2110431 INTRODUCTION TO DIGITAL IMAGING
2147329 DIGITAL IMAGE PROCESSING AND VISION SYSTEMS

Punnarai Siricharoen, Ph.D.

LECTURE 01 INTRODUCTION

Punnarai Siricharoen, Ph.D.

LECTURE 01 - OBJECTIVES

1. To introduce Digital Imaging and Image processing overview
2. To describe the example fields areas of digital imaging and applications
3. To use python for read/write/show an image

LECTURE 01 - OUTLINE

1. About this course (Syllabus)
2. Introduce Digital Imaging, applications and Challenges
3. Principal areas of digital imaging and applications
4. Python Exercises for read/write/show an image

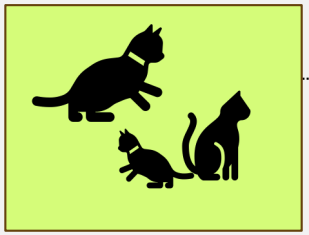
ABOUT THE COURSE

COMMUNICATIONS

- Onsite Lecture at ENG100, room 201B + recordings
- MCV course number : 2110431 or 2147329
- password : imaging2025 for course materials
- Discord : https://discord.gg/YyErcCMj

WHAT **TOPICS / IDEAS** ARE YOU WANT TO LEARN IN THIS CLASS?

IMAGE PROCESSING / DIGITAL IMAGING



cat

Image Generation

Caption: There are three cats in the image.

Image Captioning

How many cat in the image?

Visual Question Answering

Image Improvement

- Enhancement
- Superresolution
- Restoration

Image Understanding

- Feature Extraction
- Object detection
- Object Recognition
- Image Segmentation
- Interpretation/Analysis

Storage

Compression

Printing/display

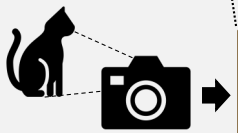


Image Acquisition

Challenges?????

Medical imaging

Image from Factory

Entertainment

Not to mention image processing in terms of computer graphics which is another topic.

TENTATIVE SCHEDULE

Week / Date	Topic	Assignment HW (2 weeks due)
1 – 6/8	Introduction to image processing and software preparation, Examples of fields using digital image processing	
2 – 14/8	Image Acquisition / Sampling and Quantization / Arithmetic Operations	
3 – 20/8	Histogram and Spatial Filtering	HW #1 Spatial Domain
4 – 27/8	Filtering in Frequency Domain	
5 – 3/9	Wavelet and Multiresolution Processing	
6 – 10/9	Image Restoration	HW #2 Frequency & Research paper
7- 17/9	Image Compression	
8-25-29/9	Midterm exam (1 A4 paper, writing) Wed 24/9/2568 13-16 (CP – Lecture room, AI&R – Faculty)	
9- 1/10	Graduation (No class)	
10 - 8/10	Fundamental Image Segmentation I	Project Assignment
11-15/10	Morphology	
12-22/10	Object Recognition (traditional /modern)	
13-29/10	Object Recognition II	HW #3 Segmentation + Image Augmentation
14-5/11	Image Segmentation II / Object Detection	
15-12/11	Image Generation and other Imaging Applications – discussion (read and discuss)	
16-19/11	Invited Speaker	
17	- Final exam Wed 3/12/2568 13-16 (CP – Lecture room, AI&R – Faculty) - Project Presentation 11/12/2568 AI & Robotics:: 9:00-10:30 (?) 11/12/2568 Com Eng: 10:30-16:00 (?)	

IMPORTANT DATES

- Midterm – Wed 24/9/2568 13-16 (CP – Lecture room, AI&R – Faculty)
- Final exam – Wed 3/12/2568 13-16 (CP – Lecture room, AI&R – Faculty)
- Project presentation – Tue 11/12/2568

You cannot take the exam outside the predefined date & time!

TEACHING METHODS

- Lecture (+ in-class Lab) : Wednesday 13:00 – 16:00

- Research discussion
- Projects

DIGITAL IMAGING

- **Textbook:**

  - Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing, Addison-Wesley

- **Grading Policy**
  - In-class Exercises: **5% (0.5 each class + bonus)**
  - Homework = **10%**
  - Project + presentation : **25%**
  - Midterm Exam : **30%**
  - Final Exam : **30%**

Score	Grade
>= 80	A
>= 75	B+
>= 70	B
>= 65	C+
>= 60	C
>= 55	D+
>= 50	D
< 50	F

CONTACT ME

Punnarai Siricharoen, Ph.D.

Department of Computer Engineering, Faculty of Engineering

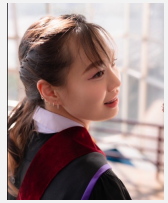
Room: 19th floor (19-05) Engineering 4 Building (Charoenvidsavakham)

Email: punnarai.s@chula.ac.th

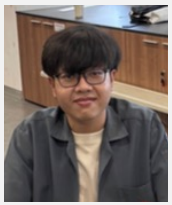
TAs:

TA Print

TA Tar



TA Liu

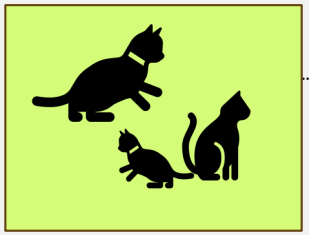


TA Nick

WHAT IS DIGITAL IMAGING / IMAGE PROCESSING FOR YOU?

Write a description / draw a picture
in a paper

IMAGE PROCESSING / DIGITAL IMAGING



cat

Image Generation

Caption: There are three cats in the image.

Image Captioning

How many cat in the image?

Visual Question Answering

Image Improvement

- Enhancement
- Superresolution
- Restoration

Image Understanding

- Feature Extraction
- Object detection
- Object Recognition
- Image Segmentation
- Interpretation/Analysis

Storage

Compression

Printing/display

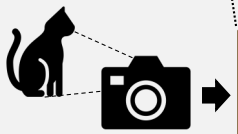


Image Acquisition

Challenges?????

Medical imaging

Image from Factory

Entertainment

Not to mention image processing in terms of computer graphics which is another topic.

DIGITAL IMAGE

170	172	171	174	178	183	185	183	182	181
170	172	171	173	178	182	184	182	182	181
170	171	171	172	177	182	184	182	182	182
170	170	171	172	176	182	183	182	183	182
170	169	171	171	175	182	182	183	183	182
169	167	171	170	175	183	182	183	183	182
168	166	170	169	174	183	181	183	183	182
168	165	170	168	173	183	181	183	182	182
171	168	169	172	176	180	181	184	181	179
168	170	172	170	174	182	183	181	181	179

What we see

What a computer sees

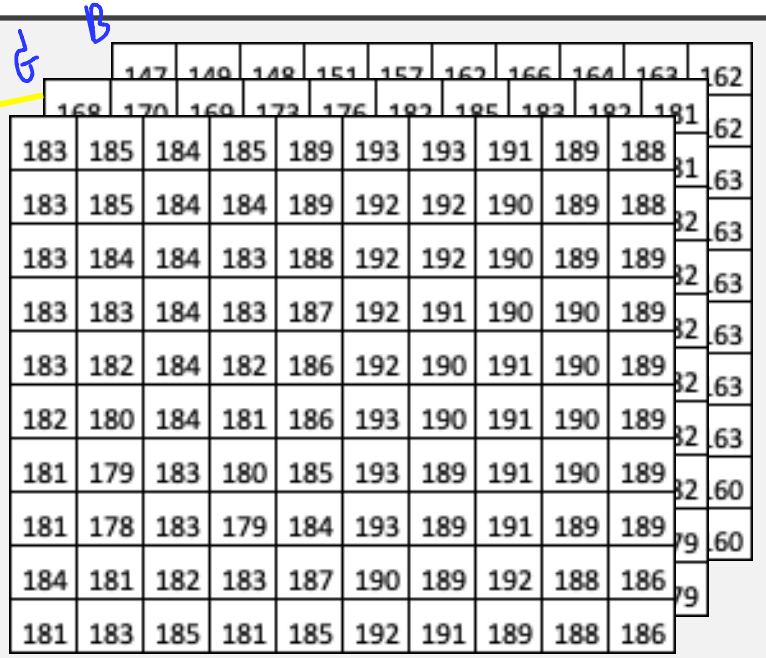
An image is denoted by two-dimensional function of the form: $f(x, y)$

f – amplitude at the spatial dimension (x,y)

color position

DIGITAL IMAGE

What we see



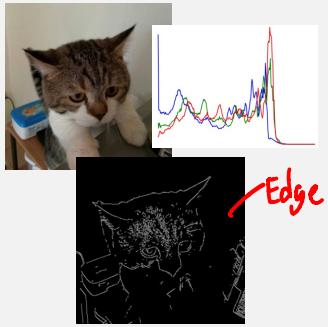
What a computer sees

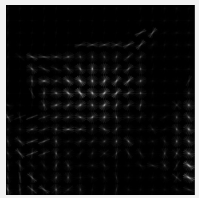
EVOLUTION

1980s

Rule-based

Edge, Threshold, Template





HOG

2000s

Feature-based

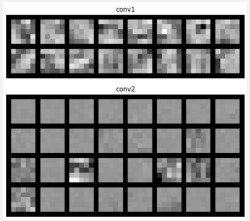
Haar, HOG, SIFT

Based on

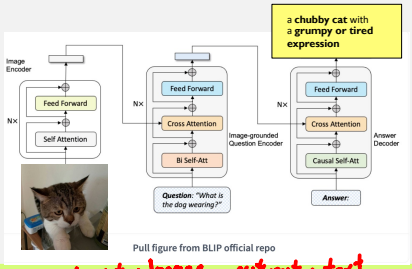
2012+

Deep Learning

CNN, ResNet, YOLO



feature



Pull figure from BLIP official repo

Input : Image output : text

2020+

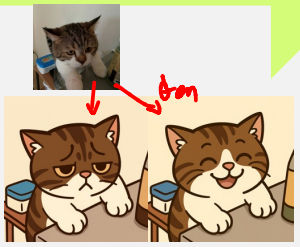
Multimodal / Vision + Language

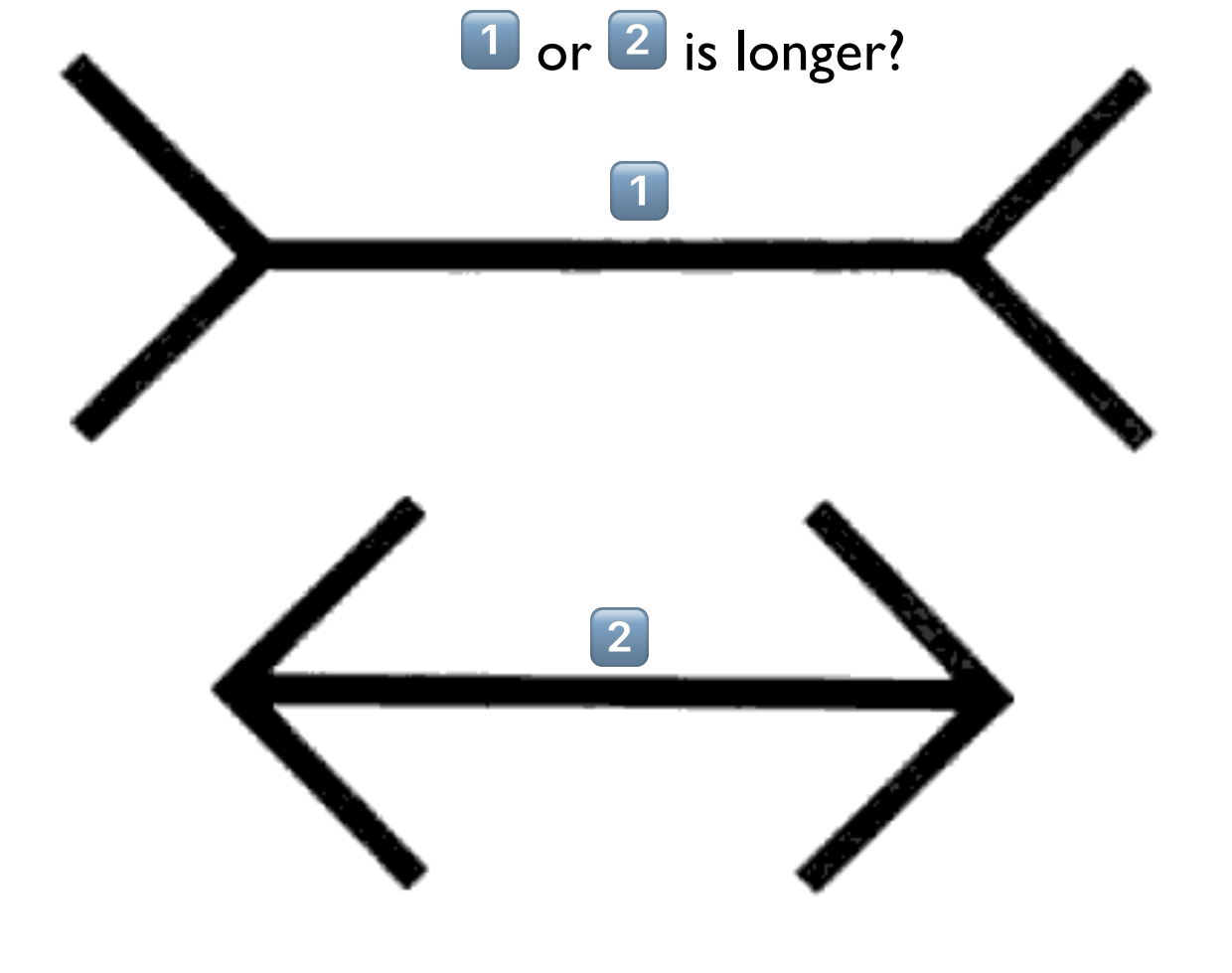
CLIP, BLIP, Flamingo, GPT-4o

2023+

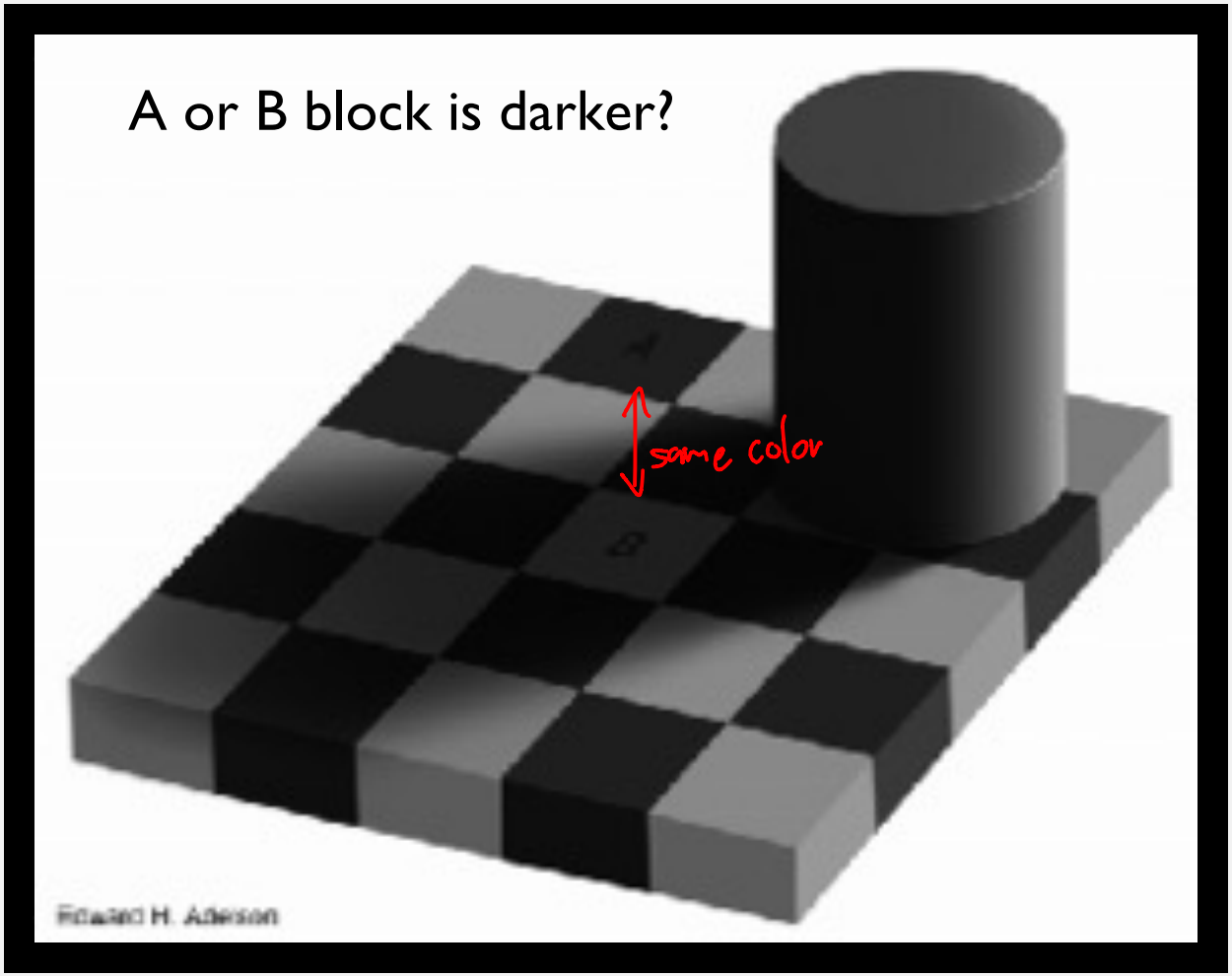
Vision + Generation + Action

DALL·E, SAM, Robotics with vision





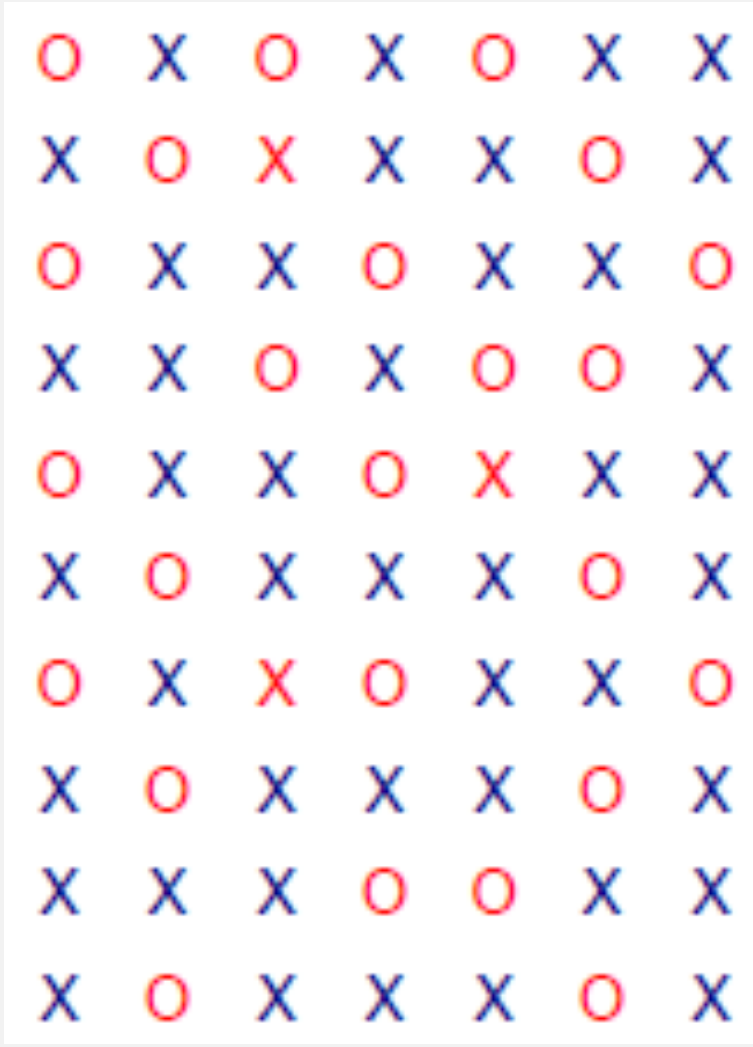
Müller-Lyer illusion



Checker Shadow Illusion by Edward Adelson

How many red X?

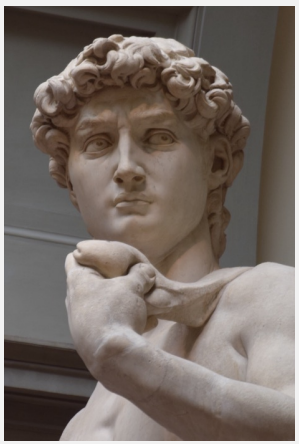
How many red X?

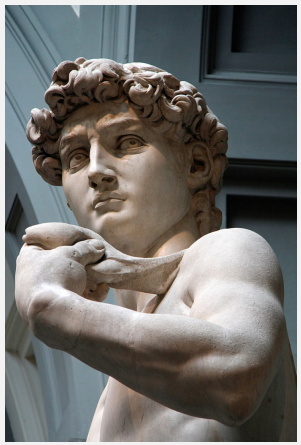


WHAT ARE CHALLENGES FOR COMPUTER TO SEE THINGS???

CHALLENGES

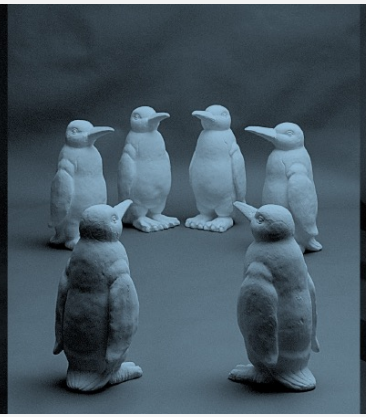
- Point of View





- Illumination

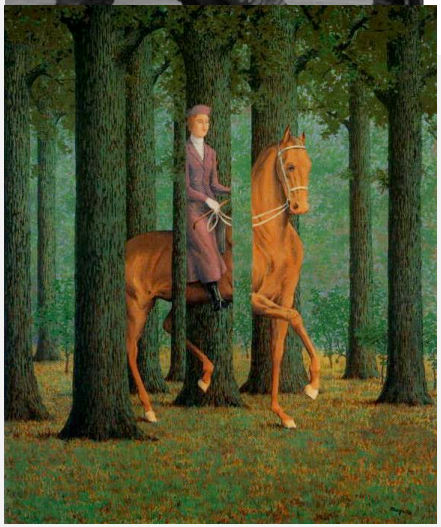
light
↓
change color

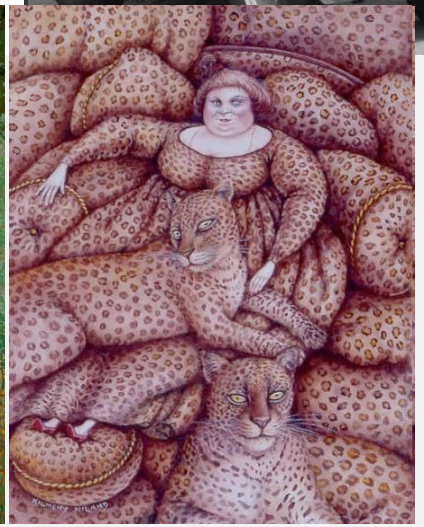




Wikipedia, Academic Gallery, Pinterest
slide credit: Fei-Fei, Fergus & Torralba

CHALLENGES

- Scale
- Deformation
- Occlusion
- Background Clutter



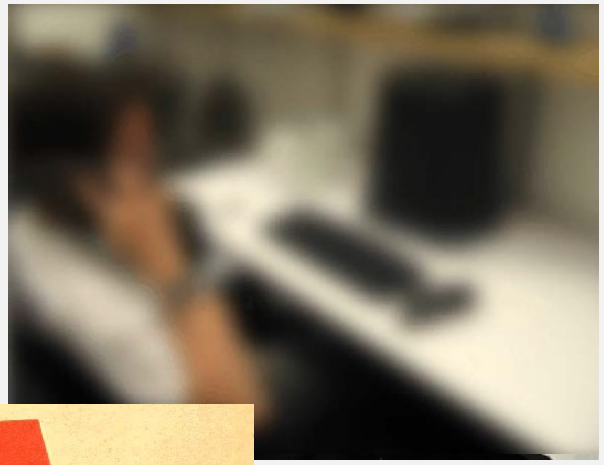


slide credit: Fei-Fei, Fergus & Torralba

CHALLENGES

- Ambiguity
- Motion
- In-class Variations

รูปหมวกแต่เป็นเก้าอี้หมด

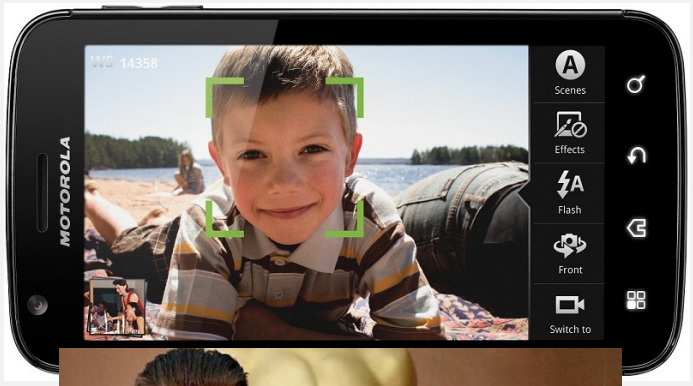


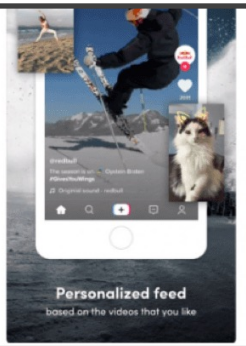
slide credit: Fei-Fei, Fergus & Torralba

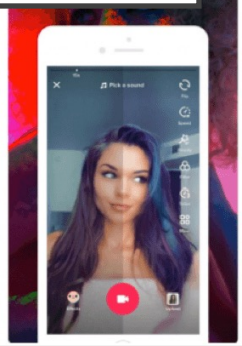
HAVE YOU EXPERIENCED IMAGE PROCESSING IN YOUR DAILY LIFE?

FACE DETECTION

- Imaging applications in Daily Life
  - Face Detection

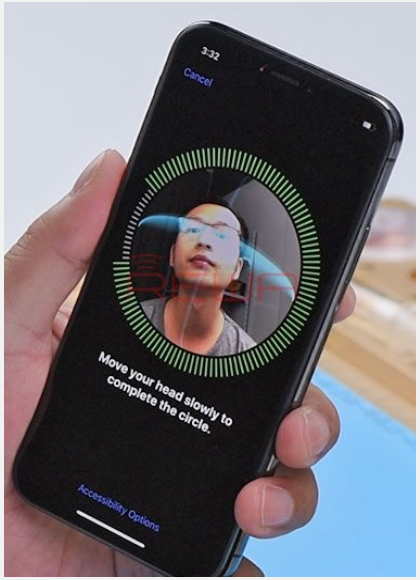






BIOMETRICS APPLICATIONS

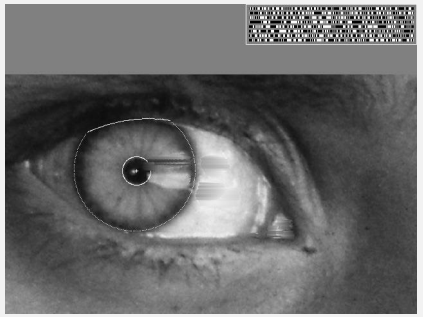
- Imaging applications in Daily Life
  - Biometrics

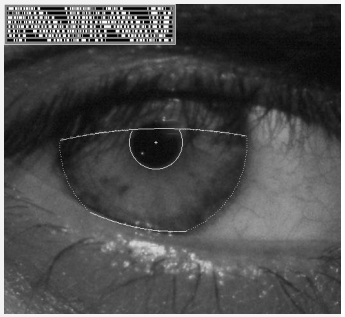


https://blog.rewatechnology.com/fix-iphone-x-face-id-not-working/

1984 2002

"*How the Afghan Girl was Identified by Her Iris Patterns*" https://www.cl.cam.ac.uk/~jgd1000/afghan.html





HANDS-FREE SELFIE

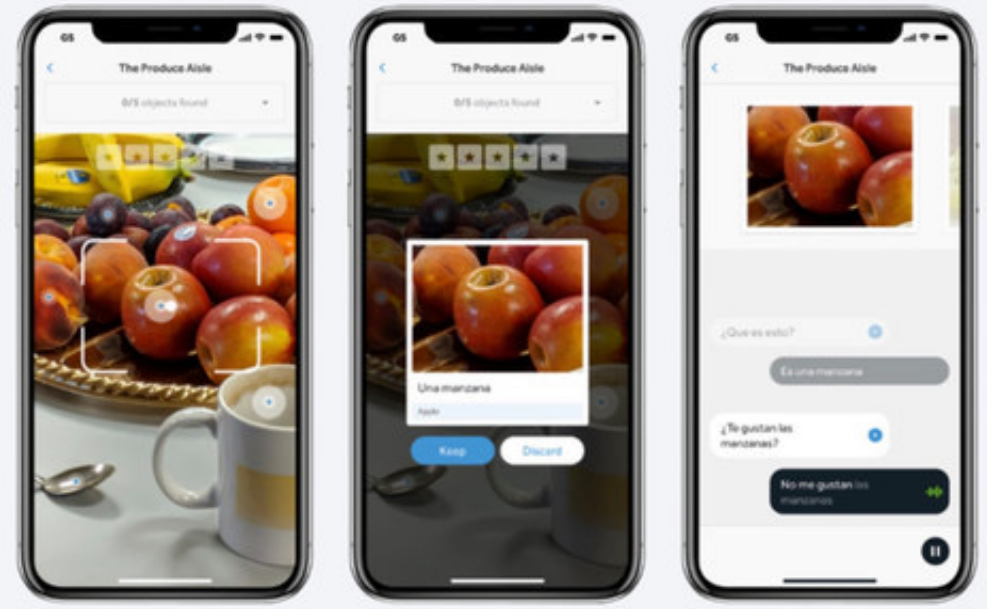
Handsome guy photo created by cookie_studio - www.freepik.com

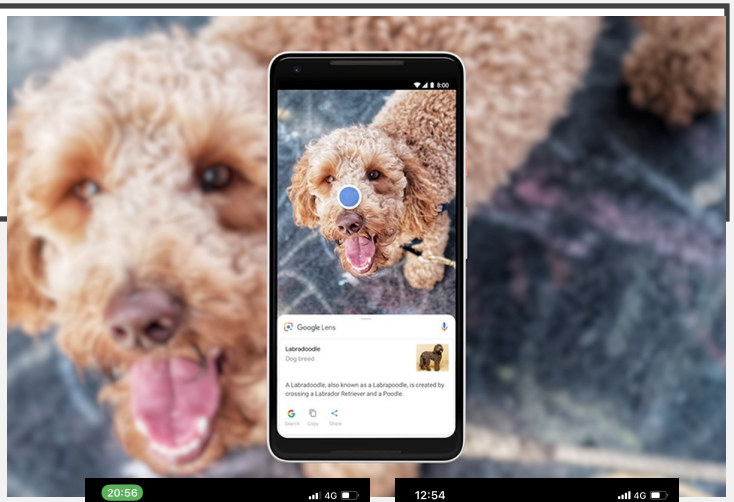
Take a HANDS-FREE selfie: App snaps photos from up to 10ft away using a simple swipe of the arm

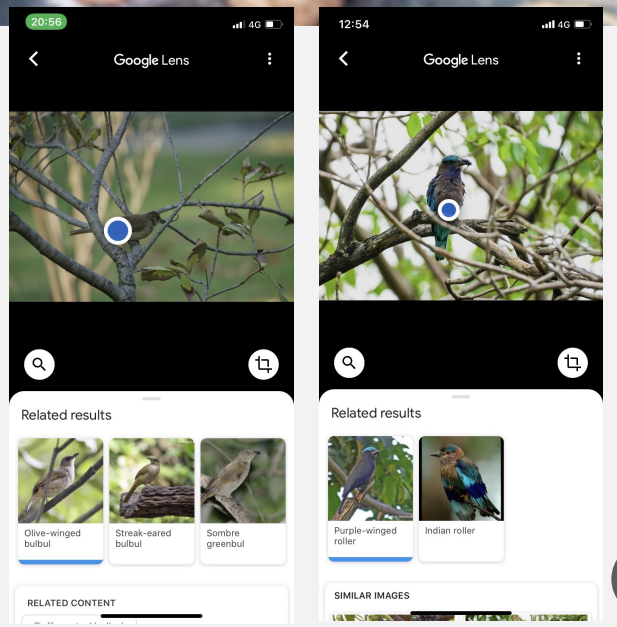
https://www.dailymail.co.uk/sciencetech/article-2674946/Take-HANDS-FREE-selfie-App-snaps-photos-10ft-away-using-simple-swipe-arm.html

OBJECT CLASSIFICATION

- Imaging applications in Daily Life
  - Object Recognition on your mobile phones







OCR – OPTICAL CHARACTER RECOGNITION

อาจารย์คณะวิศวฯ จุฬาฯ นำเทคโนโลยี AI Deep Tech พัฒนาโปรแกรมสแกนเอกสารและรูปภาพเป็นข้อความ (OCR) อ่านภาษาไทยแม่นยำกว่า 90% UTC จุฬาฯ พร้อม spin-off สู่ตลาดในนามบริษัท Eikonnex AI จำกัด

Assoc. Prof. Dr. Thanarat Chalidabhonges

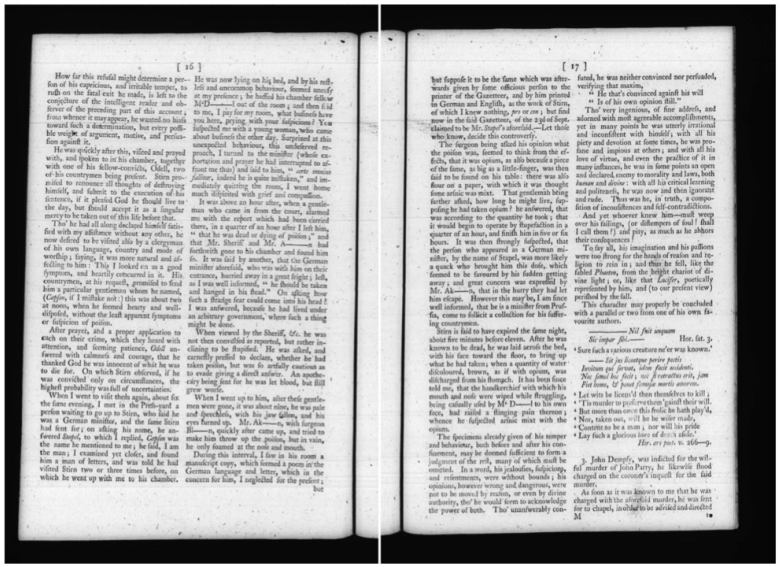
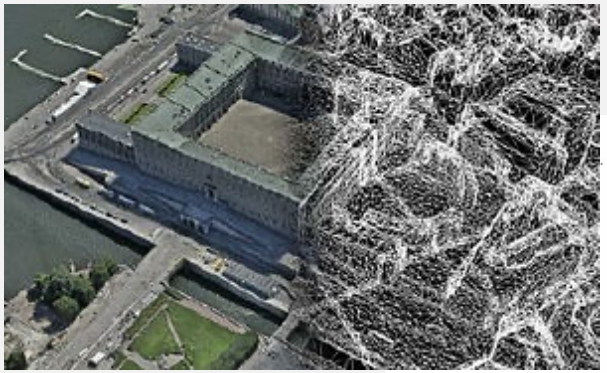


Figure 2: Ordinary's Accounts September 15, 1760 (# 3421-2)

Character Error Rate	**Left: AWS 73.97%; Azure 35.26%; GCP 76.64%**
	Right: AWS 75.90%; Azure 12.17%; GCP 77.36%

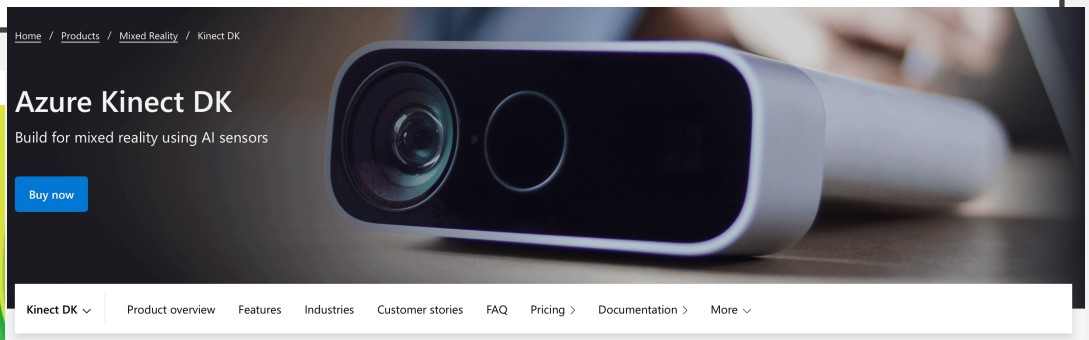
The Old Bailey and OCR: Benchmarking AWS, Azure, and GCP with 180,000 Page Images, William Ughetta, Brian W. Kernighan, Proceedings of the ACM Symposium on Document Engineering 2020

3D FROM IMAGES



https://grail.cs.washington.edu/rome/
Building Rome in a Day: Agarwal et al. 2009

KINECT



Kinect with spatial data

Azure Kinect is a cutting-edge spatial computing developer kit with sophisticated computer vision and speech models, advanced AI sensors, and a range of powerful SDKs that can be connected to Azure cognitive services.

Using Azure Kinect, manufacturing, retail, healthcare, and media enterprises are leveraging spatial data and context to enhance operational safety, increase performance, improve outcomes, and revolutionize the customer experience.

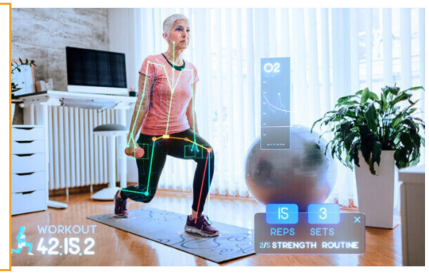
Health | Manufacturing | Media | Retail | Robotics

Xbox 360 Kinect claims a Guinness World Record for the fastest-selling cor
after selling eight million units in its first 60 days.

Release timeline

2006	PrimeSense technology shown at GDC
2007	
2008	
2009	"Project Natal" announced
2010	**Kinect for Xbox 360** released
2011	Non-commercial SDK released Commercial SDK released
2012	**Kinect for Windows** released
2013	**Kinect for Xbox One** released with console
2014	**Kinect 2 for Windows** released
2015	
2016	
2017	Discontinuation of Kinect for Xbox hardware
2018	
2019	**Azure Kinect** released

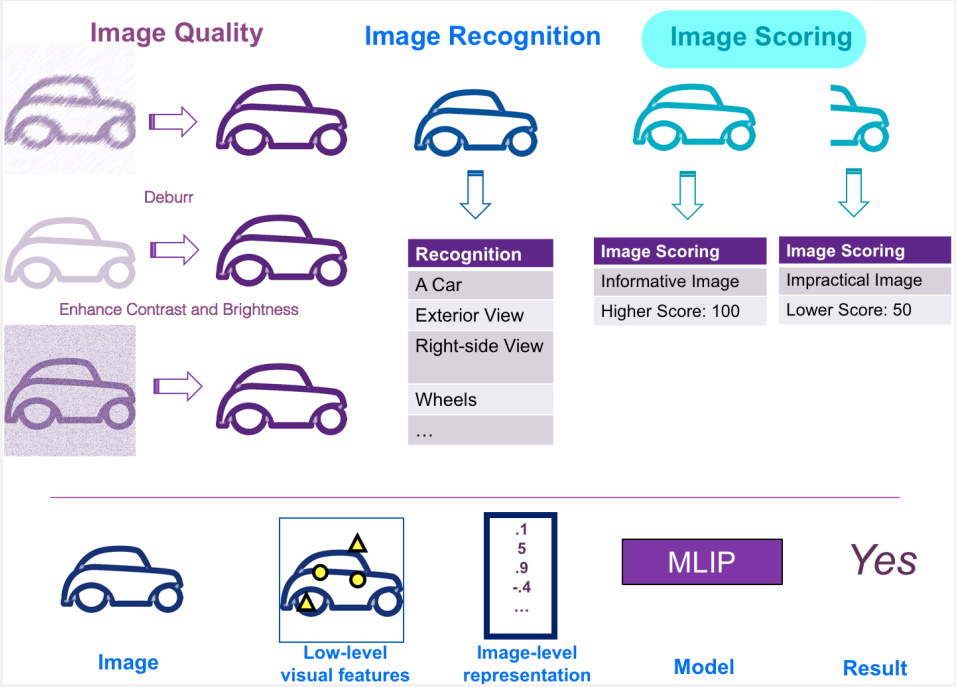
- 1-MP depth sensor
- 7-microphone array
- 12-MP RGB video camera
- Accelerometer and gyroscope (IMU) (Orientation & spatial tracking)
- External sync pins



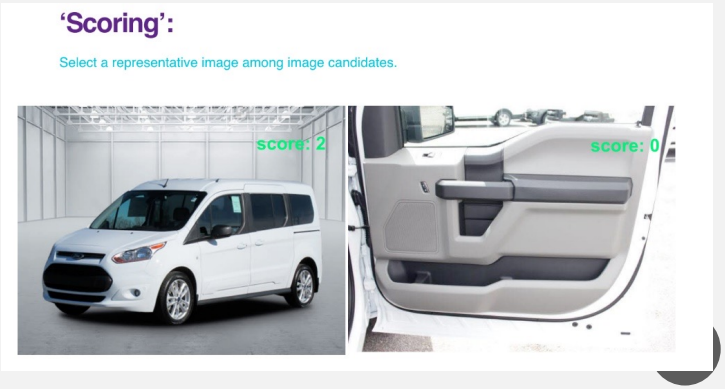
https://news.microsoft.com/en-gb/announcement/xbox-360-kinect-launched/

https://azure.microsoft.com/en-us/services/kinect-dk/#industries

DIGITAL MARKETING



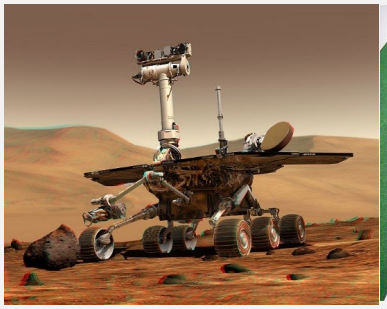
How to choose the most representative one among these image candidates needs a strategic image scoring schema

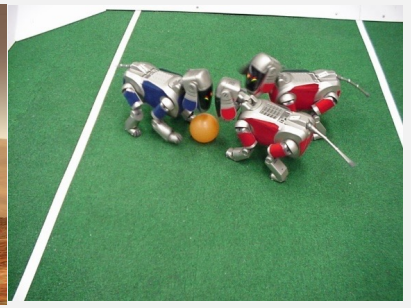


https://tech.cars.com/applications-of-machine-learning-image-processing-in-digital-marketing-982ee296dc8a

OTHERS

- Other Applications
  - Augmented Reality And Virtual Reality
  - Mobile Robots
  - Industrial Robots
  - Autocars –Tesla / Uber Bought Cmu's Lab
  - Medical Imaging





NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover
http://www.robocup.org/

Vision-guided robots position nut runners on wheels

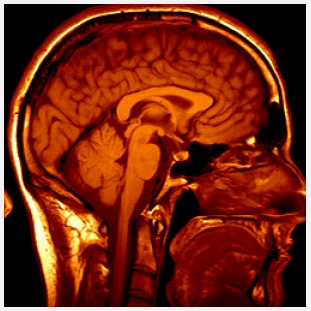


Image guided surgery Grimson et al., MIT
3D imaging MRI, CT