Welcome to Computer Vision

Computer Vision

Introduction

Dr. Muhammad Tahir

DEPARTMENT OF COMPUTER SCIENCE,

FAST-NUCES, Peshawar

Course Details

LECTURES: Tuesday & Wednesday

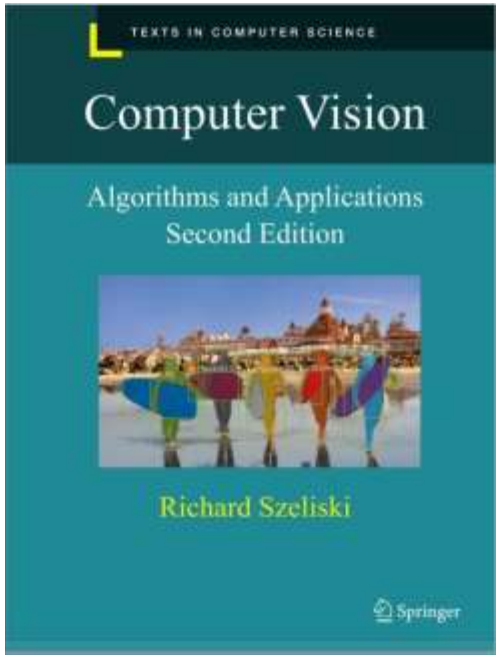
TIMINGS:

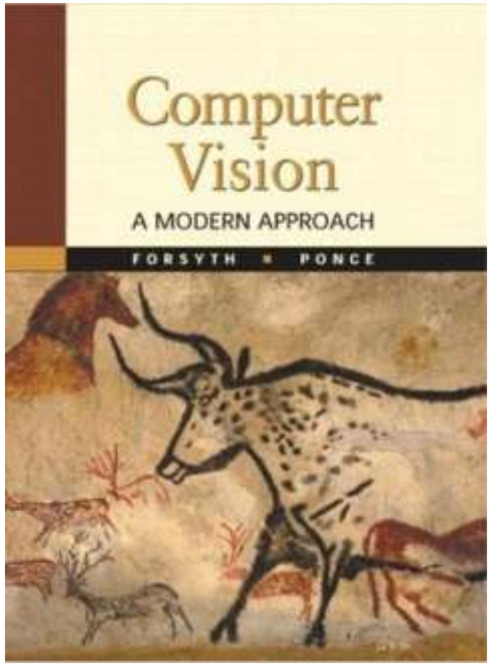
8:00 am – 9:30 am

MY OFFICE:

OFFICE HOURS:

EMAIL: m.tahir@nu.edu.pk





References

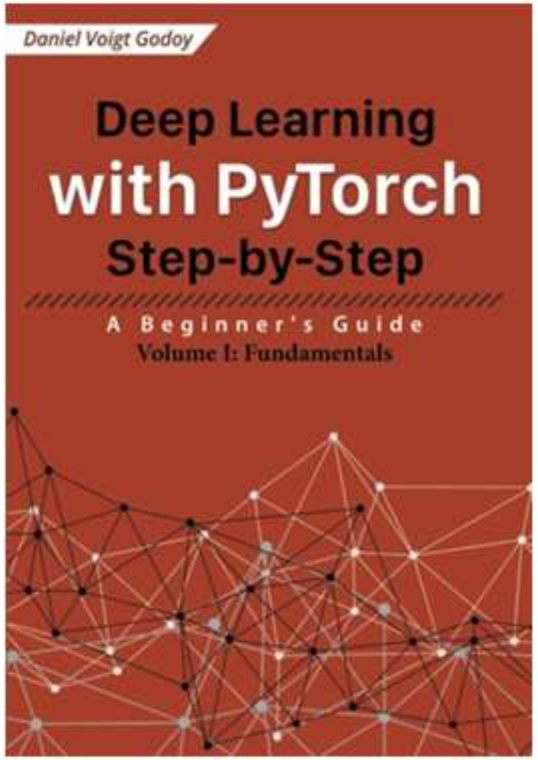
The material in these slides are based on:

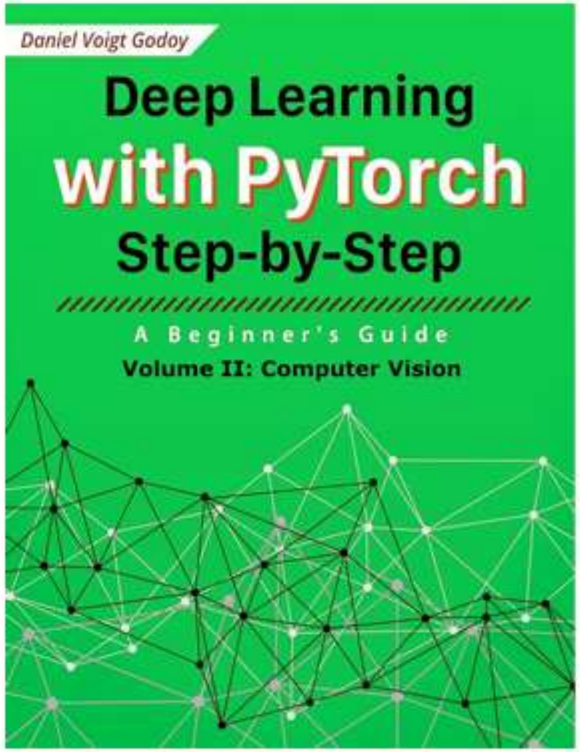
1. Rick Szeliski's book: Computer Vision: Algorithms and Applications

2. Forsythe and Ponce: Computer Vision: A Modern Approach

Recommended Books

Deep Learning with PyTorch Step-by-Step by Daniel Voigt Godoy





Course Administration

- Attendance Policy
  - ✓ FAST attendance Policy?
  - ✓ Students not making up to the 80% attendance will not be allowed to sit in the final exam
  - ✓ Latecomers should be careful
- Assignment / Quiz Policy:
  - ✓ Assignments will be hand-written EXCEPT in the case of coding task
  - ✓ Deadlines must be followed.
  - ✓ There will be no makeup of any quizzes
  - ✓ Late assignment will be accepted with negative marking, -1 for each day
  - ✓ A copied assignment/project will be simply marked zero.
- Lab Policy:
  - ✓ Home Tasks

Course Project

- The course project gives students a chance to apply Computer Vision Algorithms discussed in class to a research oriented or industrial problem
- Working in Groups
- Evaluation of article reading and project
  - ✓ Article reading: Submit a survey of at least 5 articles you read and the list of the articles (2020 to 2024, TPAMI, IJCV, ICCV, CVPR, ECCV)
  - ✓ Submit an article including introduction, methods, experiments, results, and conclusions
  - ✓ Submit the project code, the readme document, and some testing samples (images, videos, etc.) for validation
- Presentation

Course Project

- Potential Topics
  - ✓ Image-to-Image Translation (satellite images to maps, turning day scenes into night scenes, or transforming sketches into realistic images)
  - ✓ Deepfake Technology (creating realistic-looking images or videos)
  - ✓ 3D Object Recognition and Reconstruction (augmented reality)
  - ✓ Event/Anomaly Recognition in Videos
  - ✓ Multi-Modal Image Matching (optical, infrared)
  - ✓ Multi Object Tracking (under occlusions, scale changes, and appearance variations)
  - ✓ Remote Sensing and Satellite Image Analysis (land cover classification, disaster monitoring, and environmental assessment)
  - ✓ Medical Image Analysis (Pathology slides, Radiology, Retinal)

Course Learning Outcomes

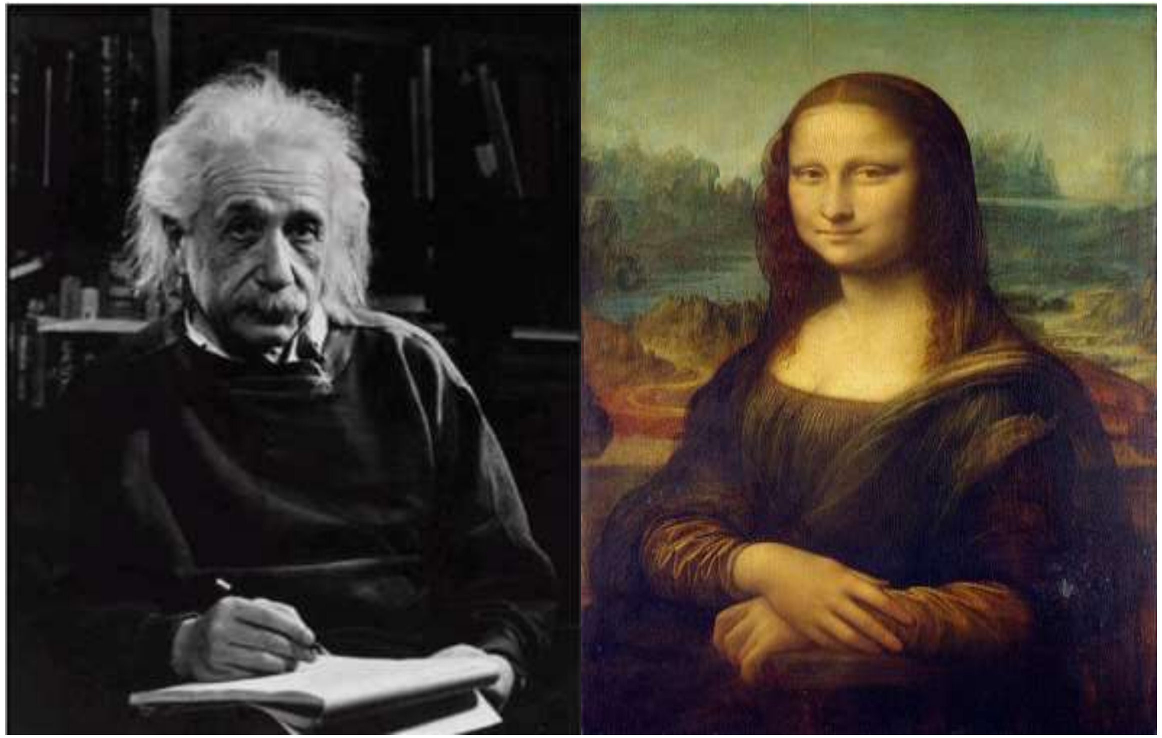
No	CLO (Tentative)	Domain	Taxonomy Level	PLO
1	Understanding basics of Computer Vision: algorithms, tools, and techniques	Cognitive	2	
2	Develop solutions for image/video understanding and recognition	Cognitive	3	
3	Design solutions to solve practical Computer Vision problems	Cognitive	3	

Outline

- What is Computer Vision
- Human Visual System
- Computer Vision Challenges
- Brief History of Computer Vision
- Computer Vision Tasks
- Computer Vision Applications

A picture is worth a 100 words.

A picture is worth 10,000 words.



A picture is worth a 100 words.

What is computer vision?

What kind of scene?

Where are the cars?

How far is the building?

What is computer vision?

Automatic understanding of images and video.

Giving the computer "eyes" to see and identify as humans would.

The study of understanding the world through visual perception.

It is also called:

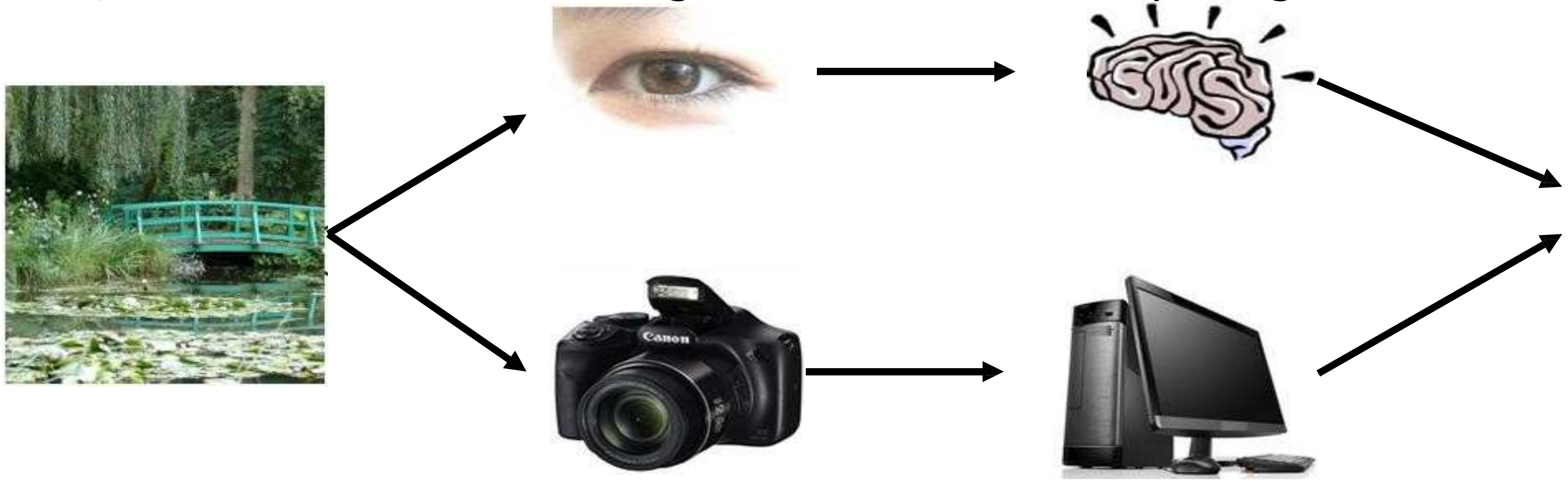
- ***Image understanding***
- *Robot vision*
- *Image Analysis*

- What kind of scene?
- Where are the cars?
- How far is the building?

What is (computer) vision?

Image (or video) Sensing device Interpreting device Interpretations

Image (or video) Sensing device Interpreting device Interpretations

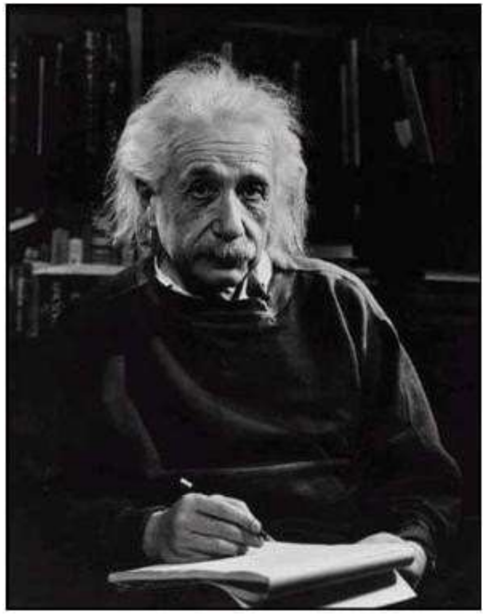


garden, spring, bridge, water, trees, flower, green, etc.

- Computer Vision emulates human vision
- Process images and extract information using context
- Make decisions using this information

What is computer vision?

- The goal of computer vision is to give computers <u>human-like perception</u>
  - ✓ bridge the gap between pixels and "meaning (semantic)"

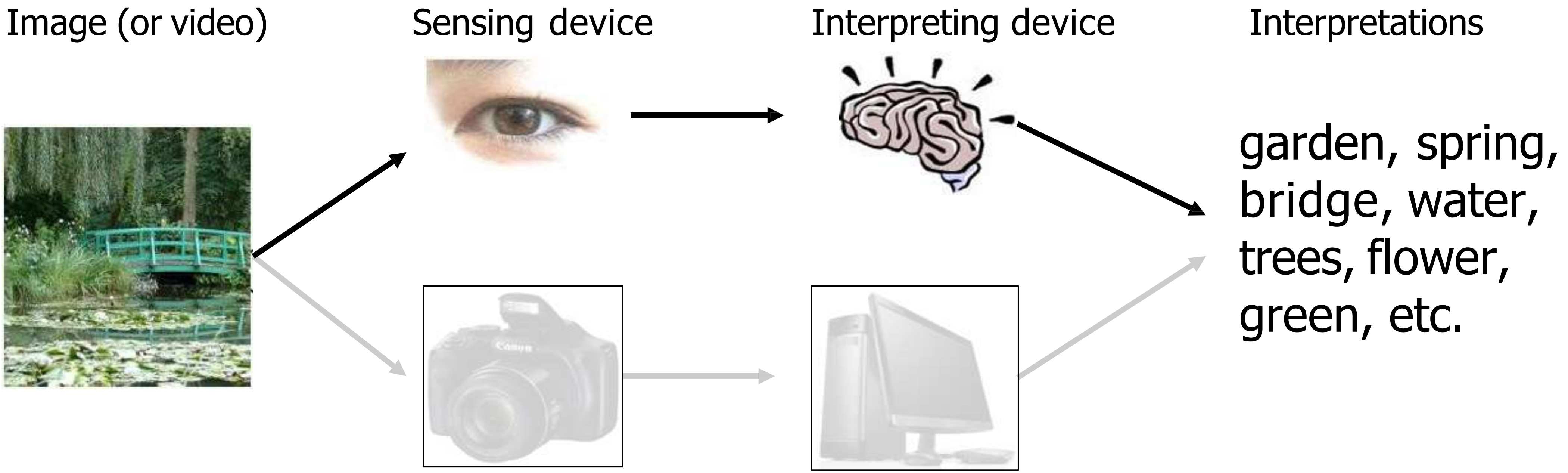


What we see

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

What is (~~computer~~) Vision?



Human vision

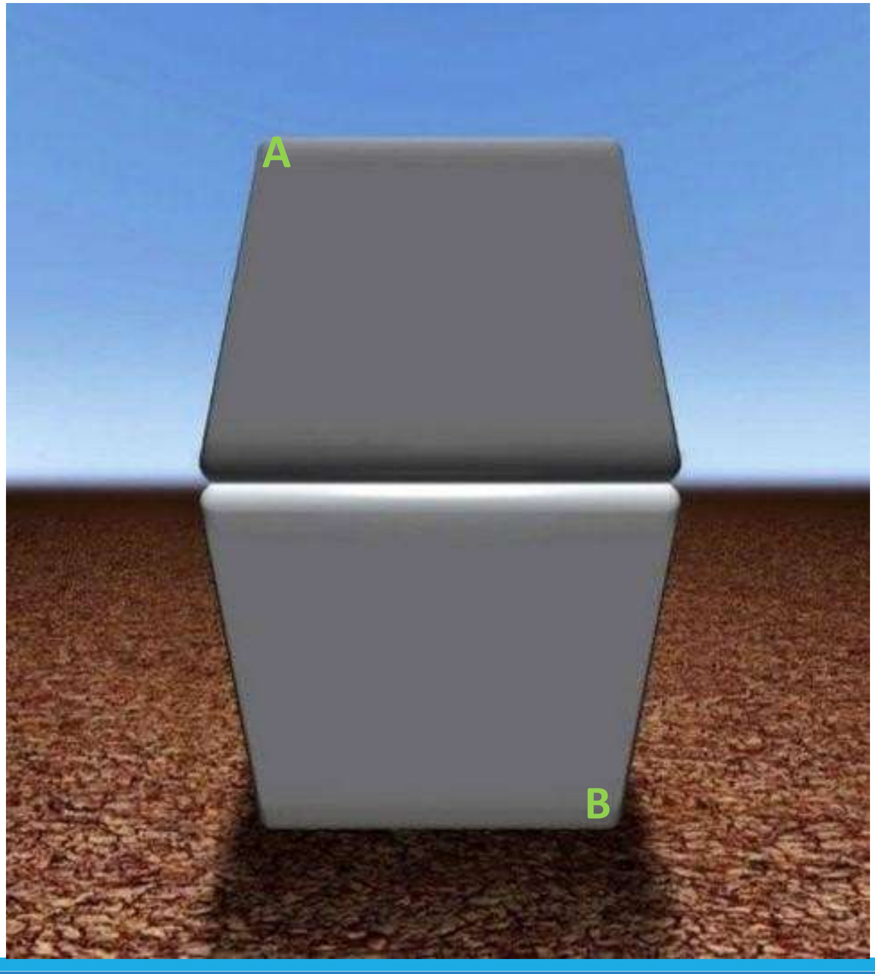
CAN DO AMAZING THINGS LIKE:

- ✓ Recognize people and objects
- ✓ Navigate through obstacles
- ✓ Understand mood in the scene
- ✓ Imagine stories

BUT STILL IS NOT PERFECT:

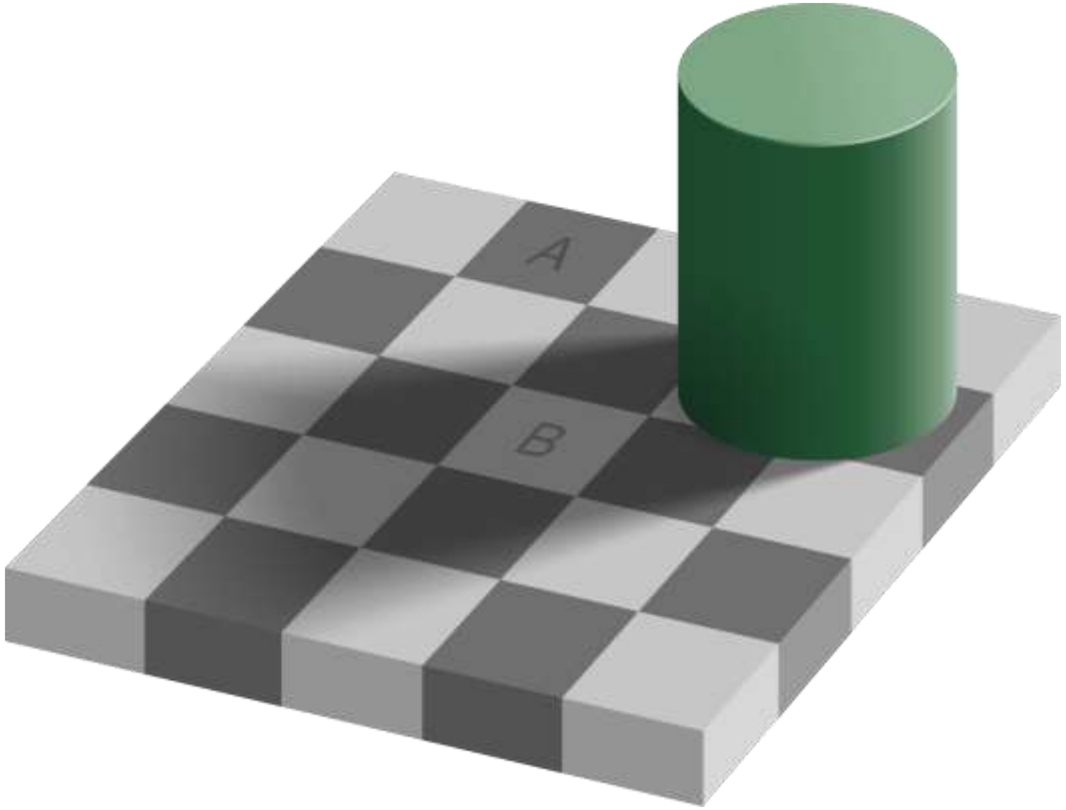
- ✓ Suffers from illusions
- ✓ Ignores many details
- ✓ Ambiguous description of the world

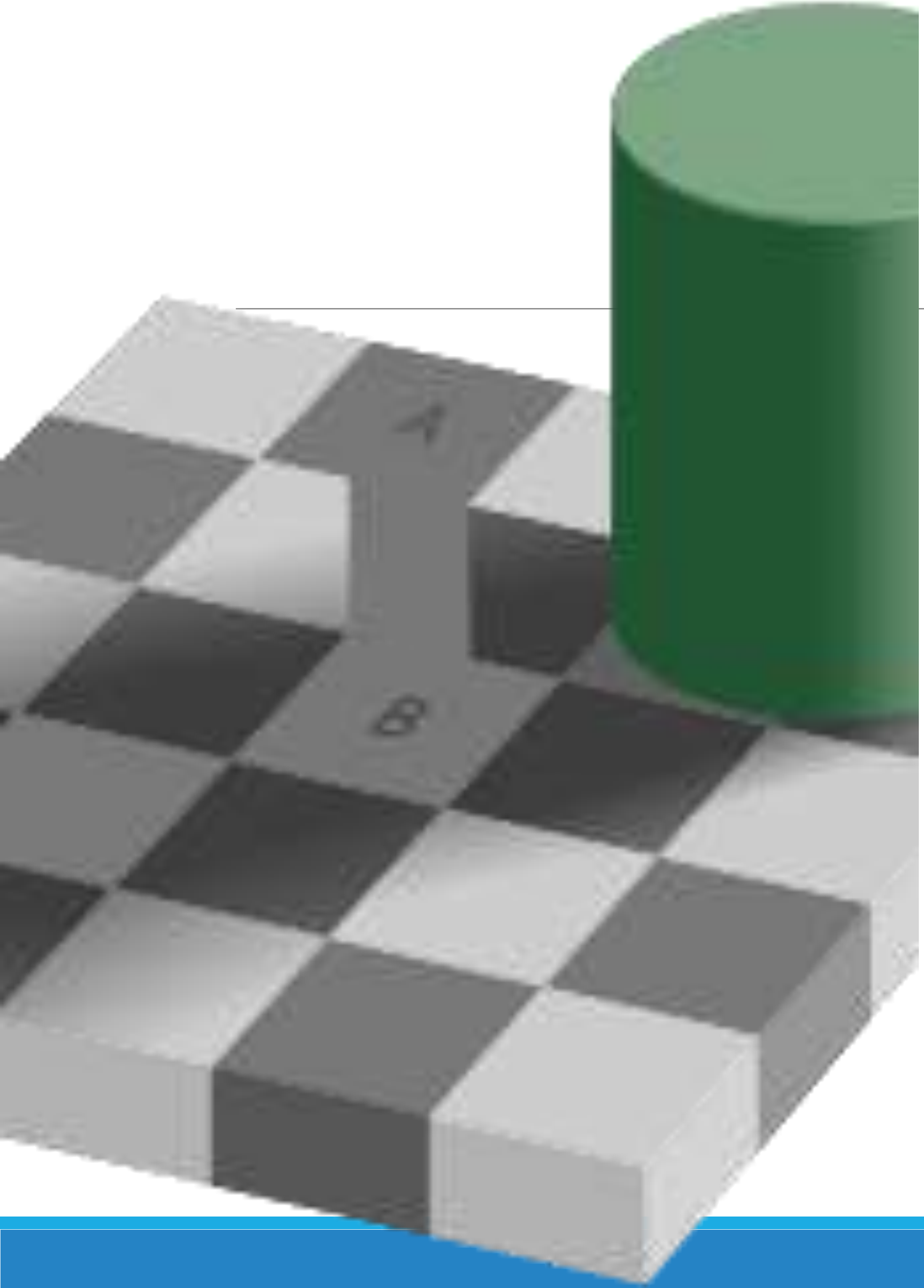
Measurement vs Perception



<u>Same Color</u>

A appears markedly darker than B, even though they are the <u>**exact same**</u> shade of gray.





Measurement vs Perception

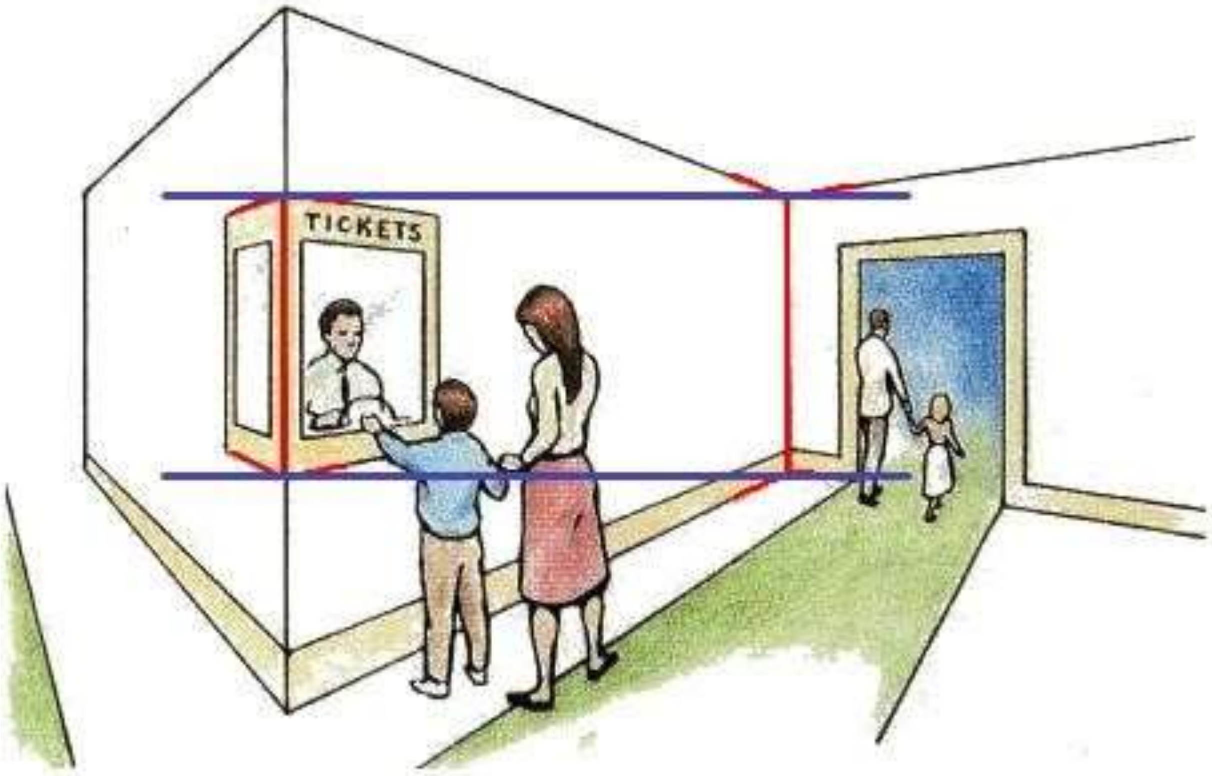
Lighter squares surrounded by darker squares appear lighter than average based on local contrast.

Conversely, darker squares surrounded by lighter squares appear darker.

Measurement vs Perception

- Length Illusion

  The brain assumes the "angles in" configuration to be closer, computes size constancy on it, and –given identical retinal size of the two angle arrangements – concludes that the "angle in"-line is shorter



Ref -
Gregory 1968: https://michaelbach.de/ot/sze-muelue/index.html

What is (computer) vision?

Image (or video) | Sensing device | Interpreting device | Interpretations

garden, spring, bridge, water, trees, flower, green, etc.

What is Computer Vision?

Information Extraction: features, 3D structure, motion etc.

Interpretation: Recognize Objects, scenes, actions, events.

No matter what your definition is ...

Vision is hard.

But is fun...

What is not Computer Vision?

Image Processing:

Takes an image and process is to produce new, more desirable image. Image enhancement, image compression, image restoration.

Pattern Recognition :

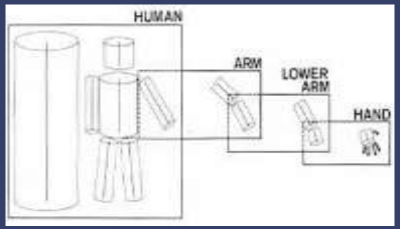
Takes a pattern and classifies it into one of predefined, finite set of classes.

Computer Graphics:

Synthesize images using powerful algorithms so that they correspond as close as possible to the real images.

Brief History of Computer Vision

- **1966**: Minsky assigns computer vision as an undergrad summer project
- **1960's**: interpretation of synthetic worlds
- **1970's**: some progress on interpreting selected images
- **1980's**: ANNs come and go; shift toward geometry and increased mathematical rigor
- **1990's**: face recognition; statistical analysis in vogue
- **2000's**: broader recognition; large annotated datasets available; video processing starts
- **2010's**: Deep learning with ConvNets
- **2020's**: Widespread autonomous vehicles?
- **2030's**: robot uprising?



1970's Example


1980's Example


1990's Example

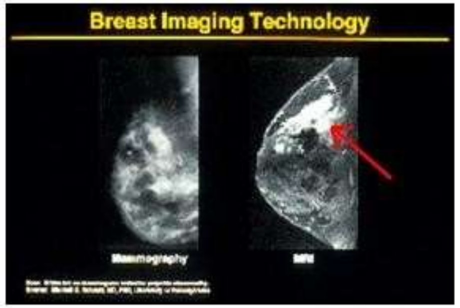

2010's Example

Why is Computer Vision so Difficult?

- It is an inverse problem, in which we seek to recover some unknowns given insufficient information to fully specify the solution.
- We resort to physics-based and probabilistic models, or machine learning from large sets of examples, to disambiguate between potential solutions.
- However, modeling the visual world in all its rich complexity is far more difficult than, say, modeling the vocal tract that produces spoken sounds

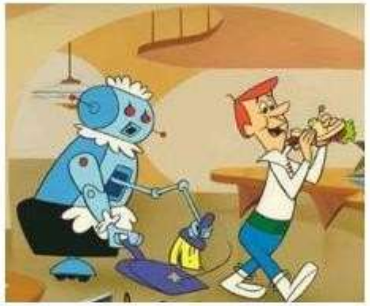
Why computer vision matters?

Safety



Health

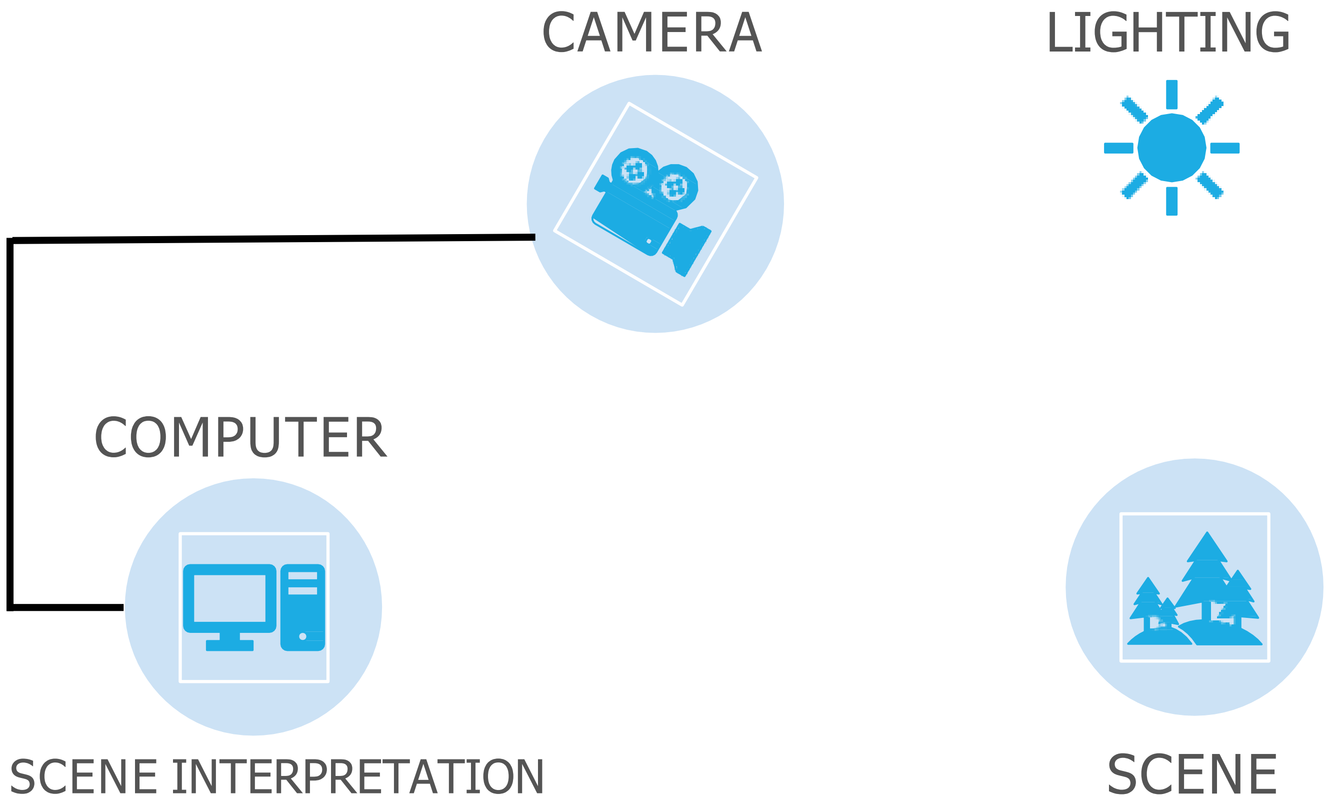
Security



Comfort

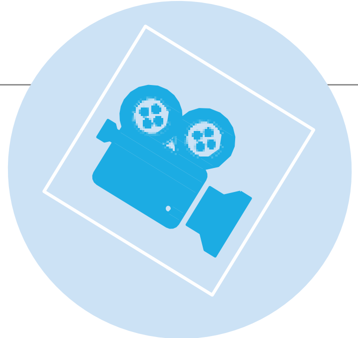
Fun

Access



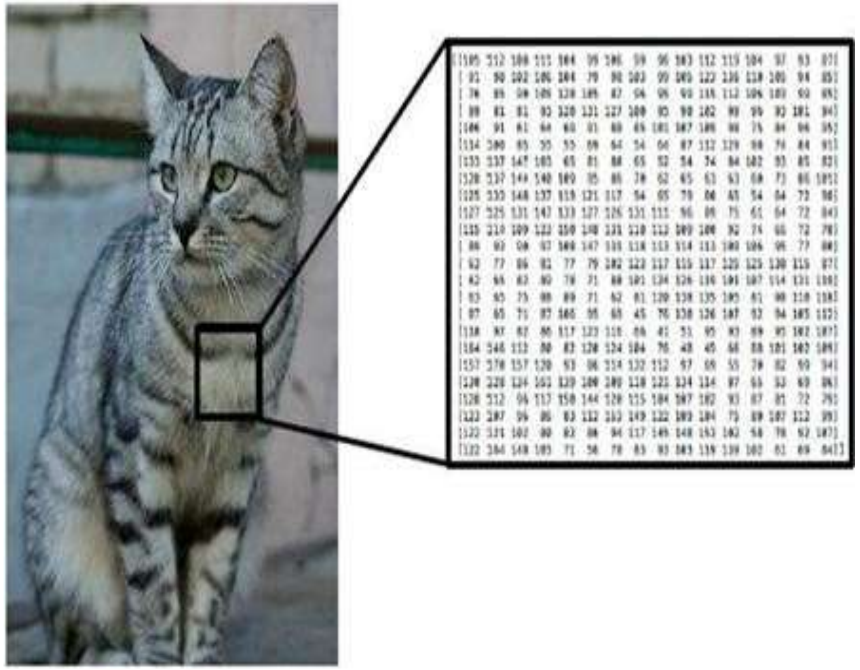
Components of a computer vision system

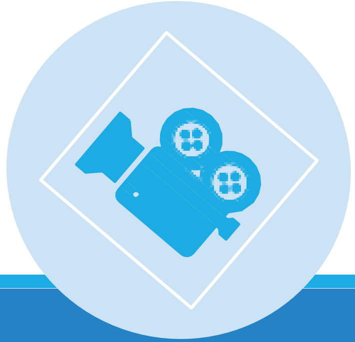
Computer Vision Challenges



- Viewpoint Variation

- All pixels change when the camera moves





Computer Vision Challenges

- Viewpoint Variation

Michelangelo 1475 –1564

Adapted from L. Fei Fei, R. Fergus, A. Torralba

Computer Vision Challenges

- Background Clutter

This image is CC0 1.0 public domain

This image is CC0 1.0 public domain

Computer Vision Challenges

- Illumination

This image is CC0 1.0 public domain

This image is CC0 1.0 public domain

This image is CC0 1.0 public domain

This image is CC0 1.0 public domain

Computer Vision Challenges

- Occlusion

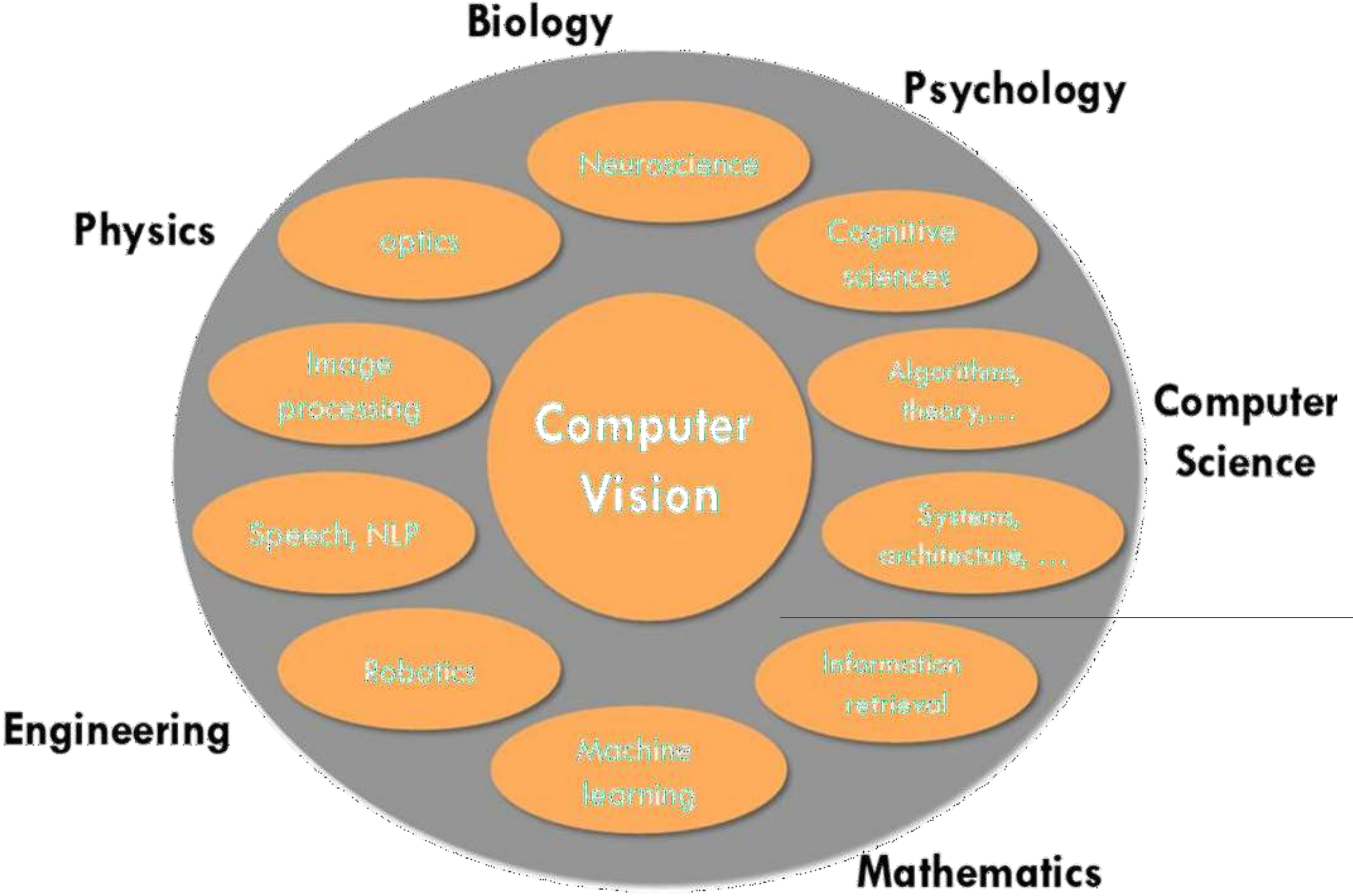
Computer Vision Challenges

- Deformation

Computer Vision Challenges

- Intraclass Variation

This image is CC0 1.0 public domain



Scope of Computer Vision

Computer Vision Tasks

Reorganization

- Segment (cluster) pixels with similar properties

Reconstruction

- Recover 3D information from data

Recognition

- Detect and identify objects

Understanding

- What is happening in the scene?

Computer Vision Tasks

Low Level Vision

Measurements

Enhancements

Region segmentation

Features

Mid Level Vision

Reconstruction

Depth

Motion Estimation

High Level Vision

Category detection

Activity recognition

Deep understandings

Pose estimation

Computer Vision Tasks

Low-Level or Early Vision

- Local image/shape properties

"There's an edge!"

Mid-Level Vision

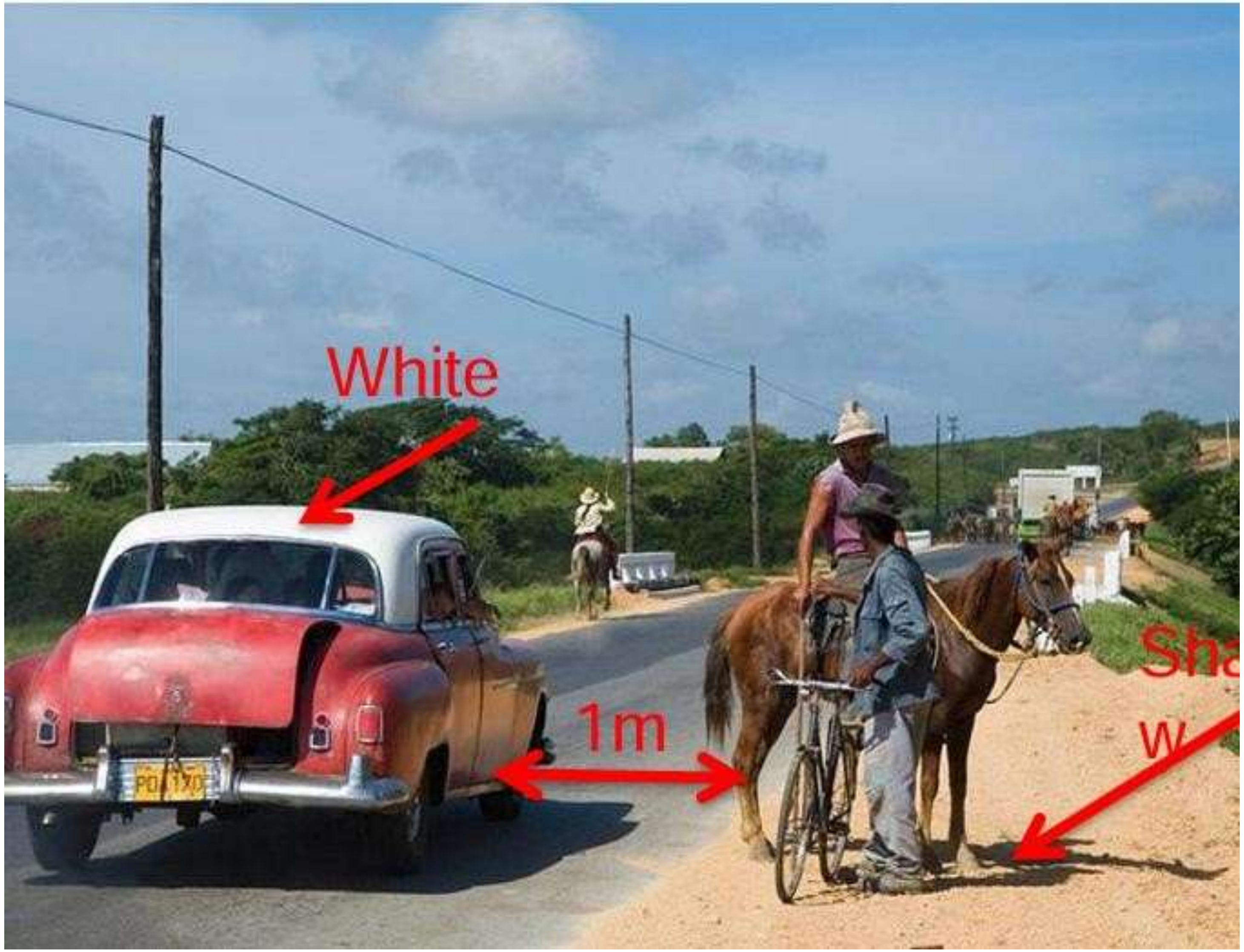
- Grouping and segmentation

"There's an object and a background!"

High Level Vision

- Recognition
- Classification

"It's a chair! It's in a room!"

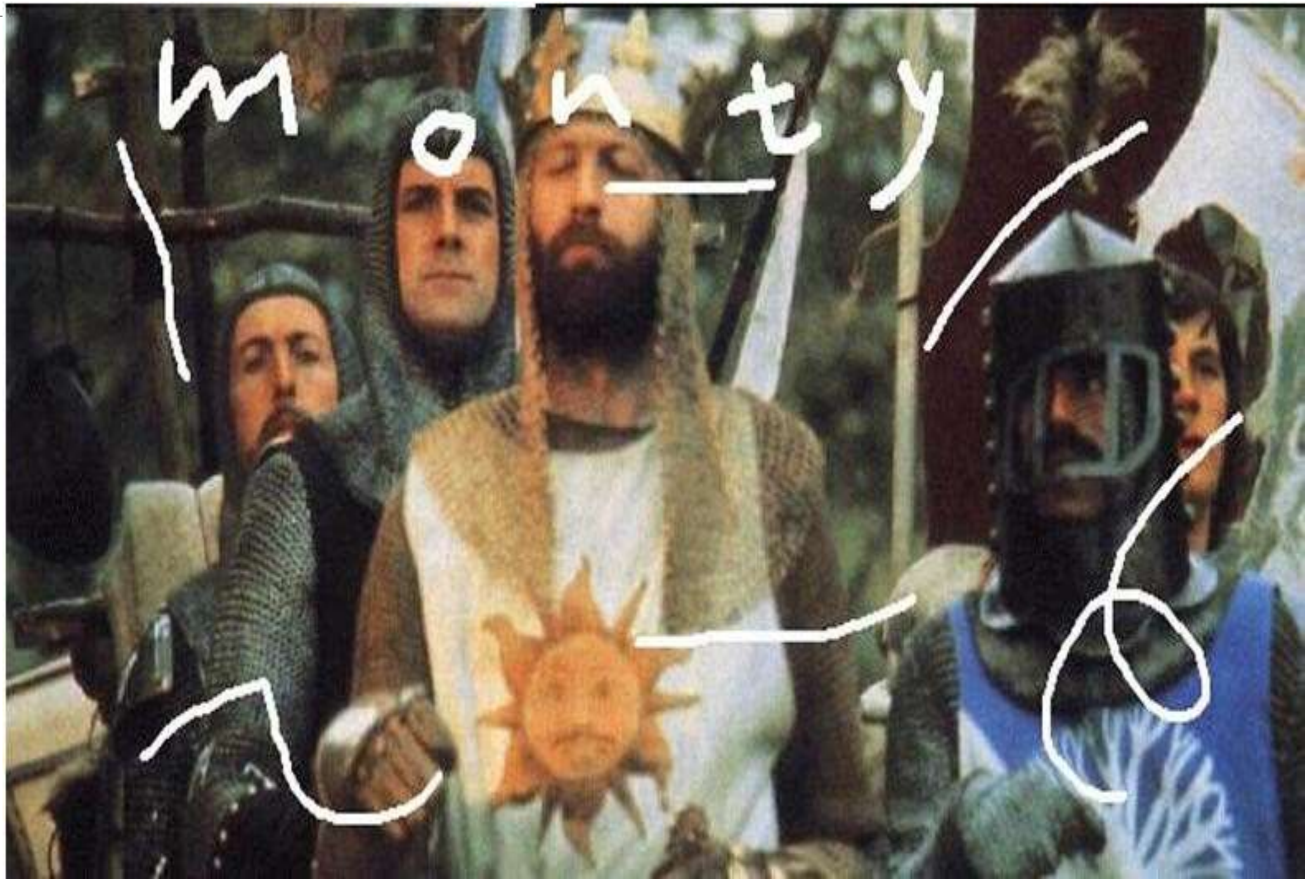


Computer Vision Tasks

- Low Level Vision
  - ✓ Measurements
  - ✓ Enhancements
  - ✓ Region segmentation
  - ✓ Features
- Mid Level Vision
  - ✓ Reconstruction
  - ✓ Depth
  - ✓ Motion Estimation
- High Level Vision
  - ✓ Category detection
  - ✓ Activity recognition
  - ✓ Deep understandings
  - ✓ Pose estimation

Image Enhancement

Replace those bad marks with its neighboring pixels so that it looks like the neighborhood



"Image Inpainting", M. Bertalmío et al. 2000.

Image Resizing

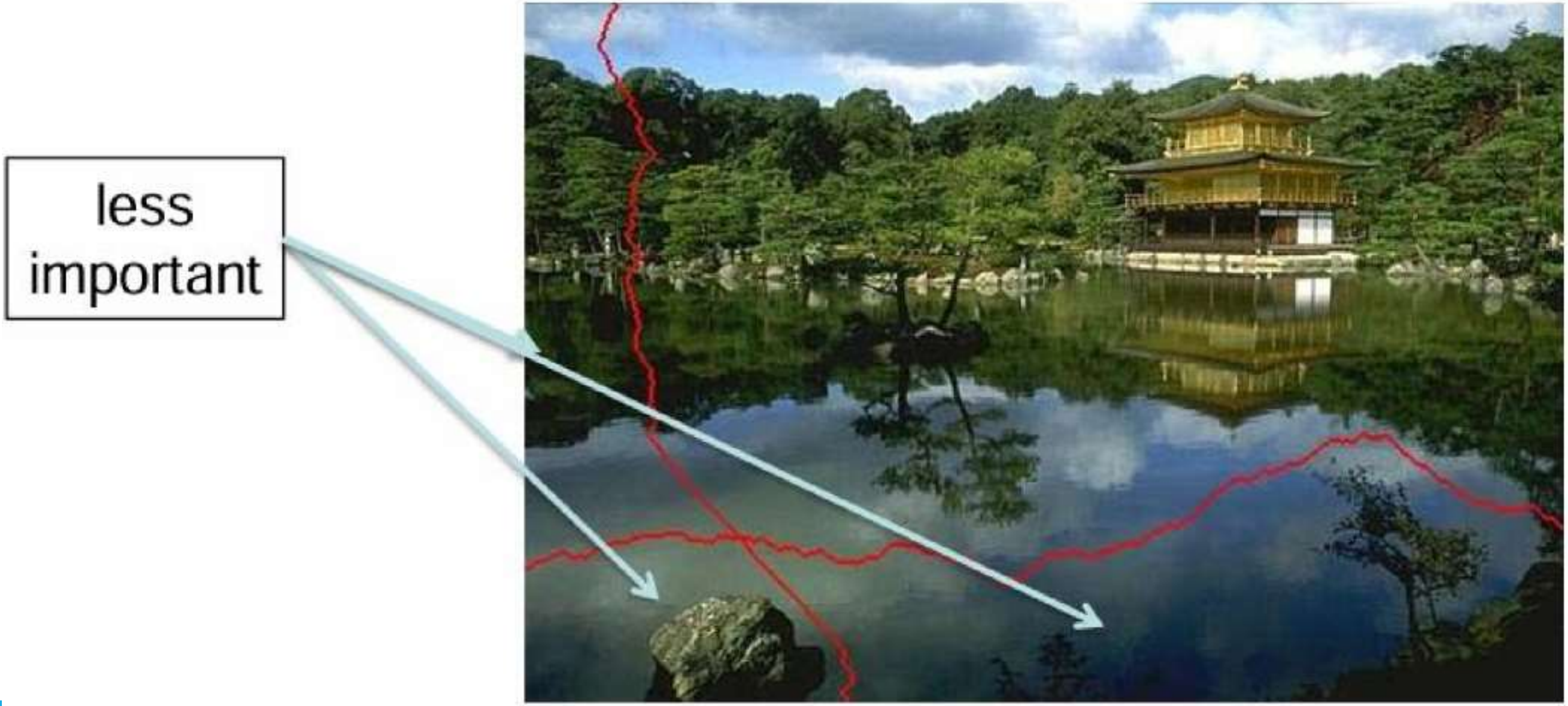


Image Resizing

Traditional resizing uses and stretches the whole image.

Seam Carving:

Content-aware resizing uses important areas. Extends in horizontal direction and reduces in vertical.

Image Segmentation

Ref -

https://gts.ai/how-do-we-solve-the-challenges-faced-due-to-semantic-segmentation/

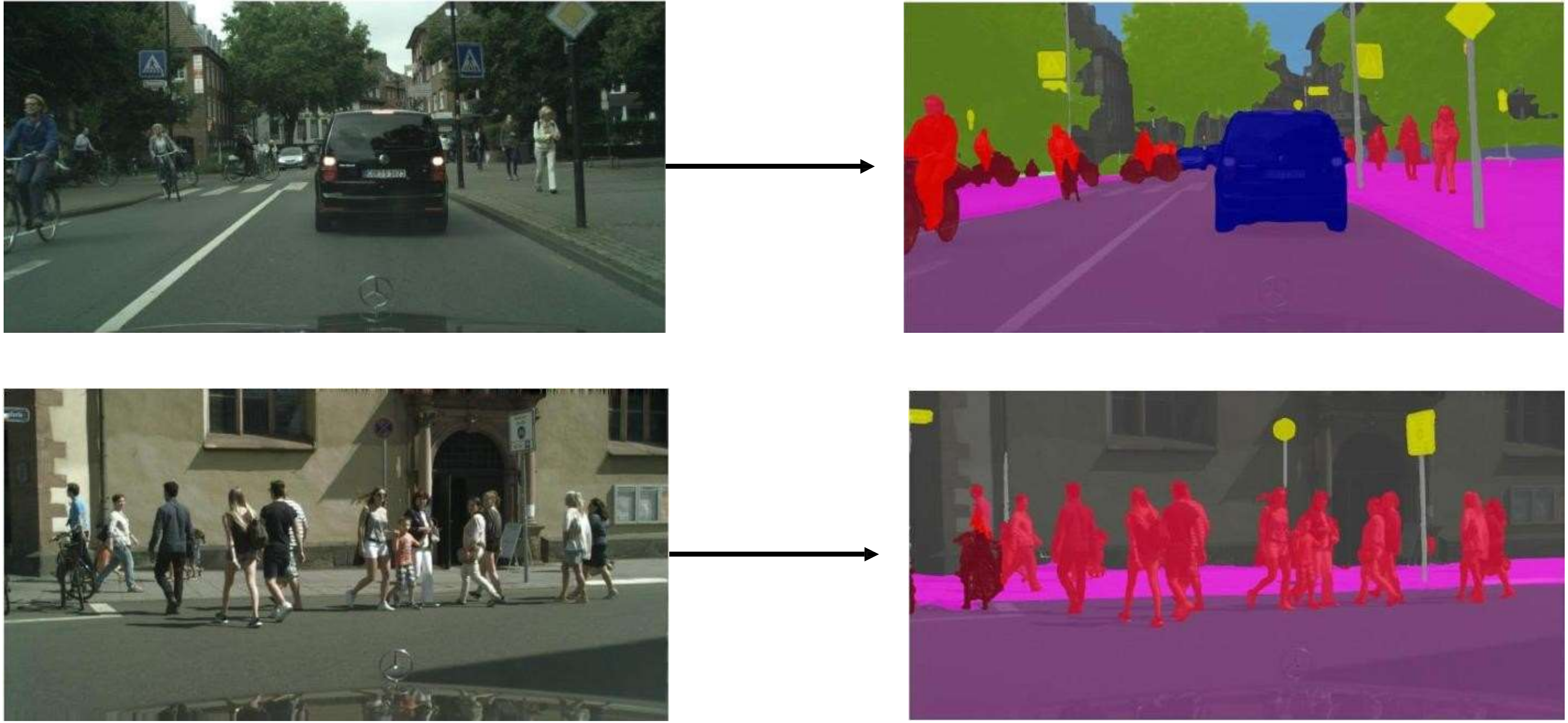
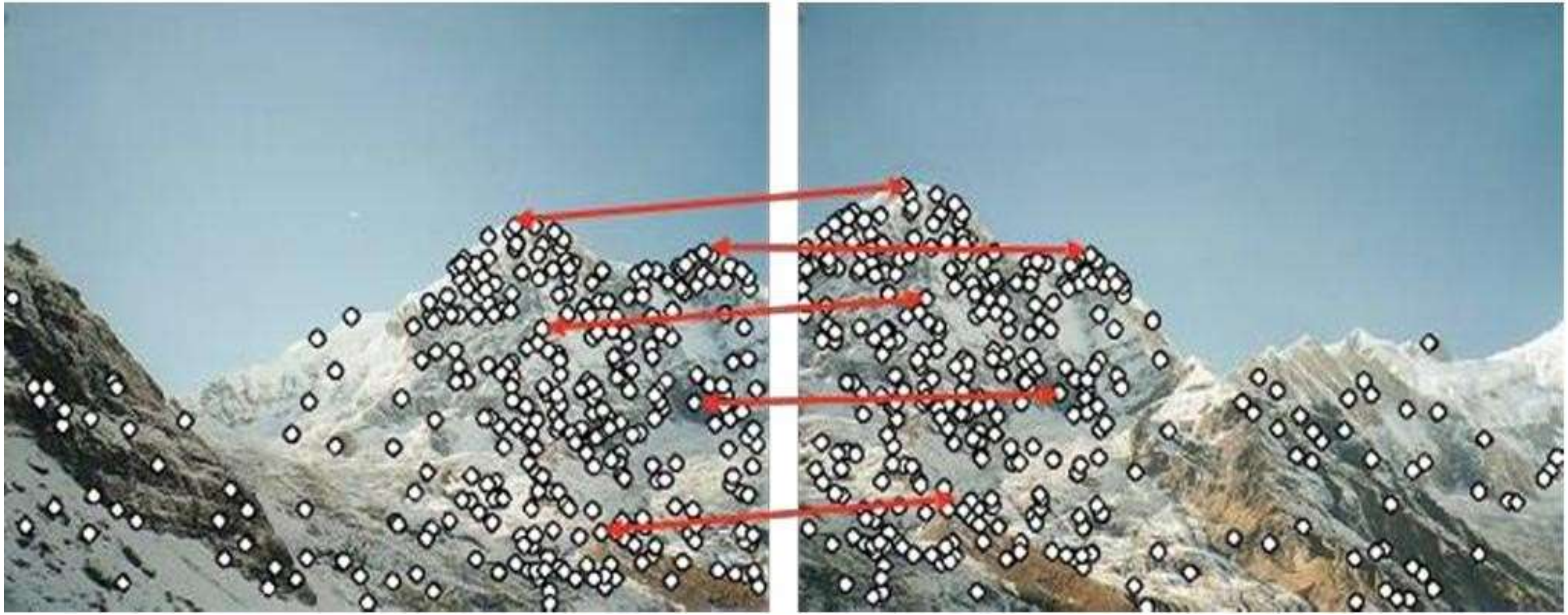
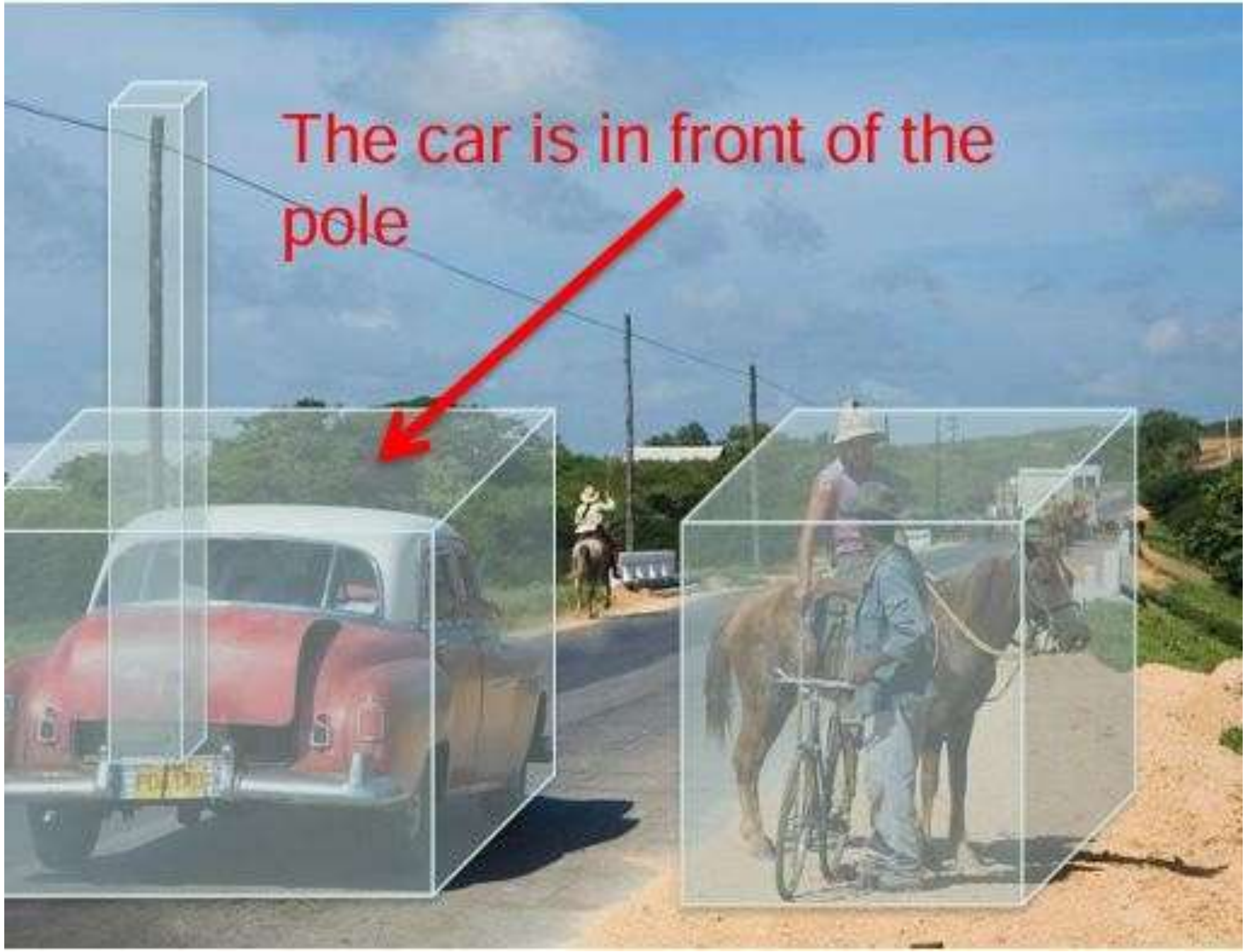


Image Segmentation

Image Features





Computer Vision Tasks

Low Level Vision

- Measurements
- Enhancements
- Region segmentation
- Features

Mid Level Vision

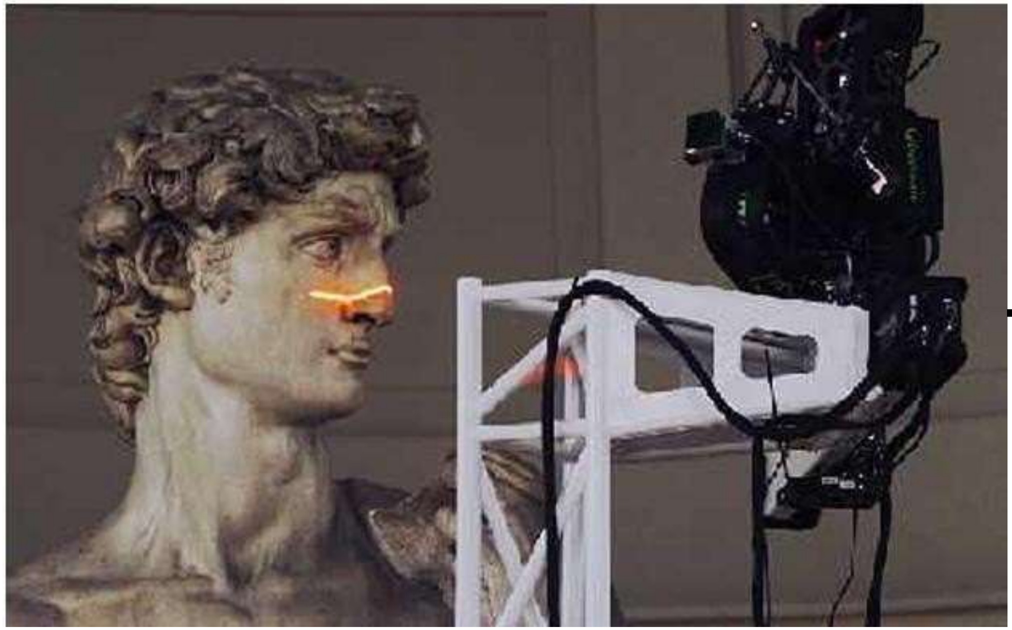
- Reconstruction
- Depth
- Motion Estimation

High Level Vision

- Category detection
- Activity recognition
- Deep understandings
- Pose estimation

Applications: 3D Scanning

- Combining multiple range and color images, allow to reliably and accurately digitize the external shape and surface characteristics of many physical objects



[The Digital Michelangelo Project](http://graphics.stanford.edu/projects/mich/) :-
http://graphics.stanford.edu/projects/mich/

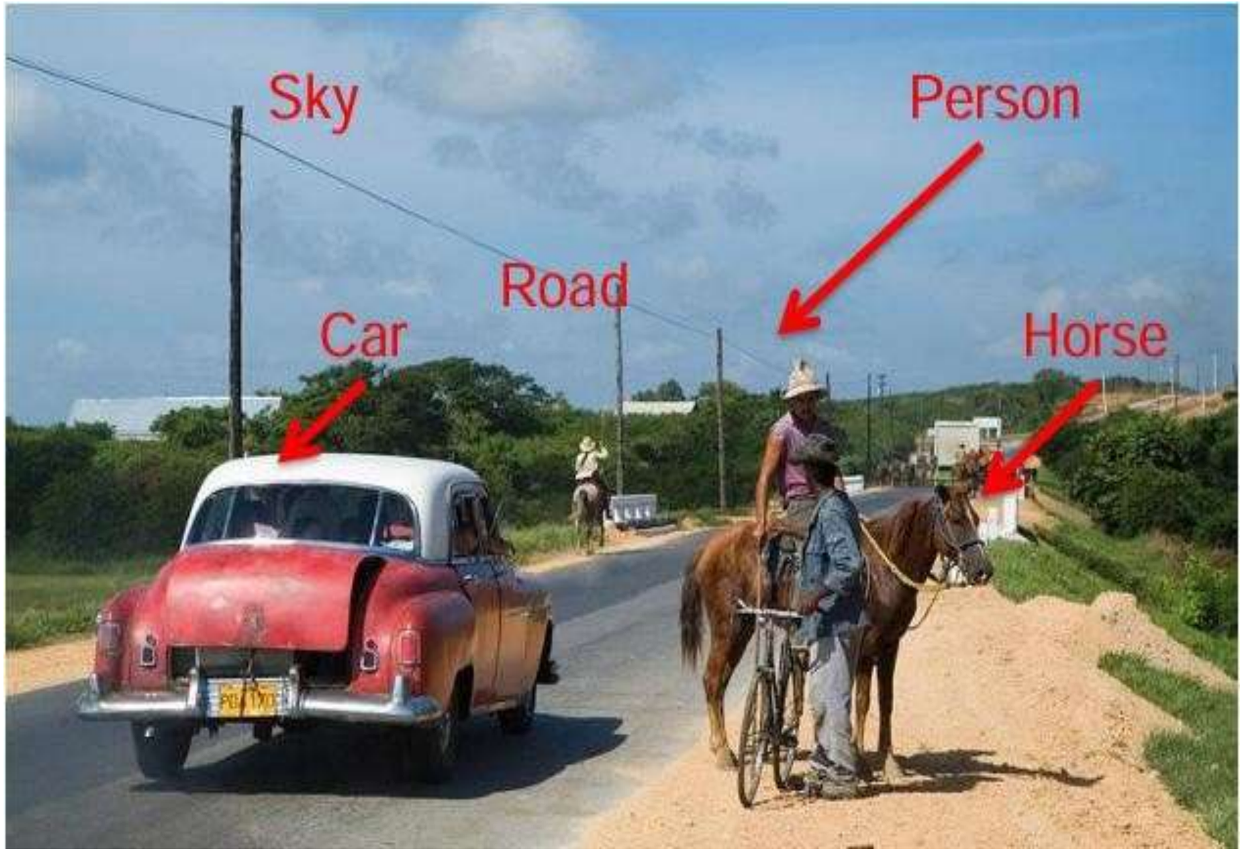
✓ 2 BILLION polygons, accuracy to .29mm

Apple’s 3D maps

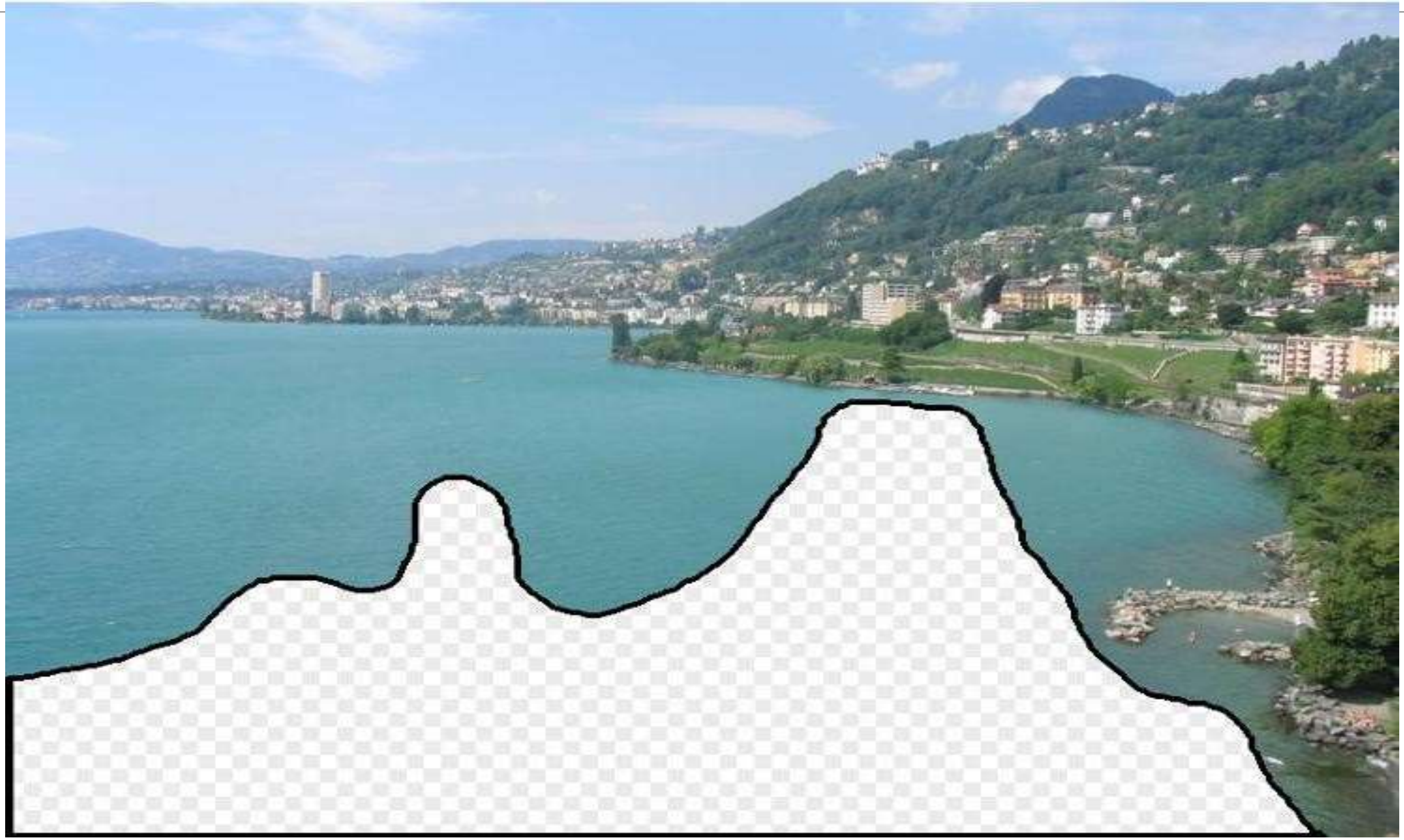
Ref - http://www.youtube.com/watch?v=InIVv-LsgZE

Computer Vision Tasks

Low Level Vision	Mid Level Vision	High Level Vision
• Measurements • Enhancements • Region segmentation • Features	• Reconstruction • Depth • Motion Estimation	• Category detection • Activity recognition • Deep understandings • Pose estimation

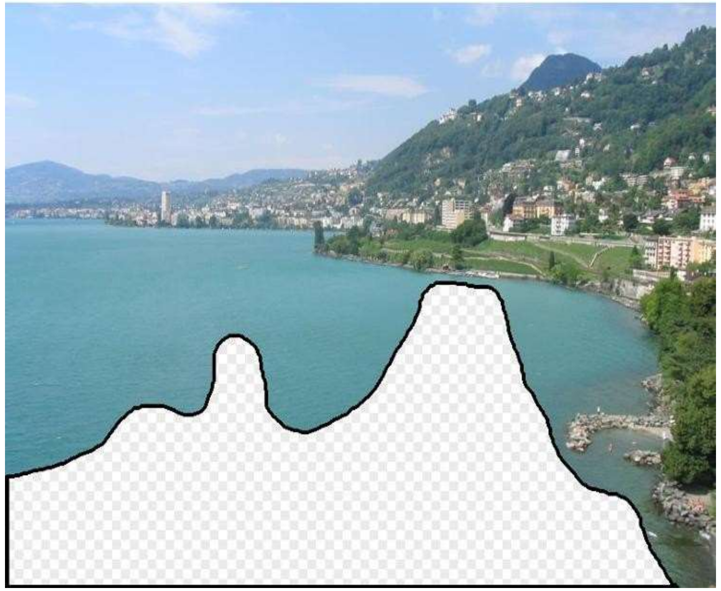


Scene Completion



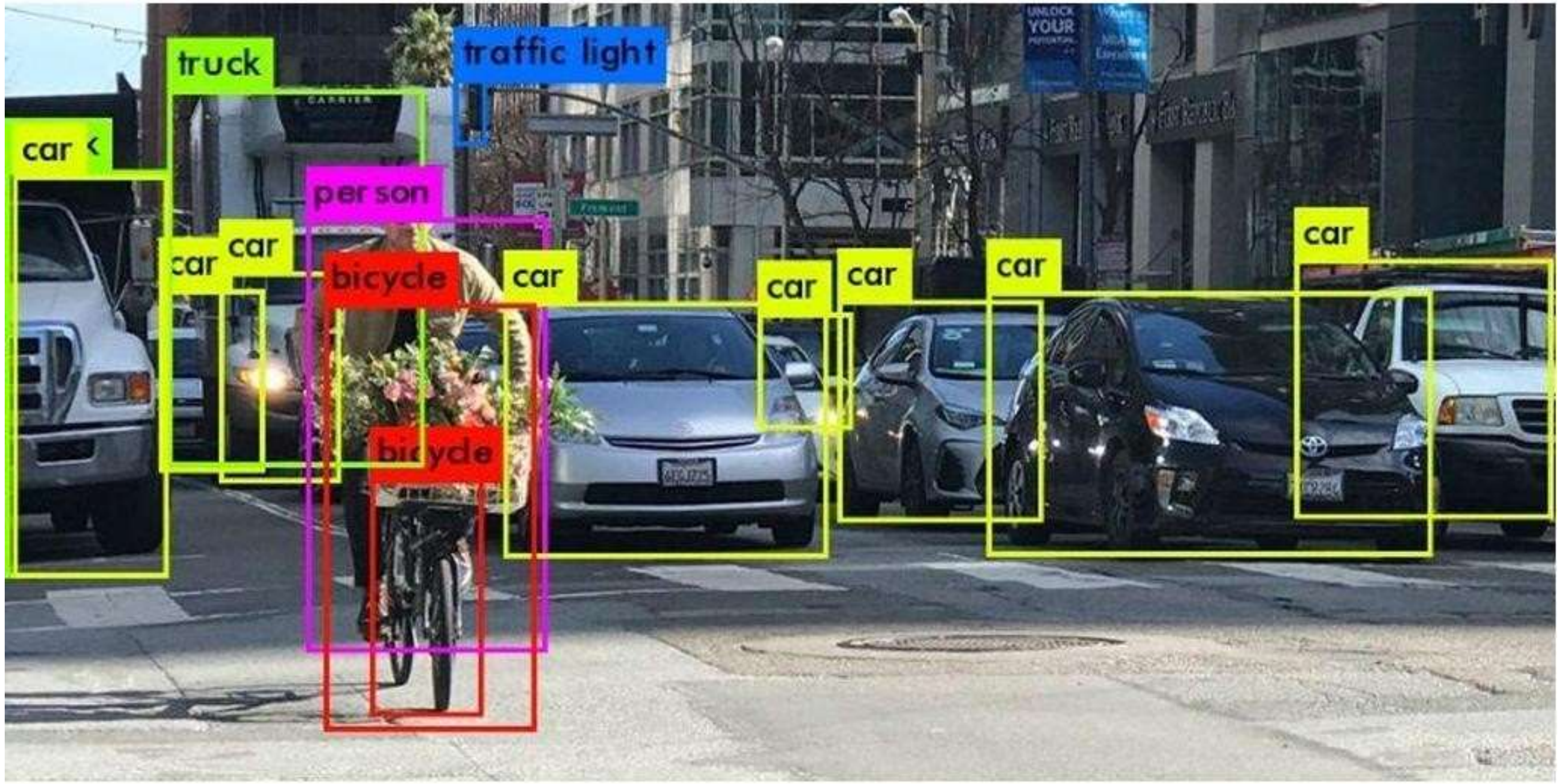
[Hays and Efros. Scene Completion Using Millions of Photographs. SIGGRAPH 2007 and CACM October 2008.]

Scene Completion



Nearest neighbor scenes from database of 2.3 million photos

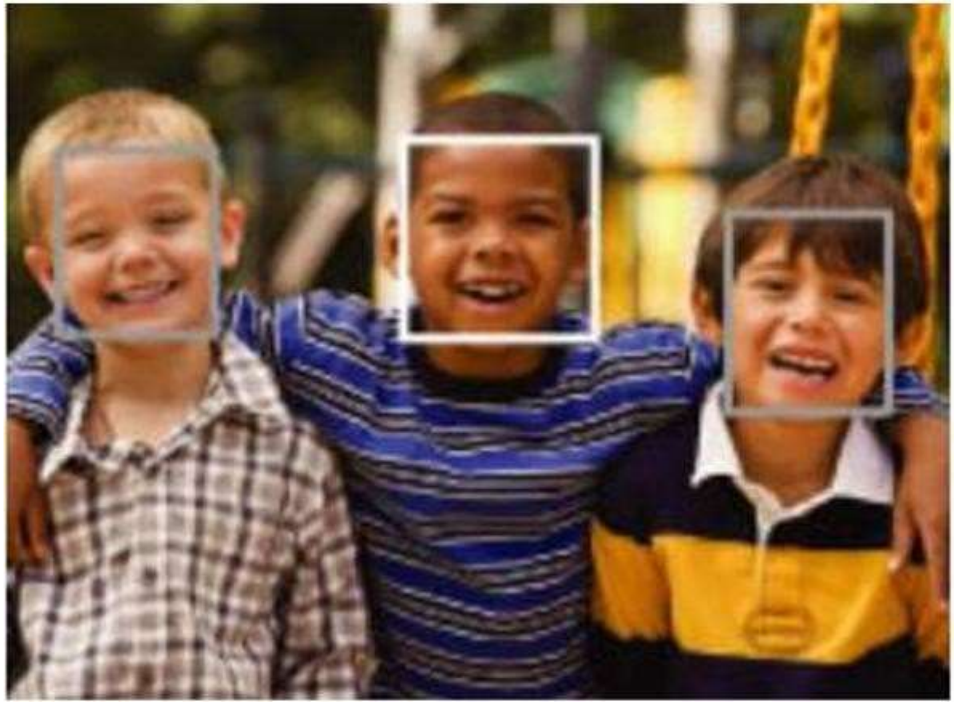
Object Detection



Ref - https://heartbeat.fritz.ai/introduction-to-basic-object-detection-algorithms-b77295a95a63

Face Detection

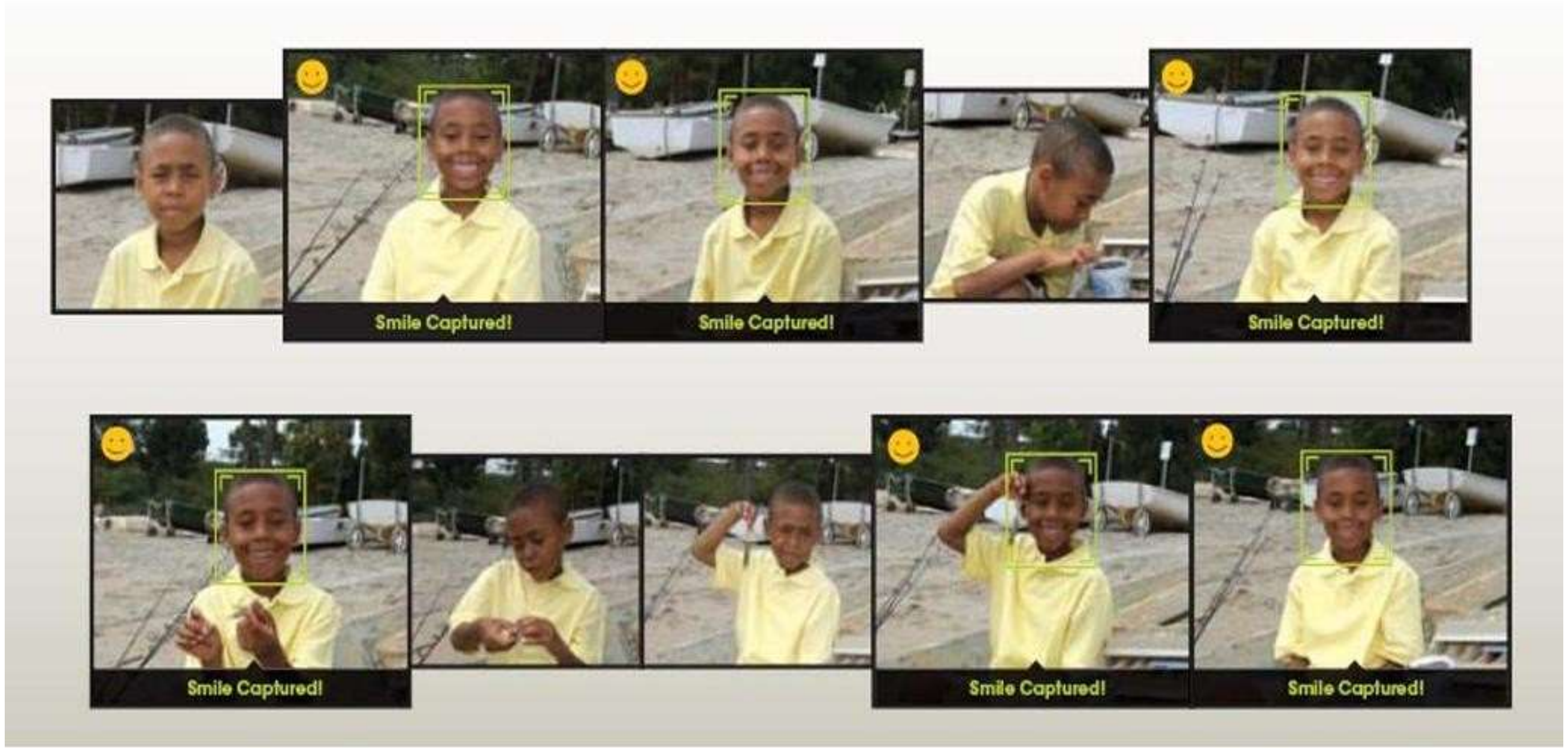
- Many new digital cameras now detect faces
  - ✓ Apple, Samsung, Canon, Sony, Fuji, …



Face Detection

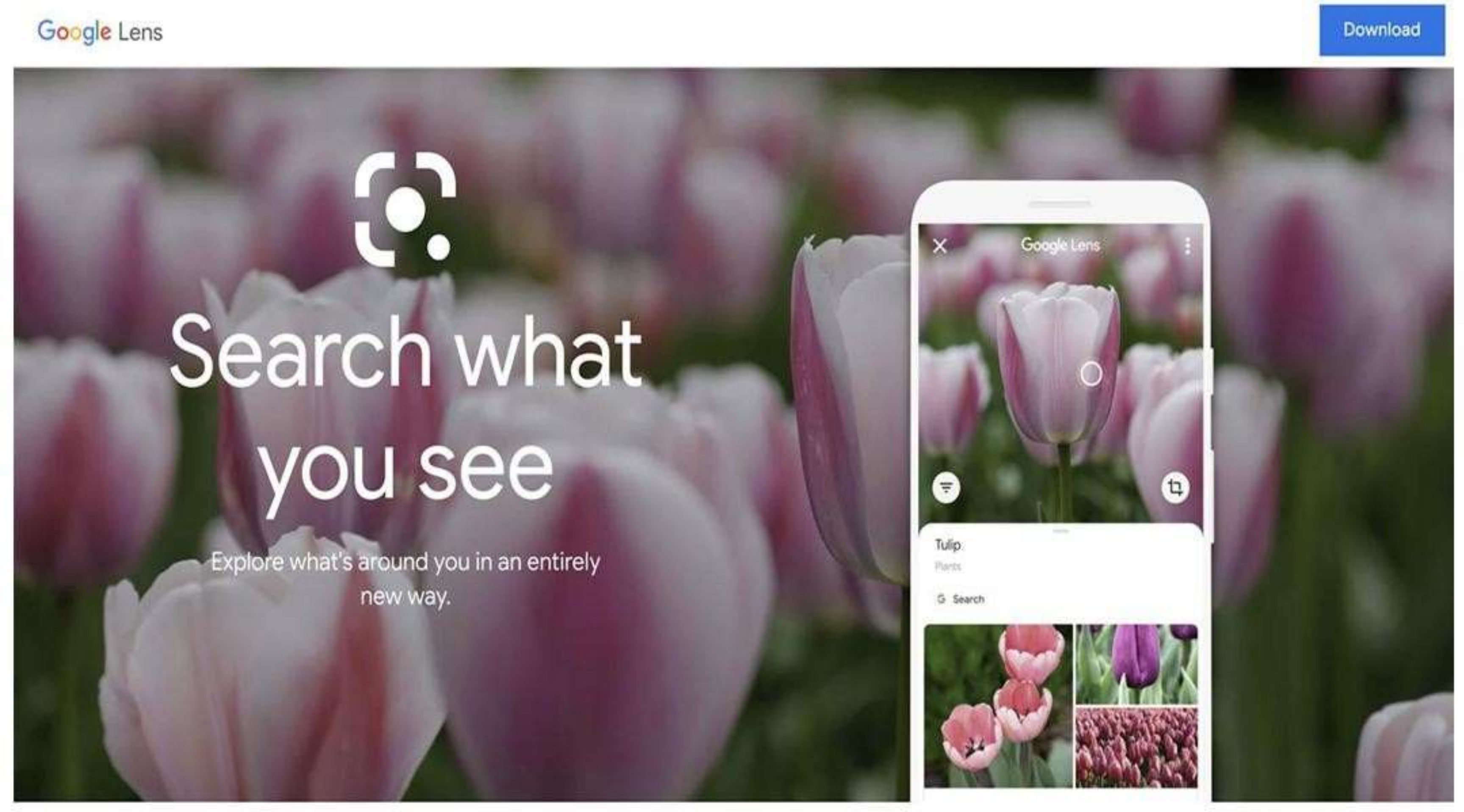
https://youtu.be/Q0M90dbqs6M

Smile Detection

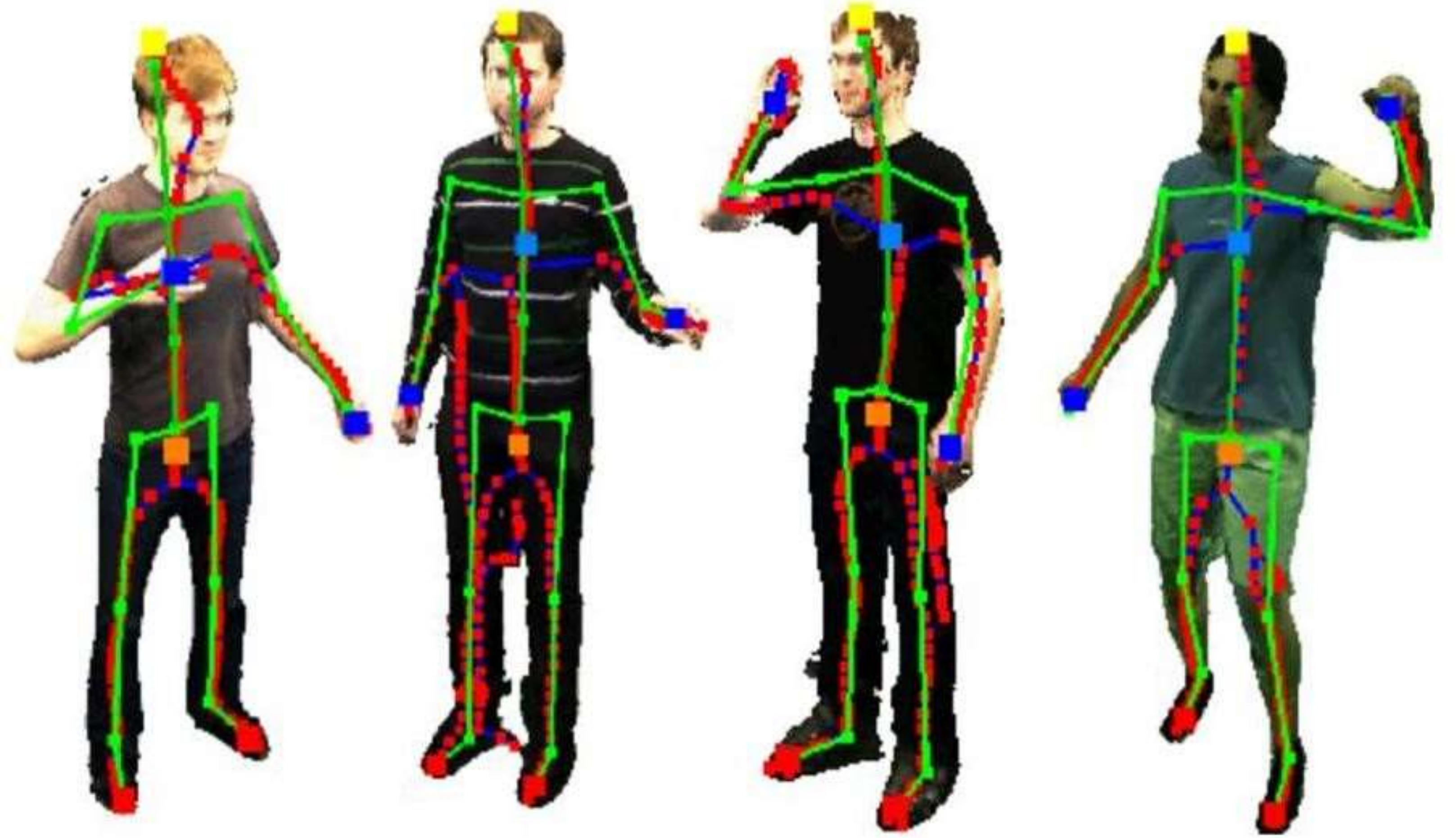


Sony Cyber-shot® T70 Digital Still Camera

Object Recognition

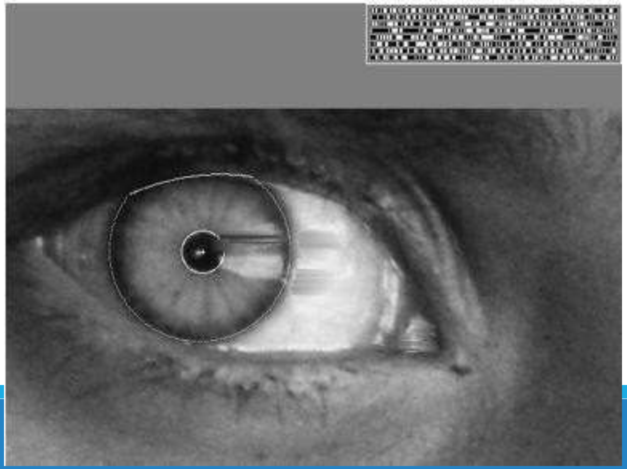


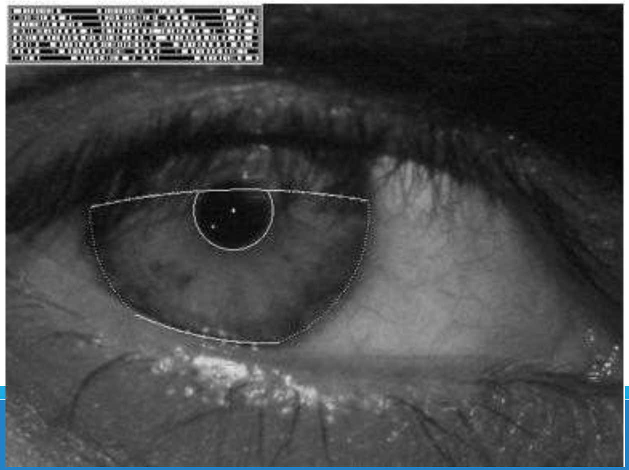
Pose Estimation



Vision-based biometrics

- "How the Afghan Girl was Identified by Her Iris Patterns" Read the [story wikipedia]





Vision-Based Interaction: Xbox Kinect

Interactive Games: Kinect

Object Recognition:

- http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o

Mario:

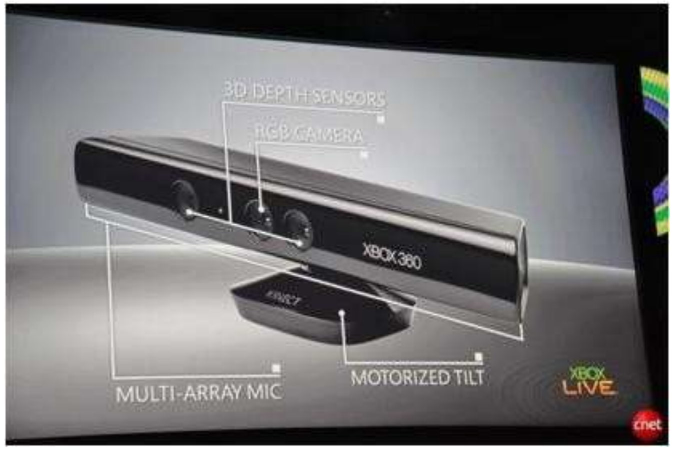
- http://www.youtube.com/watch?v=8CTJL5lUjHg

3D:

- http://www.youtube.com/watch?v=7QrnwoO1-8A

Robot:

- http://www.youtube.com/watch?v=w8BmgtMKFbY



Computer Vision Applications

Laptop: Biometrics auto-login (face recognition, 3D), OCR

Smartphones: QR codes, computational photography (Android Lens Blur, iPhone Portrait Mode), panorama construction (Google Photo Spheres), face detection, expression detection (smile), Snapchat filters (face tracking), Google Tango (3D reconstruction), Night Sight (Pixel)

Web: Image search, Google photos (face recognition, object recognition, scene recognition, geolocalization from vision), Facebook (image captioning), Google maps aerial imaging (image stitching), YouTube (content categorization)

VR/AR: Outside-in tracking (HTC VIVE), inside out tracking (simultaneous localization and mapping, HoloLens), object occlusion (dense depth estimation)

Computer Vision Applications

Motion: Kinect, full body tracking of skeleton, gesture recognition, virtual try-on

Medical imaging: CAT / MRI reconstruction, assisted diagnosis, automatic pathology, connectomics, endoscopic surgery

Industry: Vision-based robotics (marker-based), machine assisted router (jig), automated post, ANPR (number plates), surveillance, drones, shopping

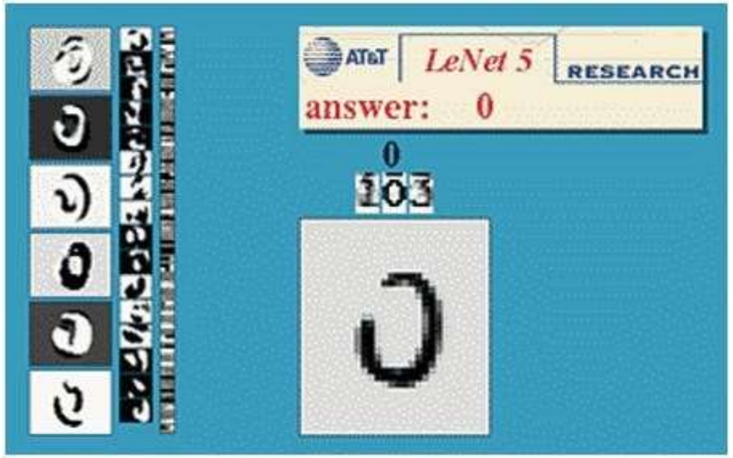
Transportation: Assisted driving (everything), face tracking/iris dilation for drunkeness, drowsiness, automated distribution

Media: Visual effects for film, TV (reconstruction), virtual sports replay (reconstruction), semantics-based auto edits

https://www.cs.ubc.ca/~lowe/vision.html

Computer Vision Applications

- Optical Character Recognition (OCR): Technology to convert images of text into text. Scanners come with OCR software.



Mail digit recognition, AT&T labs

WAYGO

Live Camera Translation

License plate readers

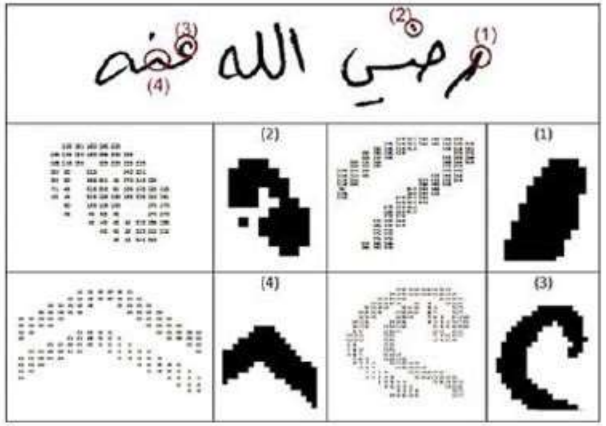
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Computer Vision Applications

- Writer Identification

Sameh Awaida and Sabri Mahmoud, “Writer Identification of Arabic Text using Statistical and Structural Features”, Cybernetics and Systems: An International Journal, Volume 44, Issue 1, pp. 57-76, 2012


السبب الأساسي في هذه المشكلات إنما يعزى إلى فشل المجتمع في التسامح
والتقبل للاختلافات و الفروق بين المعاقين من المشاركة العادية في فعاليات
وأنشطة وخبرات الحياة الاجتماعية اليومية . يمكن إكساب ذوي
الاحتياجات الخاصة مختلف المعارف والاتجاهات و القيم و المهارات التي
تؤهلهم للمشاركة الإيجابية الفعالة في مختلف أنشطة وفعاليات الحياة الإنسانية
إلى أقصى حد تؤهله لهم إمكانياتهم وقدراتهم إضافة إلى تغيير ثقافة المجتمع نحو
المعاقين



Authors	Language	No. of writers	Top-1 (%)	Top-5 (%)	Top-10 (%)
Schomaker et al. (2003)	Dutch	250	88.0	98.0	99.0
B. Zhang (2003)	English	1,000	98.0	—	—
Bulacu and Schomaker (2005)	Dutch	250	78.0	—	92.6
Bulacu, Schomaker et al. (2007)	Arabic	350	88.0	—	99.0
He et al. (2008)	Chinese	500	39.2	62.4	77.2
Siddiqi et al. (Siddiqi and Vincent 2009)	English	200	86.0	—	—
Our approach	Arabic	100	88.0	96.0	98.5
Our approach (remaining 150 writers)	Arabic	150	84.0	96.3	97.3
Our approach	Arabic	250	75.0	91.8	95.4
Our approach (with rejection)	Arabic	250	80.5	94.0	97.0

Computer Vision Applications

- Automated Visual Inspection

Computer Vision Applications

- Object Recognition

Toshiba Tech
IS-910T 2013

DataLogic
LaneHawk
LH4000 2012

Computer Vision Applications

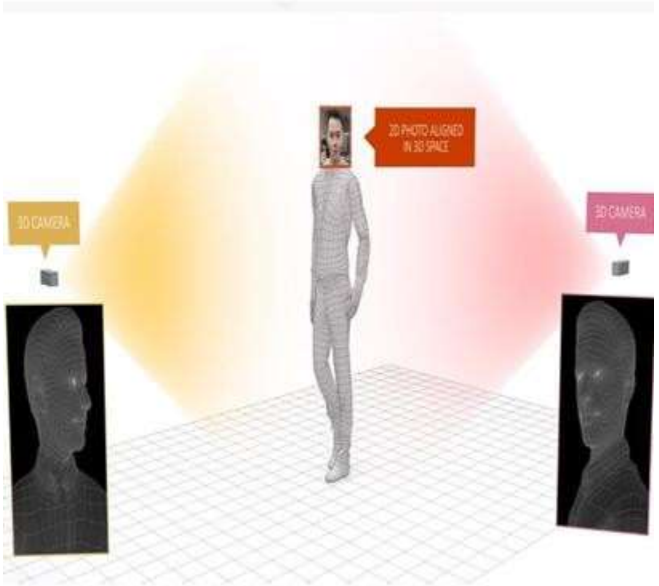
- Amazon Go

Computer Vision Applications

- Vision-based Biometrics

Fingerprint scanners on many new laptops, other devices

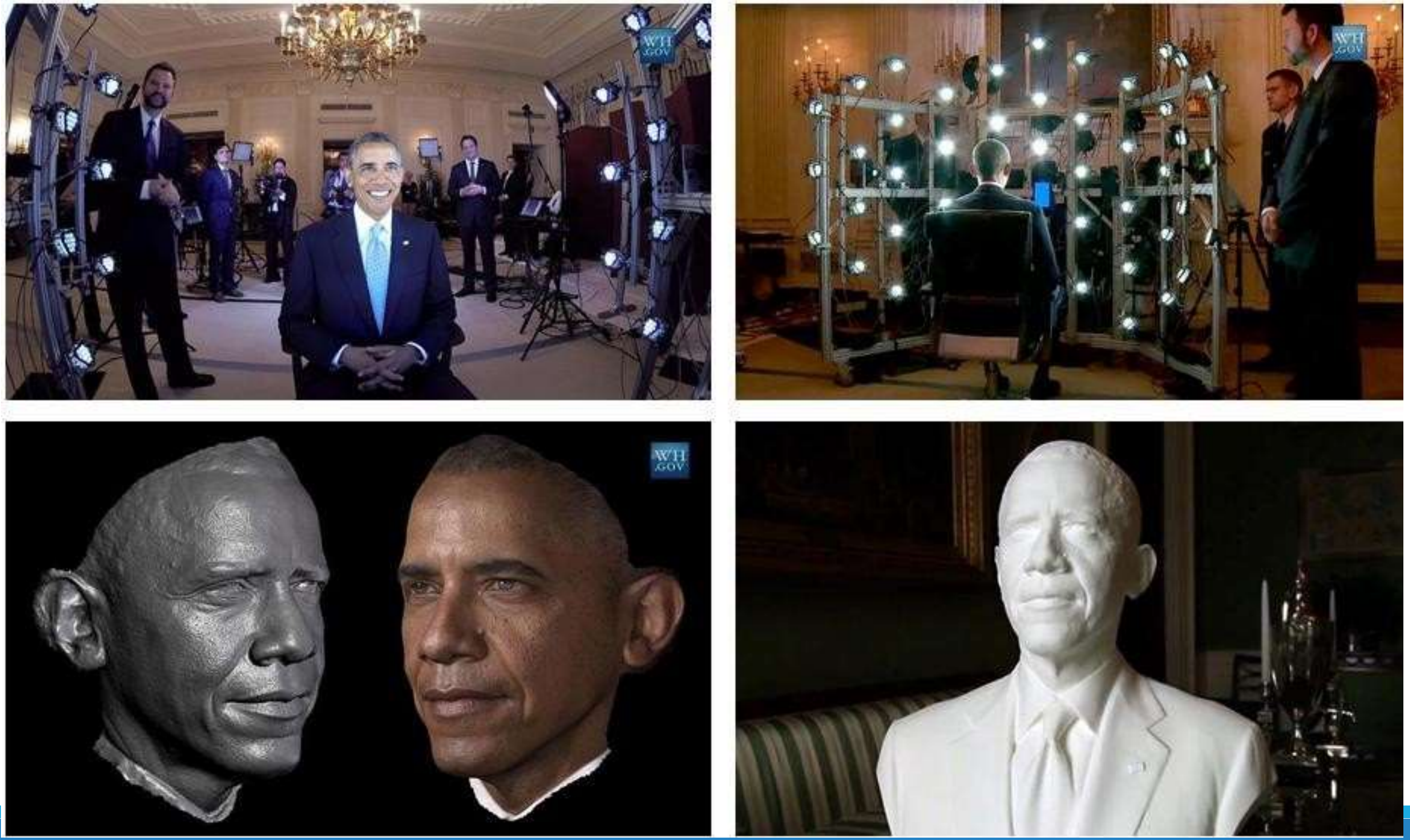
Apple Face ID



Face recognition systems now beginning to appear more widely http://www.sensiblevision.com/

Computer Vision Applications

- Human Shape Capture

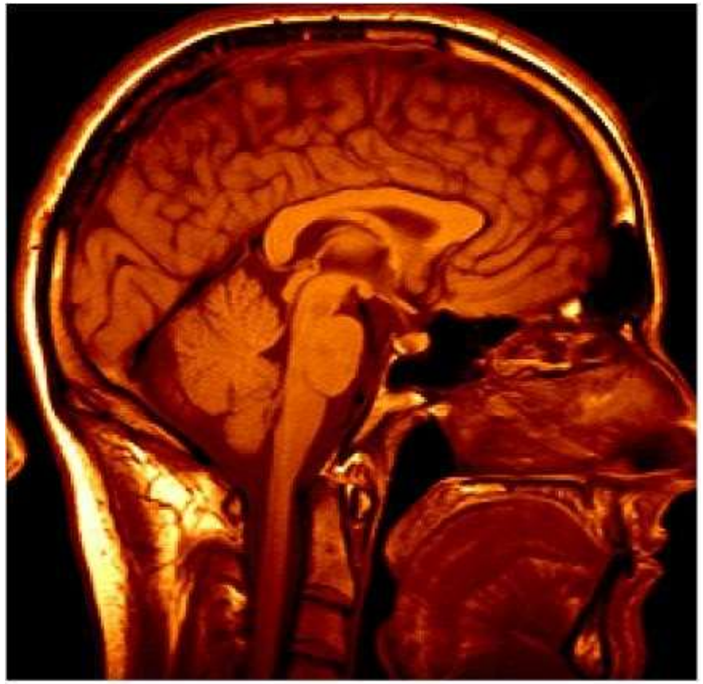


Computer Vision Applications

- Motion Capture

Computer Vision Applications

- Medical Imaging



3D imaging
MRI, CT

Image guided surgery

Computer Vision Applications

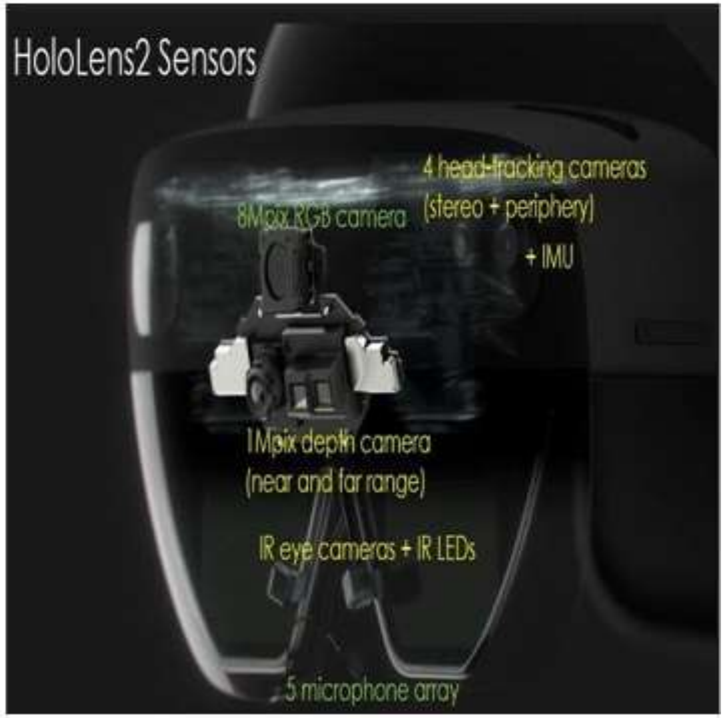
- Virtual Reality

Oculus Quest

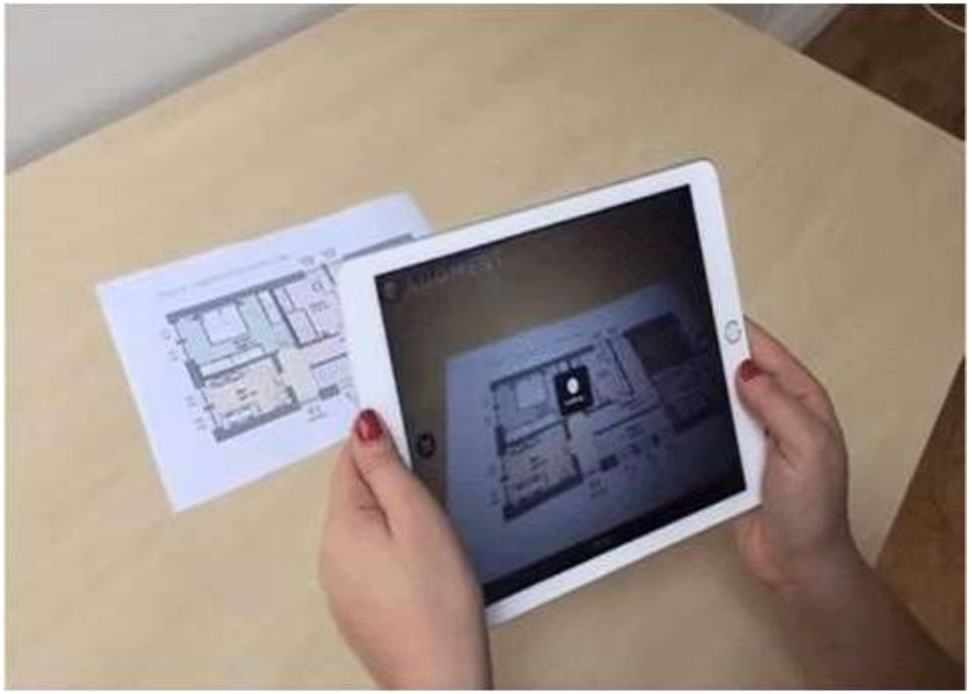
Beat Saber

Computer Vision Applications

- Augmented Reality



Microsoft Hololens 2



Computer Vision Applications

- Body Tracking

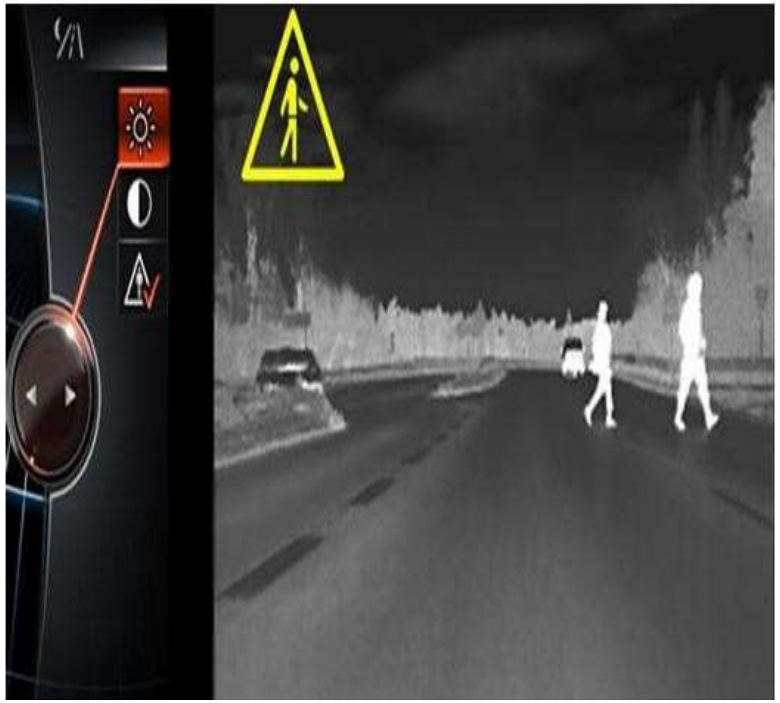
ARKit 3

Computer Vision Applications

- Smart Cars

Mobileye

Vision systems currently in high end BMW, GM, Volvo models



BMW night vision

Computer Vision Applications

- AutoCars: Self Driving Cars

https://waymo.com/tech/

Computer Vision Applications

- Industrial Robots
  - ✓ Vision-guided robots position nut runners on wheels

Computer Vision Applications

- Human Detection

Computer Vision Applications

- Video Surveillance and Monitoring

Computer Vision Applications

- Forest fire monitoring system

Early warning of forest fires

Adapted from Enis Cetin, Bilkent University

Computer Vision Applications

- Counting in Extremely Dense Crowd Images

Ground truth=634

Proposed Method by Idrees and Shah=640

Computer Vision Applications

- Counting in Extremely Dense Crowd Images

Ground Truth = 1428

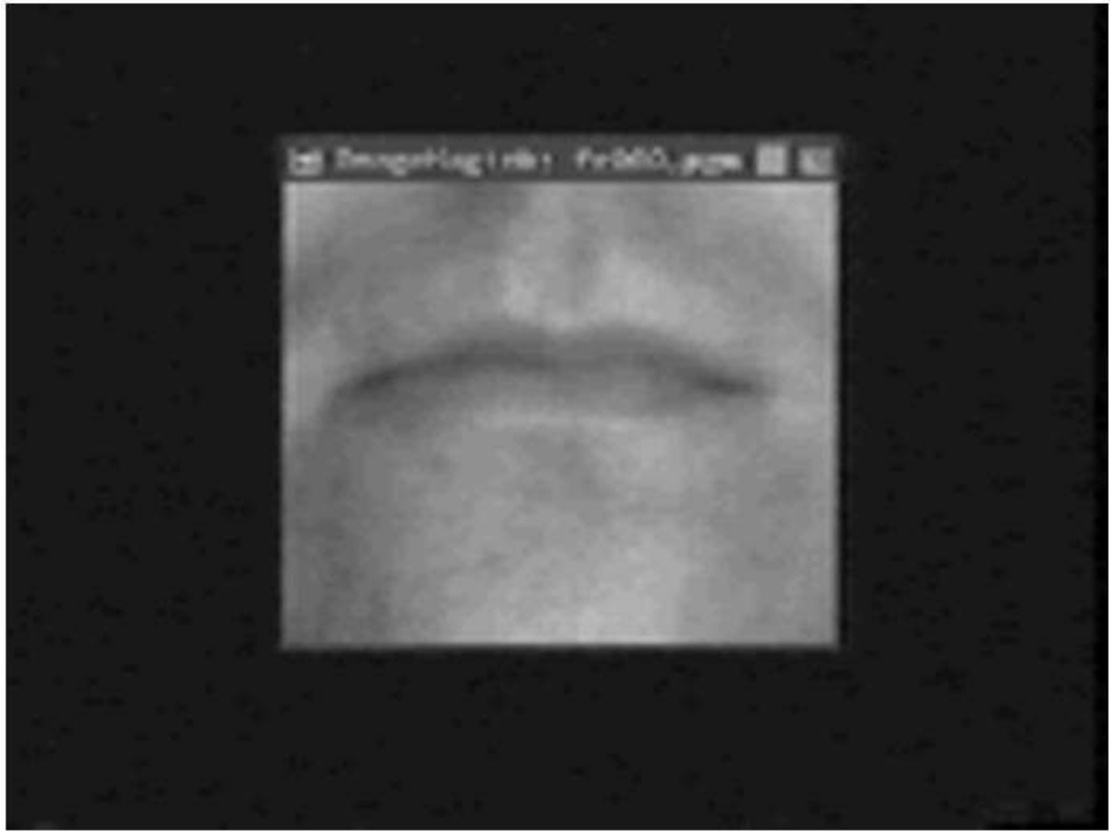
Proposed Method = 1468

Computer Vision Applications

- Fatigue detection

Computer Vision Applications

- Lip reading



Computer Vision Applications

- Human Action Recognition

9 actions, 142 videos (UCF database, Shah).

Bench Swing

Dive

Swing

Run

Kick

Lift

Ride

Golf Swing

Skate

Computer Vision Applications

- Morphing

It is a special effect in motion pictures and animations that changes (or morphs) one image or shape into another through a seamless transition.

Neural Style Transfer [Gatys et al. 2015]

Python Image Processing Libraries

OpenCV: Widely used for computer vision tasks, provides functions for manipulation, filtering, feature detection, and object recognition.

Pillow: Easy-to-use for basic tasks like resizing, cropping, rotating, and applying filters; supports multiple image formats.

scikit-image: Collection of algorithms for tasks such as segmentation, feature extraction, and image restoration.

Mahotas: Offers functions for filtering, feature extraction, morphology, and more.

SimpleITK: Interface to ITK for medical image analysis, including registration and segmentation.

PyTesseract: Wrapper for Google's Tesseract-OCR Engine, used for optical character recognition.

imageio: Supports reading/writing various image formats and animated images.

matplotlib/seaborn: Notable for data visualization but can be utilized for basic image visualization and manipulation.

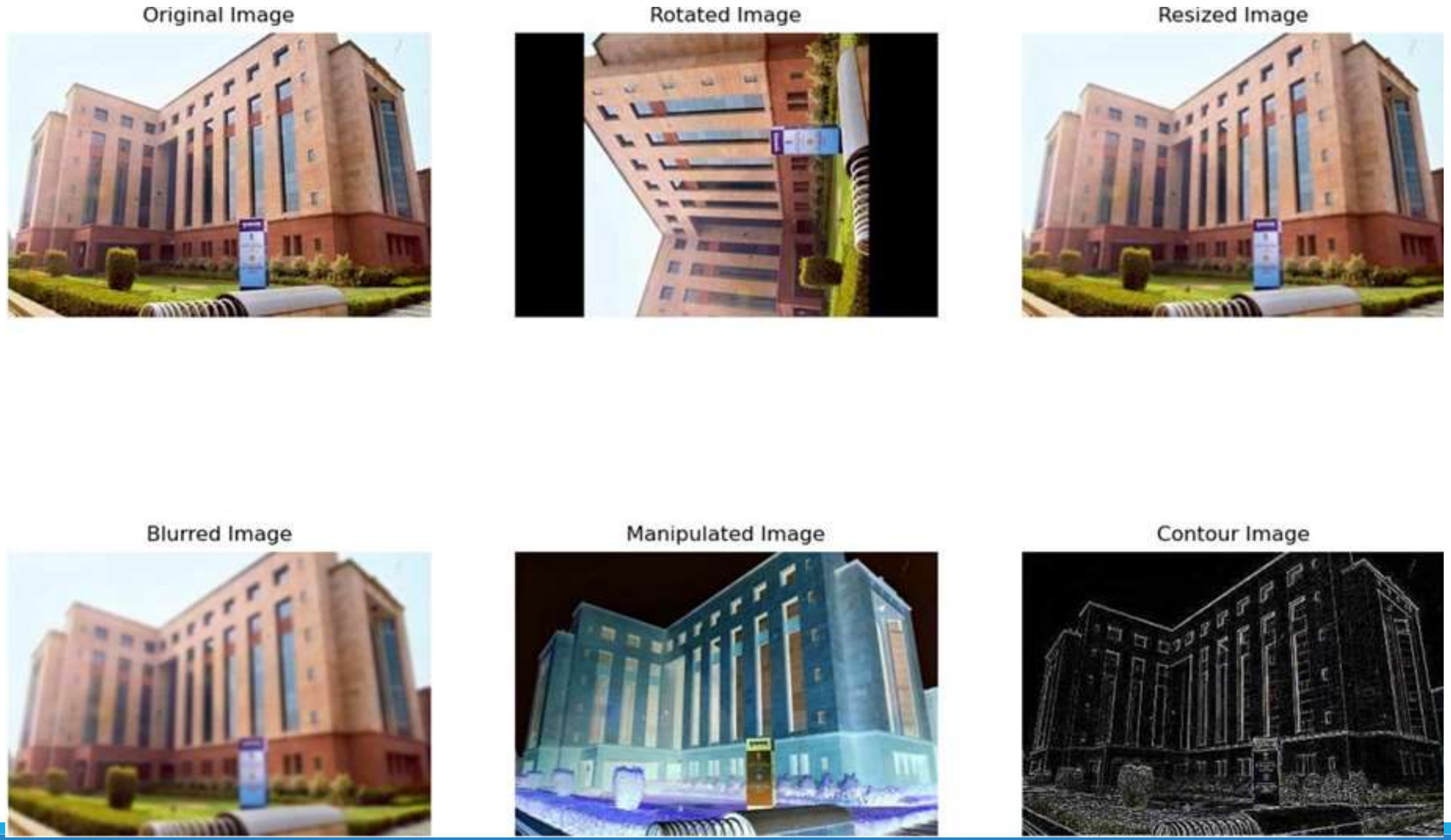
GDAL (Geospatial Data Abstraction Library): Powerful for working with raster and vector geospatial data

Working with Images

```python
from PIL import Image, ImageFilter,ImageOps
import numpy as np
import matplotlib.pyplot as plt

# Open an image file
image = Image.open("test.jpg")

# Convert the image to a NumPy array
image_array = np.array(image)

# Manipulate the pixel values (e.g., invert colors)
manipulated_image_array = 255 - image_array  # Inverts colors

# Convert the manipulated NumPy array back to an image
manipulated_image = Image.fromarray(manipulated_image_array)

# Rotate the image by 90 degrees clockwise
rotated_image = image.rotate(90)

# Resize the image to half its original size
resized_image = image.resize((image.width // 2, image.height // 2))

# Apply a Gaussian blur filter to the image
blurred_image = image.filter(ImageFilter.GaussianBlur(2))

# Apply a contour filter to the image
contour_image = ImageOps.invert(image.filter(ImageFilter.CONTOUR))
```

```python
# Display all images
fig, axes = plt.subplots(2, 3, figsize=(15, 10))
fig.suptitle('Image Processing with Pillow', fontsize=16)

axes[0, 0].imshow(image)
axes[0, 0].set_title('Original Image')
axes[0, 0].axis('off')

axes[0, 1].imshow(rotated_image)
axes[0, 1].set_title('Rotated Image')
axes[0, 1].axis('off')

axes[0, 2].imshow(resized_image)
axes[0, 2].set_title('Resized Image')
axes[0, 2].axis('off')

axes[1, 0].imshow(blurred_image)
axes[1, 0].set_title('Blurred Image')
axes[1, 0].axis('off')

axes[1, 1].imshow(manipulated_image)
axes[1, 1].set_title('Manipulated Image')
axes[1, 1].axis('off')

axes[1, 2].imshow(contour_image)
axes[1, 2].set_title('Contour Image')
axes[1, 2].axis('off')

plt.show(block=True)
```

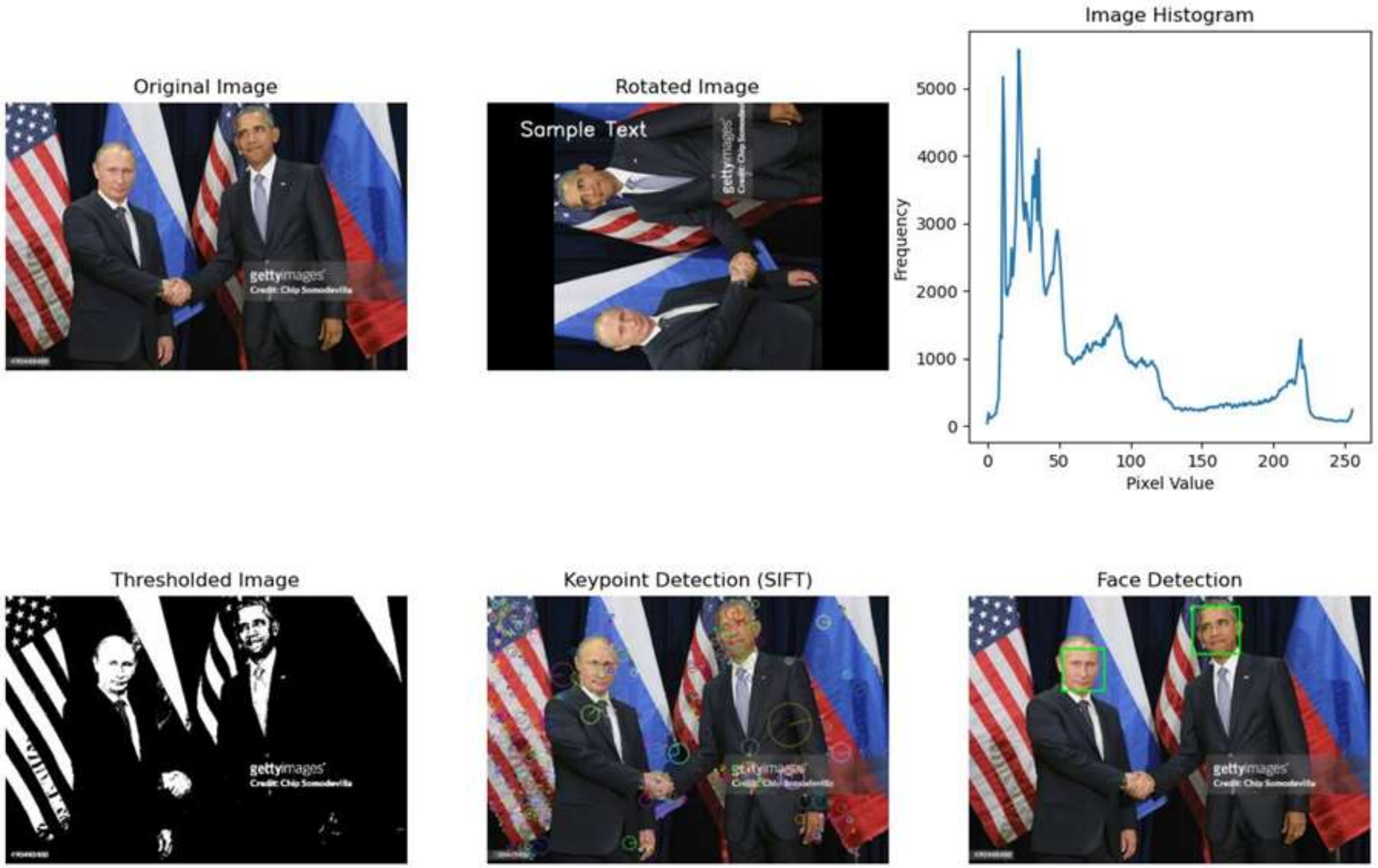
Image Processing



OpenCV

```
import cv2
import numpy as np
import matplotlib.pyplot as plt
# Open an image file
image = cv2.imread("test2.jpg")
orig_image=image.copy()
# Convert the image to grayscale
gray_image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

# Initialize the SIFT detector
sift = cv2.SIFT_create()
# Detect keypoints in the image
keypoints = sift.detect(gray_image, None)
# Draw keypoints on the original image
image_with_keypoints = cv2.drawKeypoints(image, keypoints, None, flags=cv2.DRAW_MATCHES_FLAGS_DRAW_RICH_KEYPOINTS)

# Apply thresholding to the image
_, thresholded_image = cv2.threshold(gray_image, 127, 255, cv2.THRESH_BINARY)

# Rotate the image by 90 degrees clockwise
height, width = image.shape[:2]
rotation_matrix = cv2.getRotationMatrix2D((width / 2, height / 2), 90, 1)
rotated_image = cv2.warpAffine(image, rotation_matrix, (width, height))
# Add text on top of the rotated image
image_with_text = cv2.putText(rotated_image, 'Sample Text', (50, 50), cv2.FONT_HERSHEY_SIMPLEX, 1, (255, 255, 255), 2)

# Calculate histogram of the grayscale image
hist = cv2.calcHist([gray_image], [0], None, [256], [0, 256])

# Apply face detection to the image
# Load the pre-trained face detection cascade
face_cascade = cv2.CascadeClassifier(cv2.data.haarcascades + 'haarcascade_frontalface_default.xml')
faces = face_cascade.detectMultiScale(gray_image, scaleFactor=1.1, minNeighbors=5, minSize=(30, 30))
# Draw bounding boxes around detected faces
for (x, y, w, h) in faces:
    cv2.rectangle(image, (x, y), (x+w, y+h), (0, 255, 0), 2)
```

OpenCV



Original Image
gettyimages
Credit: Chip Somodevilla
Rotated Image
Sample Text
Image Histogram
Frequency
Pixel Value
0
1000
2000
3000
4000
5000
0
50
100
150
200
250
Thresholded Image
Keypoint Detection (SIFT)
Face Detection

Pytesseract

- https://github.com/UB-Mannheim/tesseract/wiki
- pip install pytesseract


```python
import pytesseract
from PIL import Image
import matplotlib.pyplot as plt

# Path to the Tesseract executable (change this according to your system)
pytesseract.pytesseract.tesseract_cmd = r'C:\Program Files\Tesseract-OCR\tesseract.exe'

# Load the image
image_path = 'ocr.png'
image = Image.open(image_path)

# Perform OCR on the image
text = pytesseract.image_to_string(image)

# Display the image
plt.figure(figsize=(8, 6))
plt.imshow(image)
plt.axis('off')

# Display extracted text on top of the image
plt.text(0, 0, text, color='red', fontsize=12, verticalalignment='top')

plt.show()
```

Assignment-1: Working with GeoTIFF Images Using GDAL

- GeoTIFF images encode raster data with geospatial metadata, enabling precise linking of each pixel to specific latitude and longitude coordinates. These images are widely used in GIS (Geographic Information System) applications for spatial analysis and visualization.
- You are provided with a sample GeoTIFF file (or download one of your choice).

Assignment-1: Working with GeoTIFF Images Using GDAL

- Your task is to perform the following steps using the GDAL library and a GUI framework like PyQt or Tkinter:

Task-A: Load and Display the Satellite Image

- Load the GeoTIFF file using the GDAL library.
- Display the satellite image in a window.

Task-B: Interactive Mouse Tracking

- Create a GUI that displays the image. As the user moves the mouse cursor over the image:
  - Show the pixel coordinates (X,Y) of the cursor in two text boxes.
  - Display the corresponding latitude and longitude of the pixel in two additional text boxes.
  - Update this information in real time.

Task-C: Marking a Specific Location

Add two input text boxes where users can enter latitude and longitude values.

Include a "Mark" button in the GUI:

When clicked, the button should plot a cross at the specified location on the displayed image based on the entered latitude and longitude.

Task-D (Bonus): Zoom and Pan Functionality

- Enhance the GUI with:
  - Zooming functionality using the mouse wheel.
  - Panning functionality via mouse drag.

Thank you