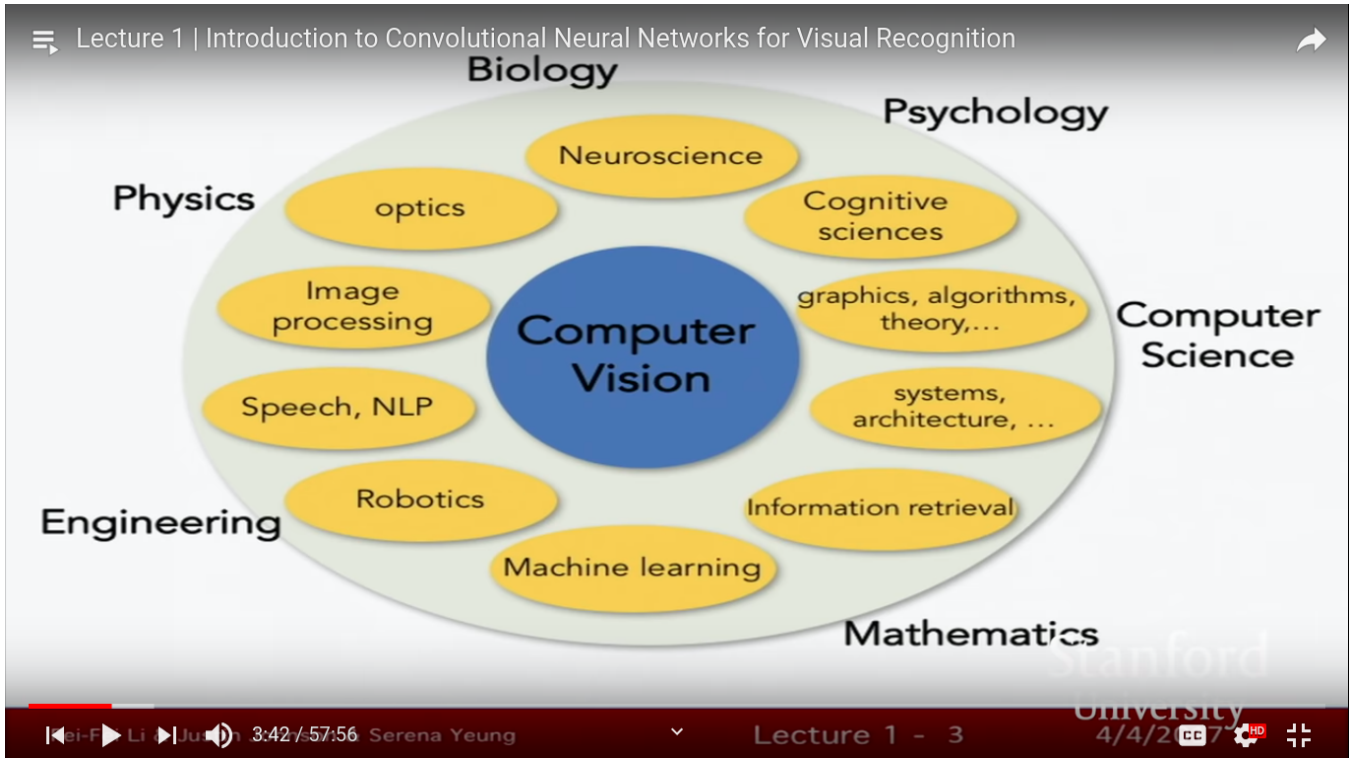


# lec1

## Lecture 1 notes on 22/10/2019 22:16

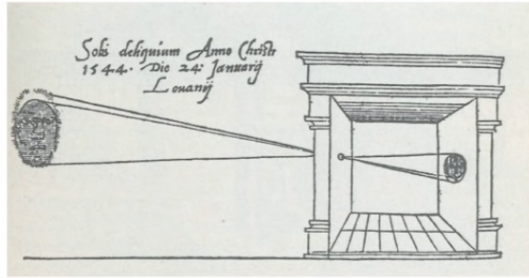
- Videos are the dark matter of the web.
- Every second there is 5hrs of video being uploaded on YouTube.
- CV needs the knowledge of various fields.



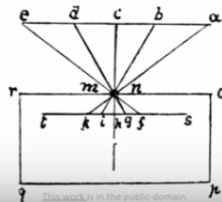
- Brief history of computer vision
  - The history of vision goes back to 543 million years.
  - Within 10 million years the number of species exploded (Evolution's Big Bang).
  - Andrew Parker, an Australian Zoologist proposed a convincing theory, from the study of fossils, it was found that the animals developed eyes and the onset of vision started.
  - So vision gave the power to explore, and this caused species to evolve to survive.
  - In humans 50% of neurons in our cortex involved in visual processing, it is the biggest sensory system.
  - Human vision helps in communicate, entertain, survive, work, move around, manipulate things and many things.

# Camera Obscura

Gemma Frisius, 1545



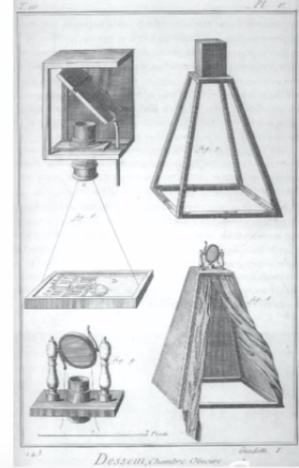
This work is in the public domain



This work is in the public domain

Leonardo da Vinci,  
16<sup>th</sup> Century AD

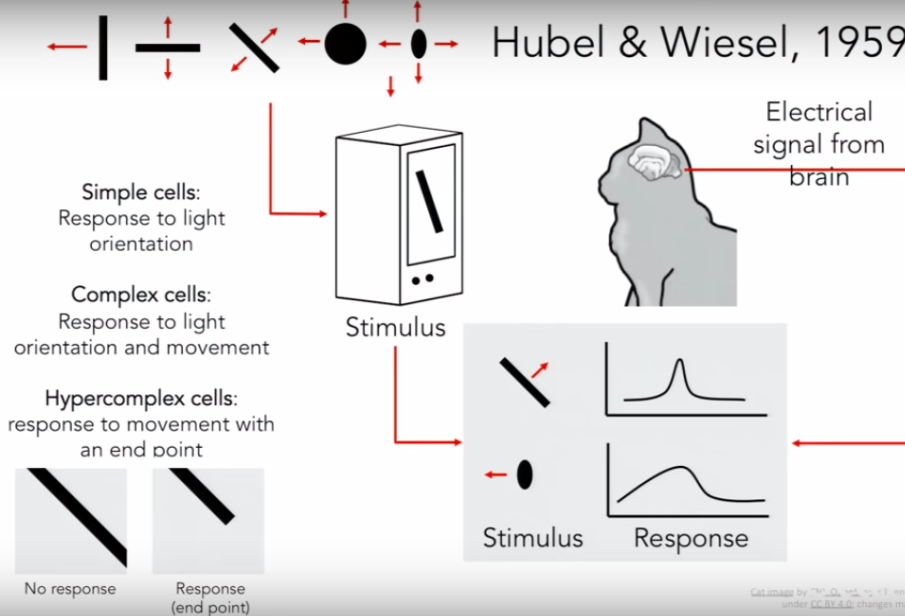
Encyclopedie, 18<sup>th</sup> Century



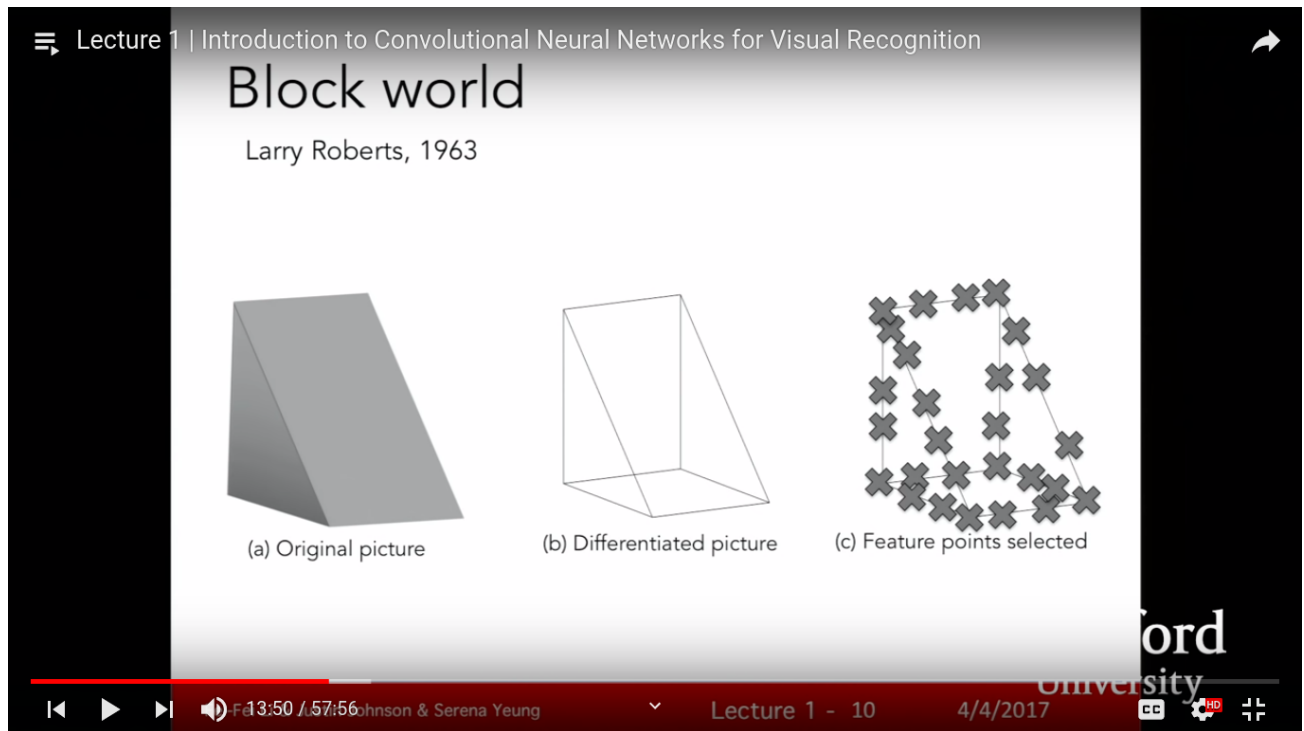
This work is in the public domain

- The eye responds to oriented edges

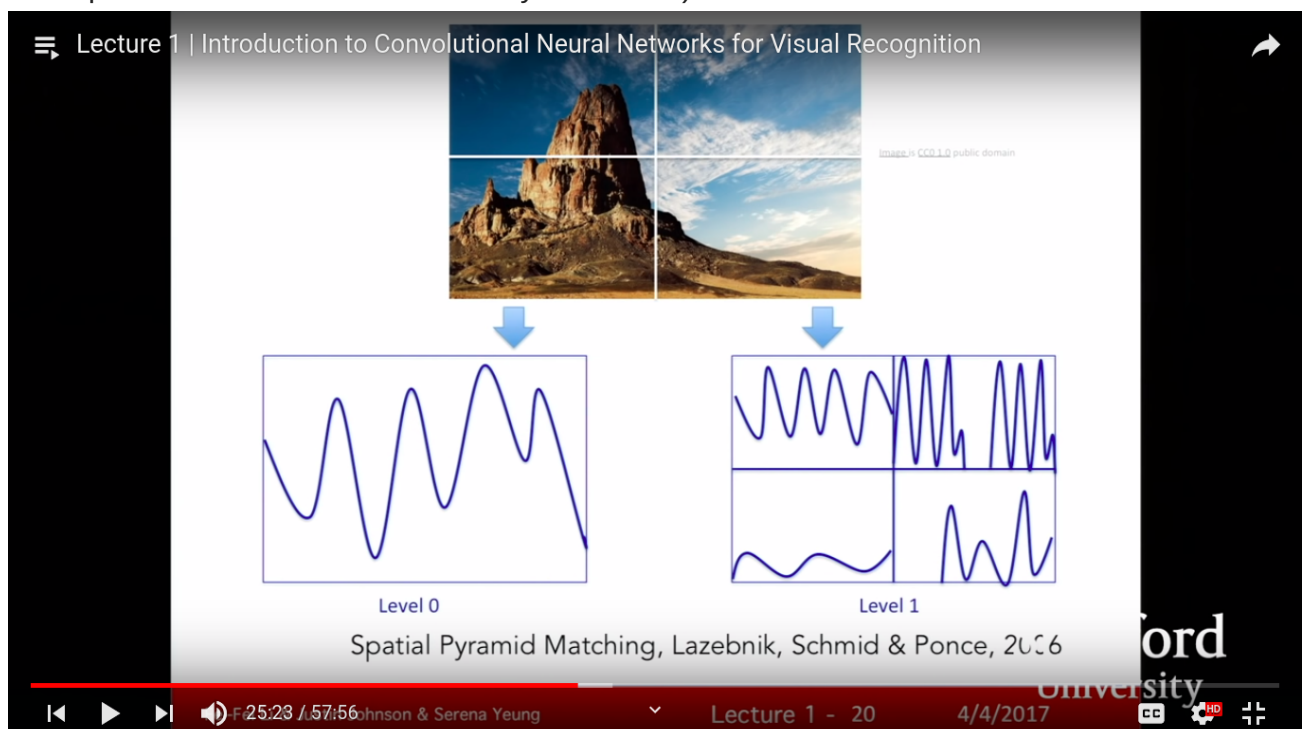
Hubel & Wiesel, 1959



- The history of CV also starts in the early 60s.



- 'Block world' can be considered the first PhD thesis on computer vision where the visual world was simplified into simple geometric shapes.
- In 1966 MIT had the 'THE SUMMER VISION PROJECT', which attempted to develop "pattern recognition"
- The task of object recognition began with identifying features in an image which are invariant to changes in the image.
- **Spatial Pyramid Matching** can be used to find the type of scene in the given image. Eg. Kitchen scene, landscape scene, traffic scene.(It uses image at different resolution to find feature descriptor and then use SVM to classify the scene.)



- PASCAL VOC (Visual Object Challenge) 2007 has 20 object categories and was one of the earliest benchmark dataset for object detection.
- ImageNet has 22K categories with 14M images, this was created in 2009 mainly to combat over-fitting issue caused by less data but large amount of features.
- There is a number of visual recognition problems that are related to *image classification*, such as *object detection*, *image captioning*, *Action classification* and so on.

Lecture 1 | Introduction to Convolutional Neural Networks for Visual Recognition

## IMAGENET Large Scale Visual Recognition Challenge

**Year 2010**  
NEC-UIUC

[Lin CVPR 2011]

Lion image by Swissfrog is licensed under CC BY 3.0

**Year 2012**  
SuperVision

[Krizhevsky NIPS 2012]

Figure copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

**Year 2014**  
GoogLeNet

[Szegedy arxiv 2014]

**Year 2014**  
VGG

[Simonyan arxiv 2014]

**Year 2015**  
MSRA

[Fei et al. 2015]

University of Cambridge 4/4/2017

- CNN were not invented overnight.

Lecture 1 | Introduction to Convolutional Neural Networks for Visual Recognition

## 1998 LeCun et al.

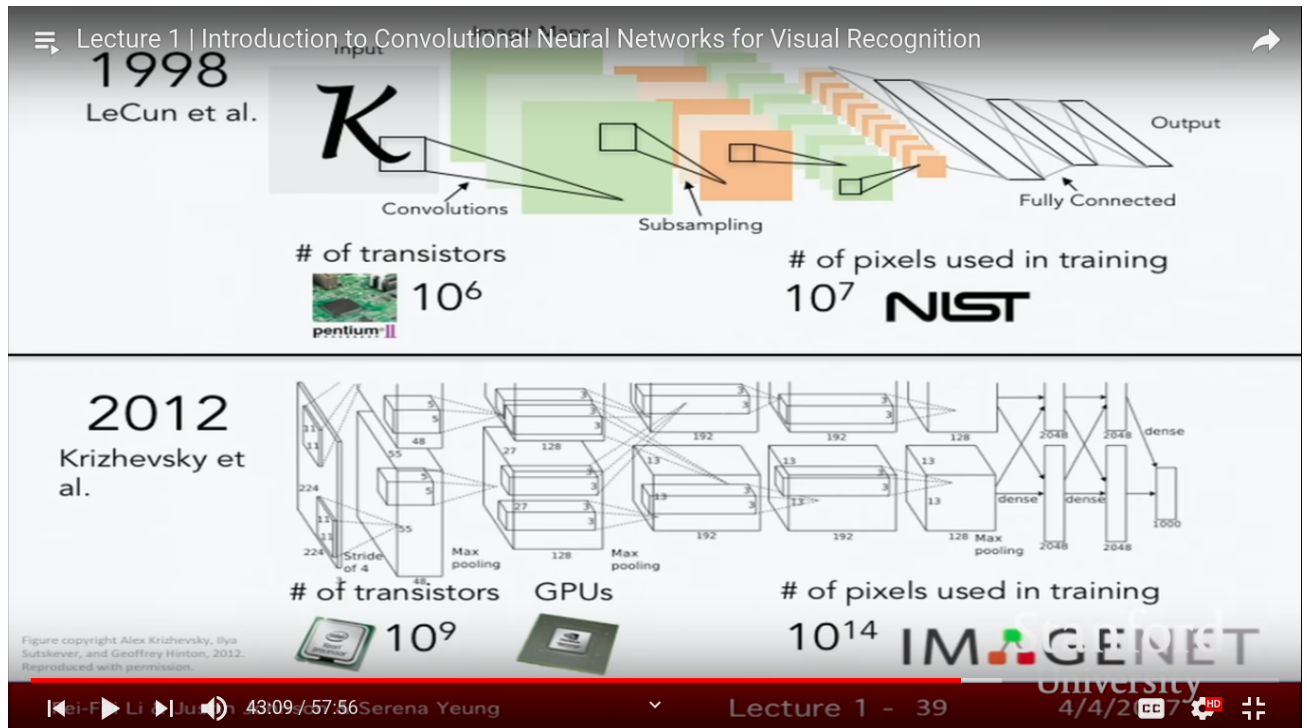
## 2012 Krizhevsky et al.

Figure copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

Stanford University 4/4/2017

- Key innovations which helped the field are:
  - Computing Power (By Moore's Law)

- GPUs (Parallel computation)
- Data availability



- The quest for visual intelligence goes far beyond object recognition
  - Semantic Segmentation
  - 3D Reconstruction
  - 3D pose construction
  - Activity Recognition

End

23/10/2019 22:18