EECS 633

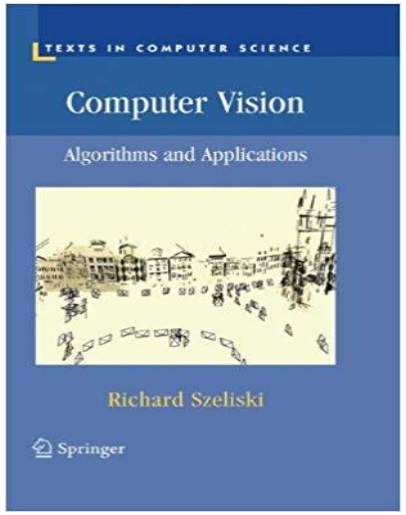
Machine Vision and Image Understanding

Electrical and Computer Engineering Department

Khalifa University

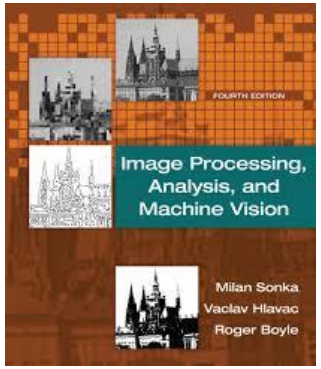
Syllabus

- Recommended Prerequisites: ENGR 112 Introduction to Computing (programming experience), and ECCE 302 Signal Processing (transforms, low-pass and high-pass filtering)
- References
  - *Richard Szeliski, Computer Vision: Algorithms and*
  - *Applications, Springer-Verlag, 2011, ISBN: 978-1848829343.*

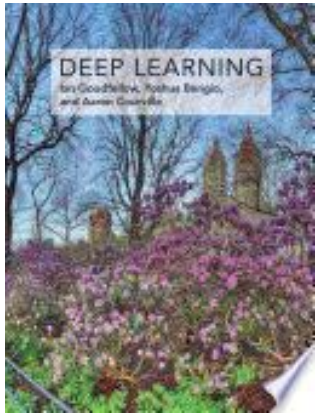


Other References

- *Milan Sonka, Vaclav Hlavac, Roger Boyle, Image Processing, Analysis, and Machine Vision, 3rd ed, CL Engineering, 2007, ISBN: 978-0495082521.*



- Goodfellow, Ian, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*. Vol. 1. Cambridge: MIT press, 2016.



- ***Plenty of online resources !!***

Syllabus

- Grading
  - 30% Coursework, 3-4 HW sets
  - 30% Project
  - 40% Final Exam

Final Project

- Open ended problem in a topic related to class content. It's expected that the project will study a topic in depth and involve a substantial amount of work
- You can choose from project ideas that will be distributed later, or you can propose a topic of your own. Project selection will occur approximately half-way through the class
- Topic can be related to your research

Imaging modalities

- Photographic images
- Medical images: MRI, CT, PET.
- Seismic data.
- Astrological images
- Radar data.

Wide range of applications

Contrast enhancement

(a)

(b)

By using a contrast enhancement algorithm, image (b) improves the appearance of image (a)

Applications

- Image denoising

Original

Noisy image

Denoised image

Edge Detection

Segmentation

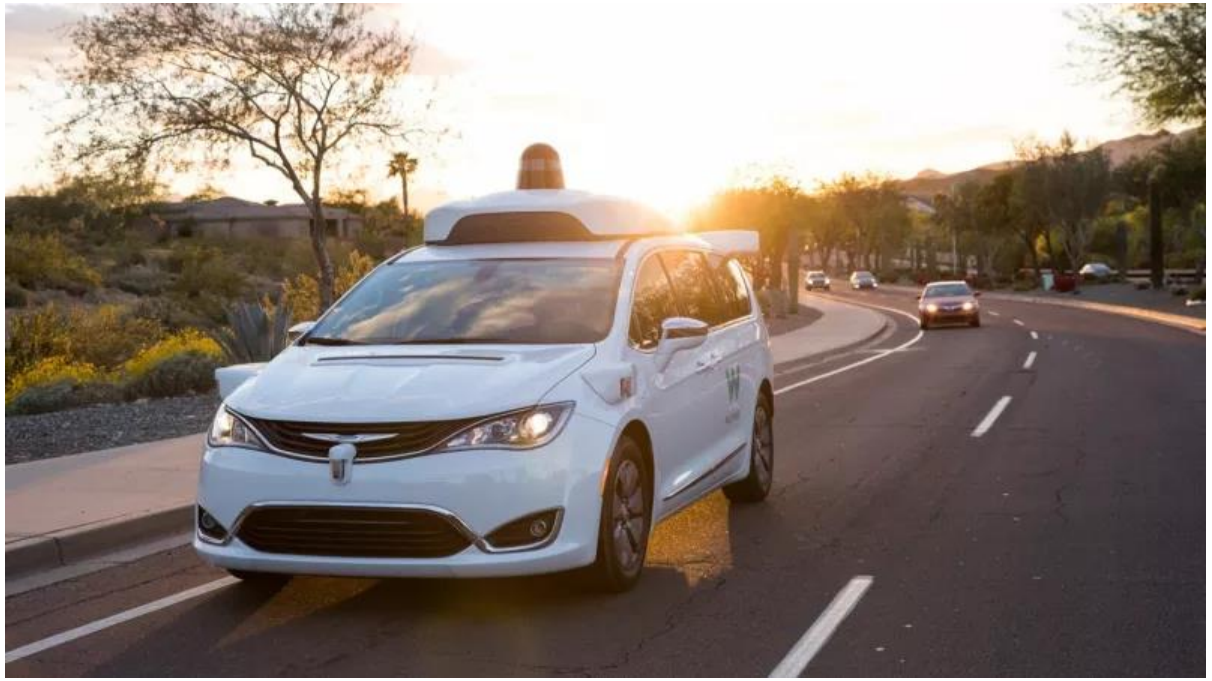
Segmentation of MRI scan of a kidney

Segmentation

Segmentation of a video sequence

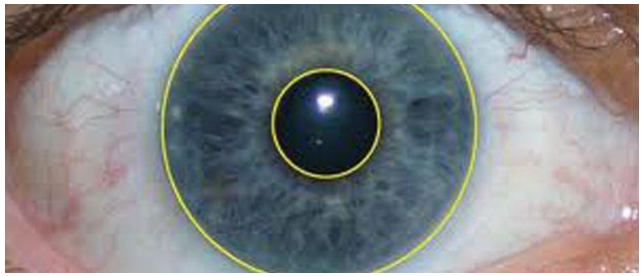
Other applications

- Self-driving cars



Other applications

- Security: biometric recognition (face, fingerprint, iris recognition)



- Drones

Other applications

- Image-guided robotics

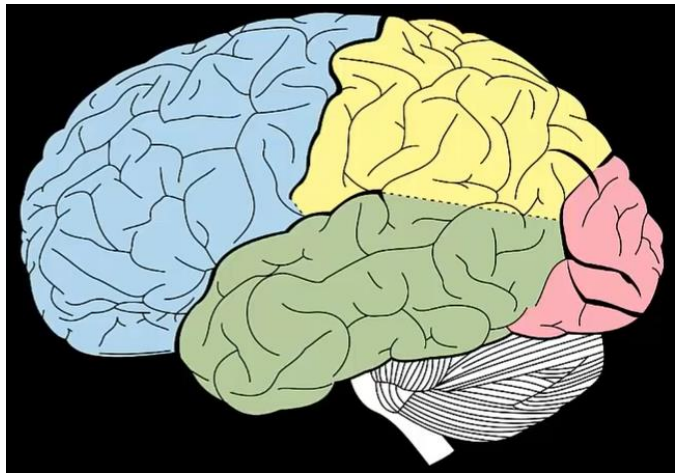
What is computer vision

Given an image or more, extract properties of the world

- Traffic scene
- Number of vehicles
- Type of vehicles
- Location of closest obstacle
- Assessment of congestion

Machine vision and human intelligence

- Vision is an essential part of how we perceive the world.
- More than 50 percent of the cortex, the surface of the brain, is devoted to processing visual information.
- Understanding how vision works may be a key to understanding how the brain as a whole works.
- One of the obstacles for solving computer problems is that we don't have a a detailed understanding of how the sensory information is processed in the brain.



Machine vision and human intelligence

- Vision involves attention

Machine vision and human intelligence

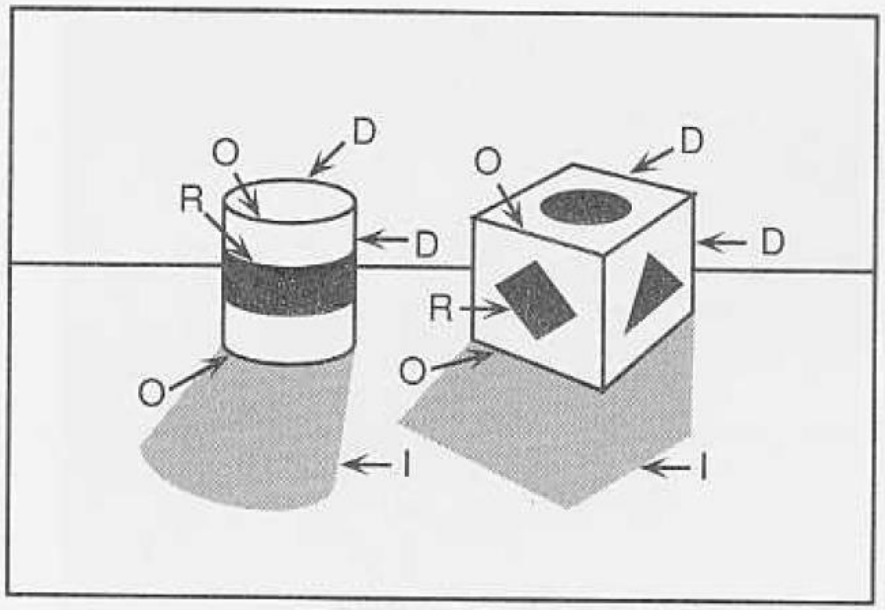


Figure 5.5.16 Four kinds of edges. This scene contains four different kinds of luminance edges: orientation edges (O) due to abrupt changes in surface orientation, depth edges (D) due to gaps between surfaces at different distances, reflectance edges (R) due to different surface pigments or materials, and illumination edges (I) due to shadows.

Challenge

- To develop human level (or better) capabilities for computers and machines.

What is in this image?

- Hand holding a man
- Hand holding a sphere
- A painting

Brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts

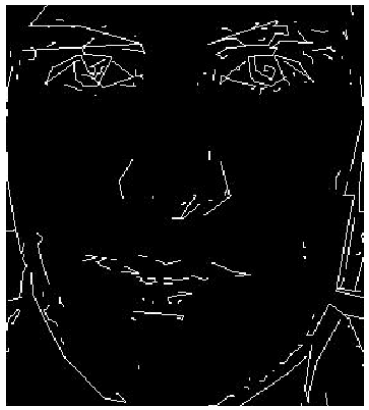
Brief history of computer vision

- 2012-2015: ANNs make a comeback. Deep learning revolution

Approaches to computer vision

- Two main paths:
  - Hand crafted features and learning rules: Systems are designed to capture visual properties with hard coded rules.

    Example: If you see an edge map similar to this, then its is a face.



Main challenges in this approach:

How to define/find edges, what to use as reference, and how to measure similarity.

Approaches to computer vision

- Training based methods (e.g., deep learning)
  - Get a large database of samples (datasets now in the size of millions!)
  - Label the dataset
  - Train a neural network to automatically learn the function that detects faces
- Challenges in this approach
  - Availability of the datasets and difficulty in manual labelling
  - Computing resources
  - Solutions are difficult to explain

Approaches to computer vision

- Successful solutions often hybrid utilizing both approaches

For example, the state-of-the-art in action recognition utilizes a deep neural network trained on optical flow images (manually designed feature).

Action recognition in videos

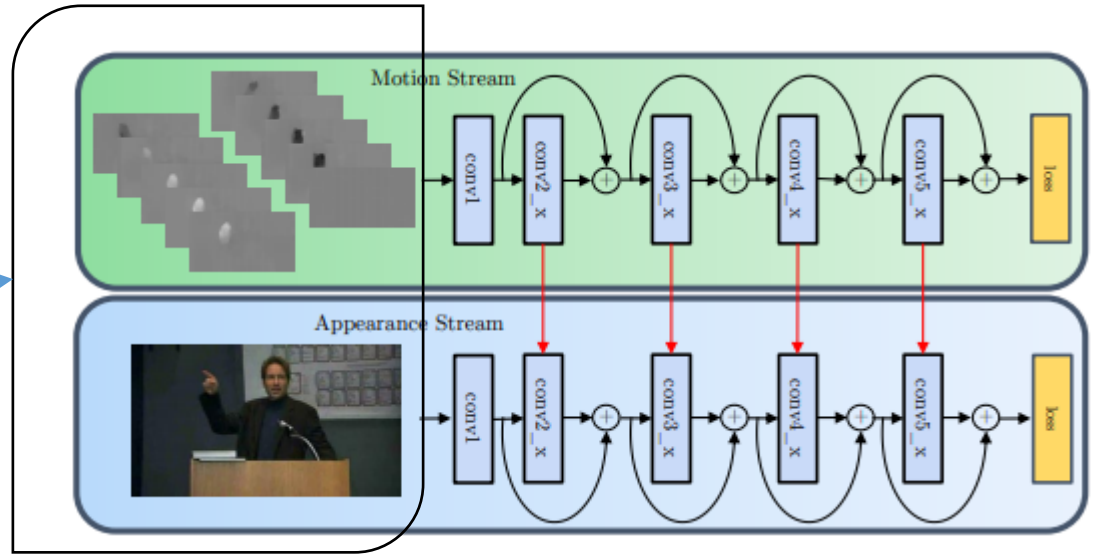
- Videos present a sequence of moving frames (images) at high frequency ~60 Hz.
- Action recognition and tracking are important problems in video analysis.

Optical flow

- Provides vectors that describe the direction of motion

Action recognition

Optical flow and
Image data are used as
inputs



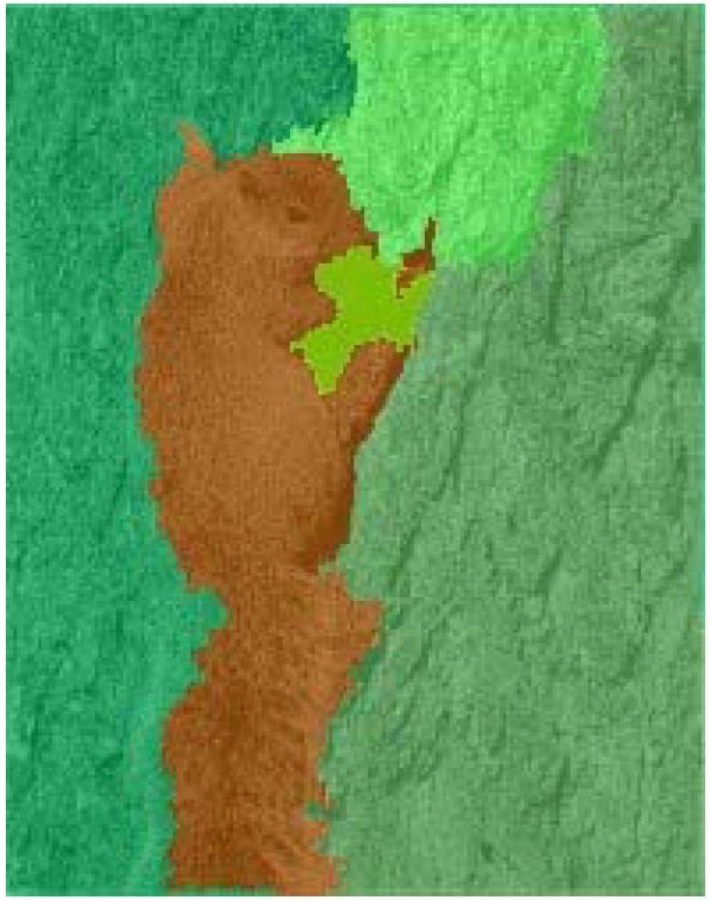
- Two streams are used for action recognition: The upper branch deals with optical flow, and the lower one on image data.

Feichtenhofer, Christoph, Axel Pinz, and Richard Wildes. "Spatiotemporal residual networks for video action recognition." Advances in neural information processing systems. 2016.

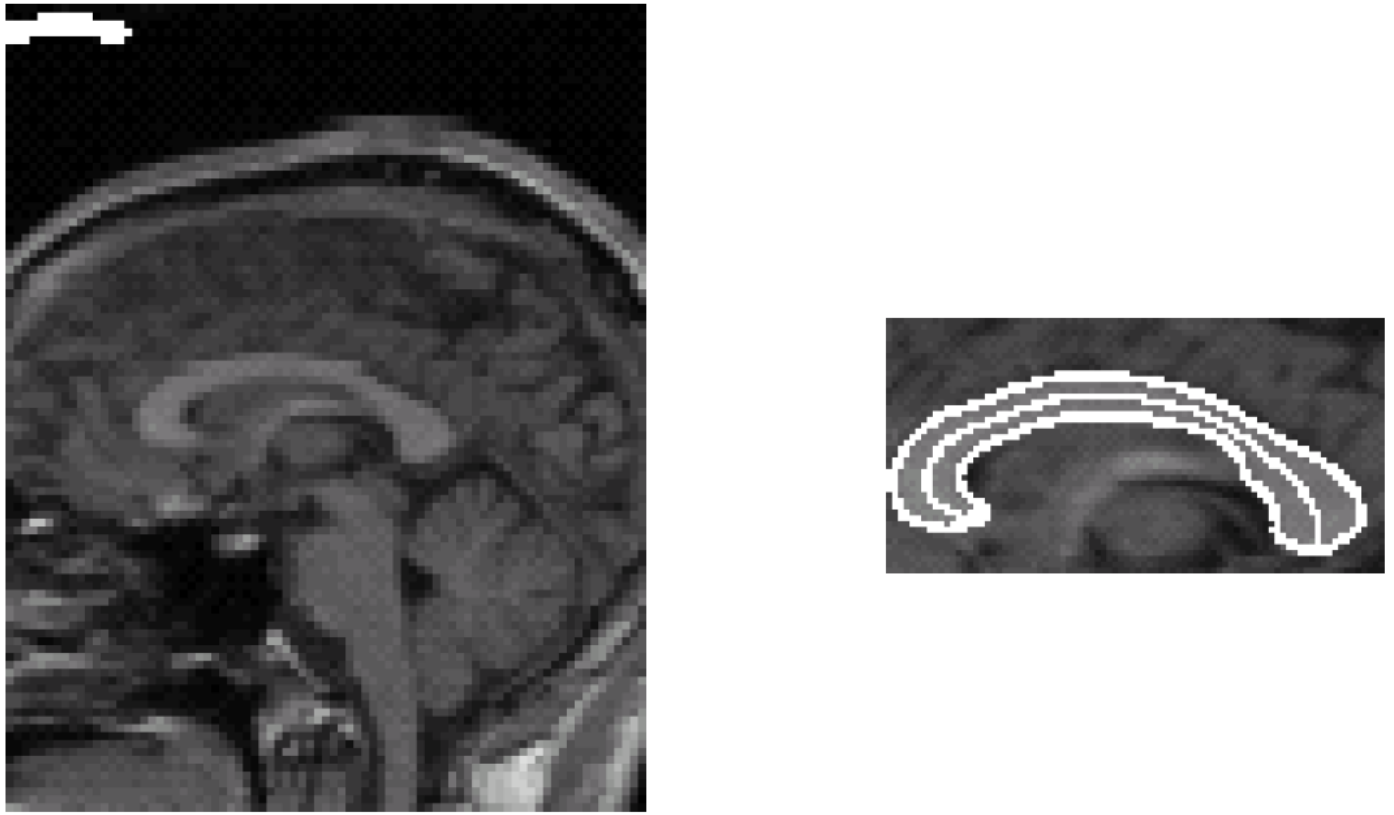
Segmentation

(Sharon, Balun, Brandt, Basri)

Segmentation



Segmentation

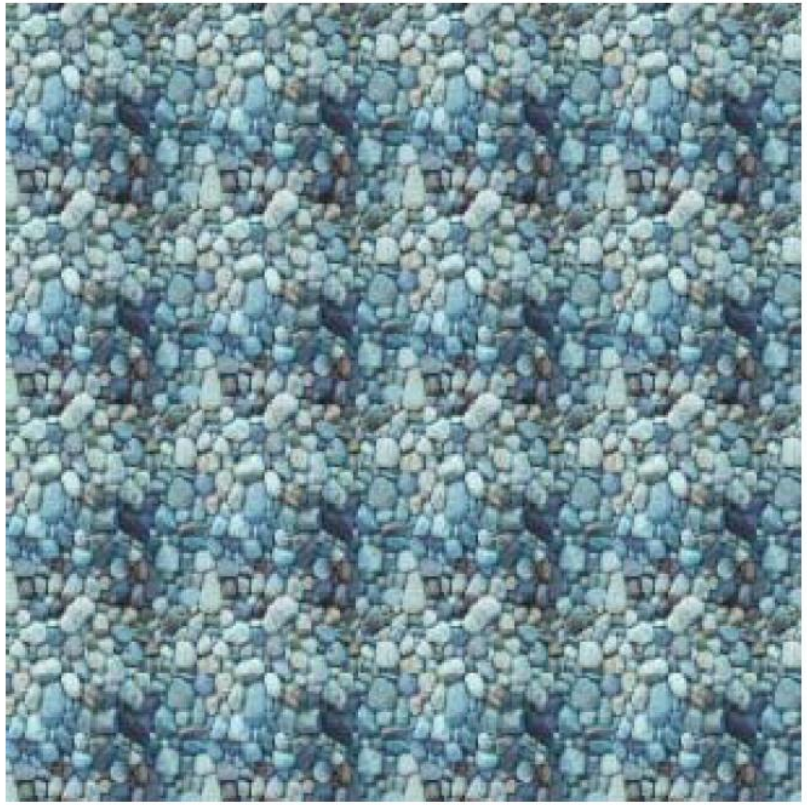


Finding the Corpus Callosum

Texture

- Textures are repeated patterns in images.

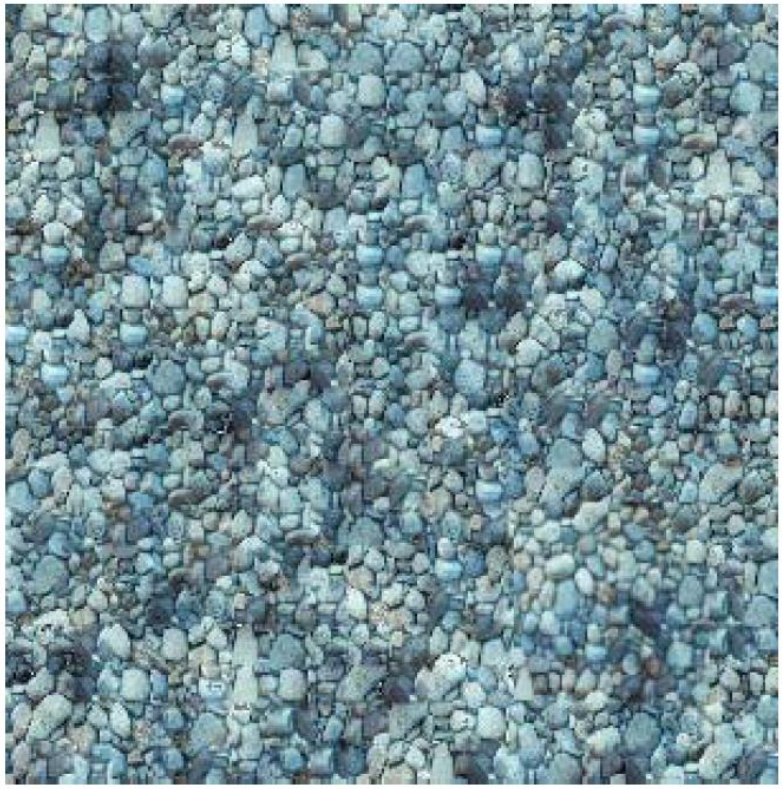
Photo



Pattern Repeated

Texture

Photo



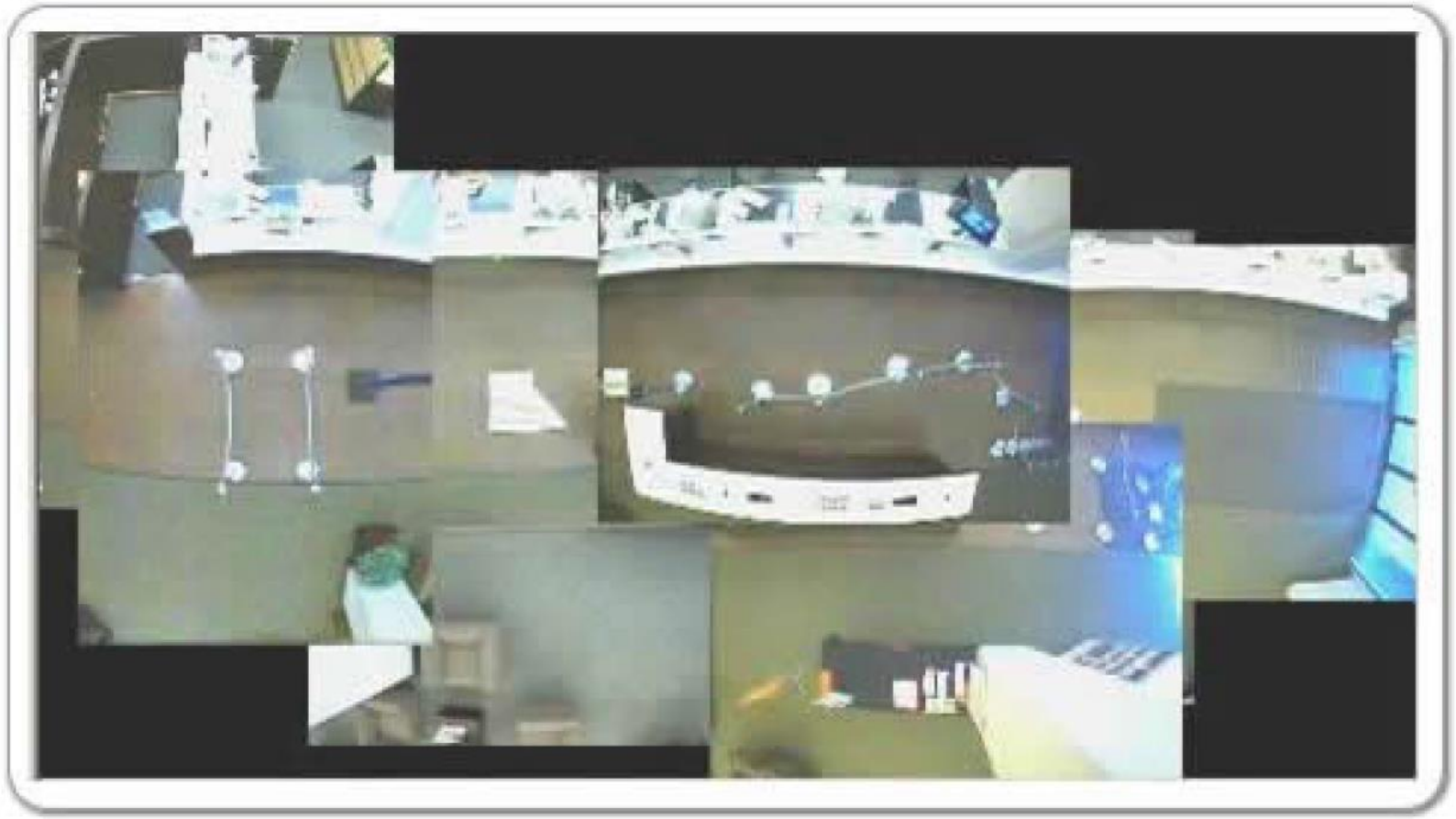
Computer Generated

Image Inpainting

Image Inpainting

Boundary detection

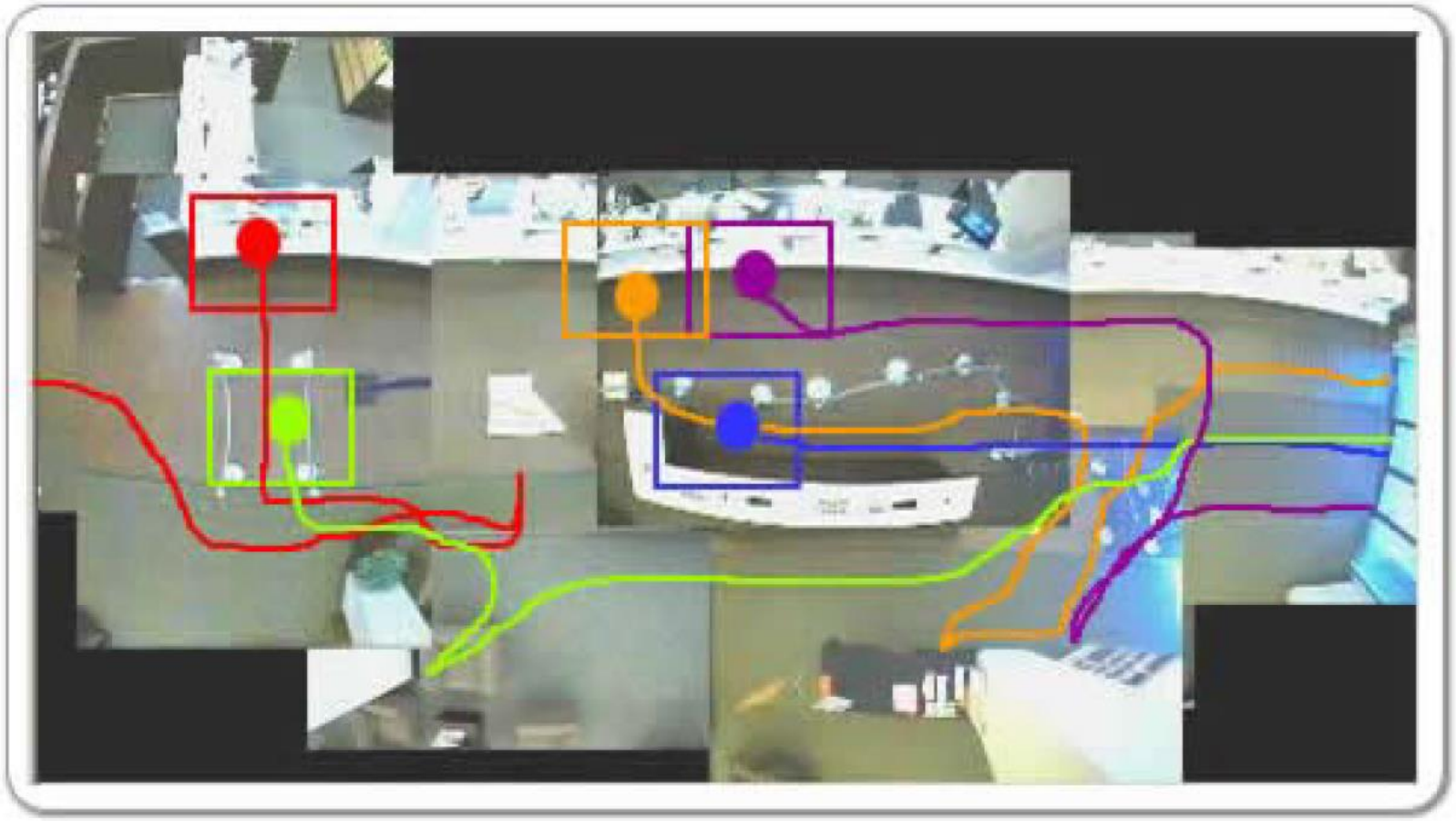
Motion tracking



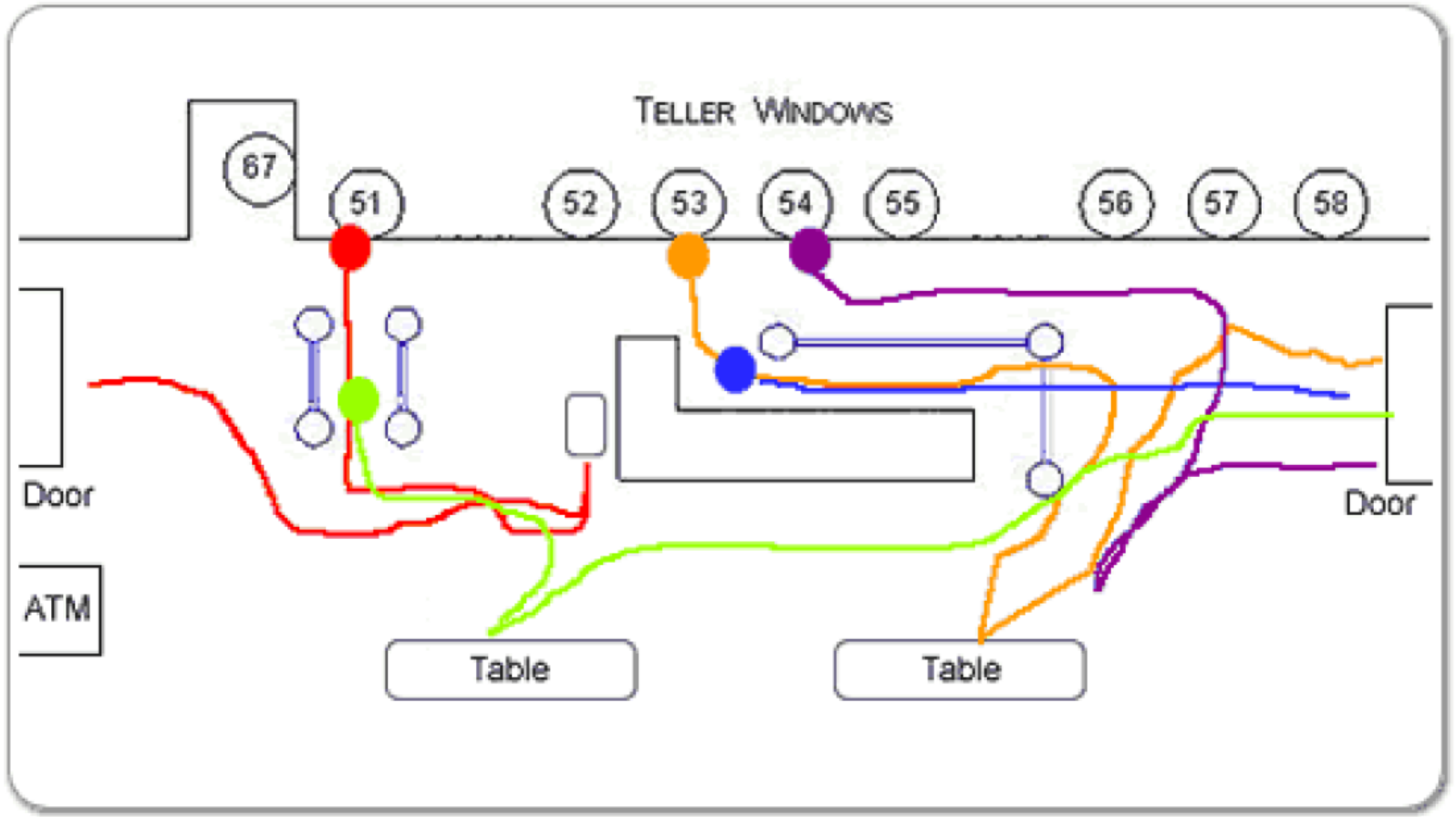
Brickstream.com

Motion tracking

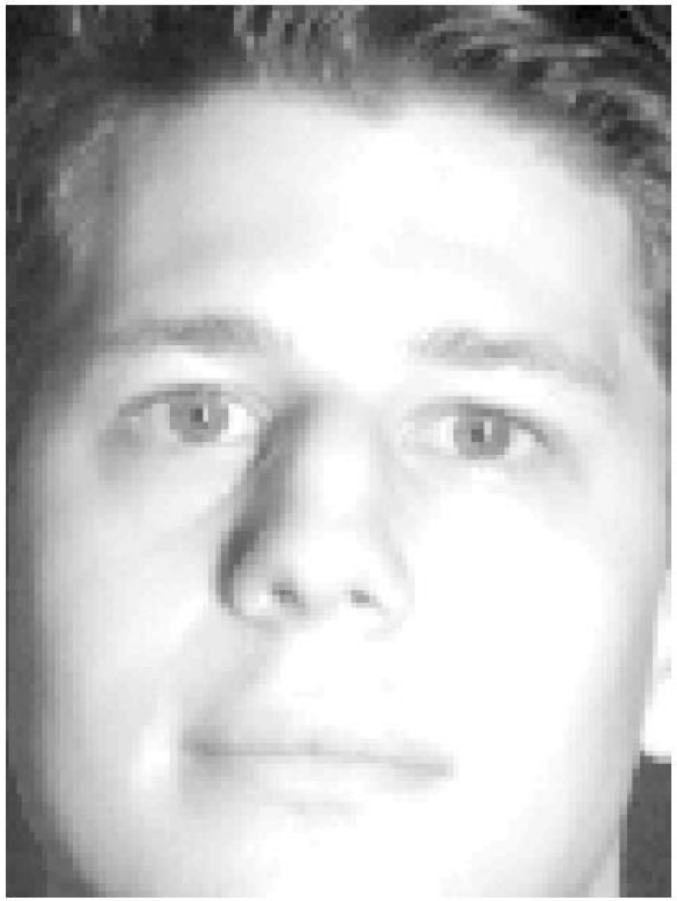
Motion tracking

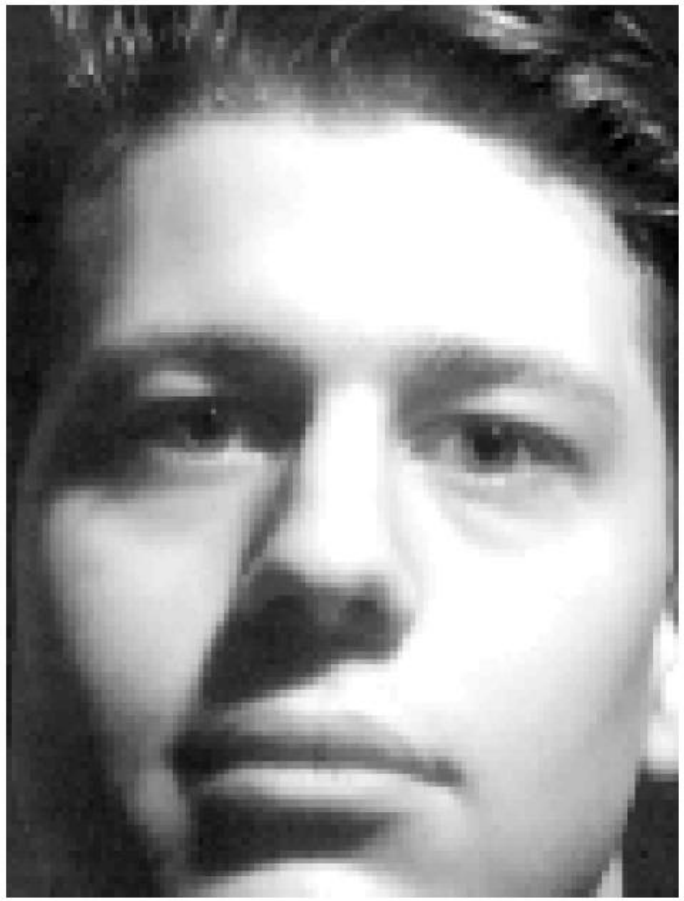


Motion tracking

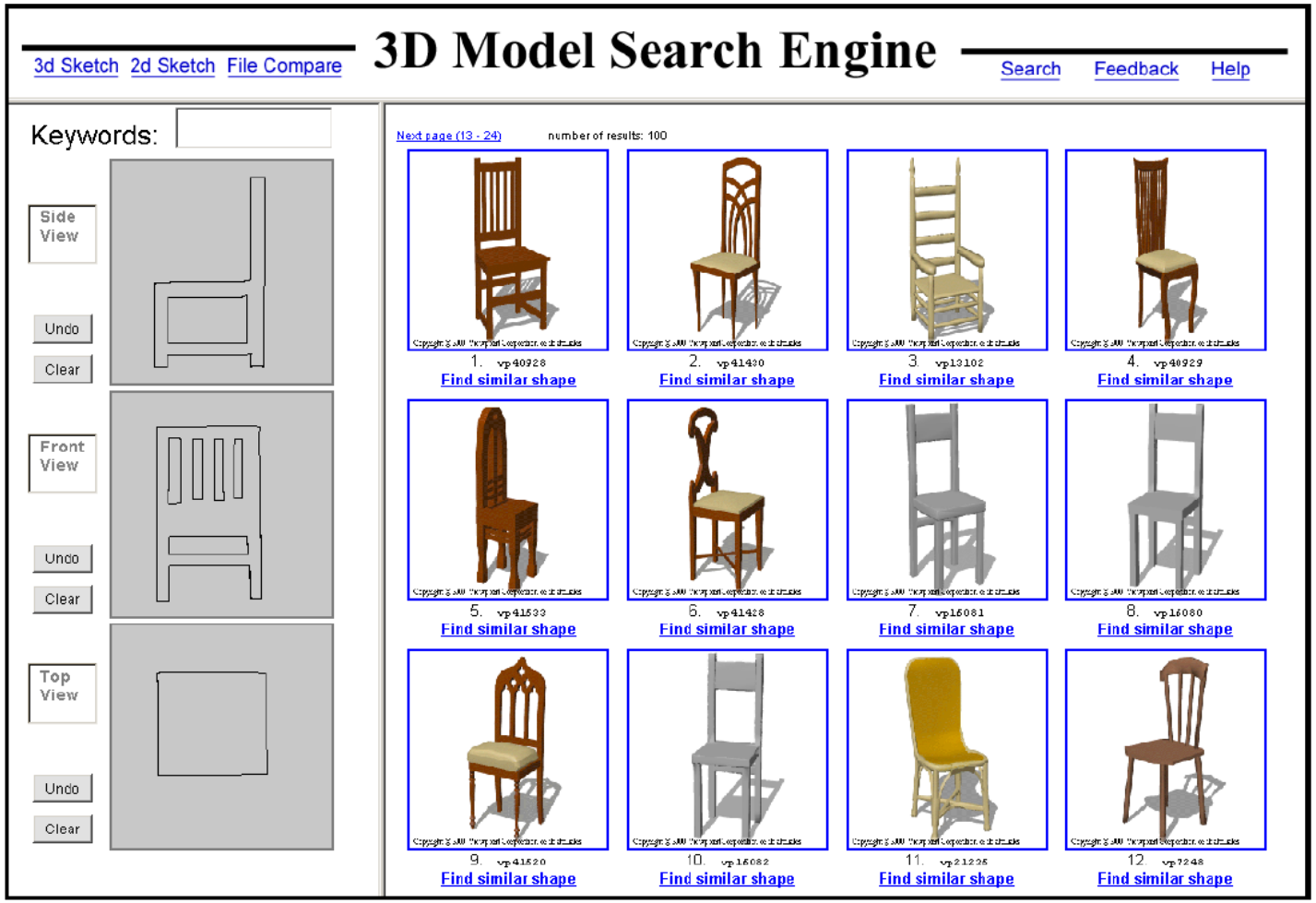


Lightening and shading play important rules



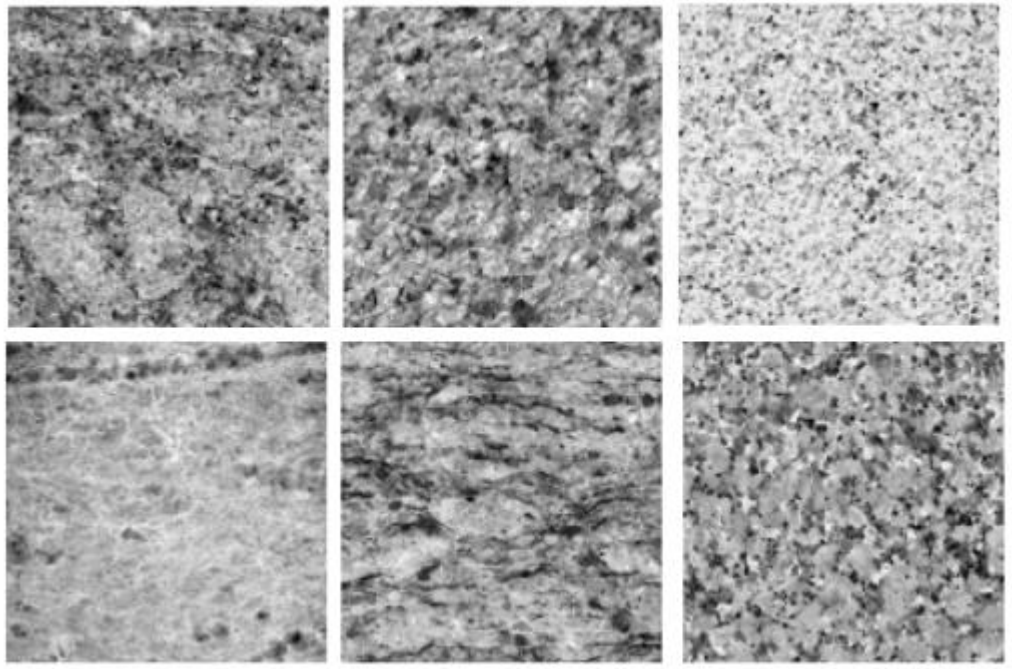


Classification from 3D shapes



(Funkhauser, Min, Kazhdan, Chen, Halderman, Dobkin, Jacobs)

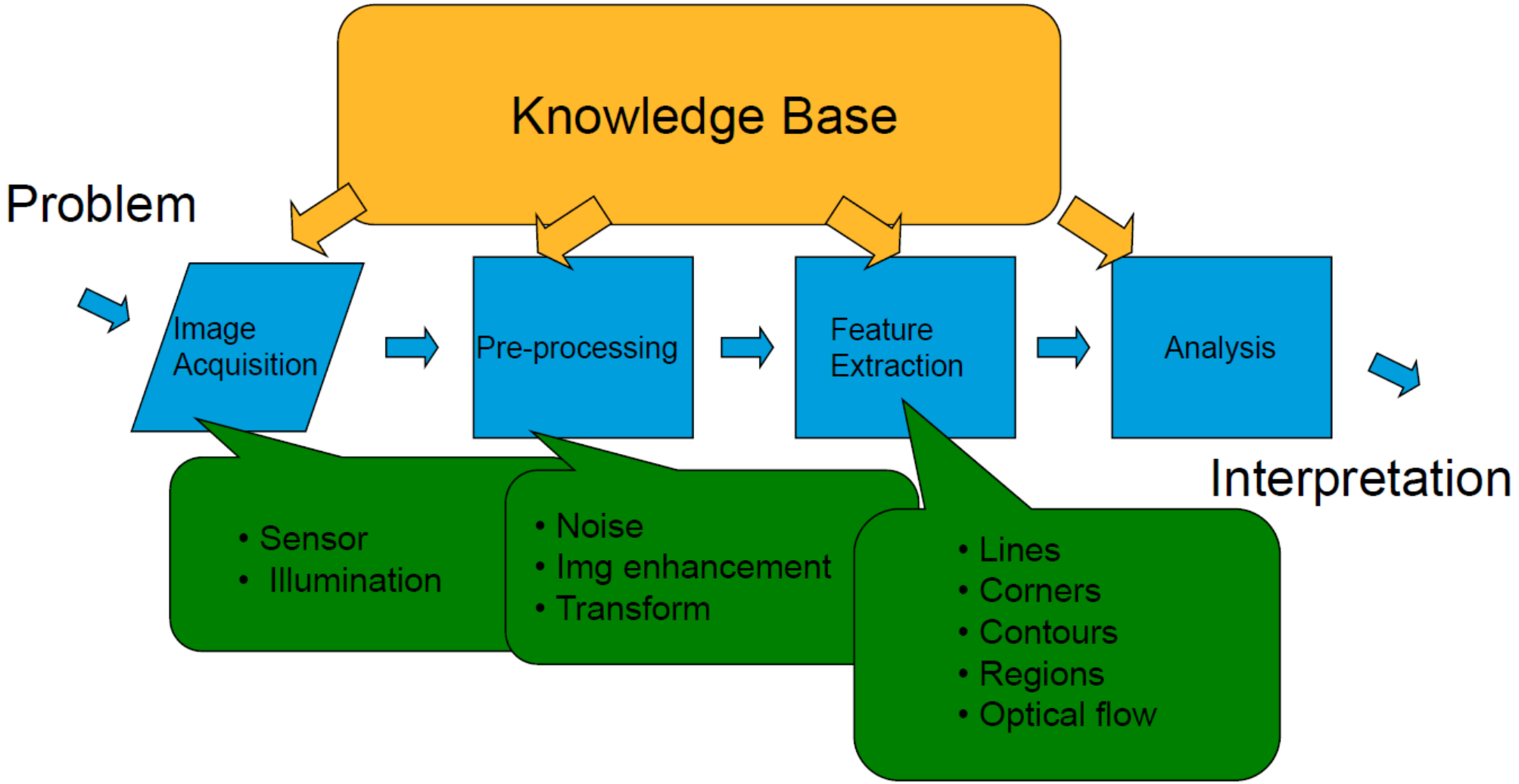
Classification using texture



Granite images

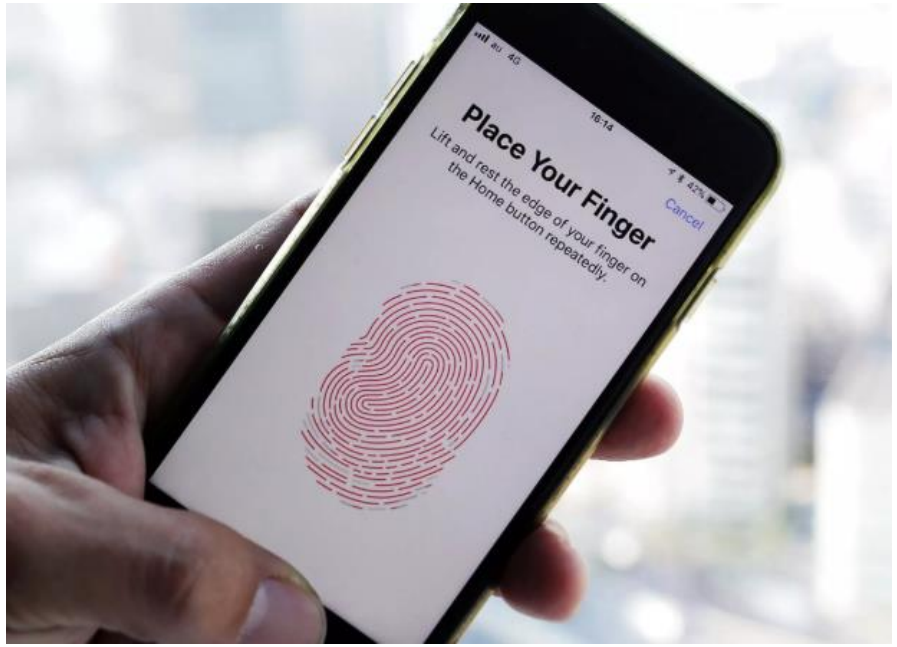
Visually guided surgery

General Structure of a CV problem

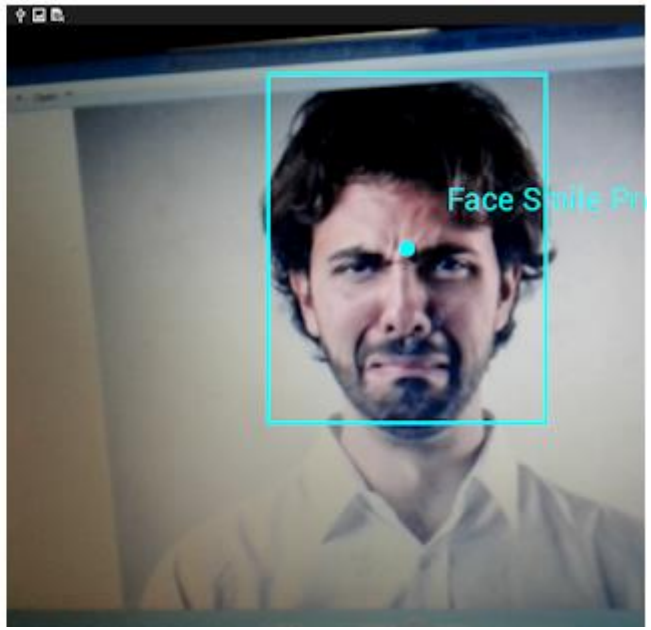


Login without a password

- Face recognition
- Fingerprints



Face/smile detection apps



Interactive video games

- https://www.youtube.com/watch?v=8CTJL5lUjHg

Self-driving cars

- Self-driving on highways. Driver attention required

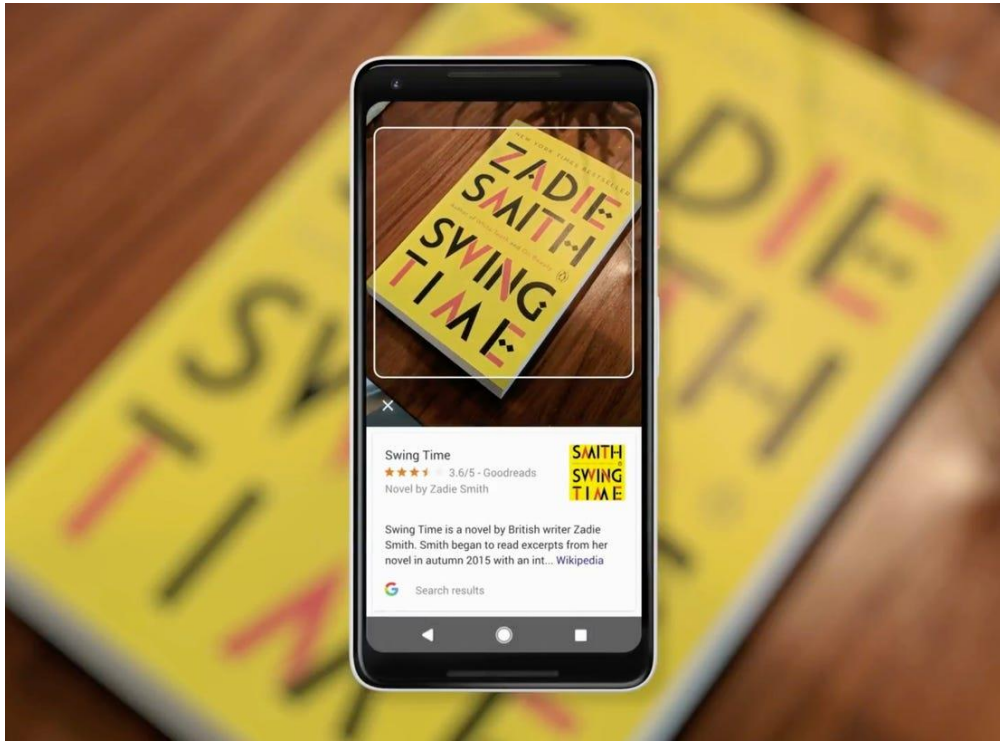
https://www.youtube.com/watch?v=tlThdr3O5Qo

- Smart Summon feature

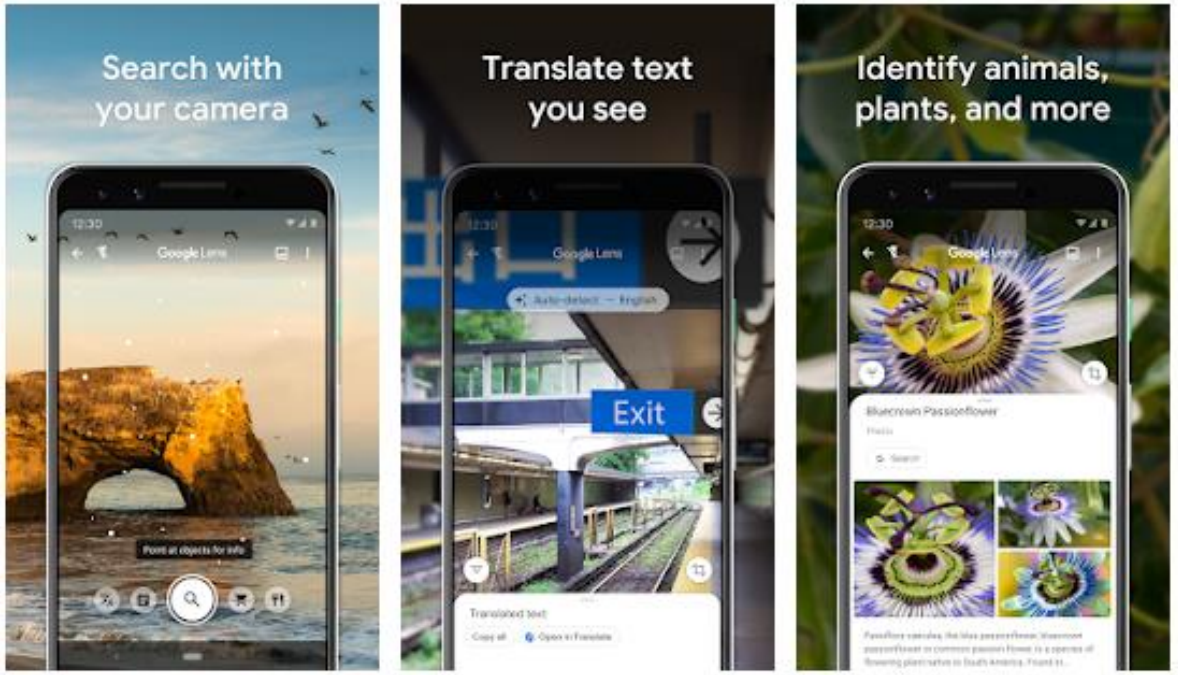
https://www.youtube.com/watch?v=nlCQG2rg4sw

- Technical presentation
  https://www.youtube.com/watch?v=IHH47nZ7FZU

GoogleLens App

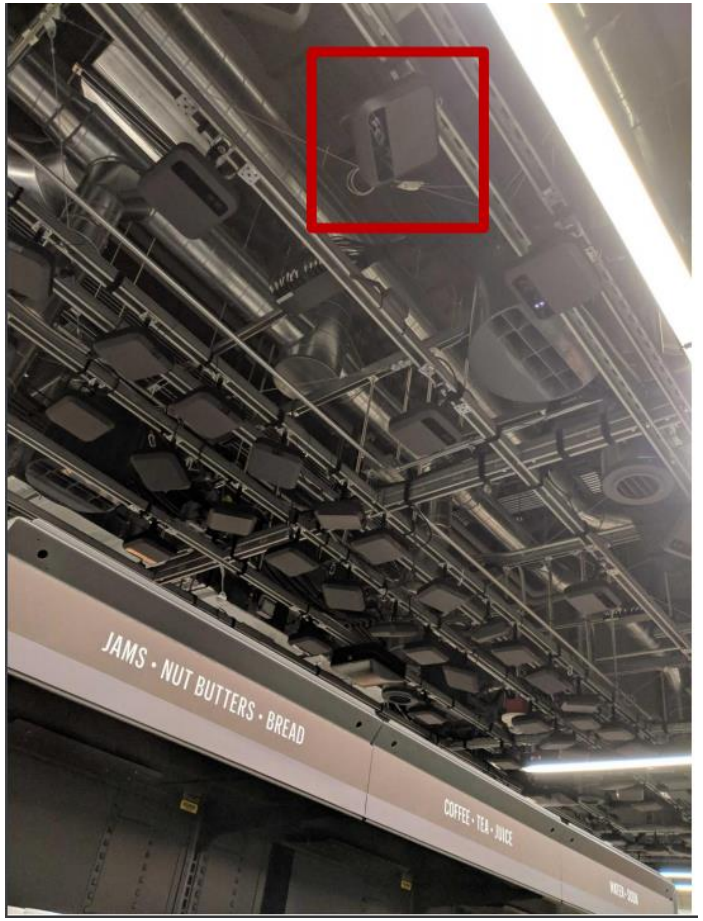


GoogleLens App



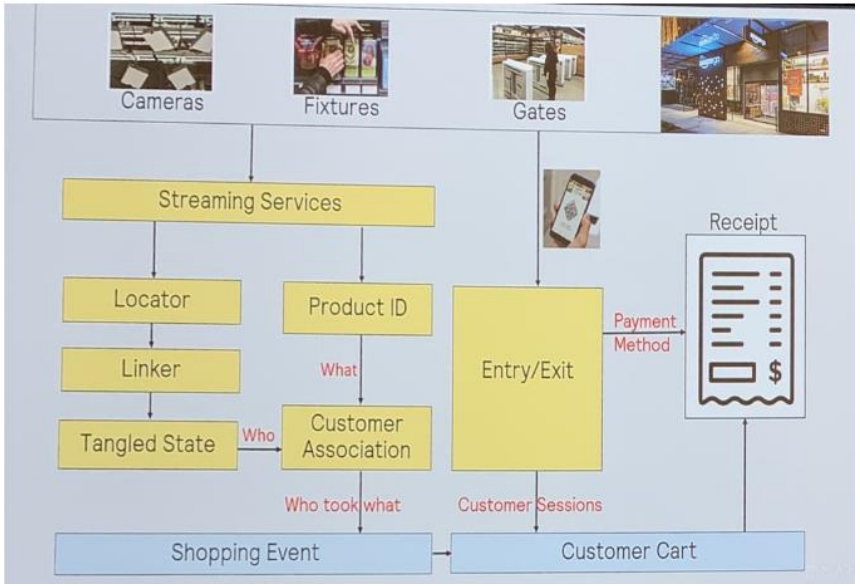
Amazon Go stores

Amazon Go stores



Amazon Go stores

- No checkout necessary. Computer vision is used to track and add your purchases to your virtual cart.
- When you get out of the store, your credit card is charged.



General Remarks

On HW + Final project

- You will have to run many experiments using different settings and parameter values.
- Start early: don’t underestimate time to get the first results + time to analyze and prepare the report
- Use online sources for help
- Summarize your analysis and results in a professional manner.

Final exam

Comprehensive, practice the problems solved in class

Outline

1. Introduction
2. Image enhancement
3. Frequency domain operations
4. Image descriptors
5. Machine learning
6. Neural networks
7. Segmentation and object detection
8. Morphological processing
9. Geometric transformations
10. Motion analysis and optical flow
11. Compression
12. Other topics