Computer Vision

Tue, July 14 (Week 4)

Computer Vision

Camera Pixels->Computer->Something Useful (for Applications)

Since cameras pixels are arranged in 2D, it is basically for 2D problems.

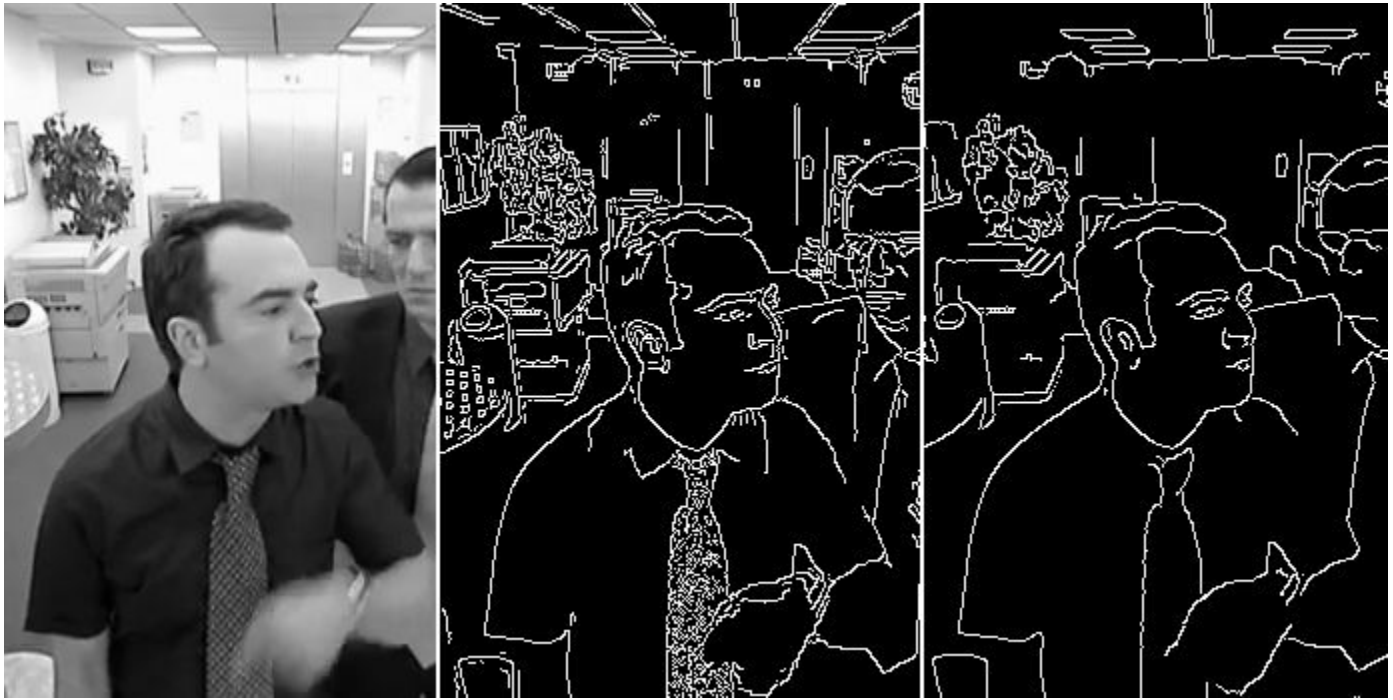
Computer Vision

The classical approach for computer vision problems:

Preprocess the input image (e.g., find edges)

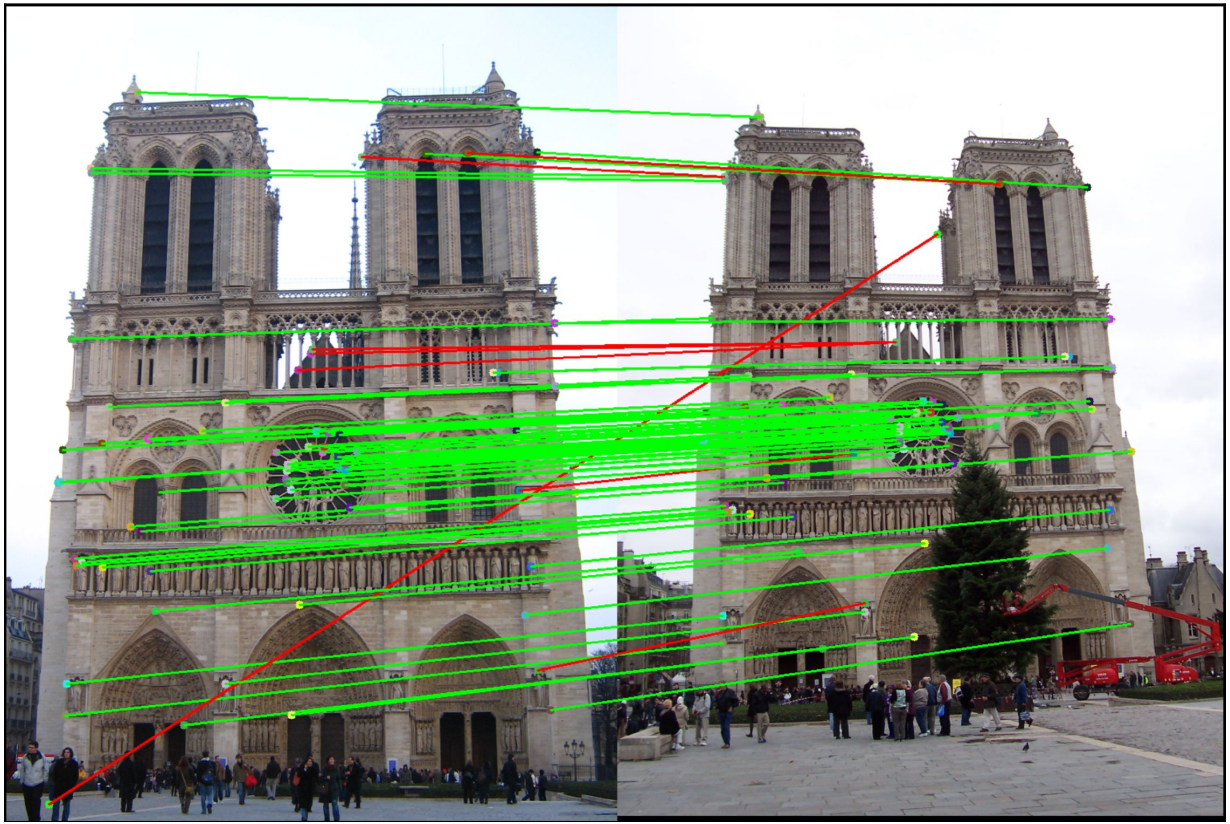
Apply assumptions that and conduct the task (e.g., find cats)

Edge Detection

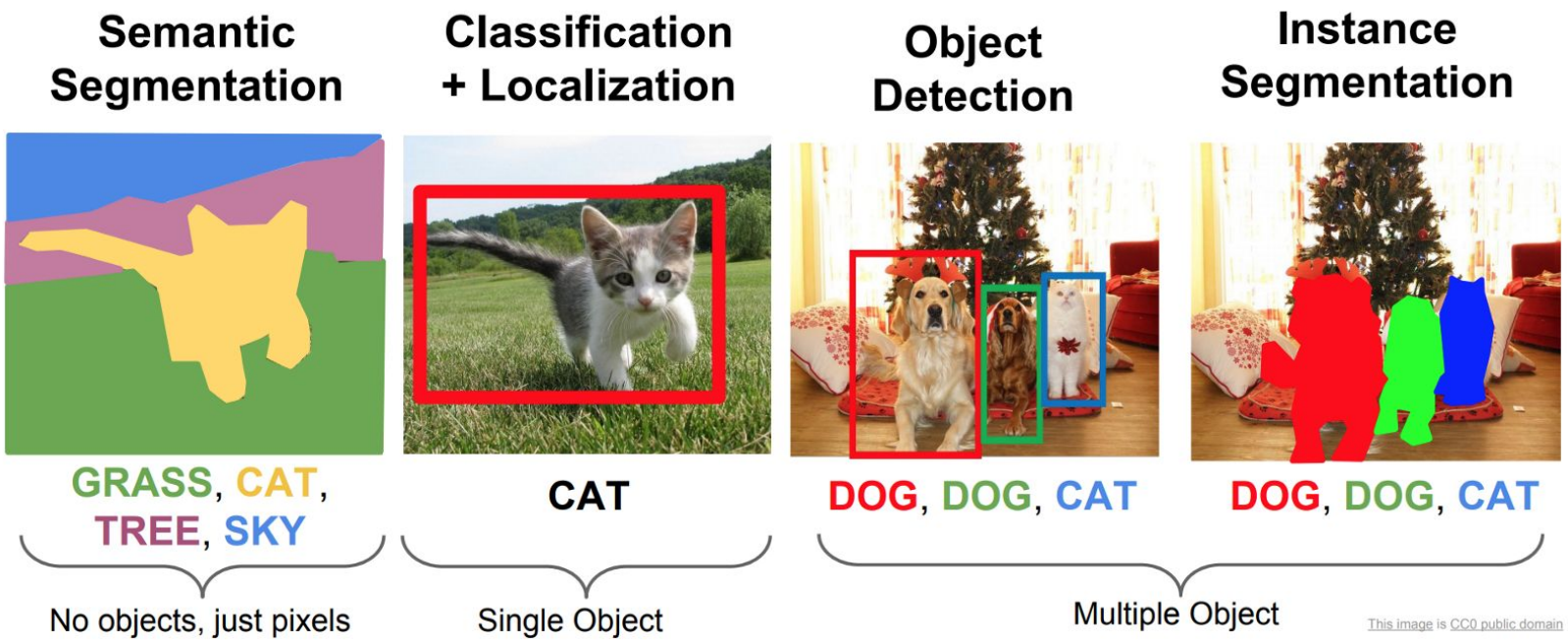


Optical Flow

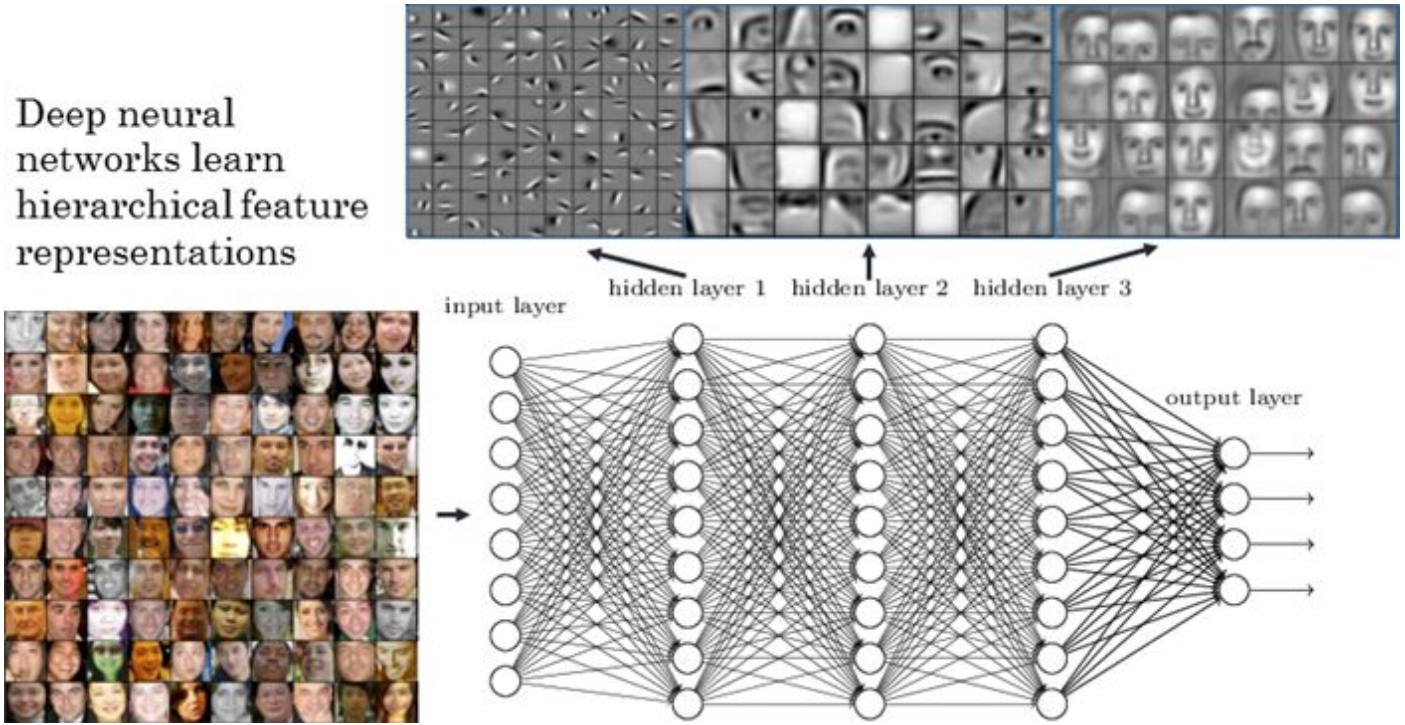
Scale-Invariant Feature Transform



Popular Computer Vision Problems



Feature Extraction from a Neural Network



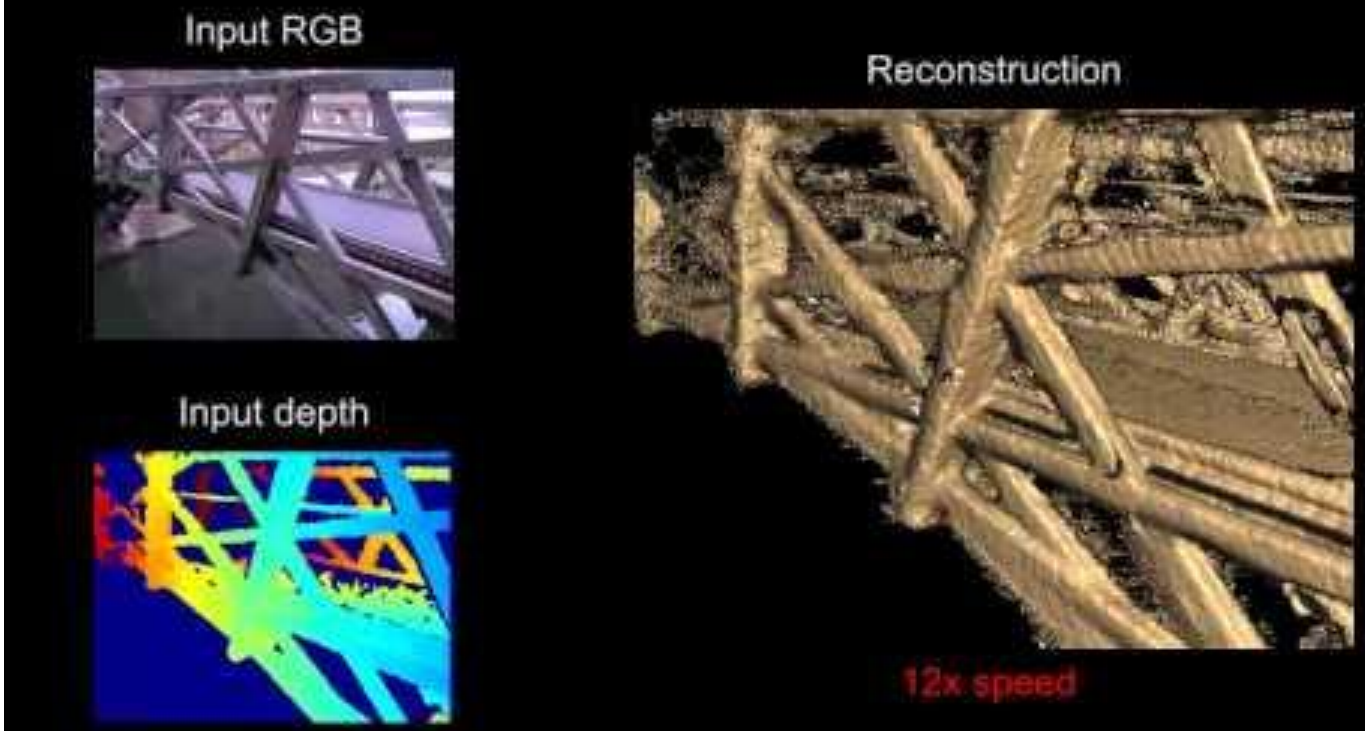
Computer Vision for AR

The meaningful things computer vision can provide for AR:

- What and where are the objects
- Where is the camera in the real world
- What is the overall structure of the surrounding environment

Usage of Depth Cameras

Per-pixel color vs. per-pixel distance



https://youtu.be/NsrmniEvO4s

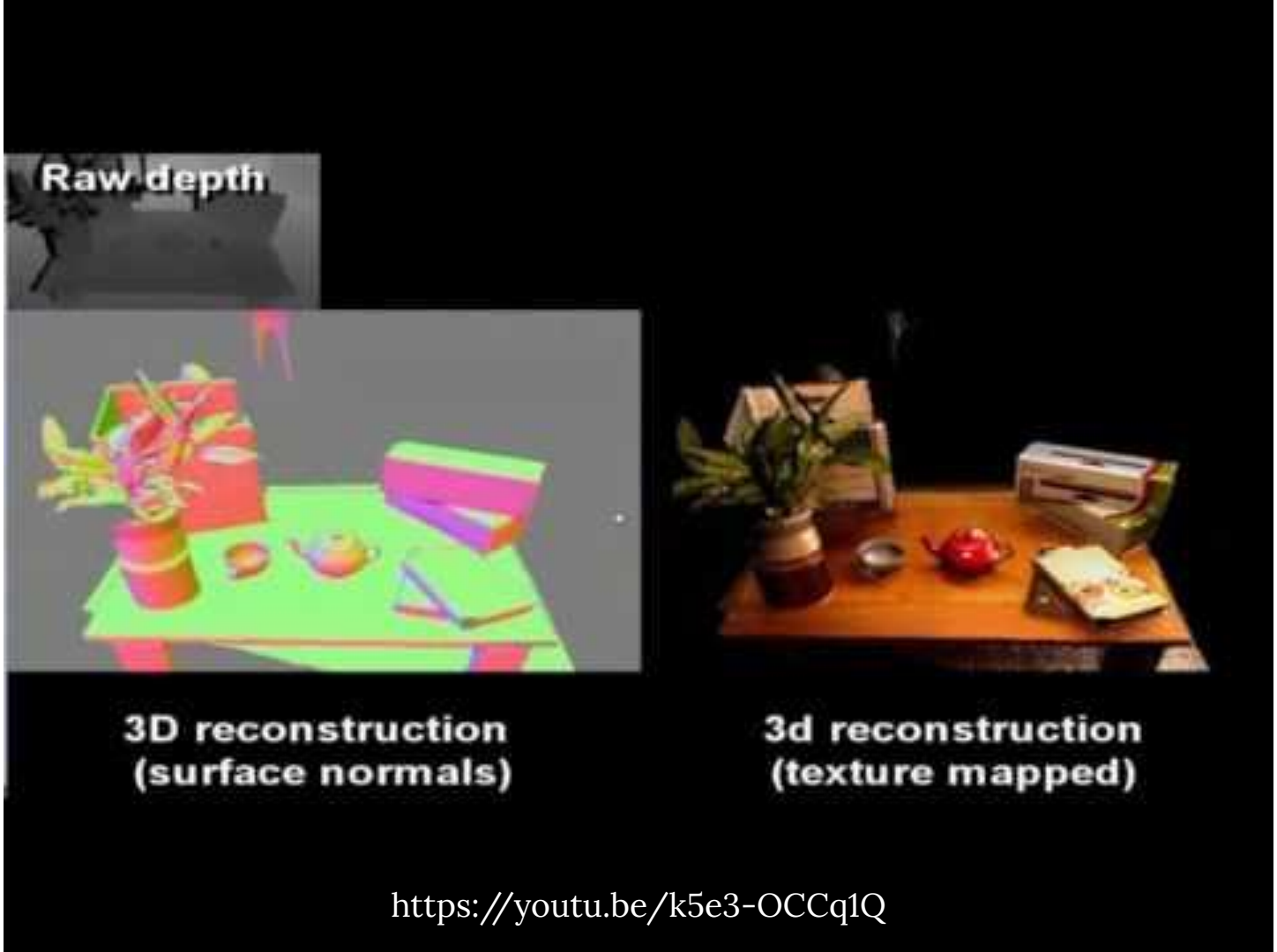
3D Reconstruction vs. SLAM

3D Reconstruction: creating the mesh of the real world, which requires knowing where the sensor is.

SLAM: figuring out where the sensor is at and creating a map for the external world while doing such.

Sounds similar, but notice their goals are different.

As a result, for example, when 3D reconstruction creates a detailed mesh, SLAM creates a much simpler map.



https://youtu.be/k5e3-OCCq1Q


https://youtu.be/n0fW3Uh_kGA

How These Works

Based on an Assumption:

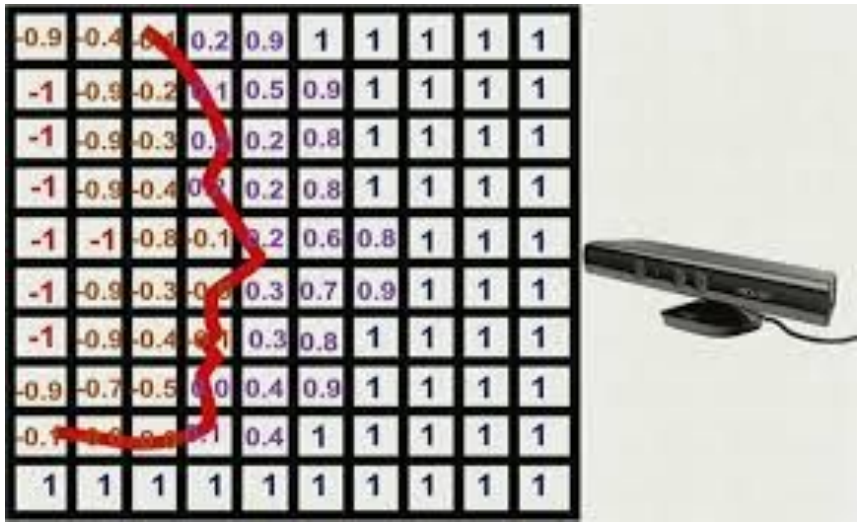
Our world is static, therefore is many things move at the same time to a certain direction, it is likely that the camera is moving to the opposite direction.

+ Usage of Data from the IMU Sensor

Further Details of 3D Reconstruction

TSDF (Truncated Signed Distance Function) Grid

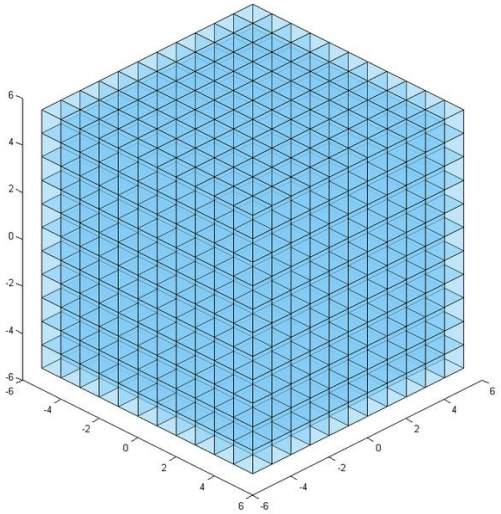
: A number per a cell that describes how far the cell is from a surface detected from the camera.



Further Details of 3D Reconstruction

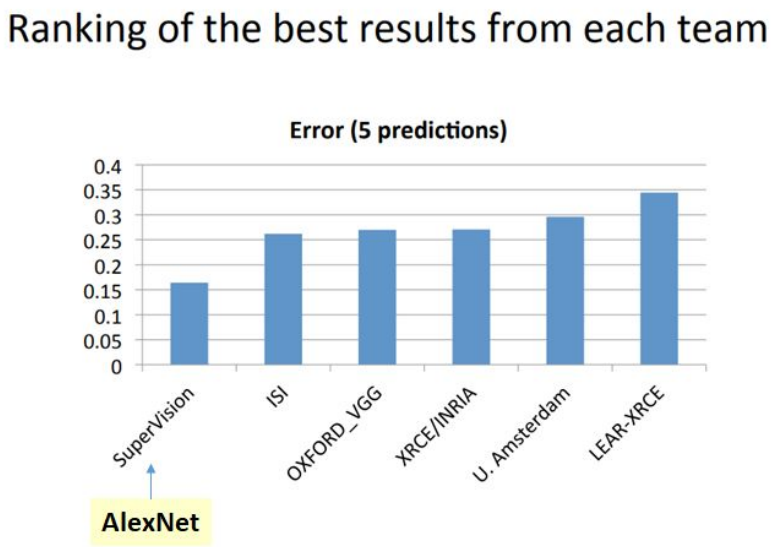
Imagine having a voxel (volumetric pixel) version of the TSDF grid...

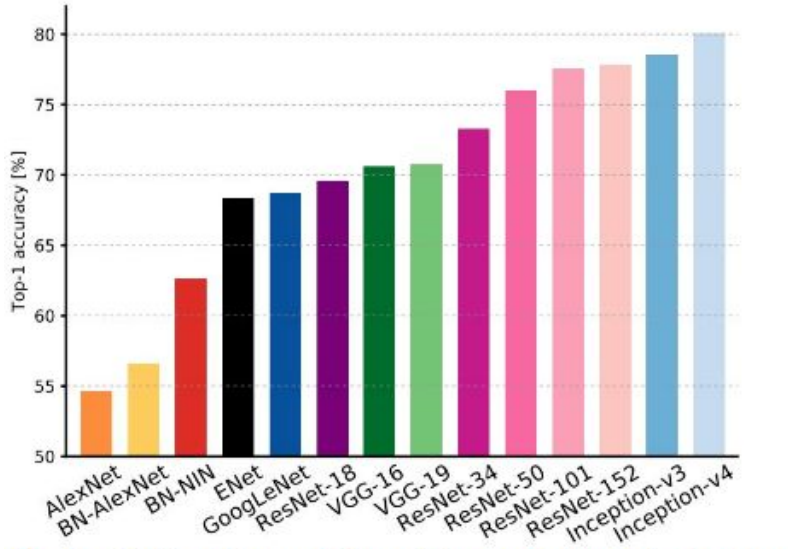
This is the first key idea and further details will be discussed next week...



Machine Learning (since 2012)

The recent history of machine learning systems is full of surprises.



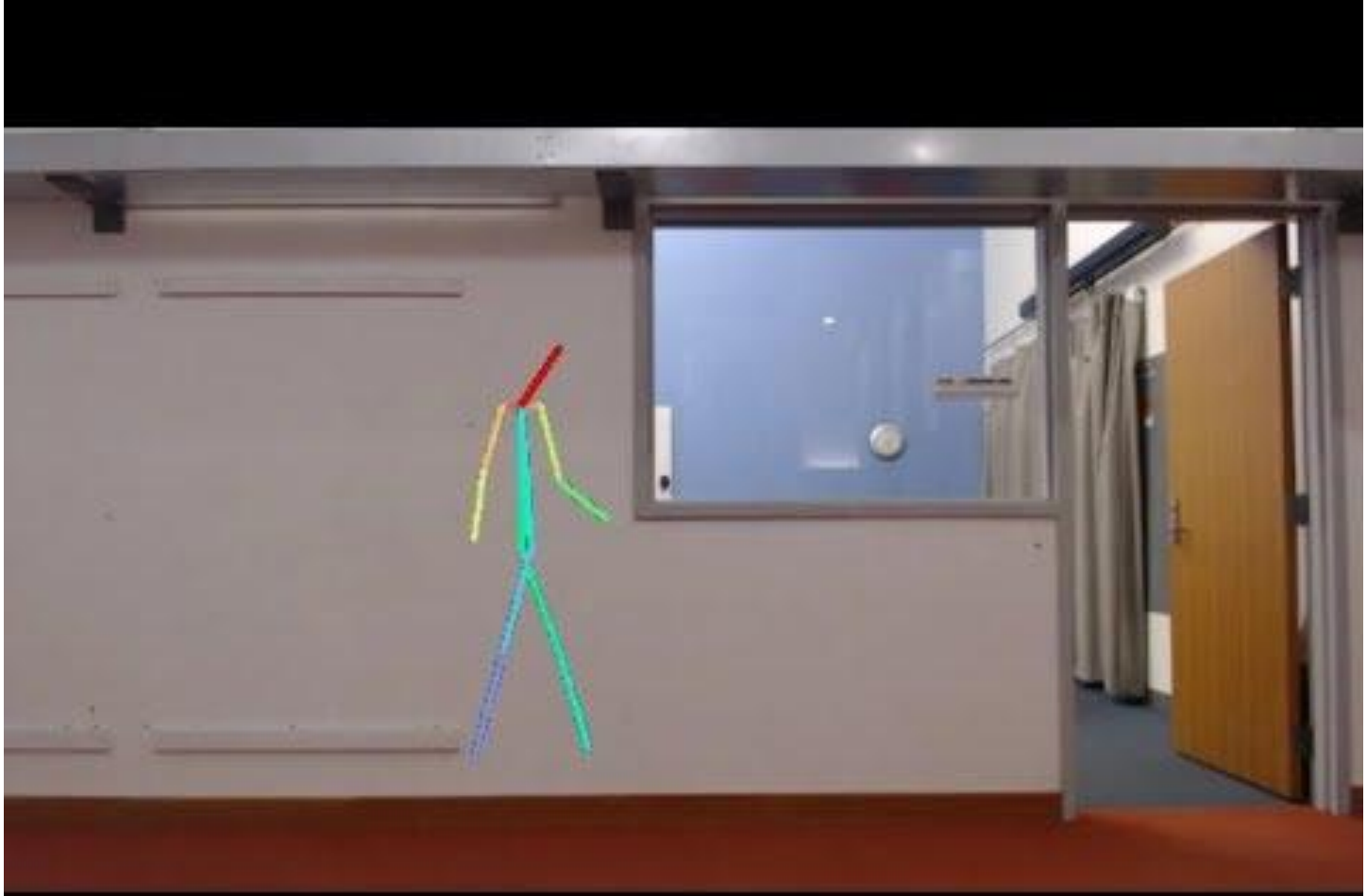


Machine Learning (since 2012)

Machine learning based solutions: good but many become a black box very hard for a human to interpret, especially when things go wrong.

Classical solutions without machine learning: harder to improve compared to machine learning; however, in case someone figures out a wonderful algorithm, it can become preferable.

Based on what happened to other fields, machine learning is likely to win.

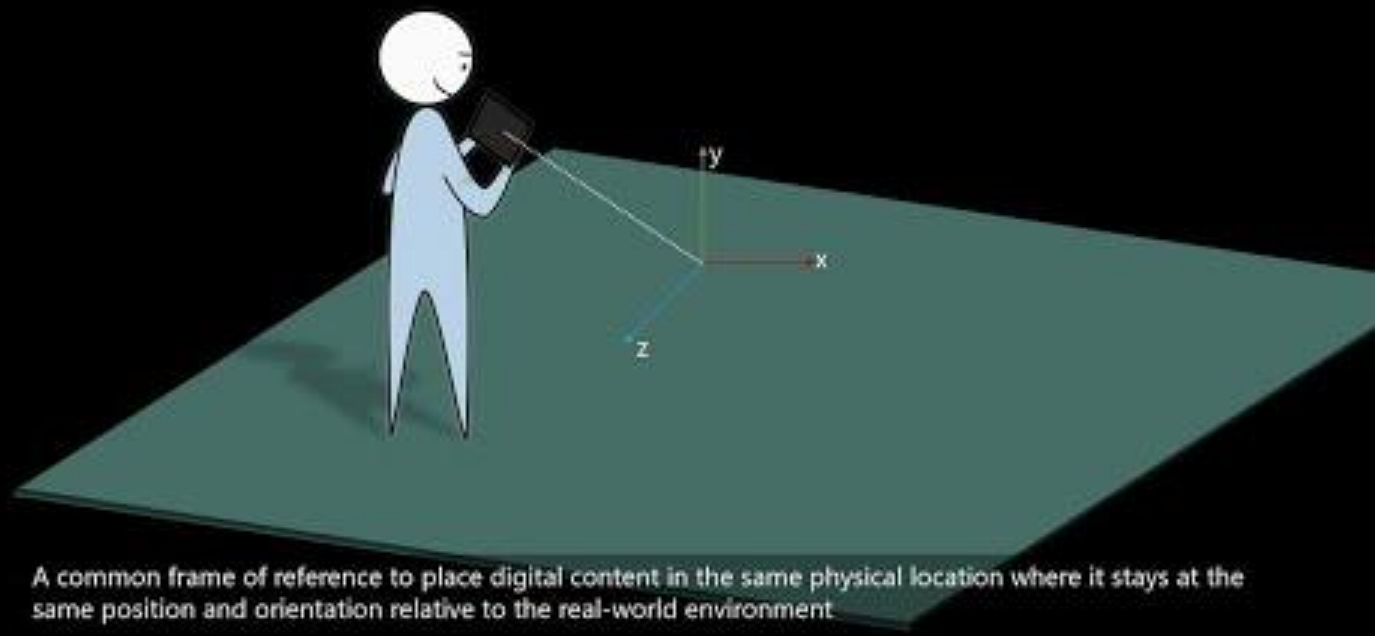


https://youtu.be/HgDdaMy8KNE

A Computer Vision Problem for AR

Getting the coordinate system: another topic for next week.

What is a Spatial Anchor?



A common frame of reference to place digital content in the same physical location where it stays at the same position and orientation relative to the real-world environment

https://youtu.be/CVmfP8TaqNU