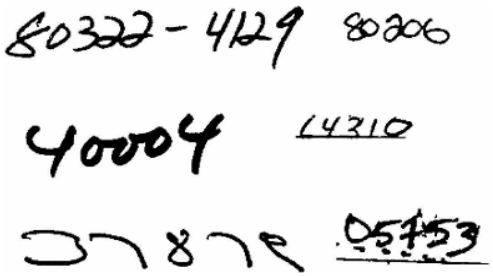
Christopher Pramerdorfer

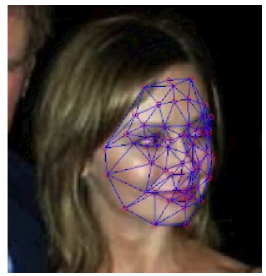
Computer Vision Lab, Vienna University of Technology

Topics

Selection of past, present, future CV applications

More detailed coverage in upcoming lectures





Images from LeCun et al. 1989, Shotton et al. 2011, Taigman et al. 2013

What is CV?

Make computers understand images and videos

- Different levels of understanding

CV is hard

- Inverse (ill-posed) problem

Still, CV has been successfully used in a variety of applications

- This lecture introduces a few in chronological order

1963: Pose Estimation

Edge-based pose estimation of polyhedra

Among first CV applications

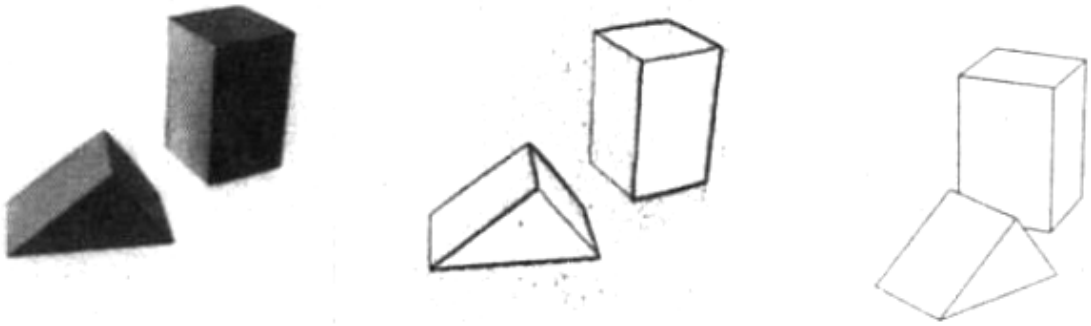


Image from Roberts 1963

1973: Part-Based Object Detection

Object representation as parts connected by springs

Known as pictorial structures or constellation models

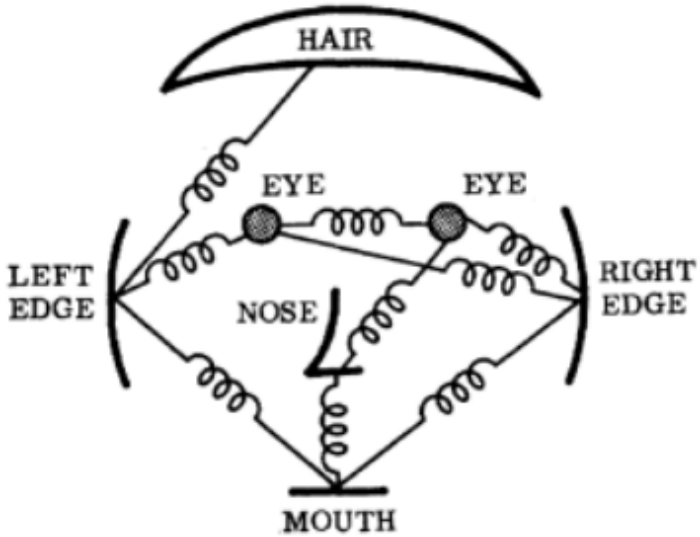


Image from Fischler and Elschlager 1973

1989: OCR via Deep Learning

Zip code recognition from images

Among first applications using convolutional neural networks

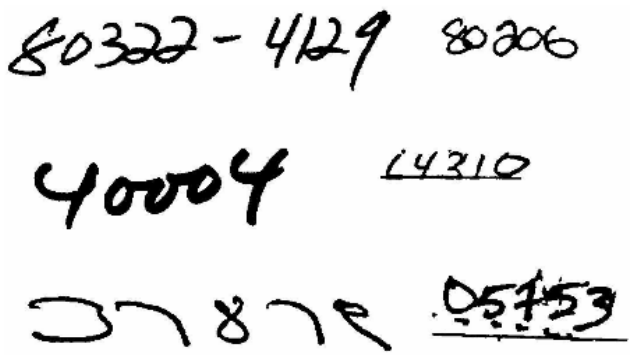


Image from LeCun et al. 1989

1989: OCR via Deep Learning

Network Architecture

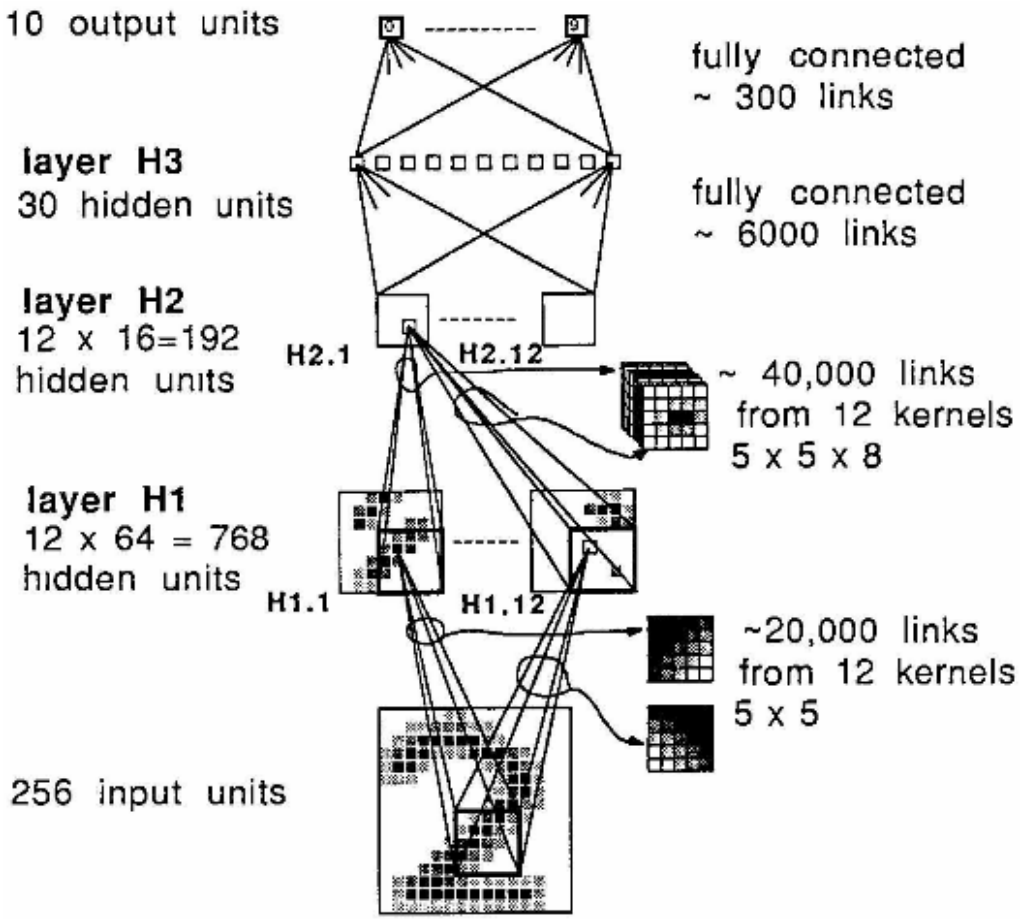


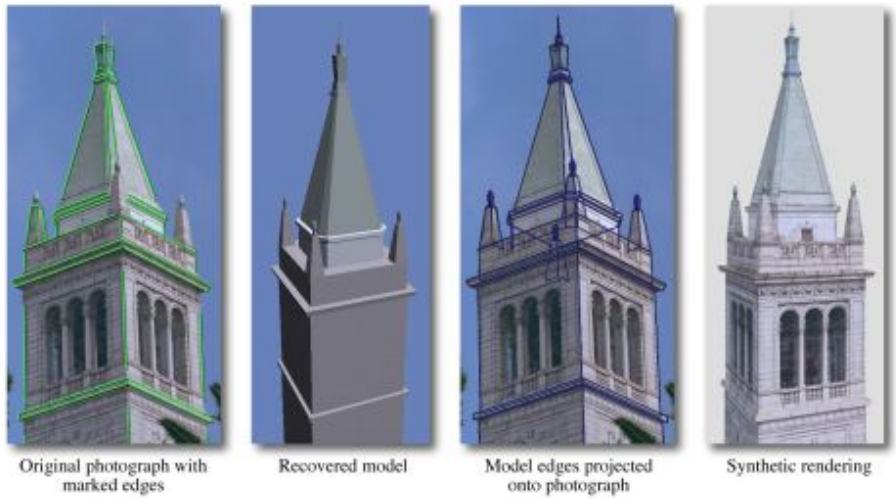
Image from LeCun et al. 1989

1996: Image-Based Modeling

Generate a 3D model from a set of images

Use this model and input images to render new images

`https://www.youtube.com/watch?v=RPhGEiM_6lM`



Images from Debevec 1996

2006: Photo Tourism

3D reconstruction from photo collections

Structure from Motion (SIFT + bundle adjustment)

Images from Snavely, Seitz, and Szeliski 2006

2006: Photo Tourism
Microsoft Photosynth

https://photosynth.net

2006: Photo Tourism
Building Rome in a Day

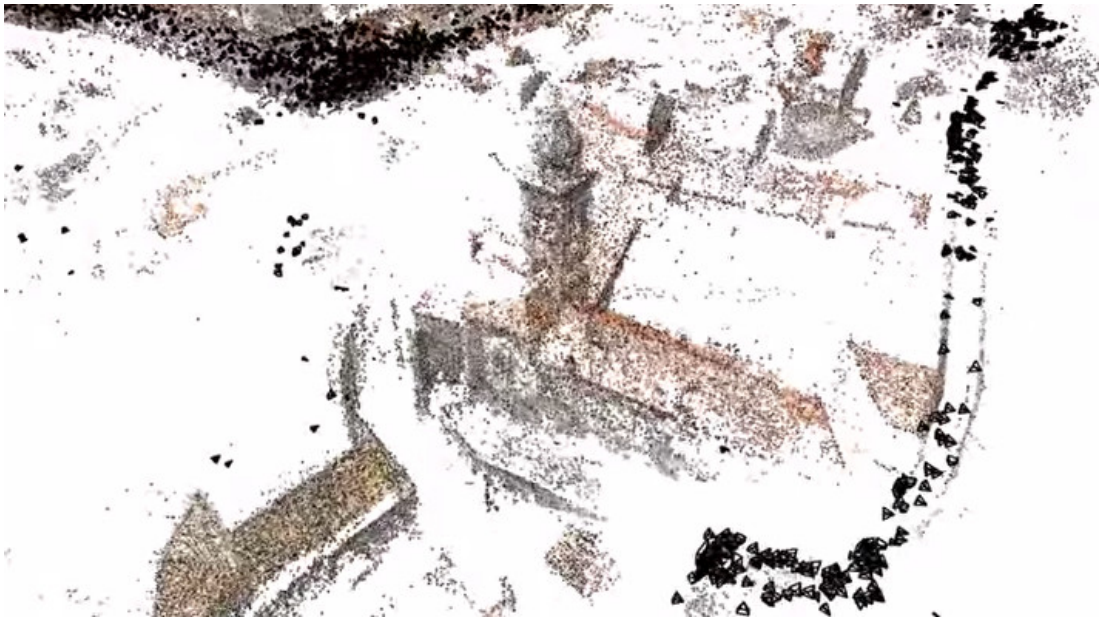


Image from https://www.youtube.com/watch?v=sQegEro5Bfo

2007: Smart Digital Cameras

Cameras with face auto focus

Technology similar to Viola and Jones 2001

Image from olympus-europa.com

2011: Kinect

Depth estimation via active stereo

Real-time pose estimation of multiple players

Image from wikipedia.org

Image from Shotton et al. 2011

2013: Human-Level Face Verification

Face verification using a deep convolutional neural network

3D face modeling and frontalization

Verification performance comparable to humans

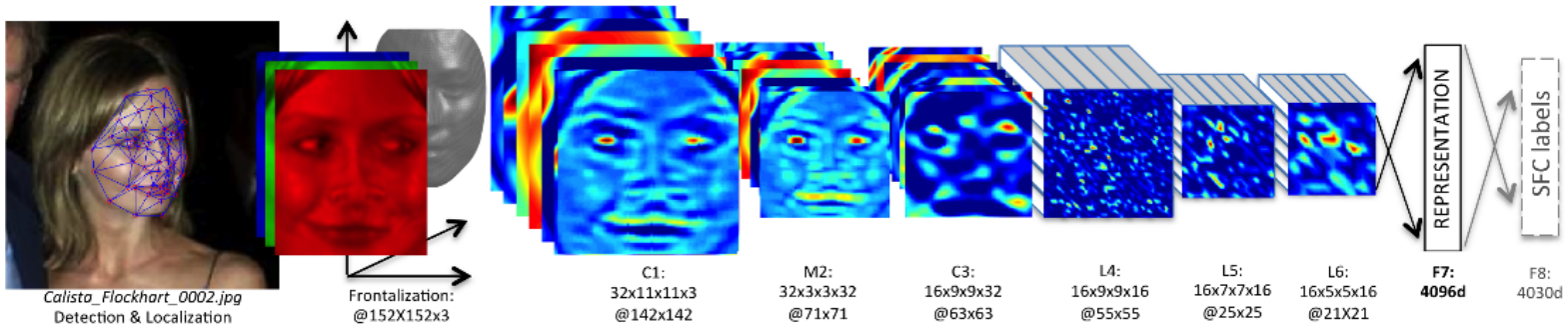


Image from Taigman et al. 2013

20xx: Human-Level Object Recognition

Object recognition without constraints

Hot research topic (`http://image-net.org/challenges/LSVRC/2014/index`)

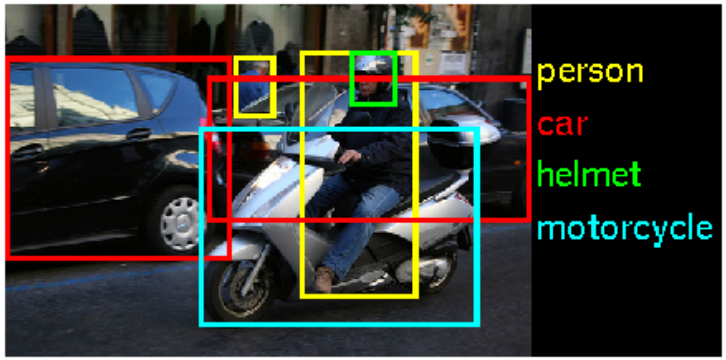


Image from `image-net.org`

20xx: Autonomous Cars

Cars that drive autonomously

Major research area (e.g. Google)

- `https://www.youtube.com/watch?v=bDOnn0-4Nq8`

Image by Google

20xx: Human-Level Scene Understanding

Object recognition and segmentation, motion, context

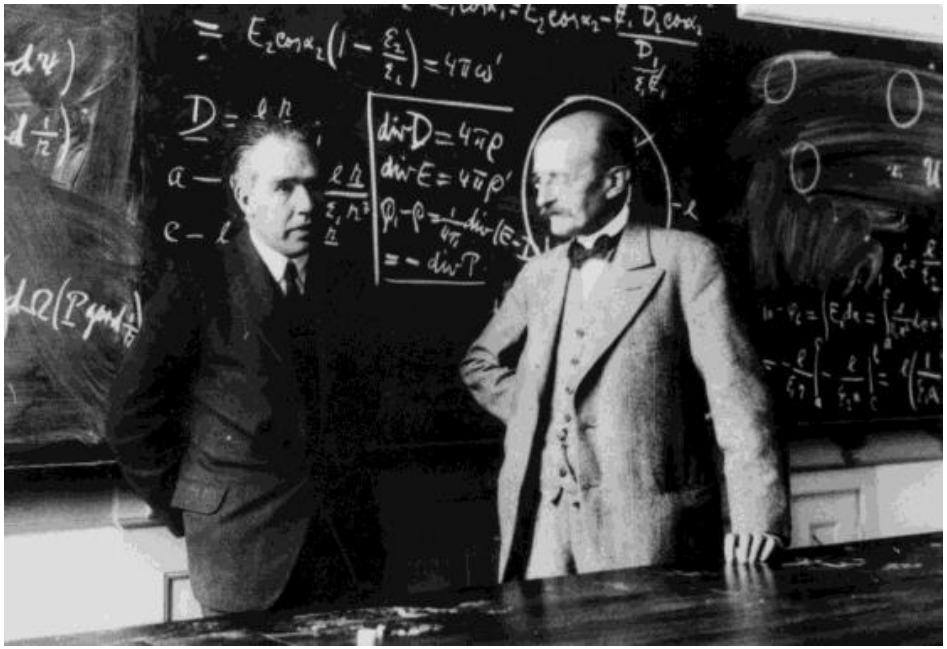


Image from Larry Zitnick's slides

Bibliography I

Debevec, Paul E. (1996). **Modeling and Rendering Architecture from Photographs**. PhD thesis. Berkley.

Fischler, Martin A and Robert A Elschlager (1973). **The representation and matching of pictorial structures**. IEEE Transactions on Computers.

LeCun, Yann et al. (1989). **Backpropagation applied to handwritten zip code recognition**. Neural computation.

Roberts, Lawrence Gilman (1963). **Machine perception of three-dimensional solids**. PhD thesis. MIT.

Shotton, Jamie et al. (2011). **Real-Time Human Pose Recognition in Parts from a Single Depth Image**. CVPR.

Bibliography II

Snavely, Noah, Steven M. Seitz, and Richard Szeliski (2006). **Photo tourism: Exploring photo collections in 3D**. SIGGRAPH.

Taigman, Yaniv et al. (2013). **Deepface: Closing the gap to human-level performance in face verification**. CVPR.

Viola, Paul and Michael Jones (2001). **Rapid object detection using a boosted cascade of simple features**. CVPR.