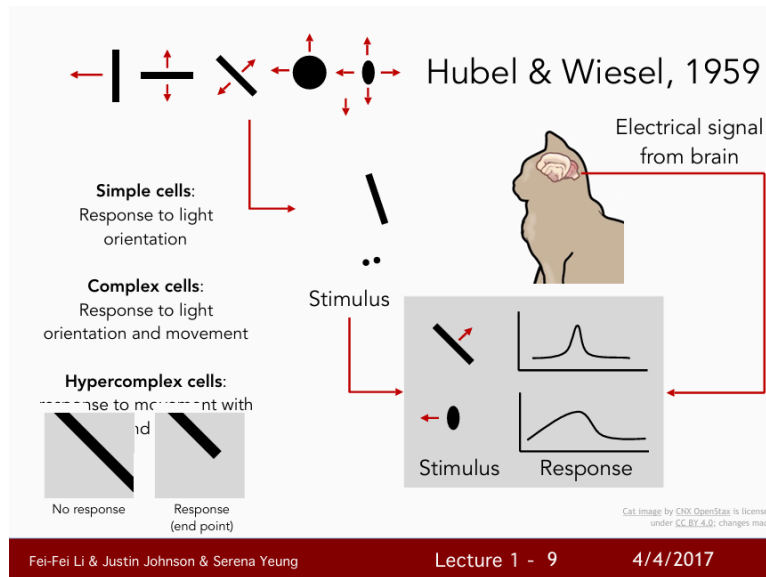


# lecture 1

## 1. A brief history of computer vision



5,60년대 전기생리학을 이용한 hubel과 wiesel의 이론:

그들이 궁금해 했던것: "What was the visual processing mechanism like in primates, in mammals"

그래서 고양이 뇌를 공부하기시작( 인간과 primary visual cortex area가 비슷한)

어떤 자극들이 고양이의 후두엽에서 뉴런을 만드는가

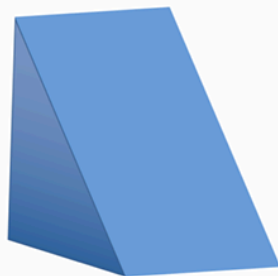
one of the most important cells, the simple cells, they respond to oriented edges when the move in certain directions

→ 뉴런들은 단순히 모양이나 선에 반응하는 것이 아니라 그 선이 특정 방향을 갖고 있고, 그 선이 특정 방향으로 움직일 때에만 반응하는 특성을 갖는다.

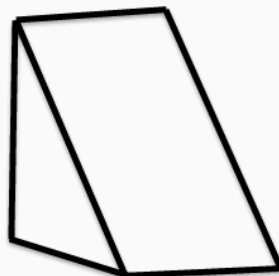
즉, 고양이의 1차 시각 피질 뉴런은 단순히 '선'을 보는게 아니라 그 선의 방향성과 움직임까지 구분하여 처리한다

# Block world

Larry Roberts, 1963



(a) Original picture

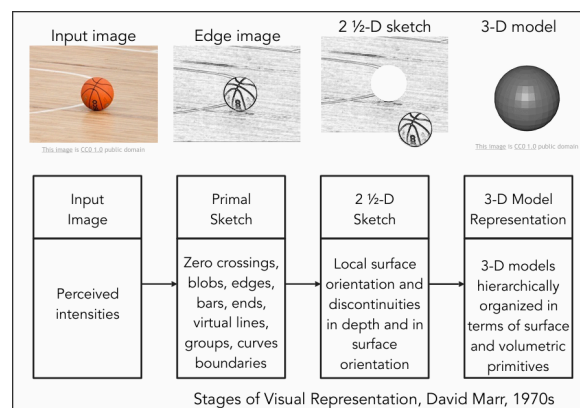
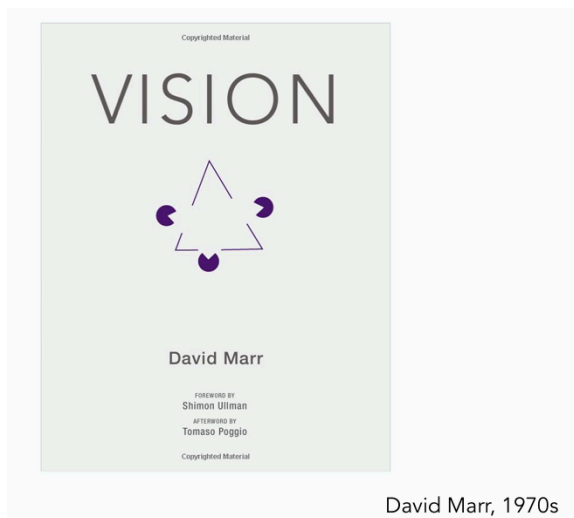


(b) Differentiated picture



(c) Feature points selected

목표는 shape를 인지하고, 어떤 shape인지 reconstruct 하는 것



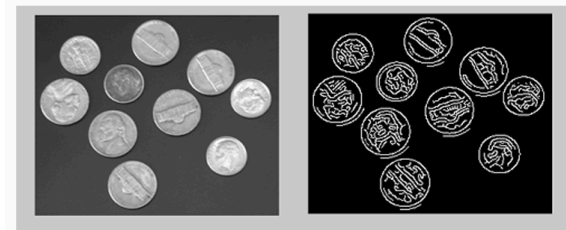
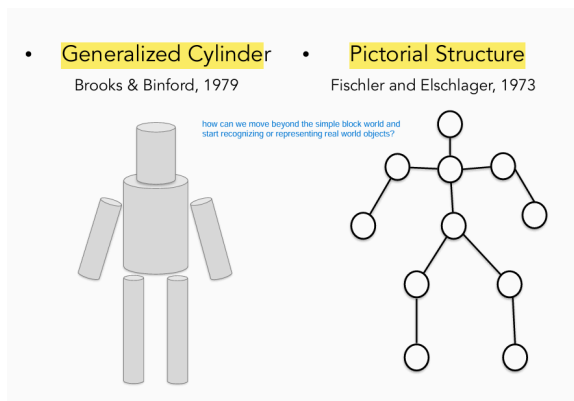
"primal sketch": mostly the edges, the bars, the ends, the virtual lines, the curves, the boundaries, 등등

"2 1/2-D sketch": 표면의 방향과 깊이의 불연속성을 표현하는 단계.

local surface orientation: 각 지점에서 어떤 방향을 향하는지 정보

discontinuities in depth and surface: 깊이나 표면의 급격한 변화 감지

"3-D Model": 계층적으로 구조화



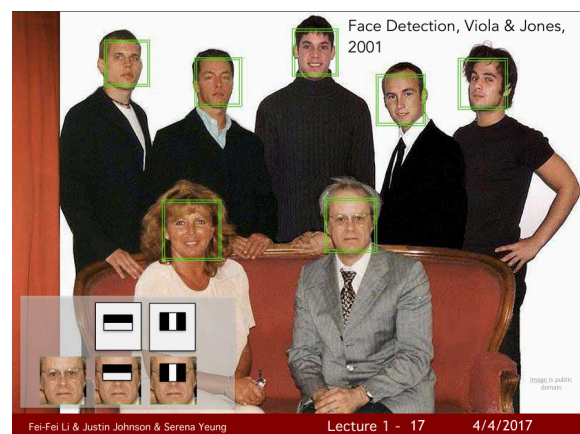
**Generalized Cylinder:** 복잡한 물체를 cylinder나 간단한 3차원 도형의 조합으로 표현하는 방법

**Pictorial Structure:** 실제 물체를 여러 개의 node와 각 부위를 연결하는 edge로 구성된 그래프로 표현.

둘 다 단순한 픽셀 집합이 아니라 구조적 관점에서 물체를 인식하려는 시도이다!

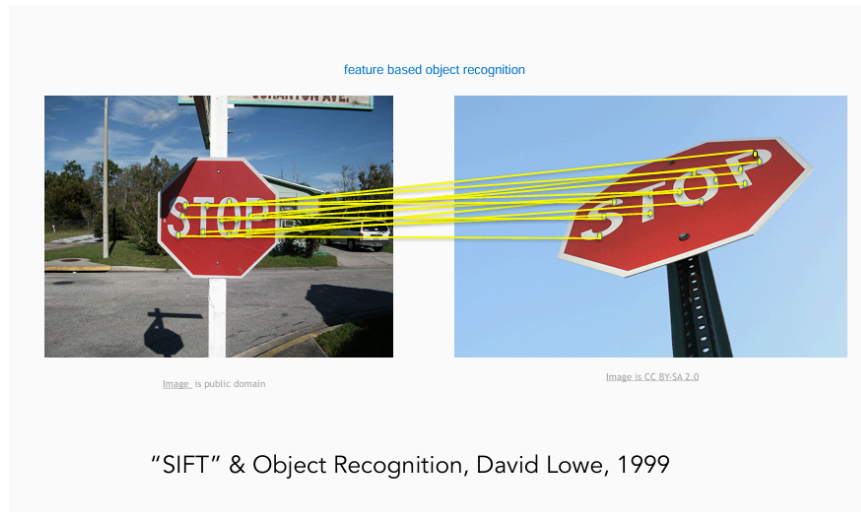
만약 object recognition이 너무 어렵다면, object segmentation부터 시도하는 것이 효과적일 수 있음(오른쪽그림). → 이미지의 각 픽셀을 의미있는 cluster로 그룹화해서, 이후에 각 영역이 무엇인지 분류하거나 인식하는 방법.

이렇게 단계적으로 접근하면 전체 인식 문제를 더 작고, 다루기 쉬운 문제로 나눌 수 있다.

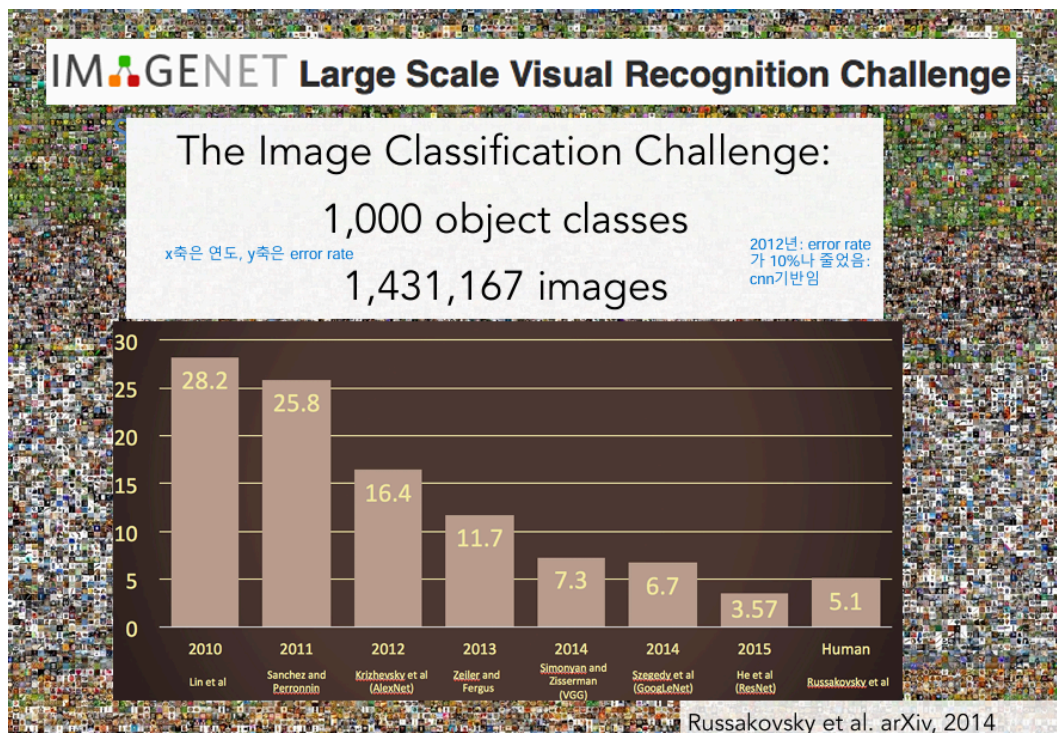


**Normalized Cut:** 이미지의 각 픽셀을 그래프로 보고, 비슷한 픽셀끼리 묶어서 이미지를 여러 부분으로 나눔→ 한 덩어리의 픽셀 집합이 어떤 의미있는 객체에 해당하도록 분리하는 것이 목적임.

그룹화된 픽셀이 인간인지 모를 수 있어도, 우리는 사람에 속한 모든 픽셀을 배경으로부터 추출할 수 있다!



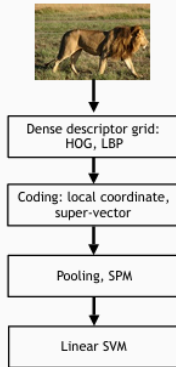
카메라의 각도, 조명, 거리 등이 달라지더라도, feature point들은 대체로 그 물체의 본질적인 정보를 담고 있기 때문에 잘 변하지 않음. 그래서 object recognition 문제는 전체 이미지를 그냥 픽셀 단위로 패턴 매핑하는 것 보다 이런 변하지 않는 핵심적인 feature point들만 잘 찾아서, 그 포인트끼리 매칭하는 방식이 더 효율적이다.



# 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

2012 이후로는 다 neural network를 했다

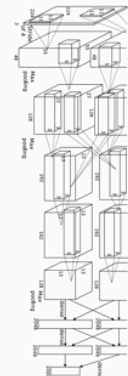


Figure copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

## Year 2014

GoogLeNet

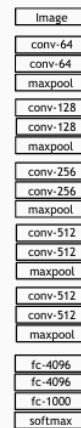
● Pooling  
● Convolution  
● Softmax  
● Other



[Szegedy arxiv 2014]

vgg는 19layer

VGG



[Simonyan arxiv 2014]

## Year 2015

MSRA

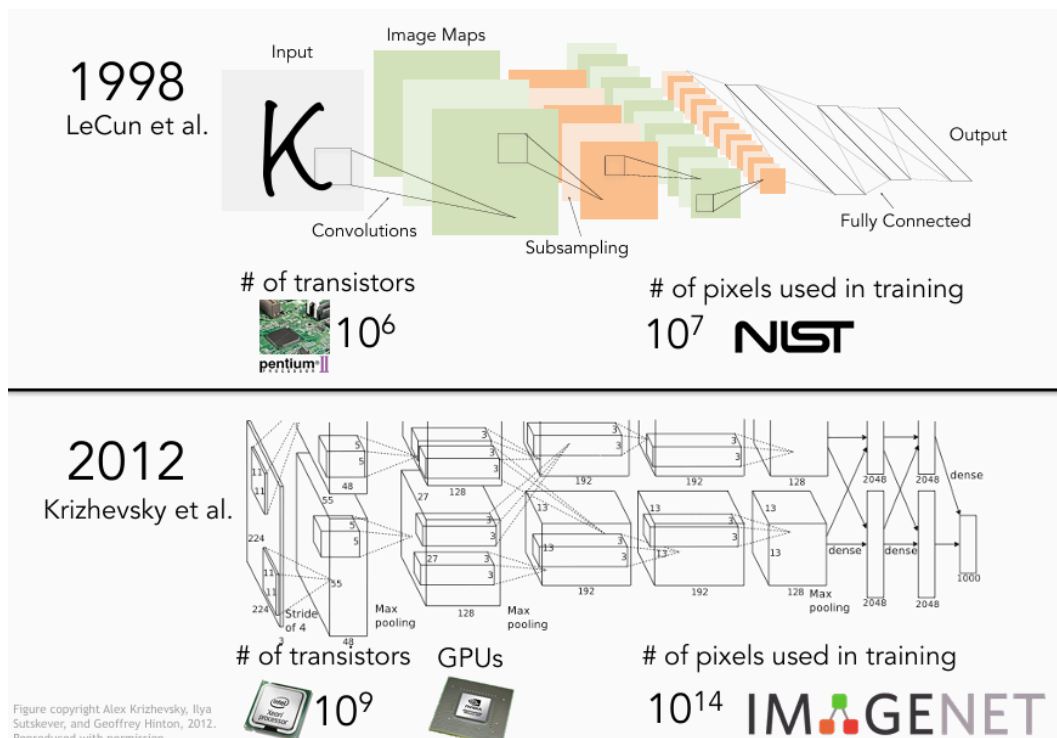
152 layer



[He ICCV 2015]

## IMAGE NET: The Image Classification Challenge

2012년 이후로는 다 neural network를 사용.



2012년에 우승한 Alexnet의 구조는 사실 1998년에도 있었는데, 오랜 시간이 흐른 후에야 각광받은 이유? → 이러한 문제는 사실 병렬 연산이 굉장히 중요한데, 시간이 지날수록 gpu가 발전하면서 더 복잡한 병렬 연산들이 가능해짐. (무어의 법칙)