

컴퓨터 비전

What is Computer Vision

- Teach machines to do what we can do with vision
- 사람으로서 눈을 가지고 할 수 있는 많은 일들을 기계 혹은 컴퓨터에게 가르쳐주는 학문
- 사진을 지능적으로 해석할 수 있는 능력
- Visual Context(대뇌에서 시각영역 담당)을 인공적으로 만들기
- Inverse optics(이미지가 주어졌을 때 빛이 어디에 존재하고 환경이 어떻게 구성되는지 연구)

Human Vision

Can do amazing things like:

- ▶ Recognize people and objects
- ▶ Navigate through obstacles
- ▶ Understand mood in the scene
- ▶ Imagine stories from pictures

But still is not perfect

- ▶ Suffers from Illusions
- ▶ Ignores many details
- ▶ Ambiguous description of the world
- ▶ Does not care about accuracy of world
- ▶ Limited memory

Computer wins!

- 물체와 사람 인식
- 장애물 인식
- 주변 환경 보고 분위기 파악
- 사진 한 장을 통한 상상

- 착시현상(Illusions)
- 세부 details 기억 못함
- 주변 환경 내 방식으로 파악 (정확히 이해 못함)
- 기억력 제한

Computer Vision

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

해상도

600x400 : 600개 행 400개 열

픽셀(0~255 정수)

엄청나게 큰 숫자 array일 뿐임

Computer Vision Hierarchy

⊗ Low-level

- ▶ Image → Image
- ▶ e.g. image processing, edge-detection, optical flow computation

이미지가 주어졌을 때 이미지 반환

⊗ Mid-level

- ▶ Image → Features
- ▶ e.g. boundary detection, segmentation, sfm

이미지가 주어졌을 때 어떤 특징 정보를 나타내는 feature(특징점) 판단

⊗ High-level

- ▶ Image → Semantics
- ▶ e.g. object recognition, scene understanding

이미지로부터 어떤 의미를 알아내는 것

Low-level

⊗ Low-level: Image → Image



Deblurring



Edge detection



Super-resolution



Colorization

- 흐릿한 부분을 없애서 선명한 이미지로 바꾸는 것
- Edge(밝기가 급격히 변하는 부분) 검출
- 저해상도 이미지를 고해상도 이미지로 변환
- 흑백 이미지를 컬러 이미지로 변환

Mid-level

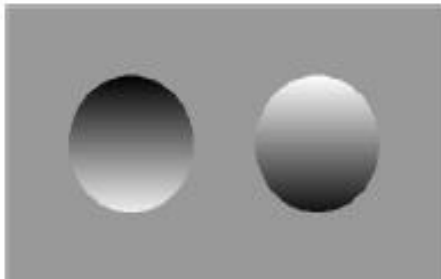
Mid-level: Image → Features



Boundary detection



Segmentation



Shape-from-shading



Alignment

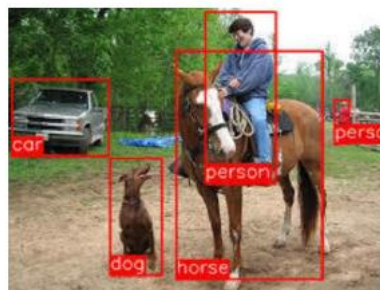
- Object의 경계 검출
- 이미지를 동일한 영역으로 분할
- 그림자로부터 모양을 알아내기
- 어떤 영역이 어떤 영역과 서로 매칭이 되는지 Align하는 과정

High-level

High-level: Image → Semantics



Image classification



Object detection



Image captioning



Pose detection

- 이미지가 주어지면 해석을 하는 것
- 우리가 관심있는 물체가 이미지 상 어디에 존재하는지 사각형으로 영역을 지정하고 어떤 object가 있는지 찾는 것
- 이미지 한 장을 하나의 문장으로 번역하는 것
- 주어진 사진에서 사람이 있는 경우 그 사람이 어떤 자세로 있는지 알아내는 것

Computer Vision vs Image Processing

⊗ Image processing

- ▶ Study image-to-image transformation
- ▶ Input and output are both images
- ▶ Image compression / restoration / enhancement

- 영상처리 : 이미지에서 이미지로 변환(low-level)
- 이미지 압축, 복원, 개선

⊗ Computer vision

- ▶ Actively use IP techniques
- ▶ The output is a description or an interpretation of image content
- ▶ High-level intelligence

- Computer vision : output이 이미지가 아니라 좀 더 그것보다 high-level의 지능적인 분석을 주로 다룸

Challenges 1: View Point Variation



3차원 물체가 있을 때 방향에 따라 사진의 결과가 달라진다.

Challenges 2: Illumination

☼ Shadow (그림자)



같은 사진인데 어느 부분은 밝고, 어느 부분은 어두움

☼ Reflection (반사)



창문에 비친 모습인데 기계는 창문을 보지 못하고 '여기 나무가 있다, 구름이 있다'고 판단함

Challenges 2: Illumination

⊗ Refraction (굴절)

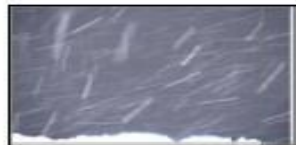


⊗ Semi-transparency (반투명)



사진이 찌그러져 보임
어디서부터 어디까지 인지(투명) 판단 어려움

⊗ Fog, rain, snowfalls...

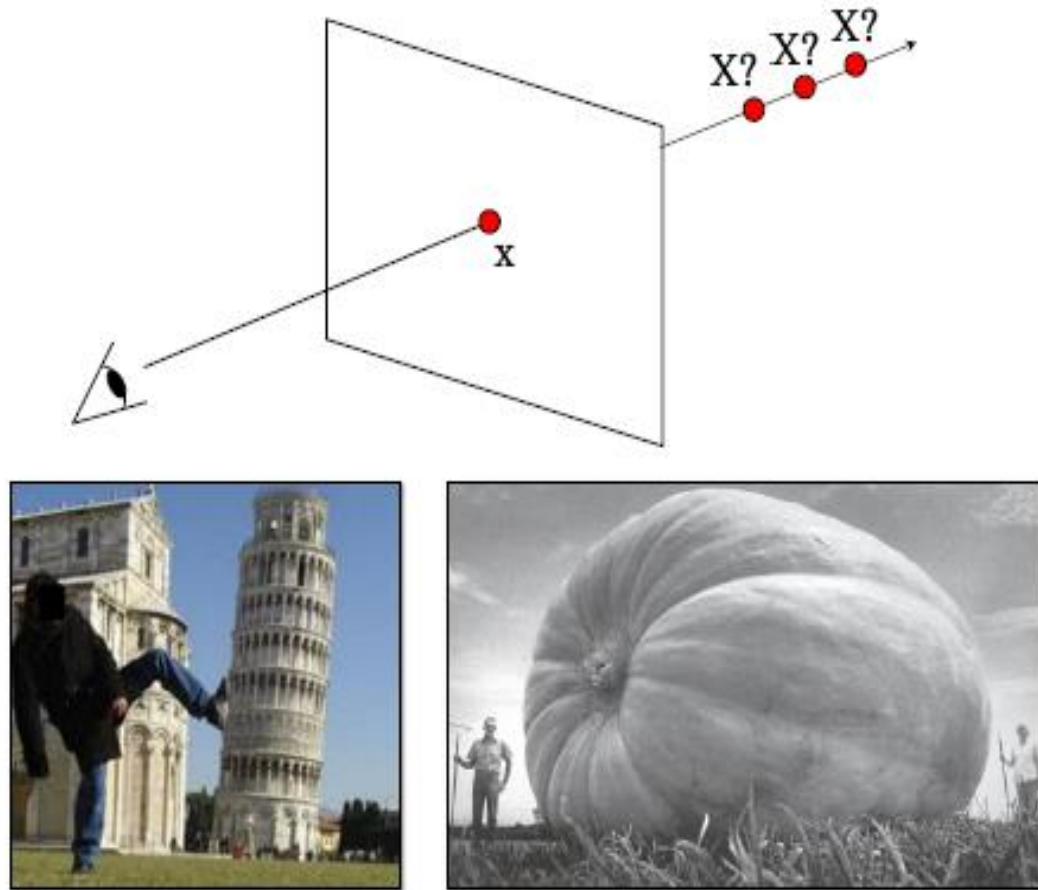


Challenges 3: Occlusion



말이 잘려 있는 것이 아님
이를 컴퓨터에게 알려줘야 함

Projection from 3D to 2D

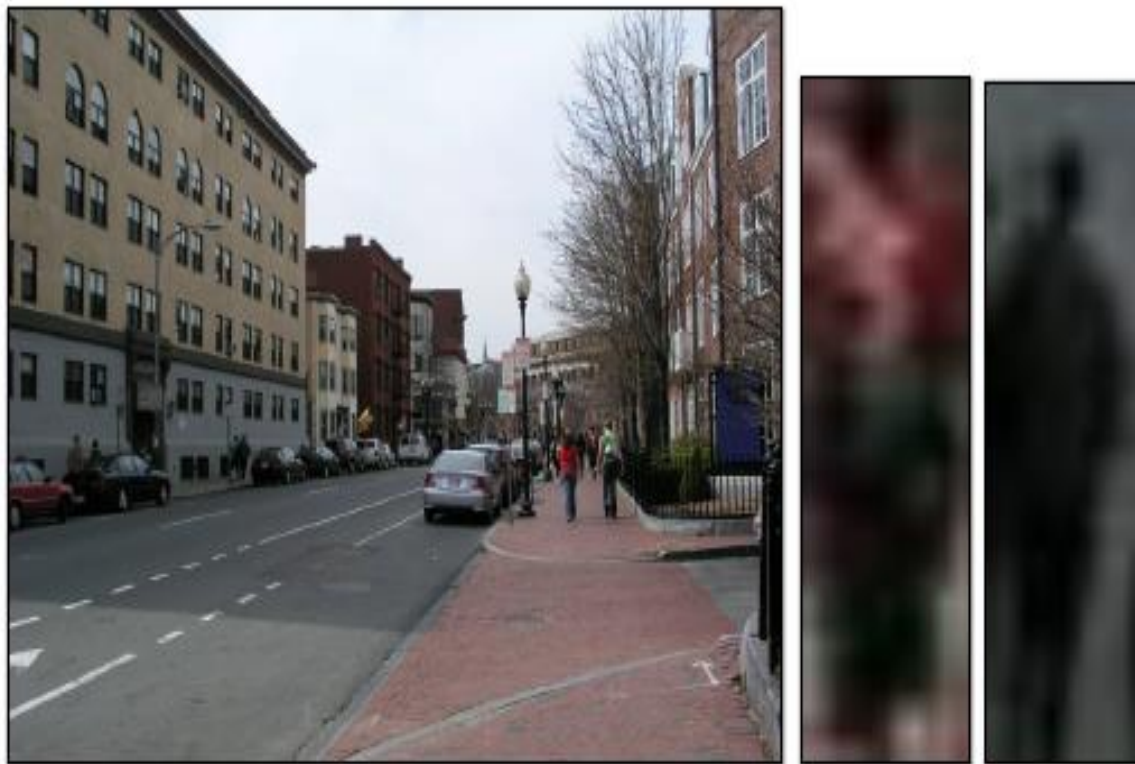


- 3차원을 사진으로 찍는 순간 2차원으로 변경된다.
- Occlusion, 가림현상 : 사람이 건물을 가림
- 멀리 있는 object는 작게, 가까이 있는 object는 크게 보임



- 사진에서 잘려 있는 부분을 보고
자동차/사람 인지를 정확히 판별
하는 것이 어려움

Challenges 4: Scale



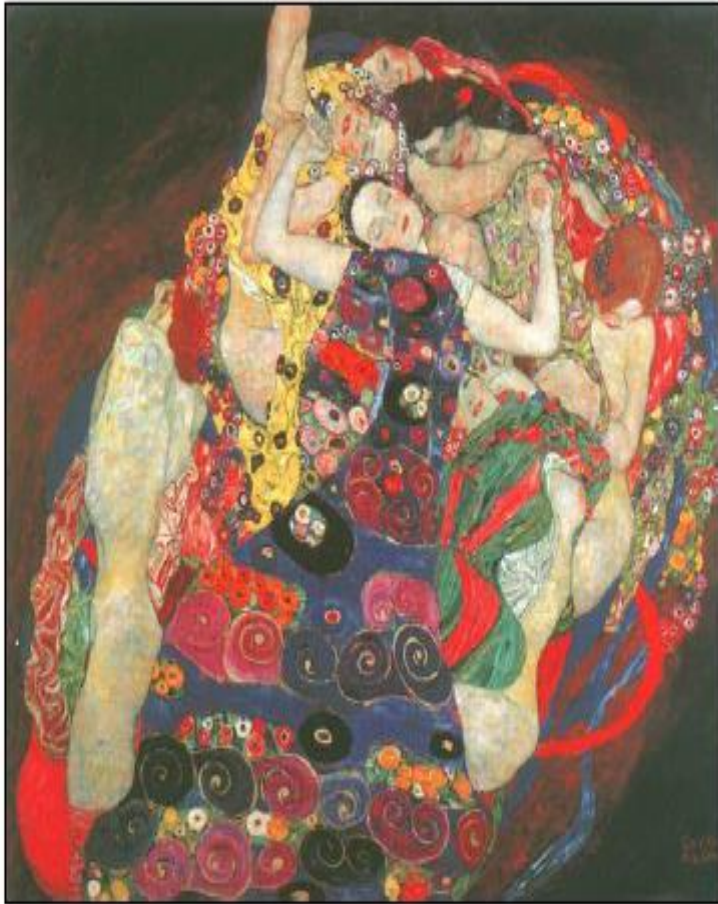
- 사진의 일부분을 확대해서 보면 사람인지 판별이 어려움
- Context(특정 영역만 보는 것이 아니라 주위 영역도 같이 봄)
- 도로, 건물, 자동차 등을 통해 사람임을 판별

Challenges 5: Deformation



- 물체는 하나의 형태만 가지고 있는 것이 아니라 여러 가지 형태를 가짐
- 움직임으로 인해 어렵게 deform이 될 수 있다.

Challenges 6: Background Clutter



Challenges 6: Background Clutter



- 배경때문에 사물을 정확히 판별하지 못함

Challenges 7: Object Intra-Class Variation



- 사진을 보고 겉모습을 통해 의자라는 것을 판별하는 것이 아니라, '앉을 수 있다!' 라는 개념을 통해 판별해야 함.