

생물은 simple cells, complex cells, hypercomplex cells 와 electronical signal 을 이용해 사물을 봄.

BUT, 기계는 인간의 눈처럼 사물을 파악하기가 쉽지 않음.

컴퓨터 비전에 관한 논의

- MIT 의 "The Summer Vision Project"가 거의 최초의 컴퓨터비전에 관한 프로젝트
- David Marr 의 Vision - 3 Steps of visual representation
 - Primal Sketch: Zero crossings, blobs, edges, bars, ends, virtual line, groups, curves, boundaries
 - 2½-D Sketch: Local surface orientation and discontinuities in depth and in surface orientation
 - 3-D Model Representation: 3-D models hierarchically organized in terms of surface and volumetric primitives

이미지를 표현/인식하는 방식

- "How can we move beyond the simple block world and start recognizing real world object?"
 - 작은 블록들 등의 간단한 기하학적 요소들을 가지고 무언가를 나타낼 수 있음
- 기울어진 하나의 이미지와 그렇지 않은 이미지를 볼 때, 같은 부분을 이어 기울어진 이미지와 그렇지 않은 이미지가 같다는 것을 확인할 수도 있음 - SIFT
- Spatial Pyramid Matching

Visual recognition problems related to image classification

- object detection
- image captioning
- action classification

90 년대와 현재(강의 기준 2017 년)의 컴퓨터비전의 차이는 매우 큼.

원인: computation 의 차이, data 양의 차이

강의에서 들어준 예시를 보면, 어떤 한 이미지를 분석 할 때 해당 이미지에 나온 것들의 행동, 장소 등의 각종 특징들을 파악한다. 그리고 그를 이용해 이미지에 나온 인물들이 어떠한 상황에 처해있는지 등을 알 수 있다. 버락 오바마가 다른 한 인물을 대상으로 장난을 치는 이미지를 이해하기 위해서는 어떤 남자가 장난을 치고 있고, 그 남자가 그냥 남자가 아닌 버락 오바마임을 파악해야 한다. 그리고 주변인들이 웃고 있는 모습을 확인한 후 어떠한 재미난 상황이 있음을 파악하게 된다. 이런 예시를 보여주면서 시각 정보가 많은 정보를 담고 있고, 이런 것들을 분석하는 것에 대해 다루려고 함을 보여주는 것 같았다.