Lecture 9

Introduction of Computer Vision

Multimedia System

Spring 2020

컴퓨터 비전(computer vision,컴퓨터 視角)

- 시각의 중요성
  - 인간의 오감 중 시각으로 얻는 정보는 전체 감각기관으로 얻는 정보의 90%이상을 차지
  - 인간은 태어나서 보기 시작한 모든 정보로부터 스스로 학습

- 컴퓨터(로봇) 비전
  - 컴퓨터(로봇)에 인간의 시각 기능을 구현하고자 하는 노력에서 비롯된 학문 분야
  - 카메라로 획득한 영상 정보에서 유용한 정보를 추출하기 위한 기술

컴퓨터비전

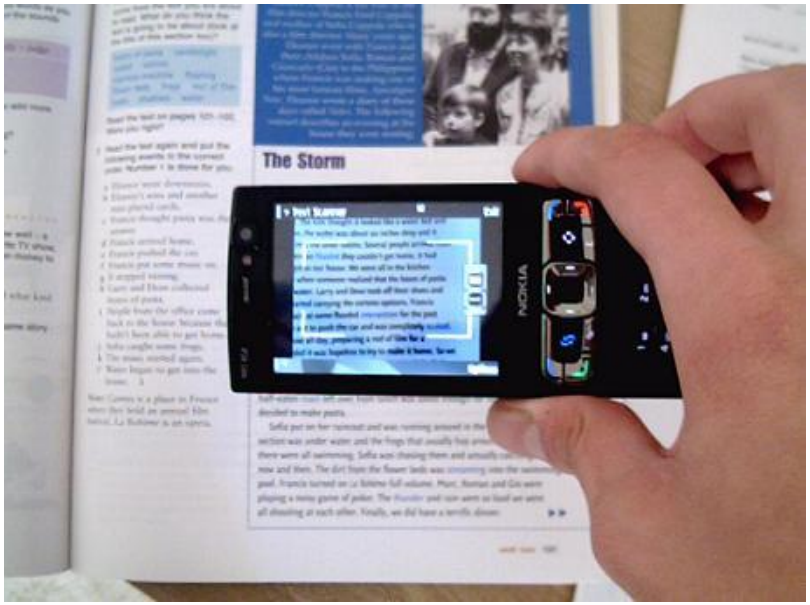
- 2차원 컴퓨터비전
  - 2차원 영상센서로부터 획득한 2차원 영상으로부터 다양한 응용분야에 적용할 수 있는 2차원 정보를 획득하기 위한 컴퓨터 기술
  - 문자인식, 얼굴인식, 물체추적 등
- 3차원 컴퓨터비전
  - 2차원 영상정보로부터 3차원 정보를 획득하기 위한 컴퓨터 기술
  - 3차원 스캐닝, 3차원 TV, 3차원 로봇비전 등

Applications

- Visual Inspection
- Robotics
- Intelligent Image Tools
- Image Compression (MPEG 1/2/4/7)
- Document Analysis (OCR)
- Image Libraries (DL)
- Virtual Environment Construction
- Environment
- Media and Entertainment
- Medicine
- Astronomy
- Law Enforcement
  - surveillance, security
- Traffic and Transportation
- Tele-Conferencing and e-Learning

Recognition (1)

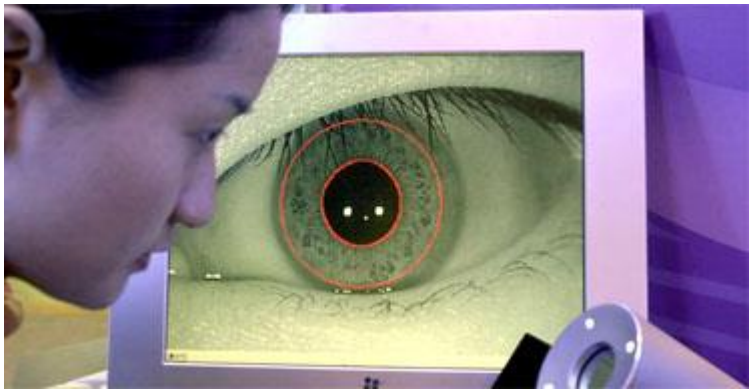
- Character recognition
- Car plate recognition

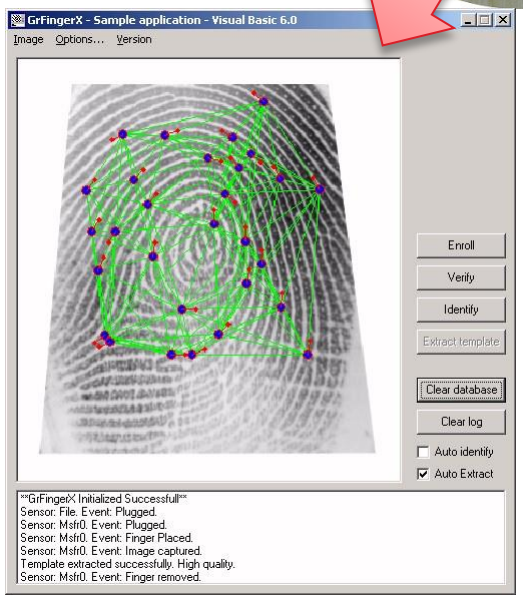


- “패턴인식” 은 확률, 통계와 밀접한 관련이 있음
- 일반적으로 별도의 과목으로 강의

Recognition (2)

- Bio recognition (fingerprint, iris, palm vein ..)





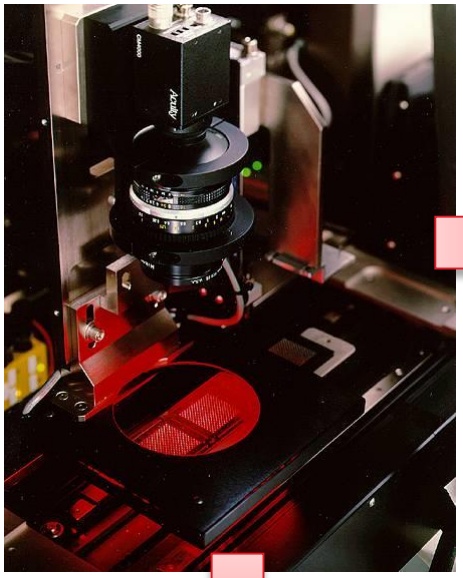
Face detection & recognition

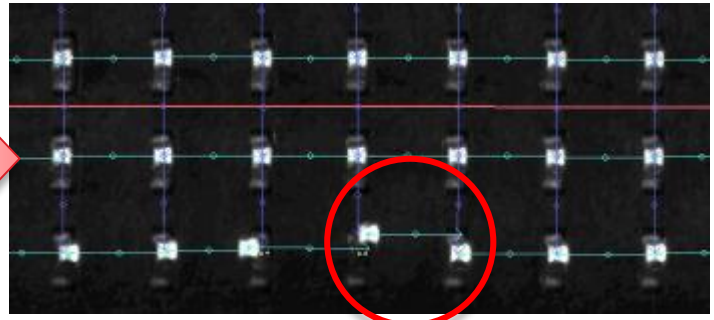
- Many application areas
  - Security
  - Smart device
  - Smile interface
  - Human robot interface

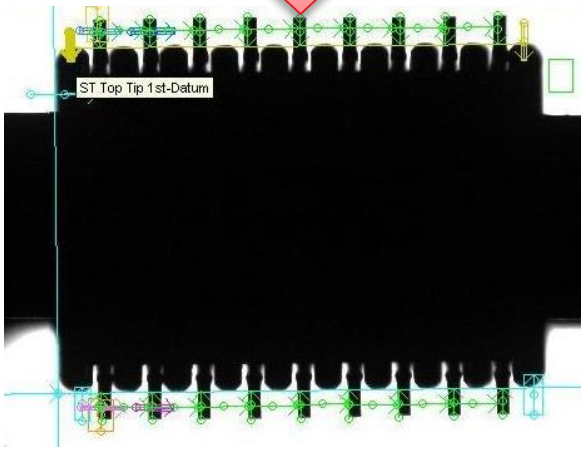
https://www.youtube.com/watch?v=sPmXZGSsiYM

Inspection

- Defect inspection





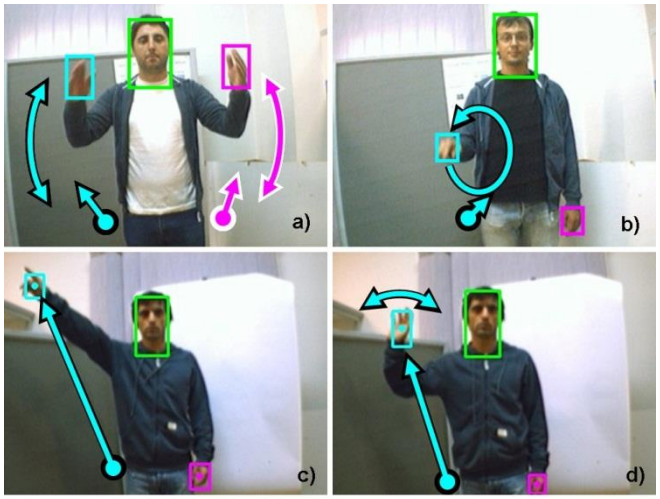


Semiconductor chip inspection

Bottle inspection

Gesture Recognition

- Human computer(robot) interaction
- Smart wall
- Smart mobile device (phone, pad..)
- Game

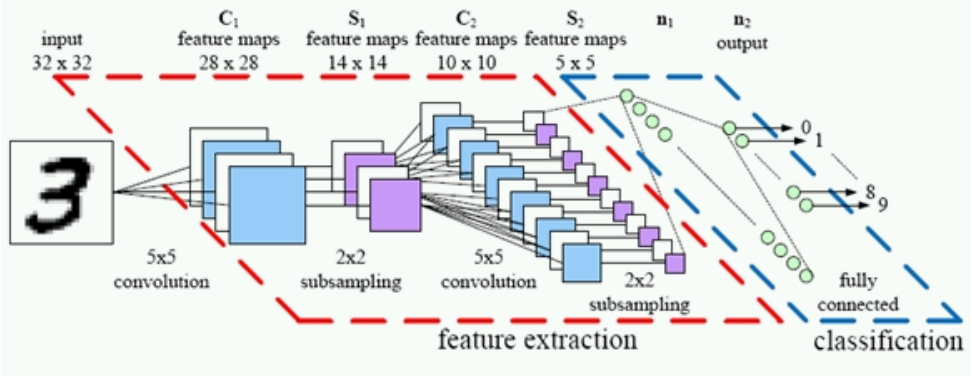


University of Tsukuba, Japan

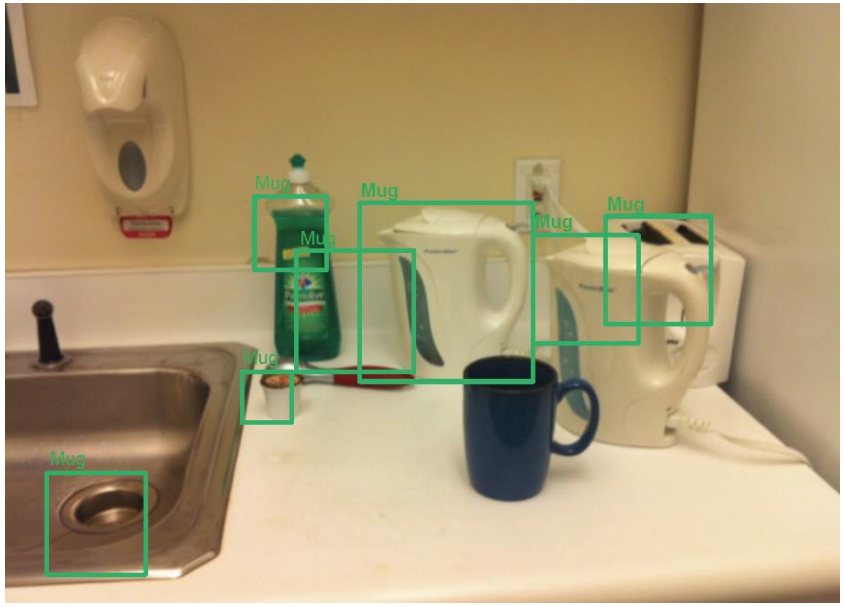
Interactive 3D TV and you're watching a Haynes manual?

Deep learning

- Neural network with specialized connectivity structure
- Stack multiple stages of feature extractors
- Higher stages compute more global, more invariant features
- Classification layer at the end

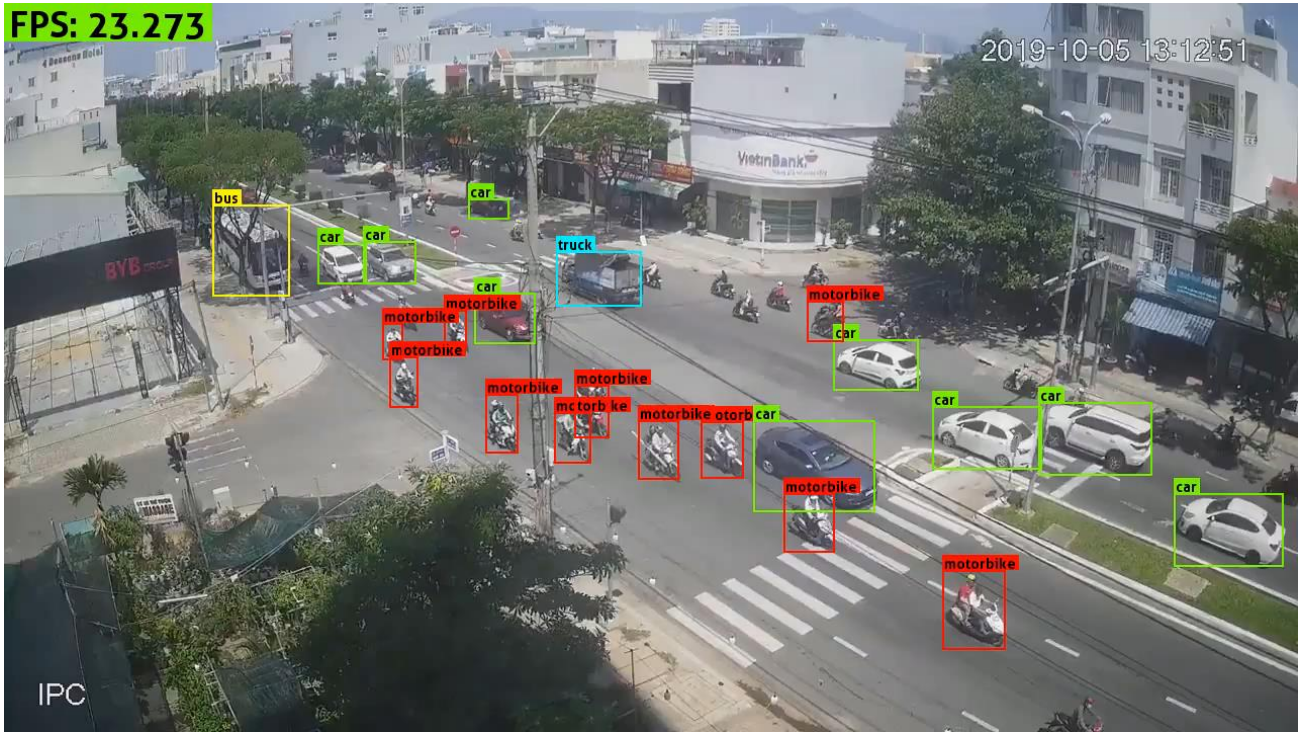


Example of Convolutional Neural Network



Detection and Tracking

- Detection and tracking of objects in a video sequence is an important technique in computer vision
- Many application areas
  - Traffic monitoring
  - Security
  - Entertainment



https://www.youtube.com/watch?v=24ZjHGx4rdQ

Detection and Tracking

- National defense is an important application area of the vision tracking technique
- UAV is a good application target of object tracking

UAV

Quadcopter

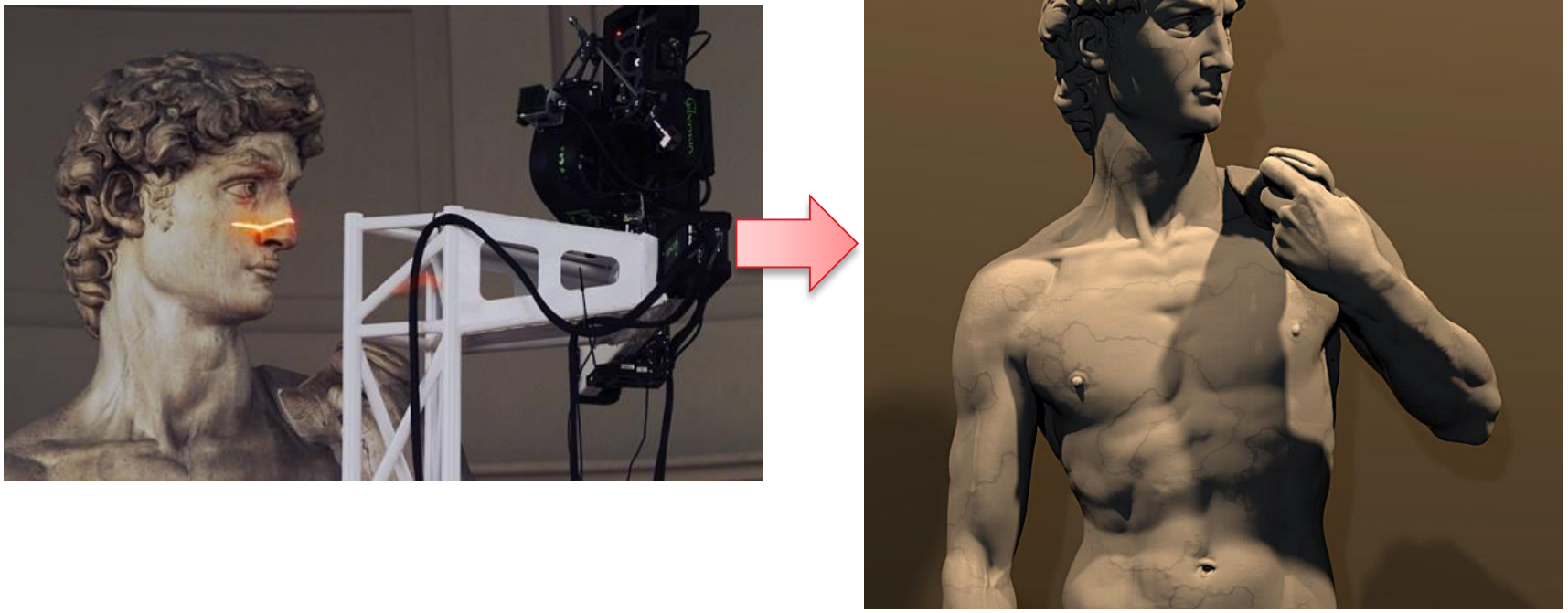
- Quadcopter examples
- http://www.youtube.com/watch?v=vHpw8zc7-JQ&list=UUrZt485fBMeNtCDlgEZAHVA&index=1
- http://www.youtube.com/watch?v=IMSozUpFFkU

3D Modeling

- 3D model generation one or more camera views.
- 3D models are saved in 3D graphic format such as PLY, OBJ, 3DM, OFF.

3D Modeling: Offline

- Accurate 3D model reconstruction
- Application: Reverse engineering, cultural heritage, city modeling…



The Michelangelo project by Stanford U.

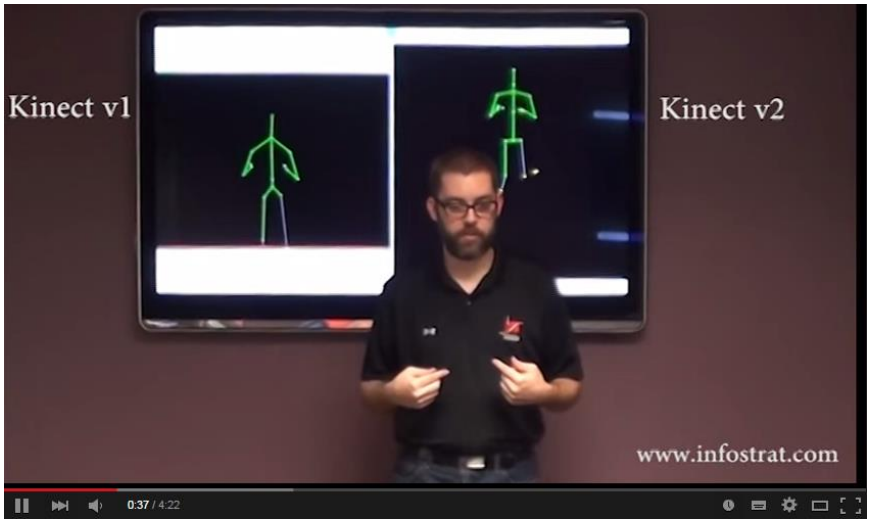
3D Modeling: Online

- Real-time 3D modeling
- Application: Game, entertainment, security…

3D modeling : Kinect

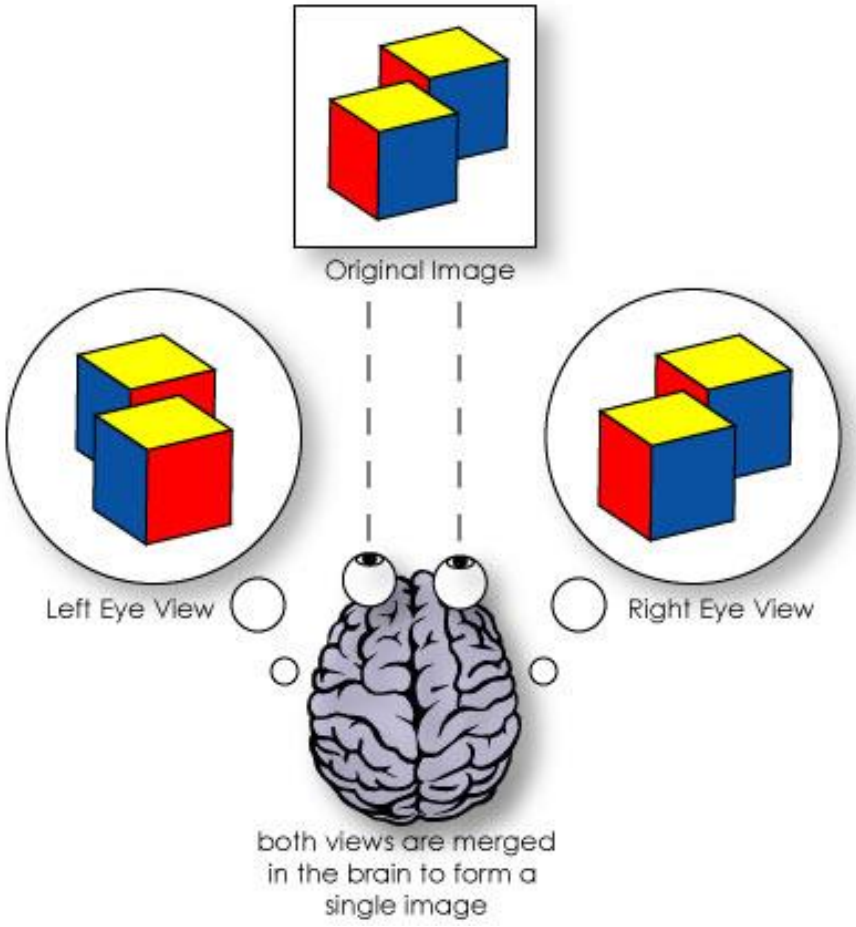
- Microsoft의 Xbox용 motion recognition device

KINECT
for XBOX 360



Stereo Vision

- Human beings have two eyes positioned side by side.
- Each eye has a view of the same area from a slightly different angle.
- The brain takes the information from each eye and unites them into one picture, interpreting the slight differences between each view as **depth**.



Stereo Vision

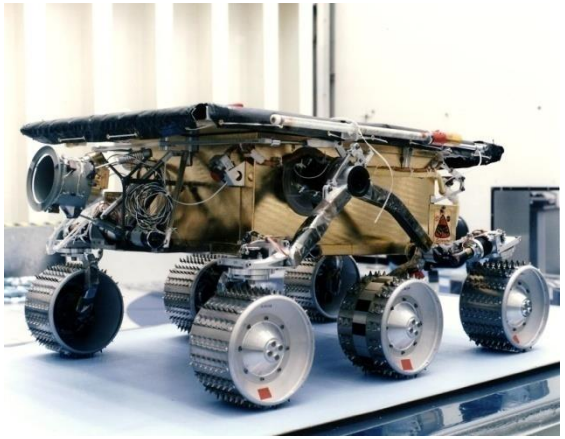
- Stereo vision is a part of the field of computer vision
- Stereo vision technique compares two or more images to find the shift between the correspondences in the images. The shifted amount is called the disparity.

Left image

Right image

Disparity image

Robot Vision



Path finder

Honda Asimo

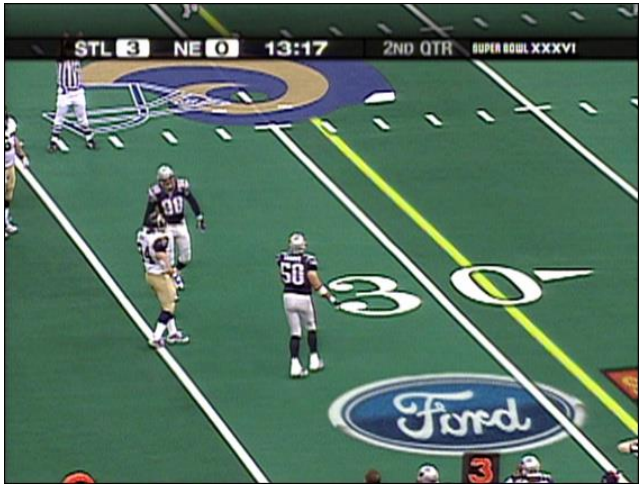
KAIST Hubo

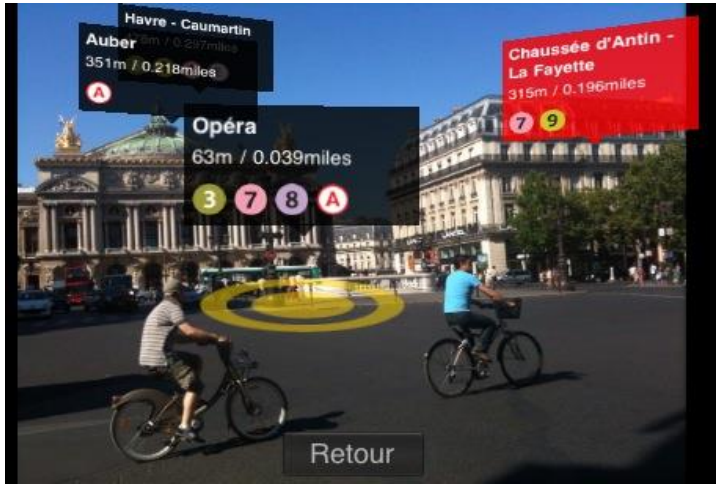
HCI, HRI

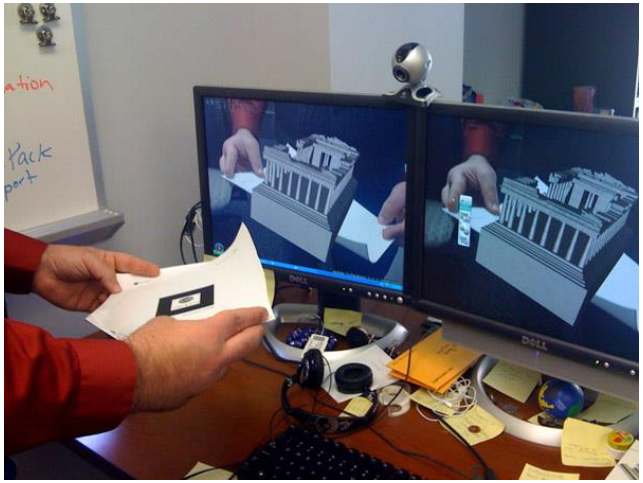
Digital Special Effects

https://www.youtube.com/watch?v=NKZUo0NaKT4

Augmented Reality

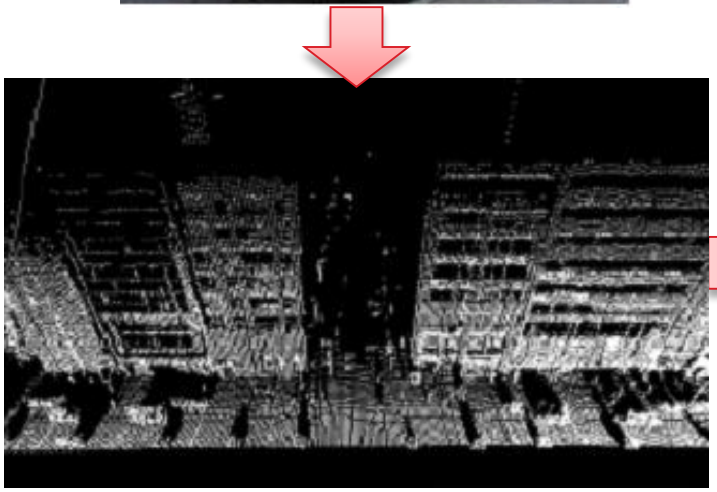




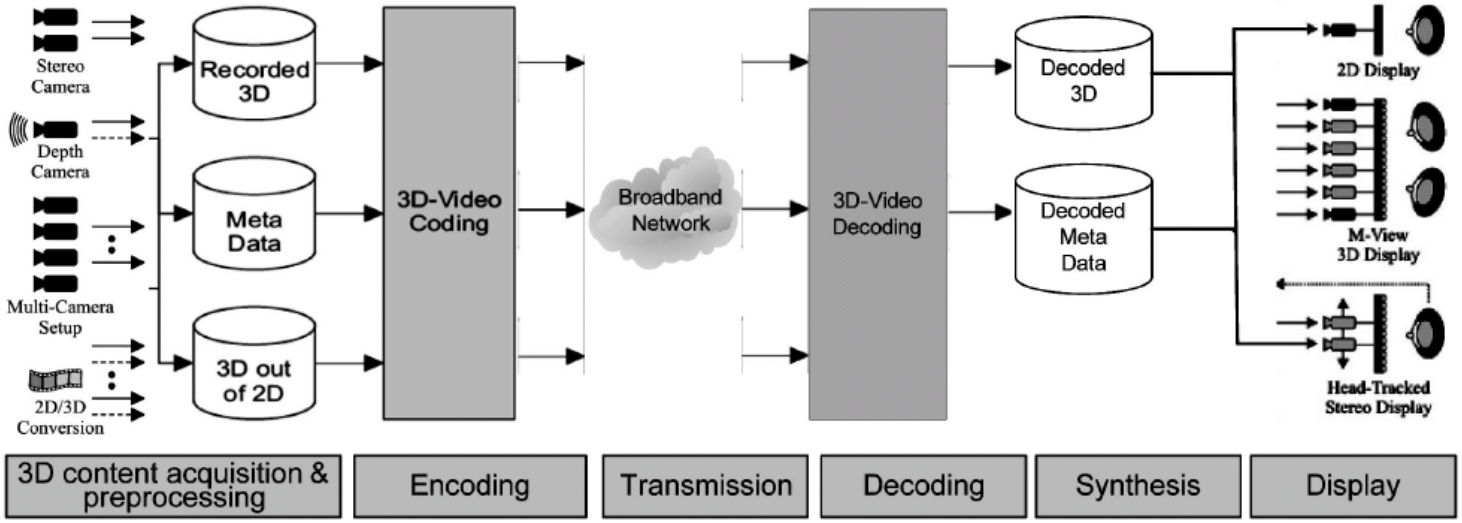


City modeling

- 3차원 센서를 사용한 도시 규모의 모델링



3DTV

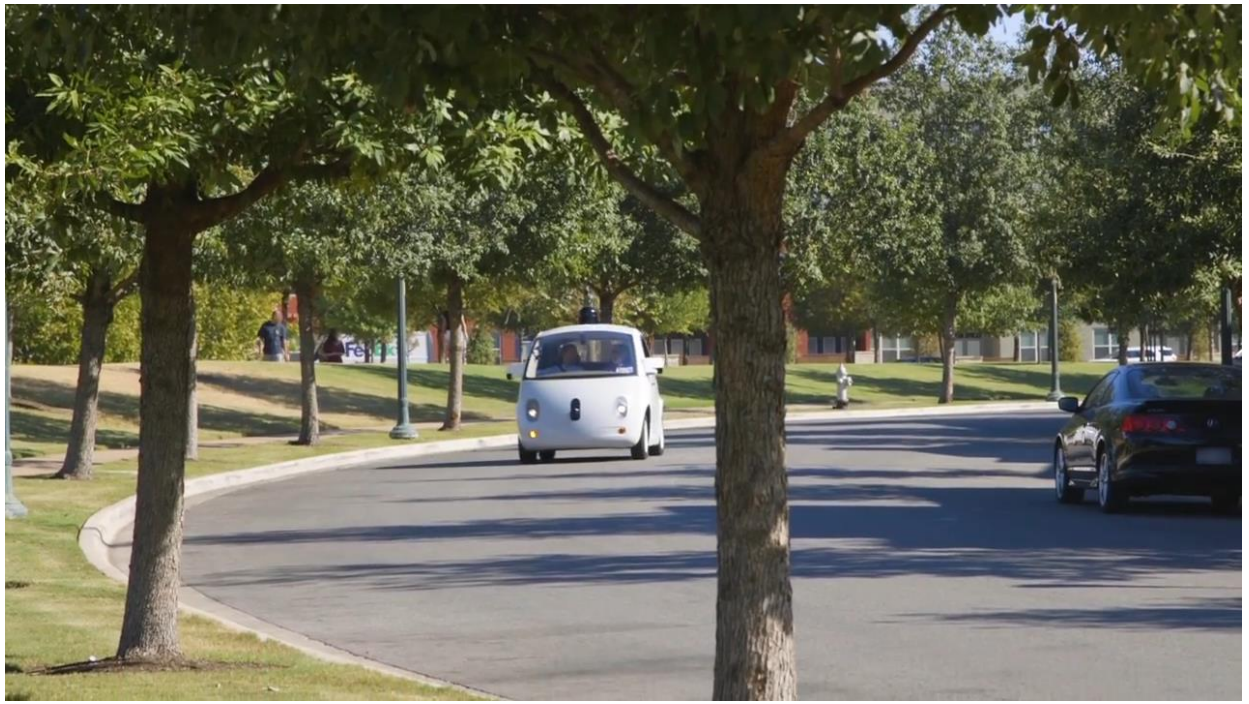


Contents Acq. and Proc. | Coding & Transmission | 3D Display

Vehicle Vision

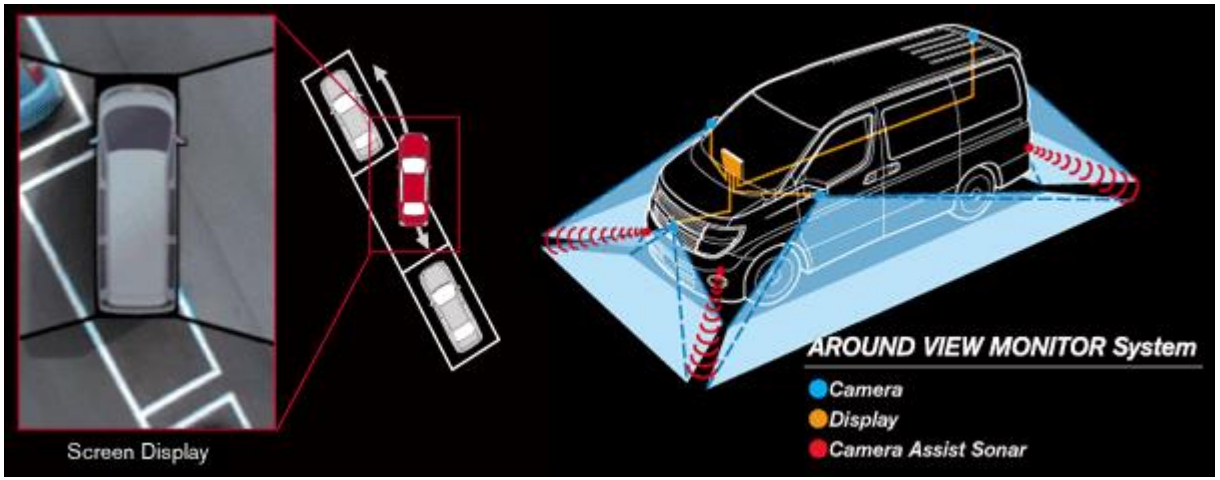
- Vehicle vision is a hot research topic in computer vision
- Vehicle in the future will be equipped with several vision, IR and Lidar sensors
- Google car

https://www.youtube.com/watch?v=X_d3MCkIvg8



Vehicle Vision

- Around view (Bird-eye view) monitoring
  - Transformation of different camera views to a common (top) view
  - Assists drivers to park more easily by better understanding the vehicle's surroundings
  - Example
    - http://www.youtube.com/watch?v=ScyMY7eqzXs

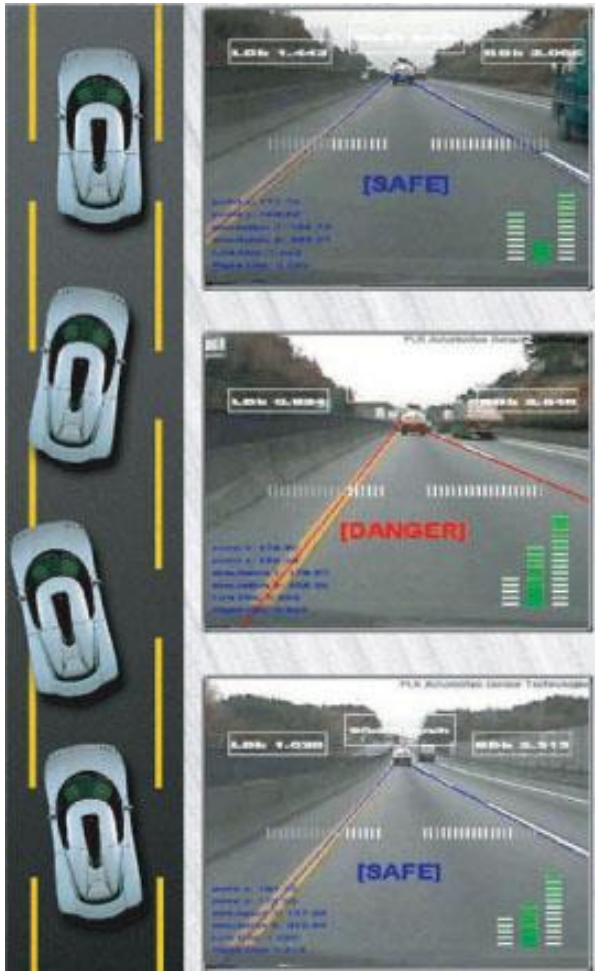


From Nissan homepage

Vehicle Vision

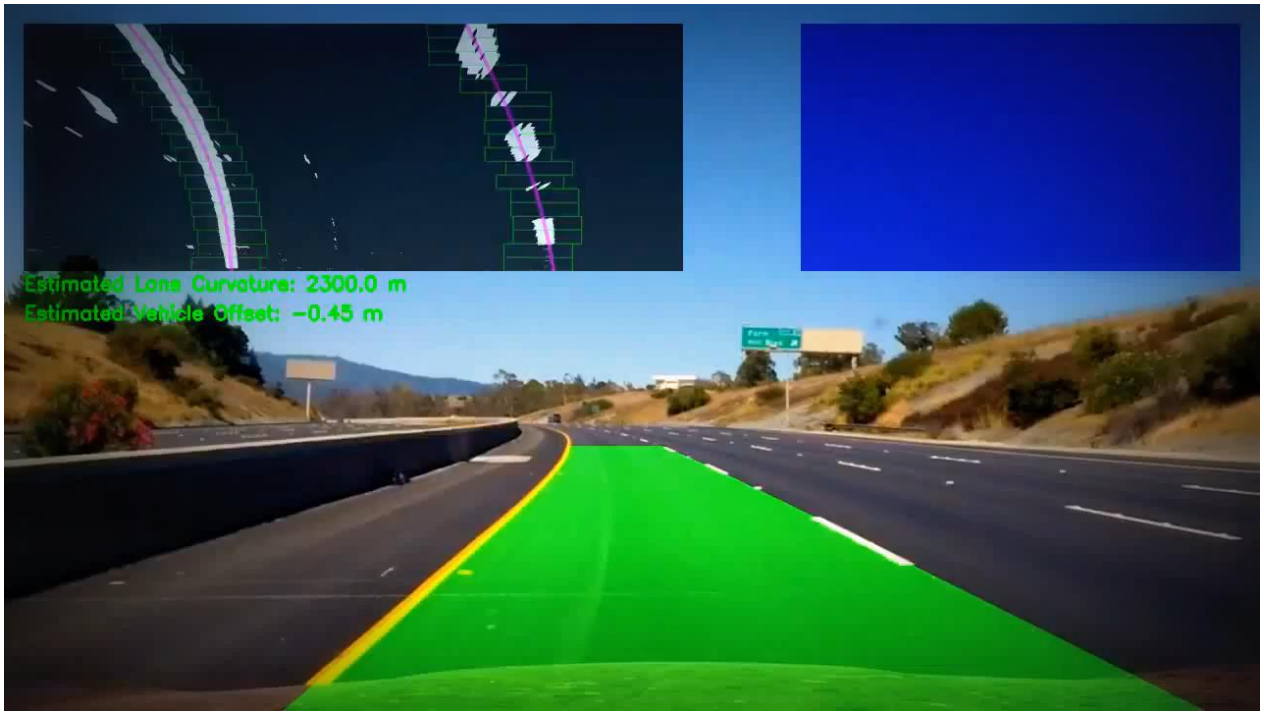
- LDWS (Lane Departure Warning system)
- A mechanism designed to warn a driver when the vehicle begins to move out of its lane

- Example
  - http://www.youtube.com/watch?v=PvnGG673ZPo



Vehicle vision

- On-road vehicle detection
- Alert a driver about driving environments and possible collision with other vehicles
- Robust and reliable vehicle detection is needed



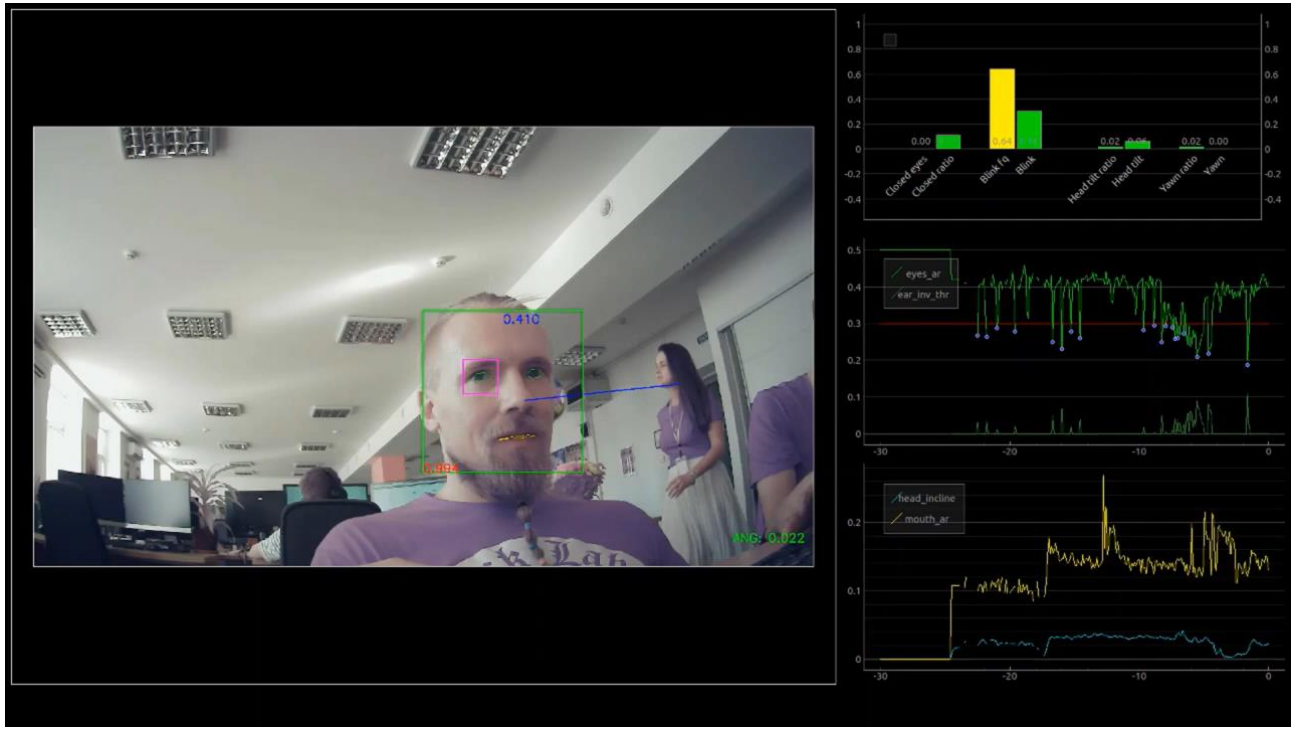
Vehicle vision

- Pedestrian detection
  - Pedestrian detection is vert important for safety issue
  - HOG (Histogram of Gradient) feature
  - Deep learning

https://www.youtube.com/watch?v=0hWW6FVcFAo

Vehicle Vision

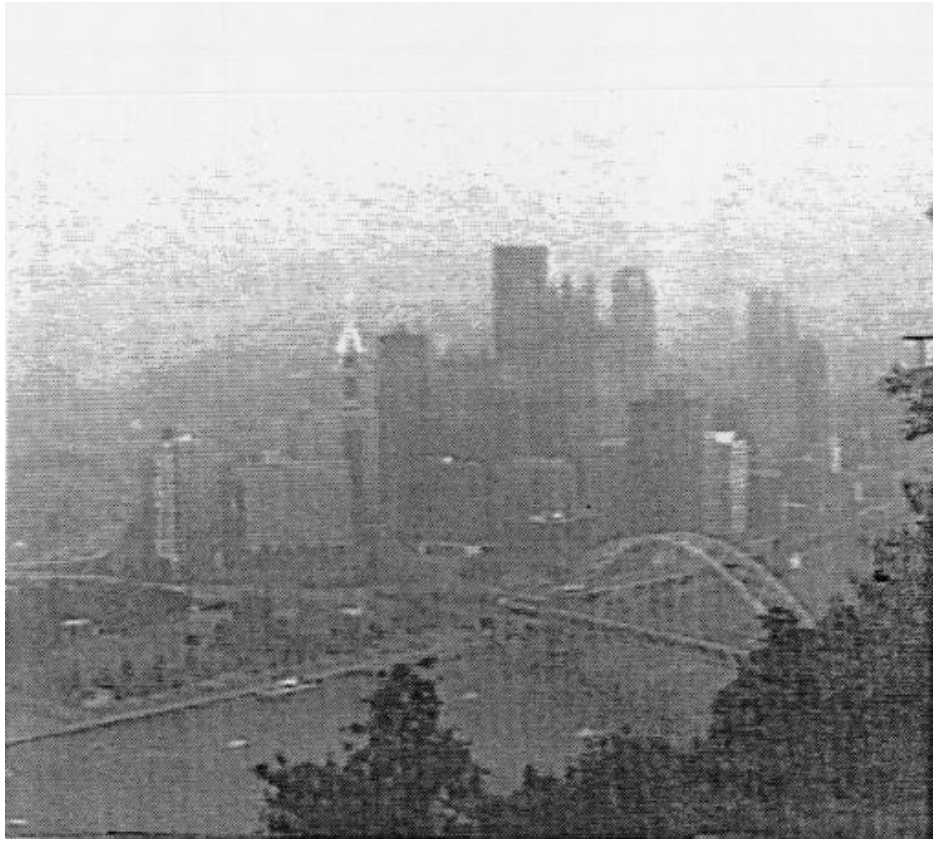
- Driver Drowsiness Monitoring
- Driver drowsiness is responsible for about 30% of severe traffic accidents
- Based on face/eye detection and eye opening monitoring



https://www.youtube.com/watch?v=4QCtDyUK2RY

Why computer vision ?

Image



Foggy golden triangle in Pittsburgh

But …

123	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
143	157	123	255	111	116	114	112	99	128	131	125	137	137	129	116
137	116	198	122	152	127	203	117	155	161	167	149	143	159	116	93
143	157	123	255	111	116	114	112	99	128	131	125	137	137	129	116
149	103	111	238	172	154	110	108	134	162	129	119	137	141	109	125
123	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
165	133	176	98	128	111	173	201	152	172	165	97	113	125	122	129
172	103	111	238	172	154	110	108	134	162	129	119	137	141	109	125
134	175	132	255	111	161	114	112	99	128	131	125	137	137	129	116
171	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
152	172	165	97	113	125	122	129	165	133	176	98	128	111	173	201
117	155	161	167	149	143	159	116	93	137	116	198	122	152	127	203
116	114	118	116	154	201	134	133	99	123	123	131	171	121	119	127
110	108	134	162	129	119	137	141	109	125	149	103	111	238	172	154
119	127	116	114	118	116	123	121	154	201	134	133	99	123	123	131
159	116	93	137	116	198	122	152	127	203	117	155	161	167	149	143
123	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
143	157	123	255	111	116	114	112	99	128	131	125	137	137	129	116
137	116	198	122	152	127	203	117	155	161	167	149	143	159	116	93
143	157	123	255	111	116	114	112	99	128	131	125	137	137	129	116
149	103	111	238	172	154	110	108	134	162	129	119	137	141	109	125
123	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
165	133	176	98	128	111	173	201	152	172	165	97	113	125	122	129
172	103	111	238	172	154	110	108	134	162	129	119	137	141	109	125
134	175	132	255	111	161	114	112	99	128	131	125	137	137	129	116
171	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
152	172	165	97	113	125	122	129	165	133	176	98	128	111	173	201
117	155	161	167	149	143	159	116	93	137	116	198	122	152	127	203
116	114	118	116	154	201	134	133	99	123	123	131	171	121	119	127
110	108	134	162	129	119	137	141	109	125	149	103	111	238	172	154
119	127	116	114	118	116	123	121	154	201	134	133	99	123	123	131
159	116	93	137	116	198	122	152	127	203	117	155	161	167	149	143

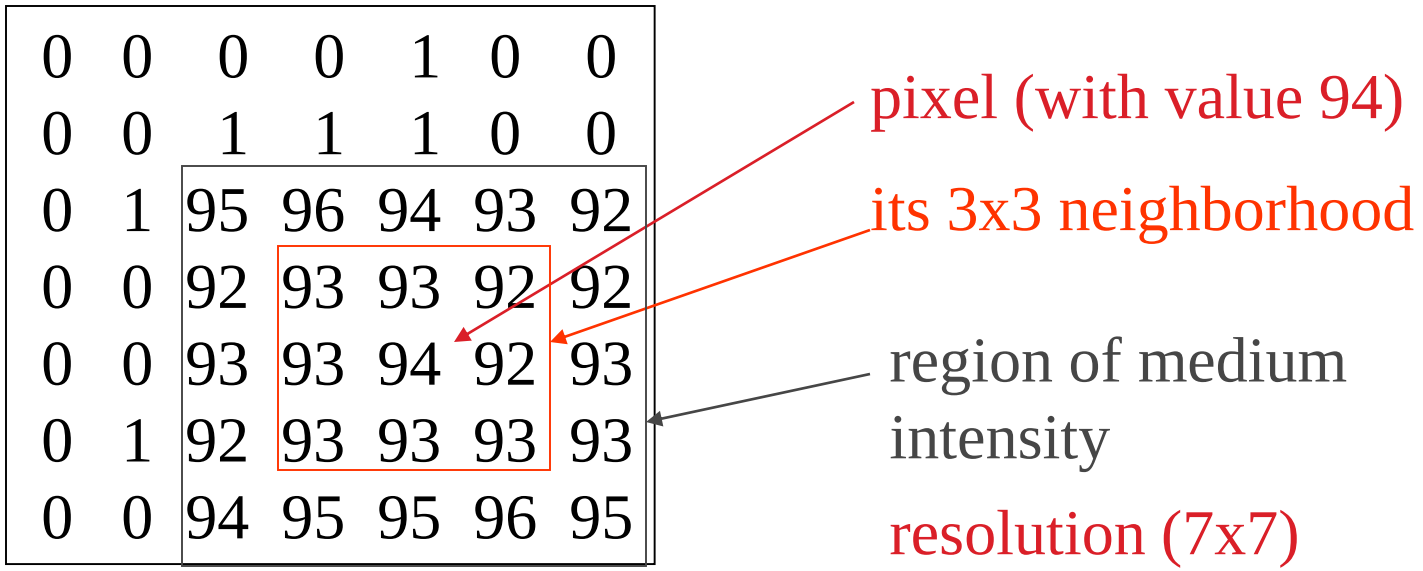
A lot of data

- Landsat image
  - 1scene: 3300 x 2300 x 4 = 30000000 bytes
  - 200 scenes/ day
- Color TV image
  - 512 x 512 x 3 x 30 = 25000000 bytes/sec

Why difficult ?

- A lot of data
- Ambiguity
  - Projection of a 3D world to a 2D image
- Many factors to influence the image
  - Illumination condition
  - Object shape
  - Camera characteristics

Digital Image Terminology:



- binary image
- gray-scale (or gray-tone) image
- color image
- multi-spectral image
- range image
- labeled image

Three Stages of Computer Vision

- low-level

  image ⟶ image

- mid-level

  image ⟶ features

- high-level

  features ⟶ analysis

Low-Level

sharpening

blurring

Low-Level

original image

Canny

edge image

Mid-Level

edge image

ORT

data structure

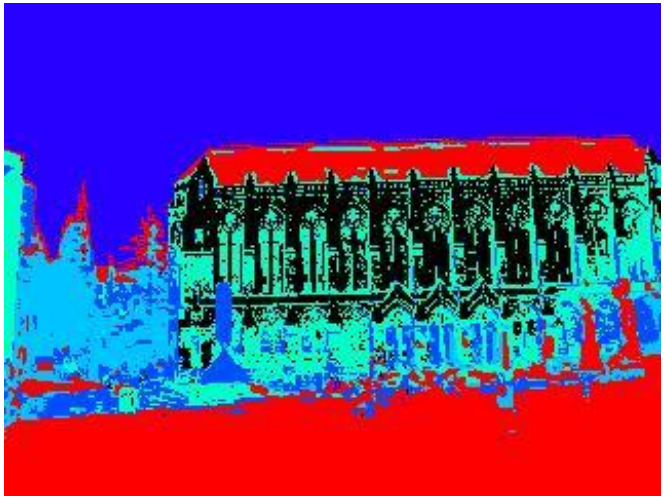
circular arcs and line segments

Mid-Level

original color image

K-means clustering

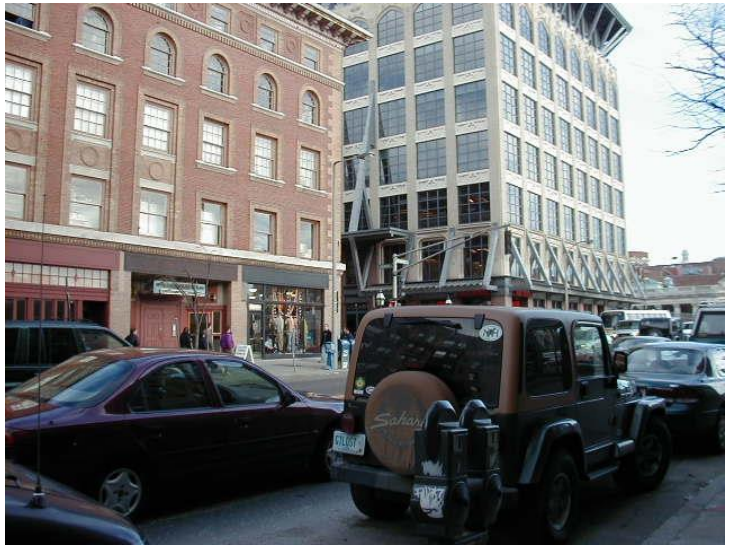
(followed by connected component analysis)



regions of homogeneous color

data structure

Low- to High-Level



low-level

edge image

mid-level

consistent line clusters

high-level