TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA KHOA HỌC MÁY TÍNH

CS231.N2x. Nhập môn Thị giác máy tính

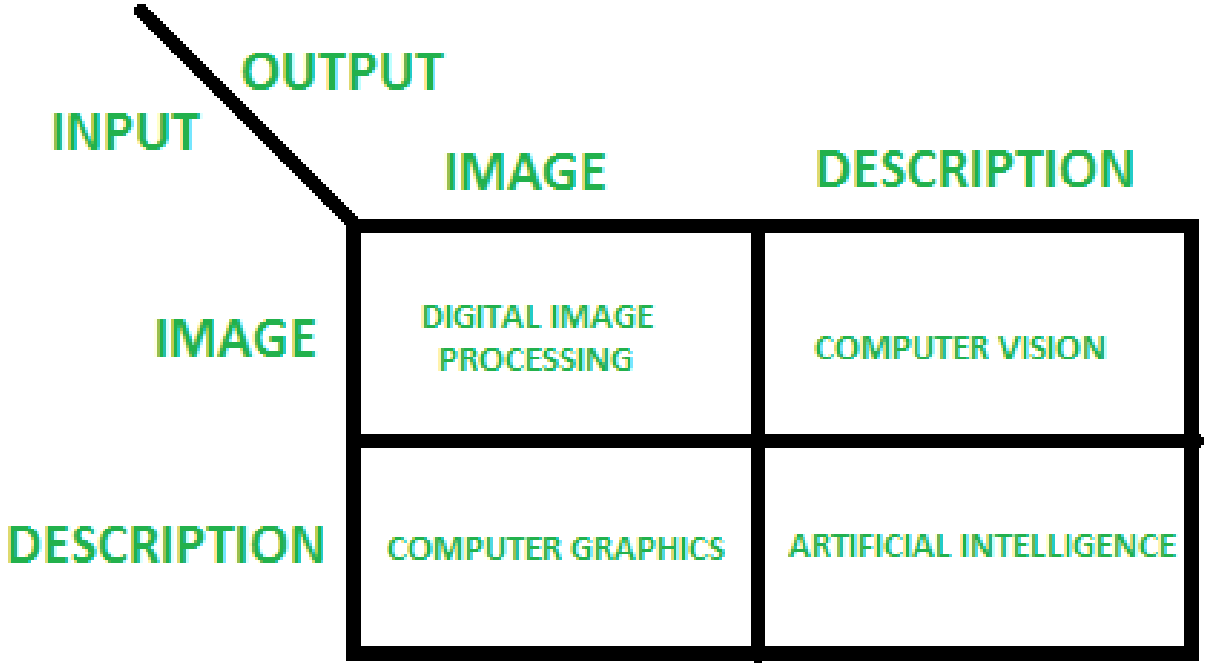
Bài 01
Giới thiệu môn học

Mai Tiến Dũng

1. Thông tin chung

Tên môn học (tiếng Việt):	**Nhập môn Thị giác Máy tính**
Tên môn học (tiếng Anh):	Introduction to Computer Vision
Mã môn học:	CS231
Thuộc khối kiến thức:	Đại cương ☐; Cơ sở nhóm ngành ☐; Cơ sở ngành ☐; Chuyên ngành ☒; Tốt nghiệp ☐

2. Mô tả môn học



2. Mô tả môn học

- Môn học này giới thiệu các nội dung căn bản trong ngành Thị giác máy tính, bao gồm các chủ đề về:
  1. low-level computer vision: image → image


Before
After

- Image Inpainting

2. Mô tả môn học

- Môn học này giới thiệu các nội dung căn bản trong ngành Thị giác máy tính, bao gồm các chủ đề về:
  1. low-level computer vision: image → image
  2. mid-level computer vision: image → feature

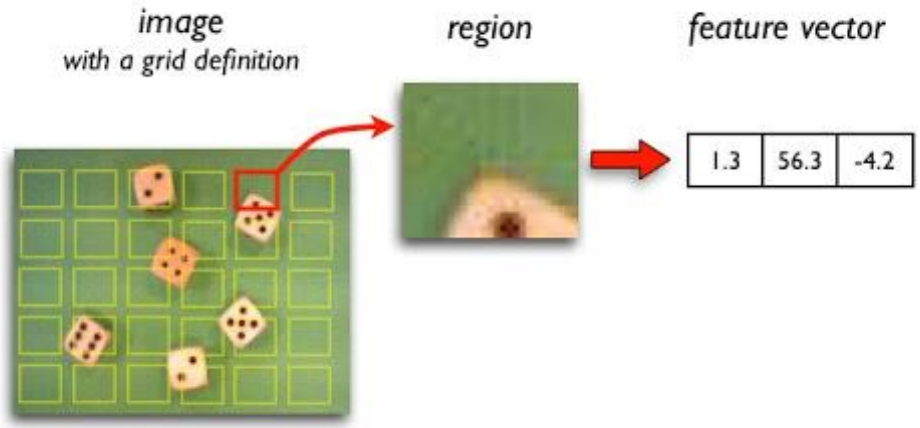


image
with a grid definition
region
feature vector
1.3
56.3
-4.2

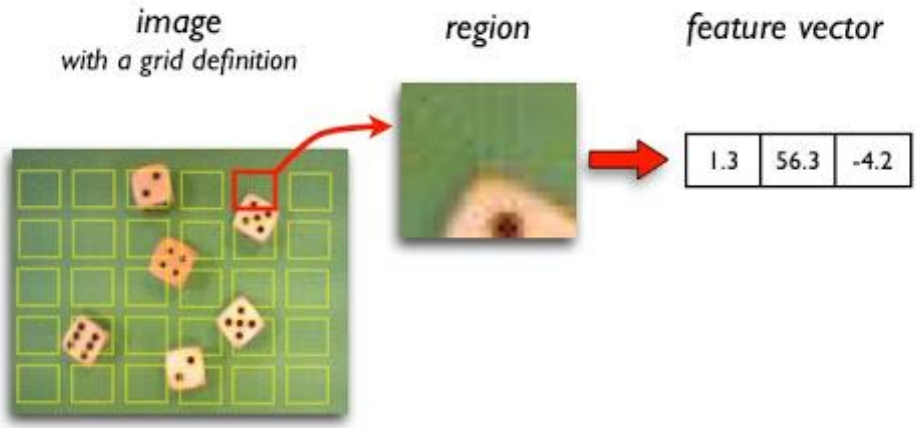
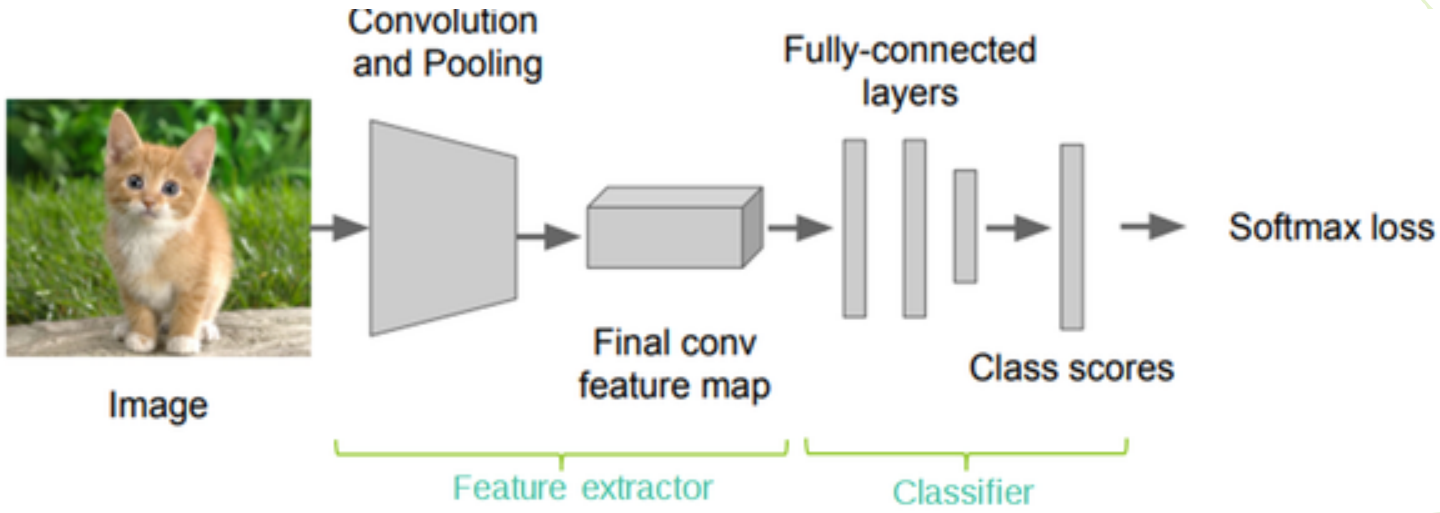


image
with a grid definition
region
feature vector
1.3
56.3
-4.2


Convolution
and Pooling
Fully-connected
layers
Softmax loss
Image
Final conv
feature map
Class scores
Feature extractor
Classifier

2. Mô tả môn học

- Các chủ đề cụ thể gồm:
  - rút trích và khai thác thông tin trên ảnh,
  - các loại đặc trưng thị giác cấp thấp và phương pháp biểu diễn đặc trưng thị giác cấp thấp,
  - các kĩ thuật so khớp ảnh,
  - các kĩ thuật phân đoạn ảnh, phương pháp theo vết (tracking).

3. Mục tiêu môn học

1. Làm việc ở mức độ cá nhân và cộng tác nhóm để trình bày và giải quyết một số bài toán liên quan tới Thị giác máy tính.

2. Hiểu và giải thích được các khái niệm cơ bản, thuật ngữ liên quan tới Thị giác Máy tính

3. Mục tiêu môn học

3. Ứng dụng các kiến thức căn bản về Thị giác máy tính để ứng dụng vào các vấn đề đơn giản trong thực tế.

4. Ứng dụng các công cụ hỗ trợ, thư viện mở về Thị giác Máy tính.

4. Nội dung

Tuần 1	Giới thiệu về ảnh/video số
Tuần 2-4	Khái niệm và các phương pháp rút trích đặc trưng cấp thấp trên ảnh • Biên, cạnh • Đặc trưng vân ảnh • Đặc trưng màu sắc • Đặc trưng hình dáng • Biểu diễn đặc trưng toàn cục, cục bộ

4. Nội dung

Tuần 5-7	Khái niệm và các phương pháp phân đoạn ảnh • Phân đoạn ảnh dựa trên gom cụm (K-Means, MeanShift) • Phân đoạn ảnh dựa trên đồ thị (GraphCut)

4. Nội dung

Tuần	Khái niệm và các phương pháp so khớp ảnh
8 - 12	• Phát hiện điểm nổi bật (Keypoint / Interest Points / Blobs) • Rút trích và biểu diễn đặc trưng cục bộ (SIFT, SURF) • So khớp ảnh • RANSAC

4. Nội dung

Tuần 13 –15	Khái niệm và các phương pháp theo vết (tracking) • Optical Flow

5. Đánh giá

A4. Cuối kỳ - **LT cuối kỳ** - **Đồ án cuối kỳ**		**70 %**
A3. Thực hành		30 %

What is computer vision ?

What is vision ?

- Human Vision

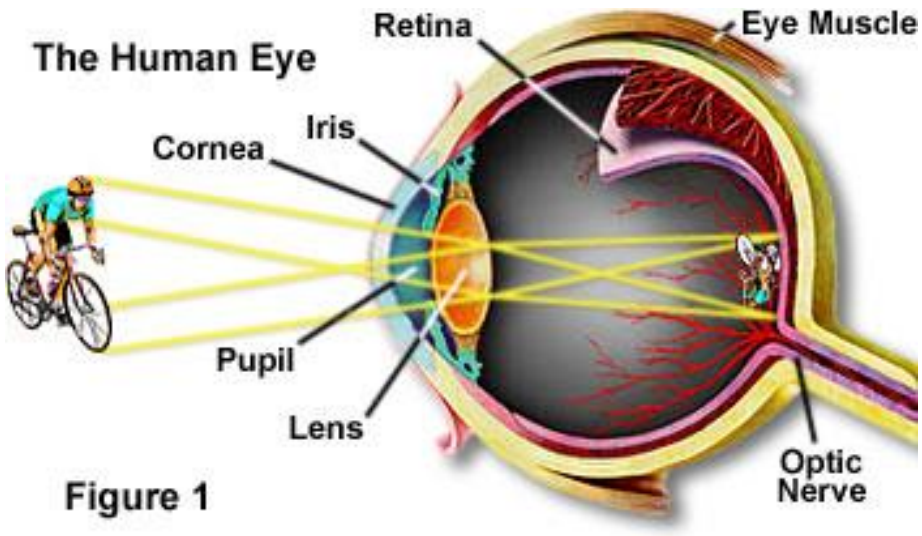


Figure 1

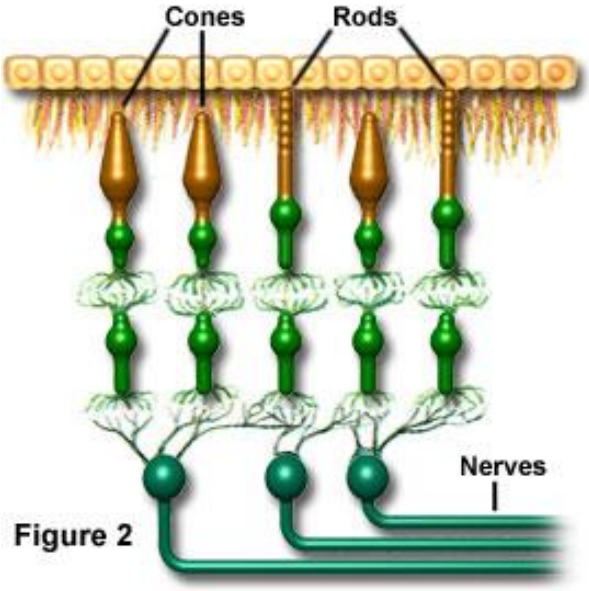
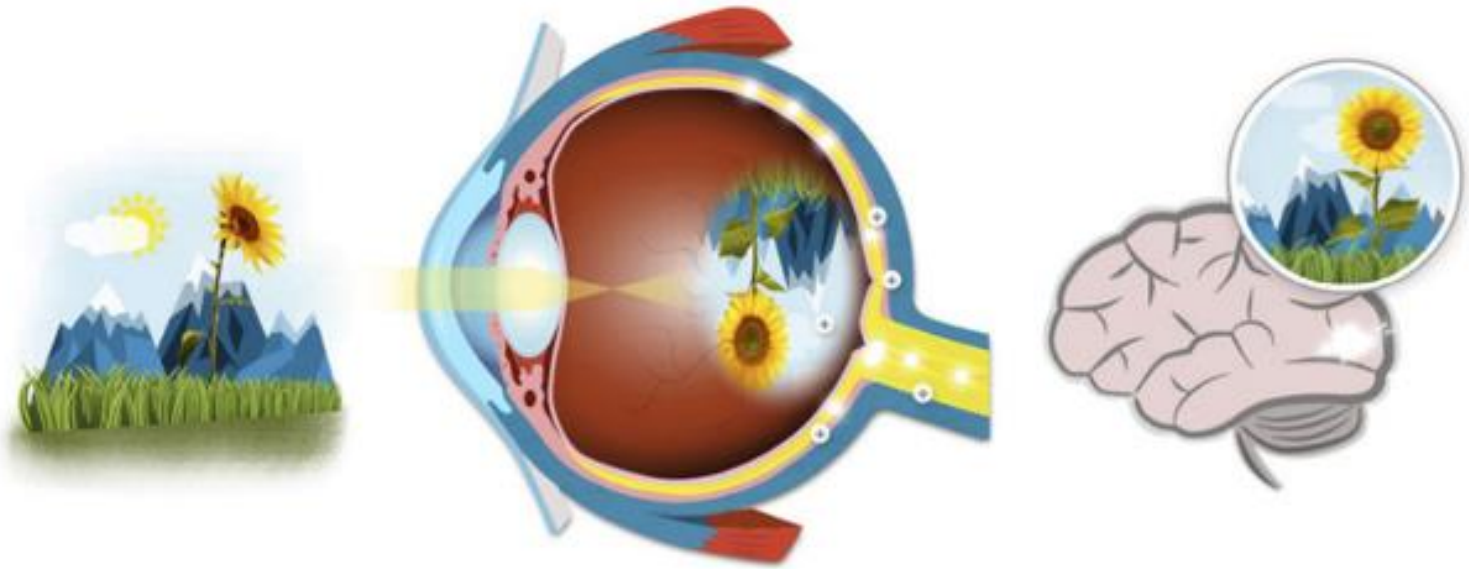


Figure 2

https://micro.magnet.fsu.edu/optics/lightandcolor/vision.html

What is vision ?

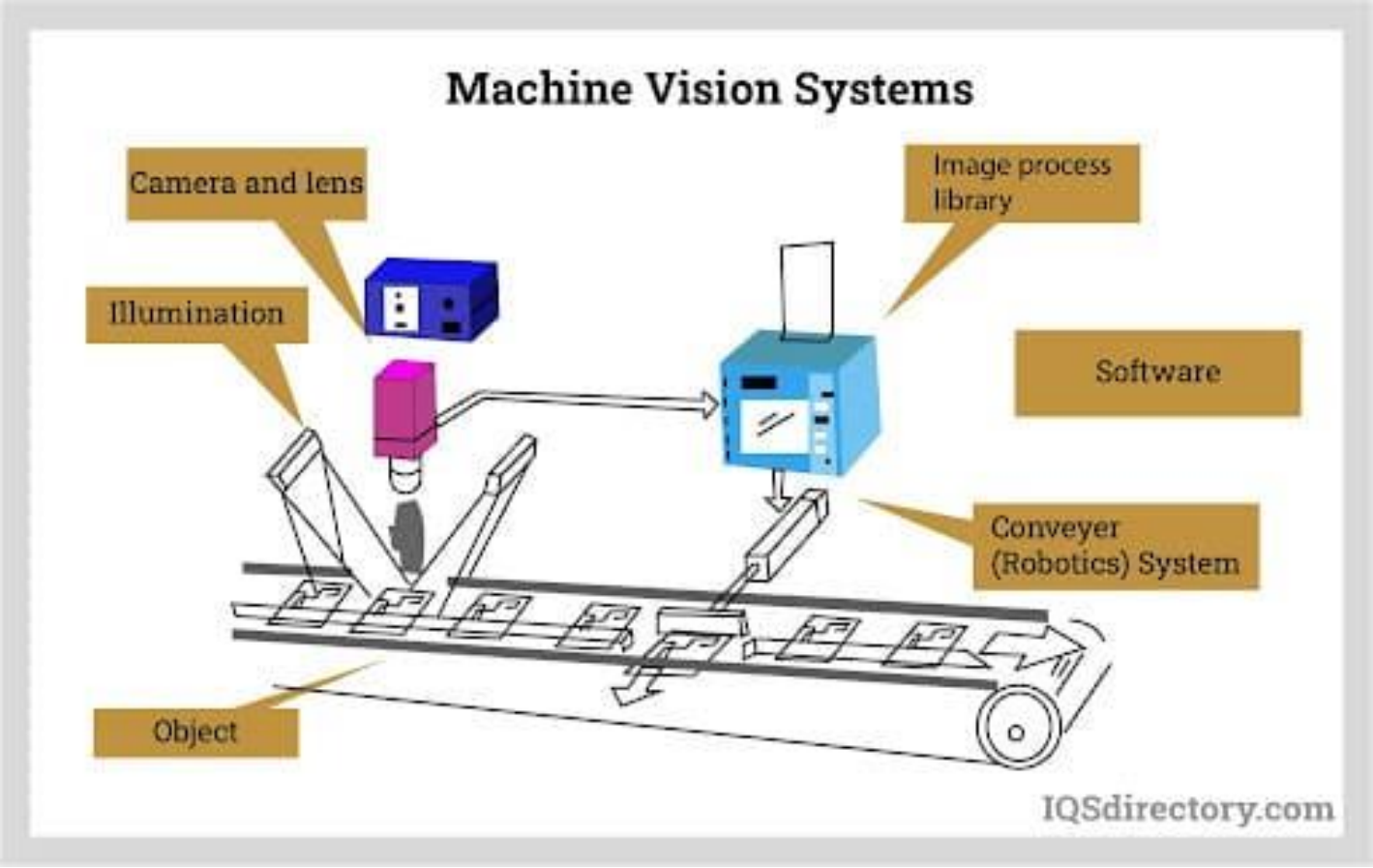
- Human Vision



https://blog.tonkabi.com/blog/post/computer-vision-vs-human-vision

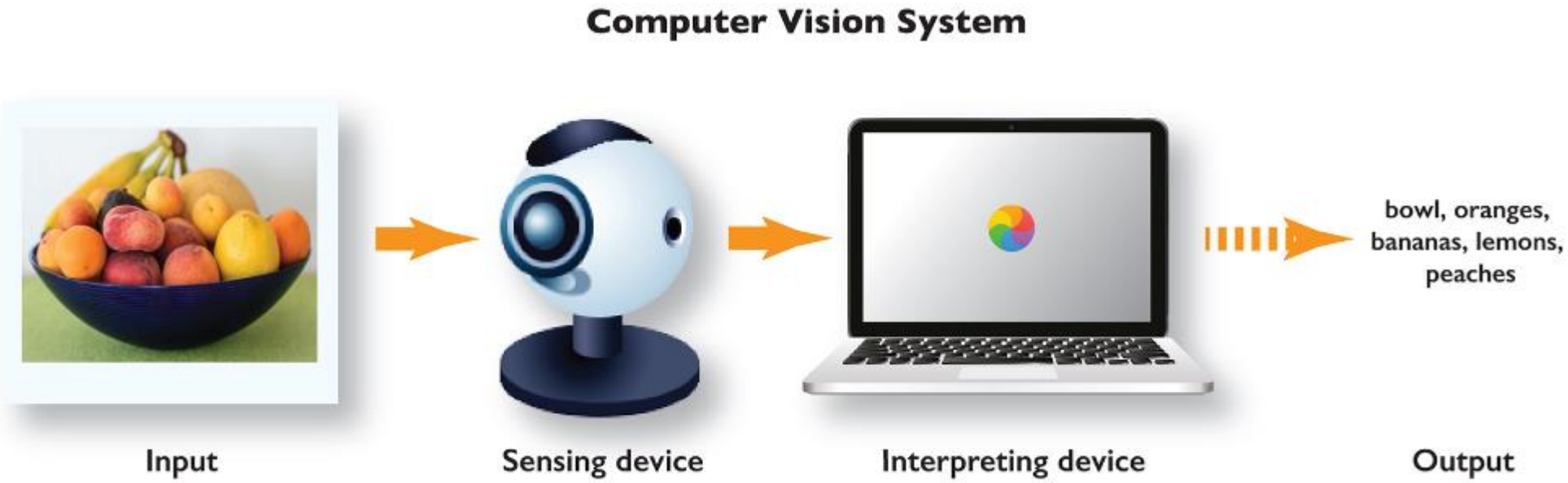
What is vision ?

- Machine Vision



What is vision ?

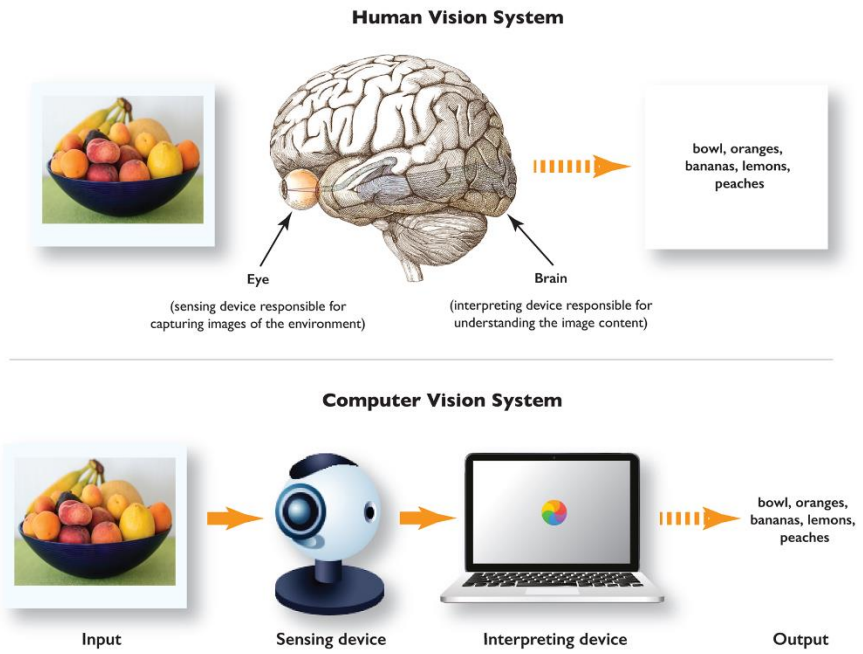
- Computer vision



https://freecontent.manning.com/computer-vision-pipeline-part-1-the-big-picture/

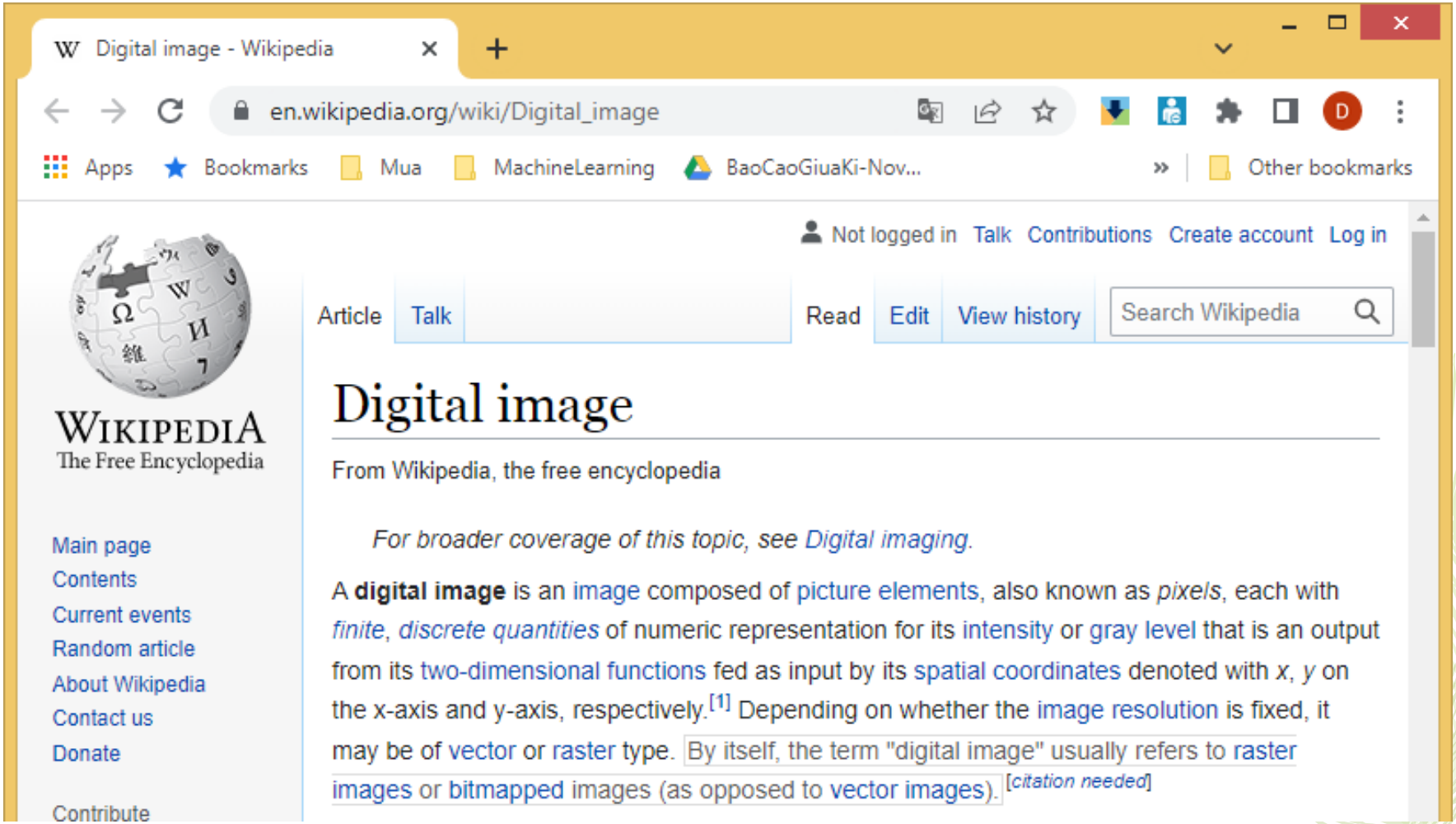
What is vision ?

- Human vs Computer vision

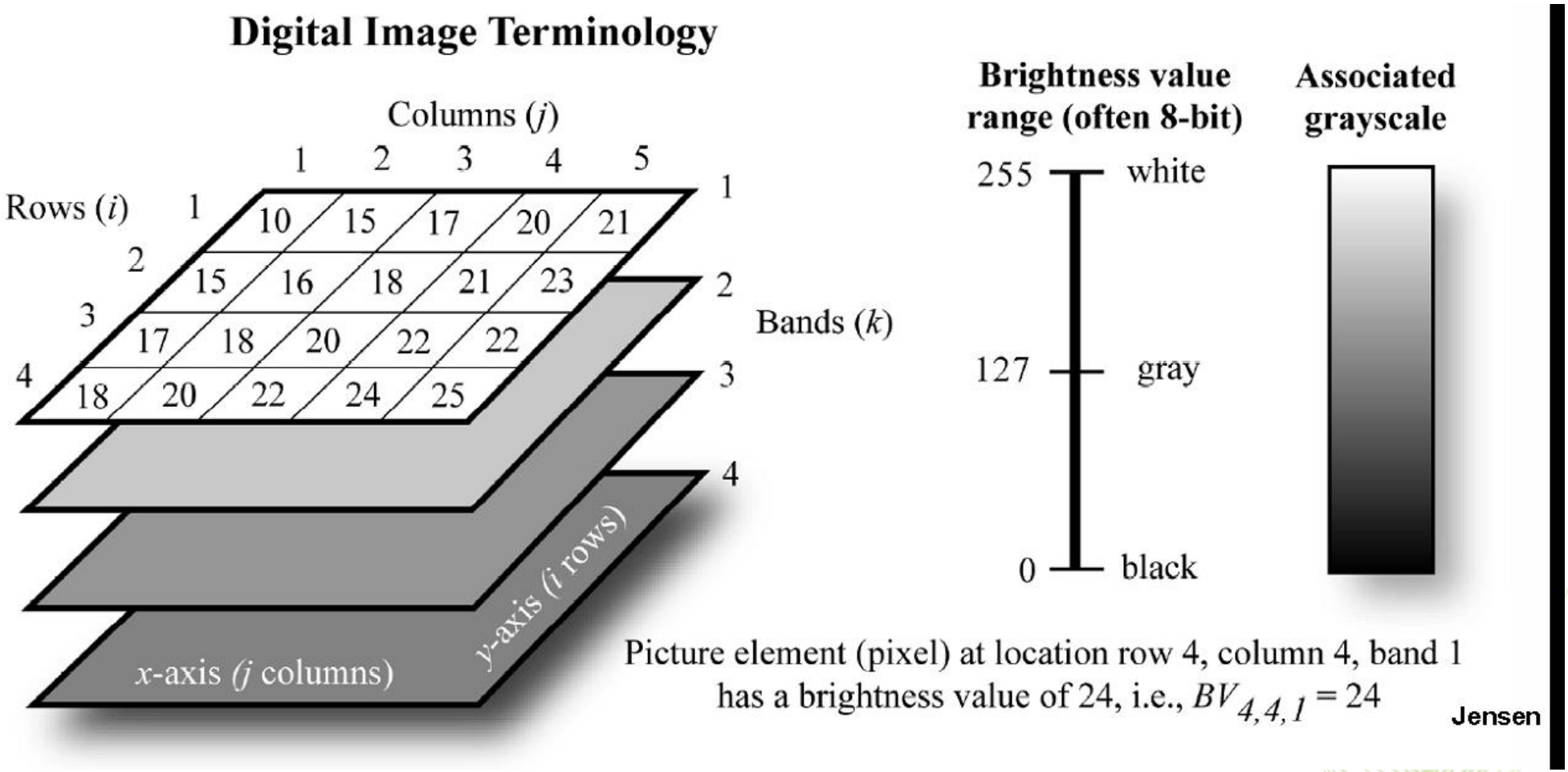


https://freecontent.manning.com/mental-model-graphic-grokking-deep-learning-for-computer-vision/

What is digital image ?

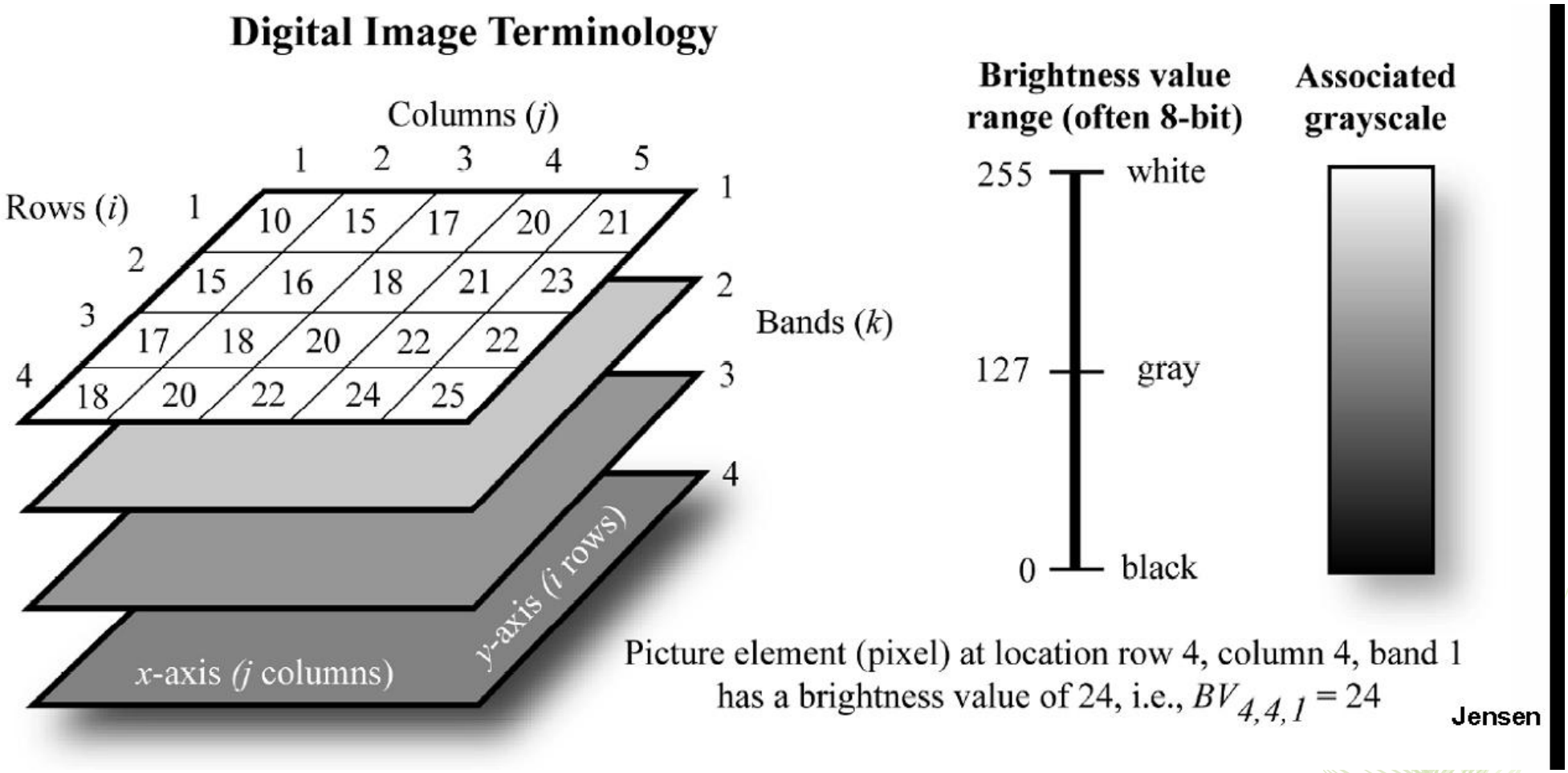


What is digital image ?



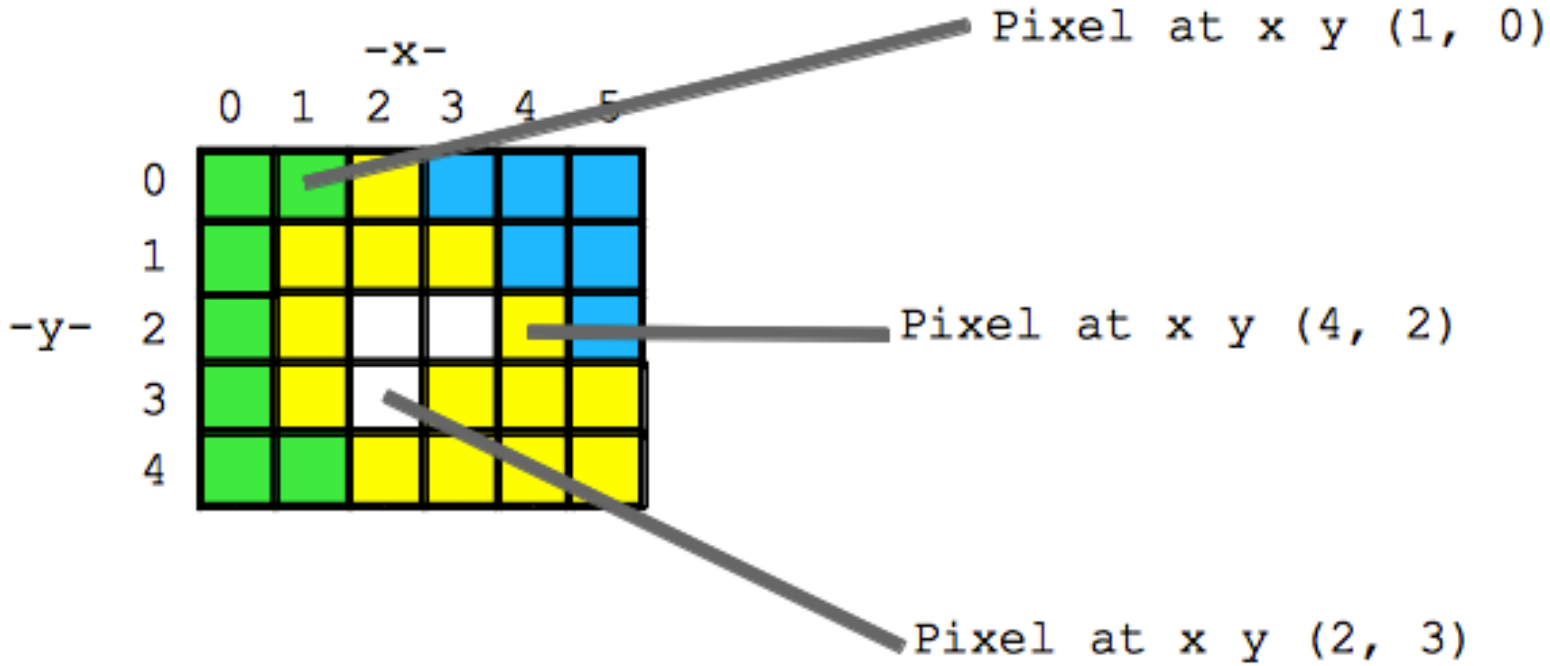
https://www2.geog.soton.ac.uk/users/trevesr/obs/rseo/what_is_a_digital_image.html

What is digital image ?



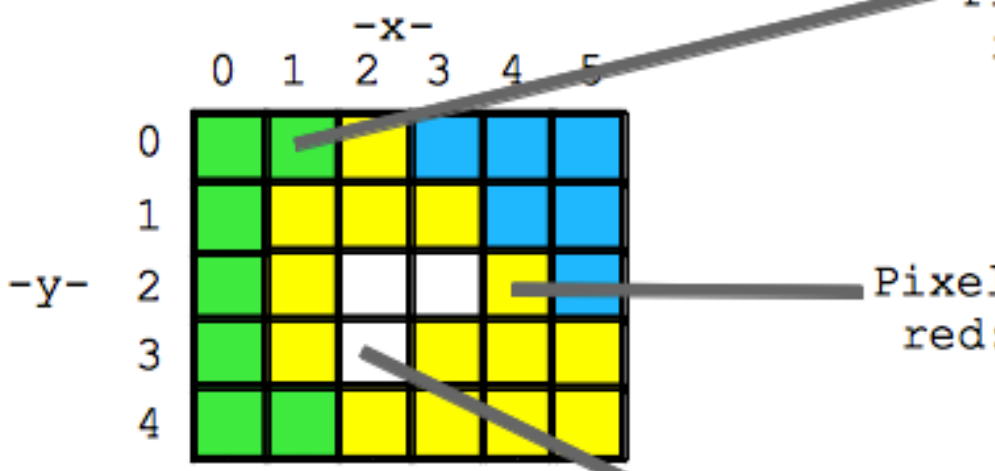
https://www2.geog.soton.ac.uk/users/trevesr/obs/rseo/what_is_a_digital_image.html

What is digital image ?



https://web.stanford.edu/class/cs101/image-1-introduction.html

What is digital image ?



Pixel (1, 0)
red:6 green:250 blue:7
(i.e. shade of green)

Pixel (4, 2)
red:241 green:252 blue:23

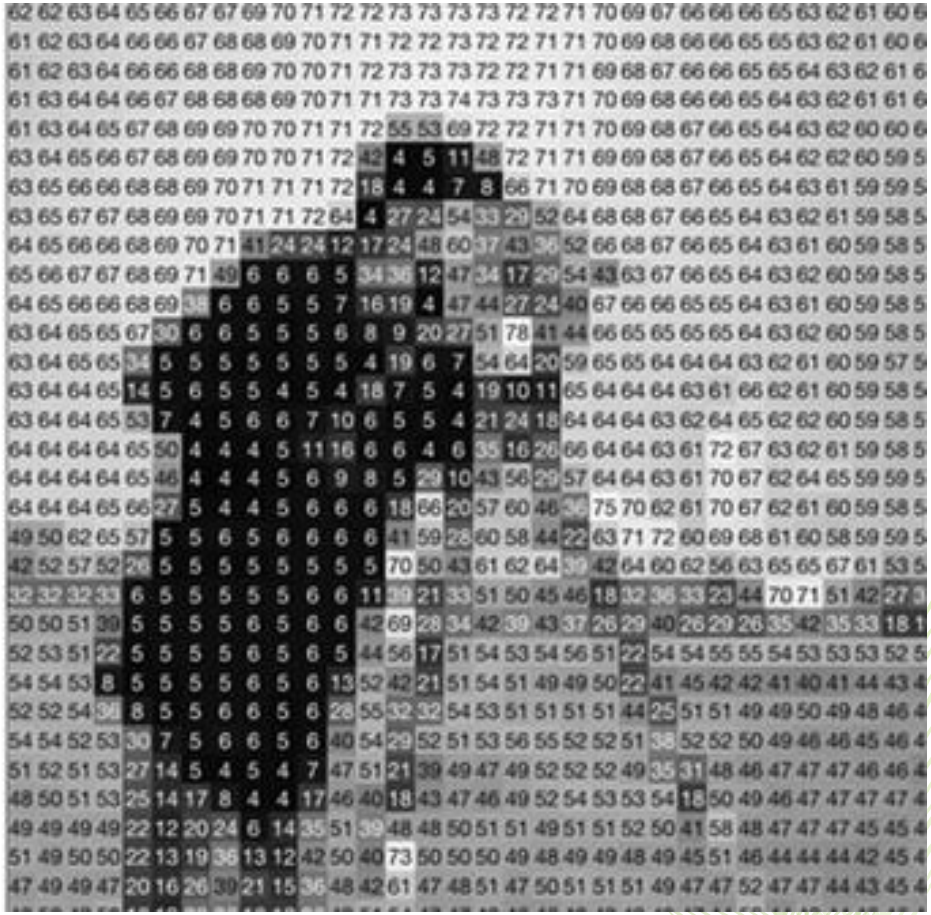
Pixel (2, 3):
red:247 green:250 blue:237

https://web.stanford.edu/class/cs101/image-1-introduction.html

What is computer vision ?

Human

What is computer vision ?



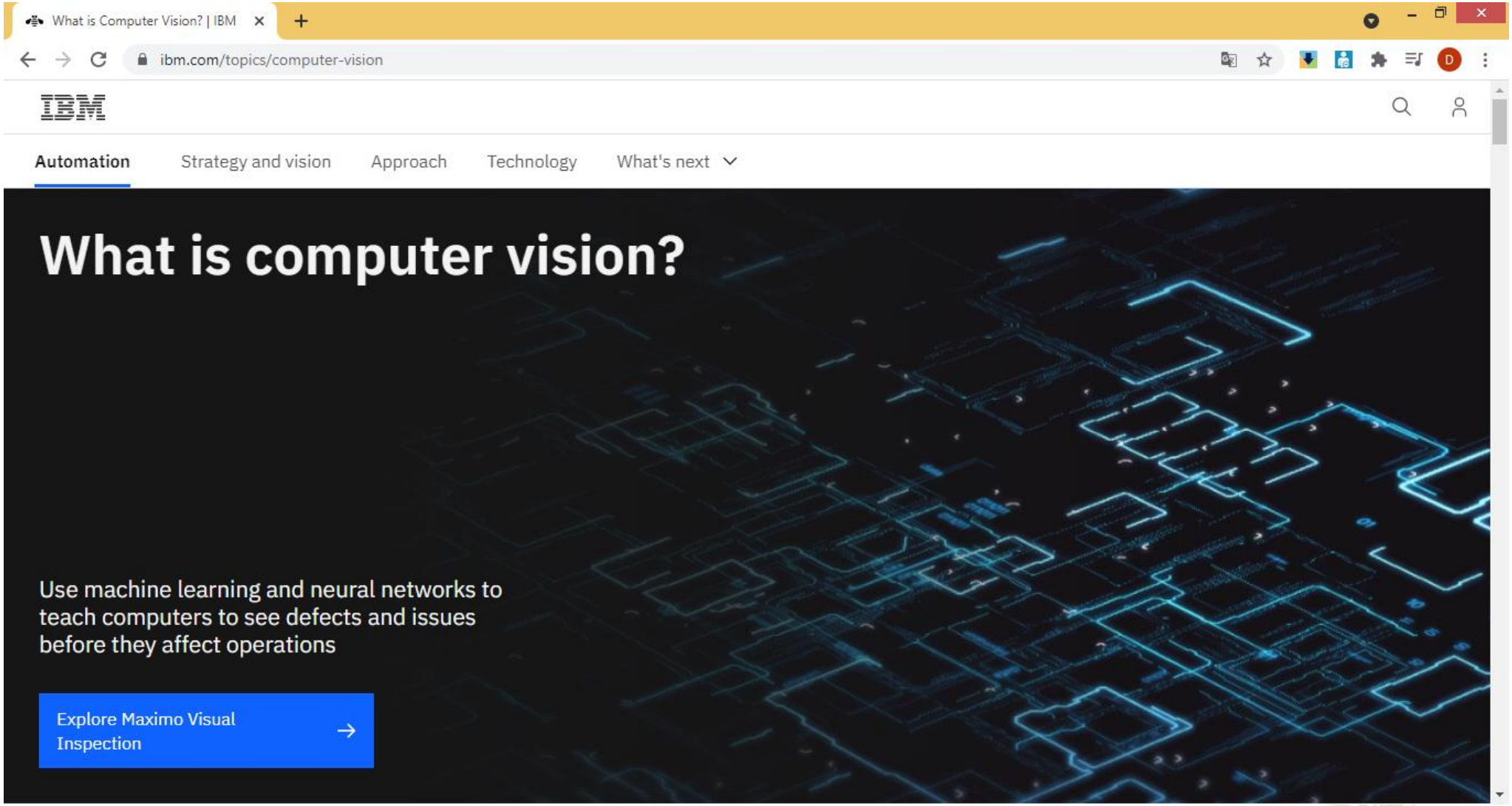
Human

Computer

What is computer vision ?

- https://www.ibm.com/topics/computer-vision
- https://www.coursera.org/learn/introduction-computer-vision-watson-opencv

What is computer vision ?



What is computer vision ?

- Computer vision is a field of artificial intelligence (AI)

What is computer vision ?

- Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs

What is computer vision ?

- Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs — and take actions or make recommendations based on that information.

What is computer vision ?

- Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs — and take actions or make recommendations based on that information. If AI enables computers to think, computer vision enables them to see, observe and understand.

What is computer vision ?

- Computer vision works much the same as human vision, except humans have a head start.

What is computer vision ?

- Computer vision works much the same as human vision, except humans have a head start. Human sight has the advantage of lifetimes of context to train how to tell objects apart, how far away they are, whether they are moving and whether there is something wrong in an image.

What is computer vision ?

- Computer vision trains machines to perform these functions, but it has to do it in much less time with cameras, data and algorithms rather than retinas, optic nerves and a visual cortex. Because a system trained to inspect products or watch a production asset can analyze thousands of products or processes a minute, noticing imperceptible defects or issues, it can quickly surpass human capabilities.

What is computer vision ?

- Computer vision is used in industries ranging from energy and utilities to manufacturing and automotive – and the market is continuing to grow. It is expected to reach USD 48.6 billion by 2022(1)

1. *https://www.forbes.com/sites/bernardmarr/2019/04/08/7-amazing-examples-of-computer-and-machine-vision-in-practice/#3dbb3f751018*

Why use computer vision?

Like all technologies,

- Slower → Faster
- Expensive → Cheap
- Manual → Automated
- Difficult → Easy
- Inconvenient → Convenient
- Unscalable → Scalable

IBM Developer
SKILLS NETWORK

Computer vision applications

Industries

- Automotive
- Food & Beverage
- Manufacturing
- Pharmaceutical
- Supply chain
- Energy & Utilities
- Hospitality
- Life Sciences
- Human Resources
- Insurance

IBM Developer SKILLS NETWORK

Computer vision examples

https://www.forbes.com/sites/bernardmarr/2019/04/08/7-amazing-examples-of-computer-and-machine-vision-in-practice/?sh=2f5b2b8f1018

Computer vision examples

1. Autonomous vehicles
2. Google Translate app
3. Facial recognition
4. Healthcare
5. Real-time sports tracking
6. Agriculture
7. Manufacturing

Computer vision examples

1. Autonomous vehicles

Computer vision is necessary to enable self-driving cars. Manufacturers such as Tesla, BMW, Volvo, and Audi use multiple cameras, lidar, radar, and ultrasonic sensors to acquire images from the environment so that their self-driving cars can detect objects, lane markings, signs and traffic signals to safely drive.

Computer vision examples

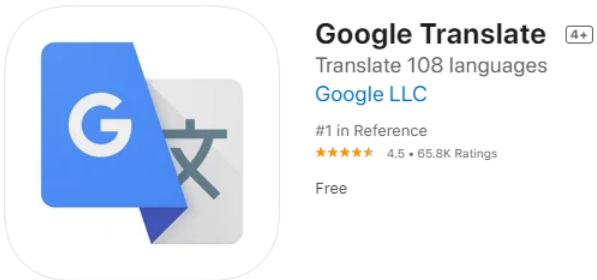
1. Autonomous vehicles

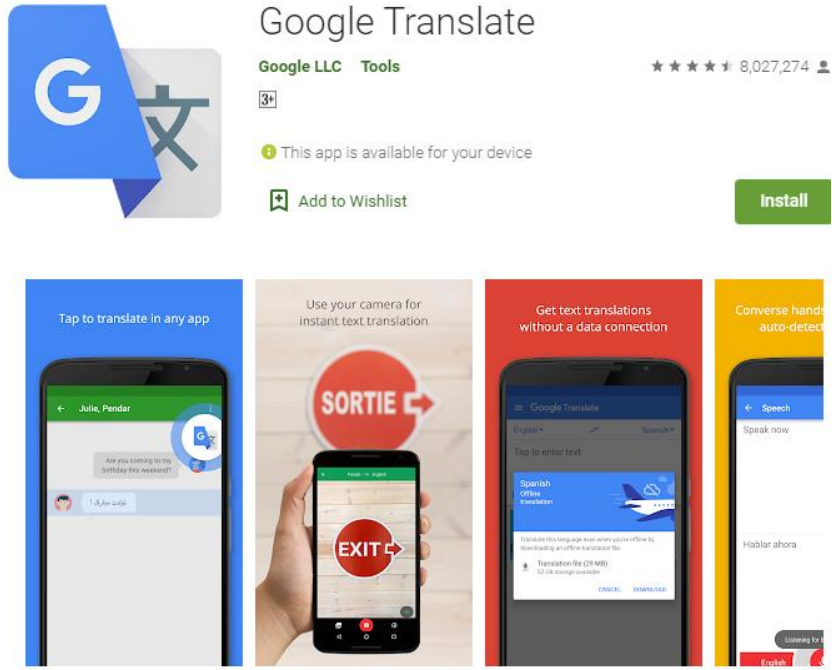
https://www.youtube.com/watch?v=tlThdr3O5Qo

Computer vision examples

2. Google Translate app

All you need to do to read signs in a foreign language is to point your phone's camera at the words and let the Google Translate app tell you what it means in your preferred language almost instantly. By using optical character recognition to see the image and augmented reality to overlay an accurate translation, this is a convenient tool that uses computer vision.

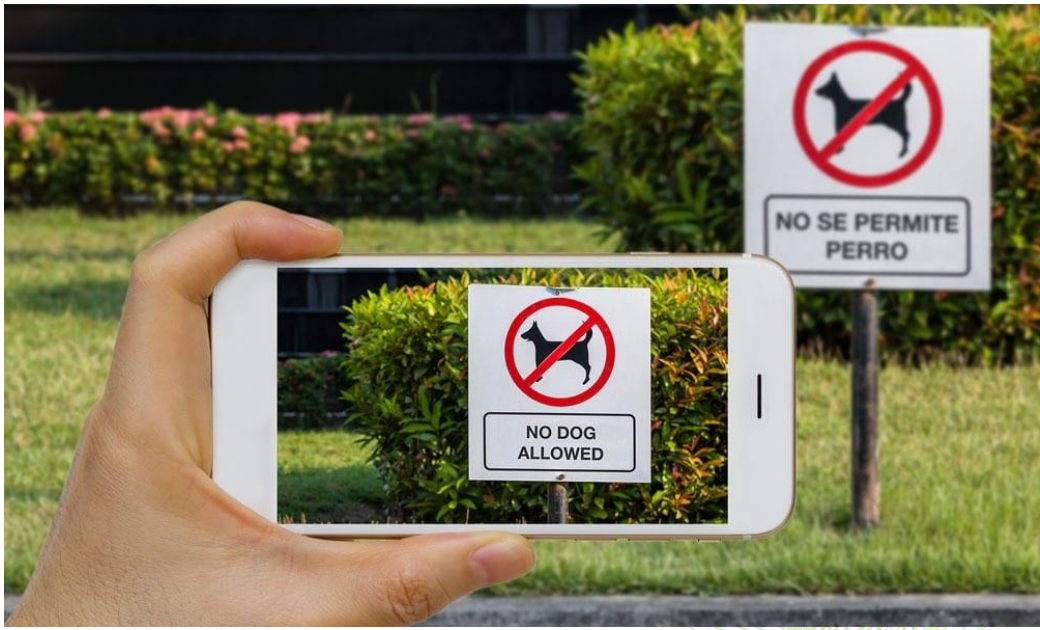




Computer vision examples

2. Google Translate app

https://mashable.com/article/word-lens-japanese



https://www.hellotech.com/blog/the-5-best-free-translation-apps

Computer vision examples

3. Facial recognition

China is definitely on the cutting edge of using facial recognition technology, and they use it for police work, payment portals, security checkpoints at the airport and even to dispense toilet paper and prevent theft of the paper at Tiantan Park in Beijing, among many other applications.

Facial recognition system

A facial recognition system is a technology capable of matching a human face from a digital image or a video frame against a database of faces, typically employed to authenticate users through ID verification services, works by pinpointing and measuring facial features from a given image. Wikipedia

Computer vision examples

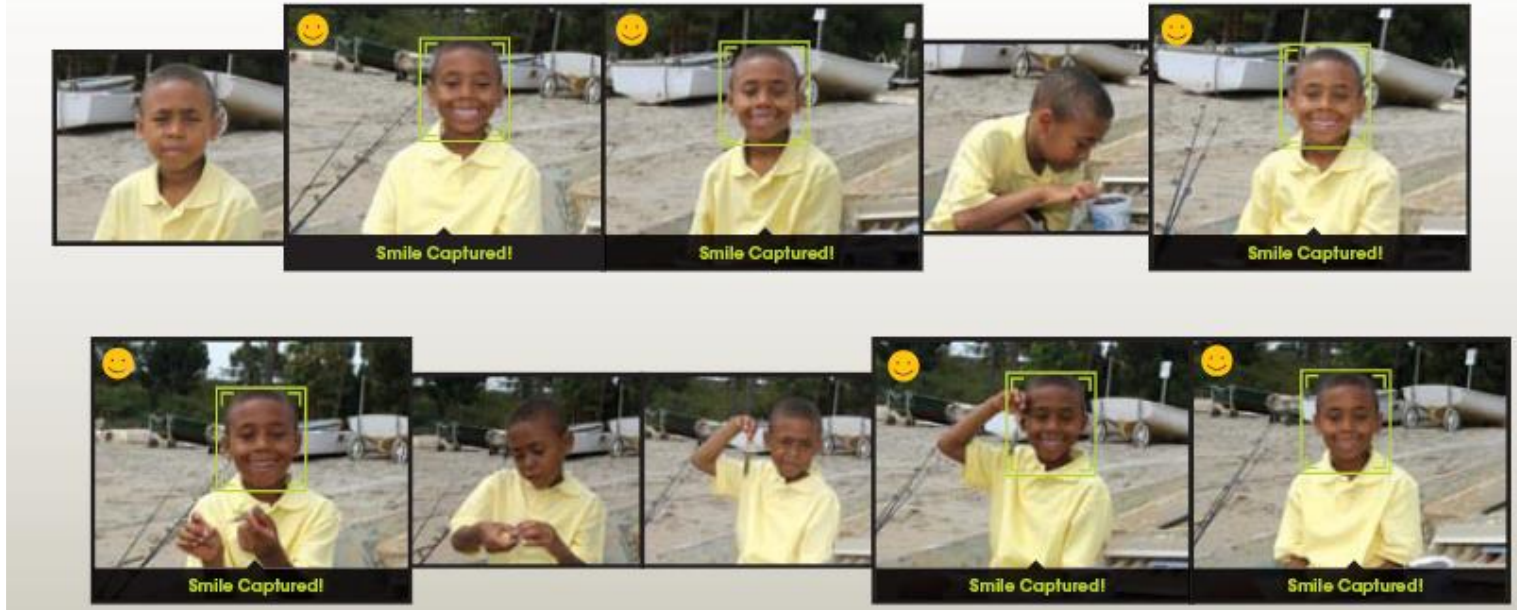
3. Facial recognition

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

Smile detection

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



Sony Cyber-shot® T70 Digital Still Camera

Face recognition: Apple iPhoto, Facebook, Google, etc

Computer vision examples

3. Facial recognition

Shanghai airport first to launch automated clearance system using facial recognition technology

https://www.scmp.com/tech/enterprises/article/2168681/shanghai-airport-first-launch-automated-clearance-system-using?module=perpetual_scroll&pgtype=article&campaign=2168681

Computer vision examples

3. Facial recognition

Alibaba launches ‘smile to pay’ facial recognition system at KFC in China

An Alibaba employee demonstrates ‘Smile to Pay’, an automatic payment system that authorize payment via facial recognition

Alex Wong | Staff | Getty Images

Computer vision examples

4. Healthcare

Since 90 percent of all medical data is image based there is a plethora of uses for computer vision in medicine. From enabling new medical diagnostic methods to analyze X-rays, mammography and other scans to monitoring patients to identify problems earlier and assist with surgery, expect that our medical institutions and professionals and patients will benefit from computer vision today and even more in the future as its rolled out in healthcare.

Computer vision examples

4. Healthcare

assisted living, patient monitoring
[Lan et al, PAMI 2012]

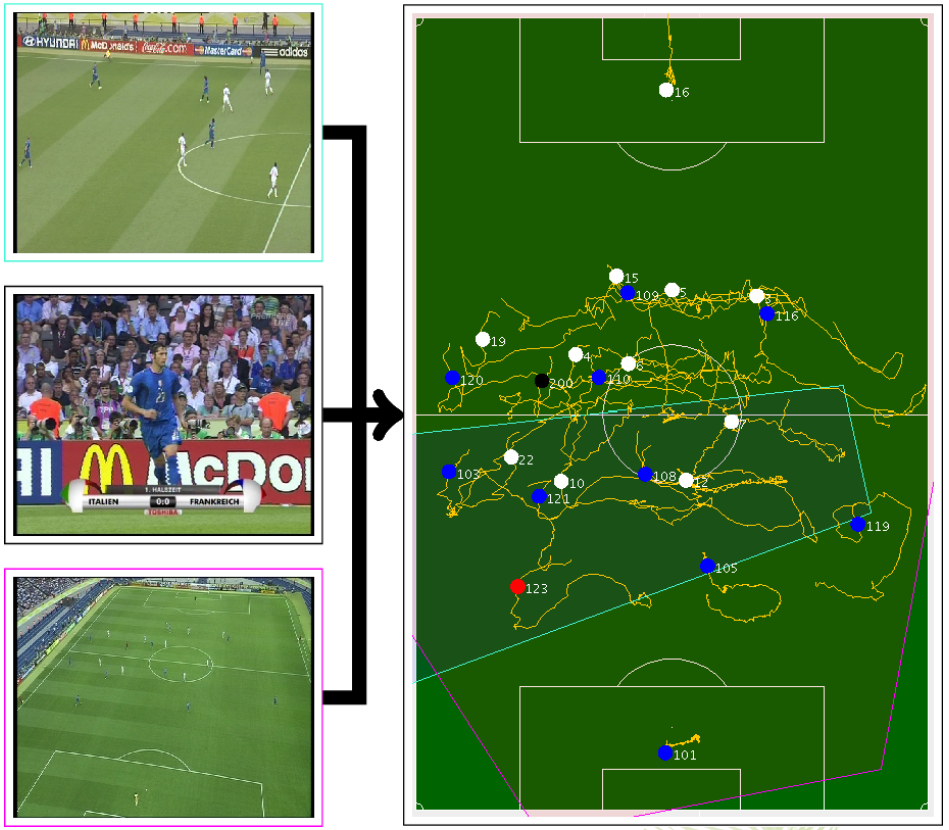
autism screening (sàng lọc tự kỷ)

http://www.gatech.edu/newsroom/release.html?nid=60509

Computer vision examples

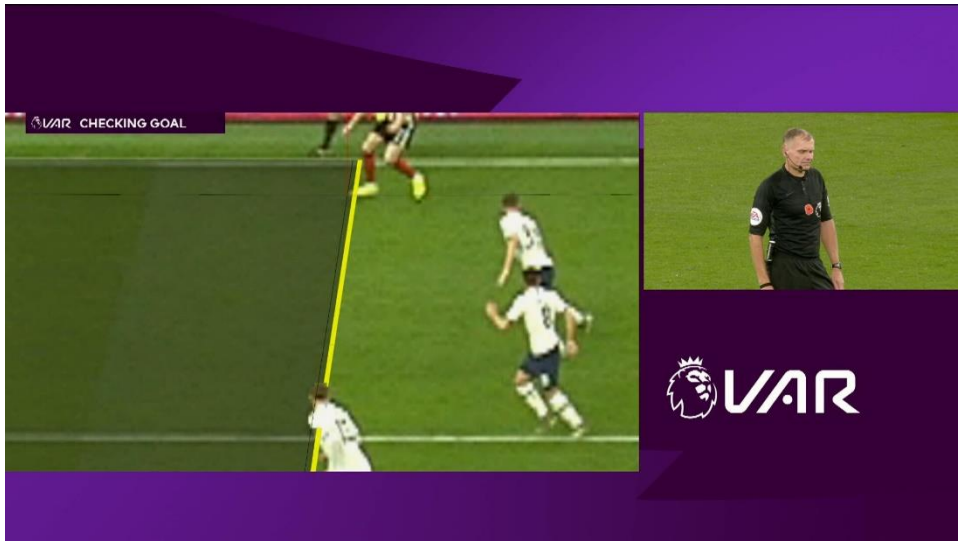
5. Real-time sports tracking

Ball and puck tracking on televised sports has been common for a while now, but computer vision is also helping play and strategy analysis, player performance and ratings, as well as to track the brand sponsorship visibility in [sports broadcasts](#).



Computer vision examples

5. Real-time sports tracking



https://www.skysports.com/football/news/11095/11928910/var-premier-league-clubs-call-for-change-in-marginal-offside-calls

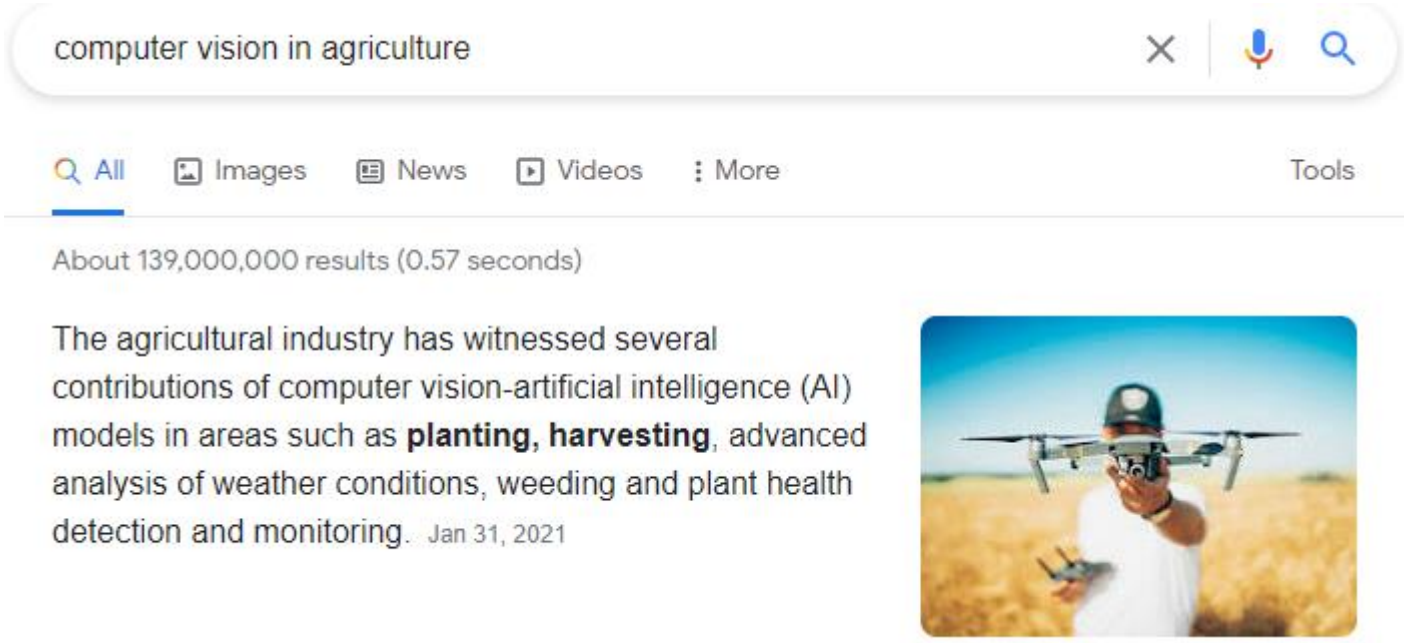
Computer vision examples

6. Agriculture

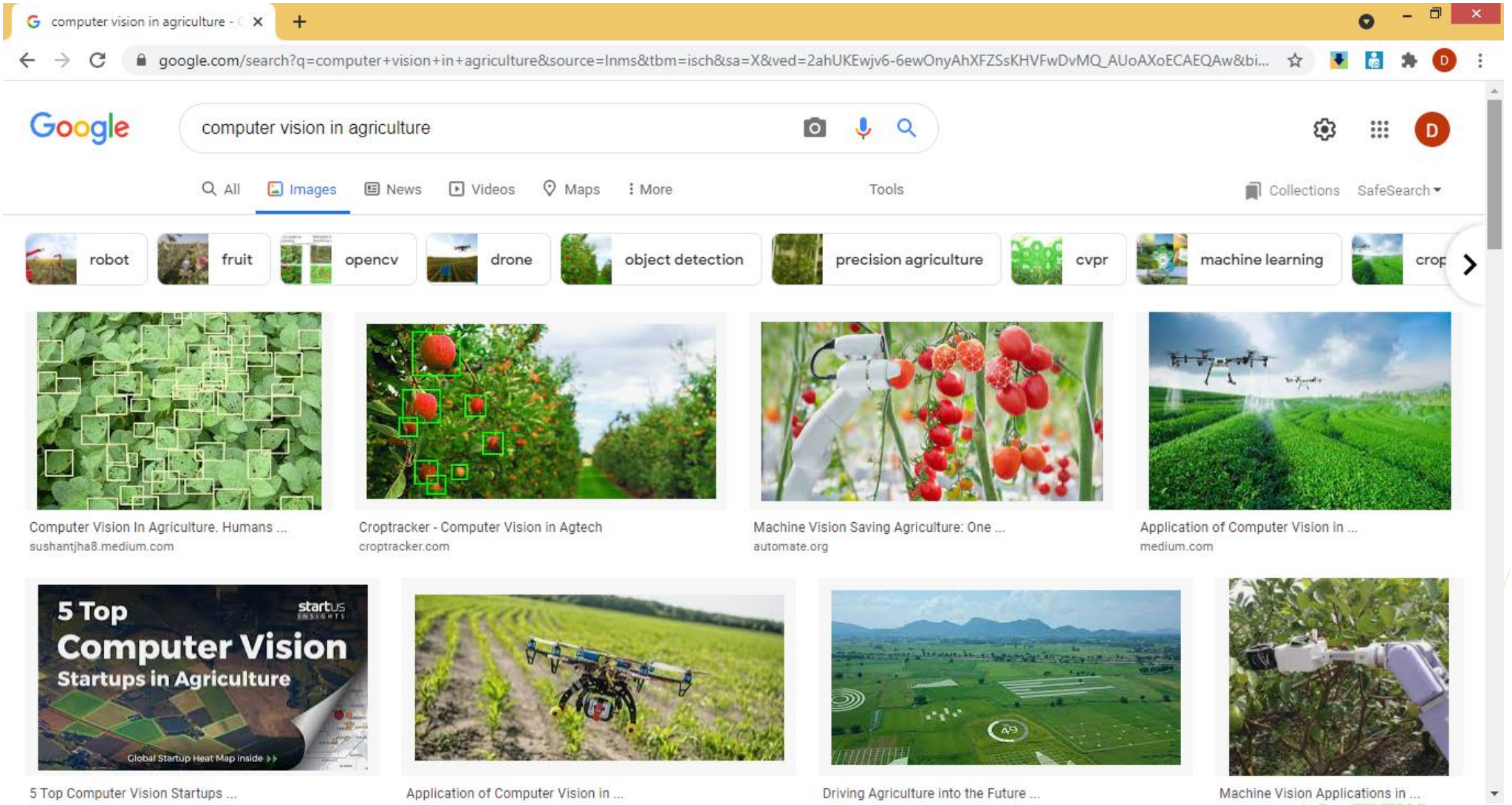
At CES 2019, John Deere featured a semi-autonomous combine harvester that uses artificial intelligence and computer vision to analyze grain quality as it gets harvested and to find the optimal route through the crops. There's also great potential for computer vision to identify weeds so that herbicides can be sprayed directly on them instead of on the crops. This is expected to reduce the amount of herbicides needed by 90 percent.

Computer vision examples

6. Agriculture



Computer vision examples



Computer vision examples

7. Manufacturing

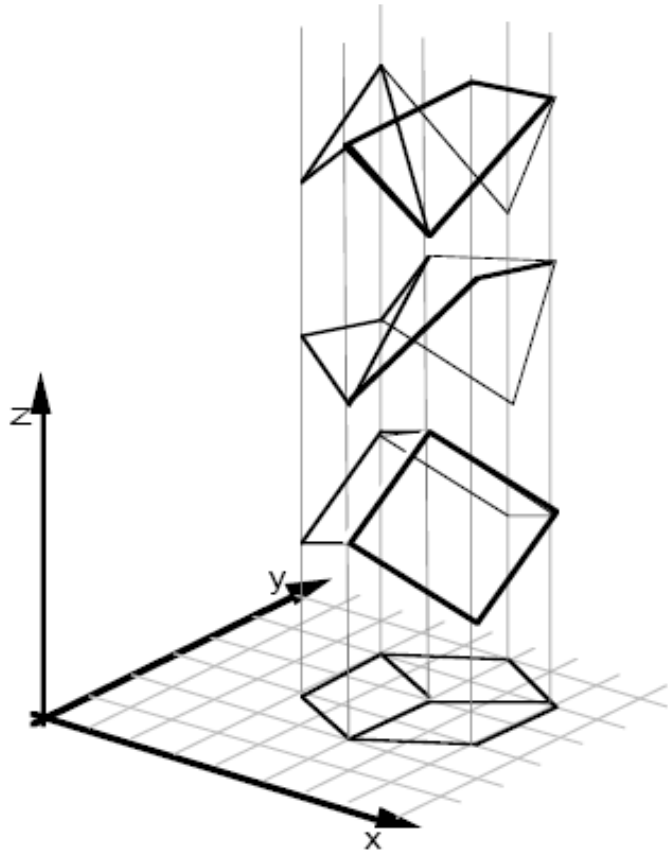
Computer vision is helping manufacturers run more safely, intelligently and effectively in a variety of ways. Predictive maintenance is just one example where equipment is monitored with computer vision to intervene before a breakdown would cause expensive downtime. Packaging and product quality are monitored, and defective products are also reduced with computer vision.

https://www.devteam.space/blog/10-examples-of-using-machine-vision-in-manufacturing/

Why vision is so hard?

Why is vision so hard?

- Ill-posed problem



Challenges 1: view point variation

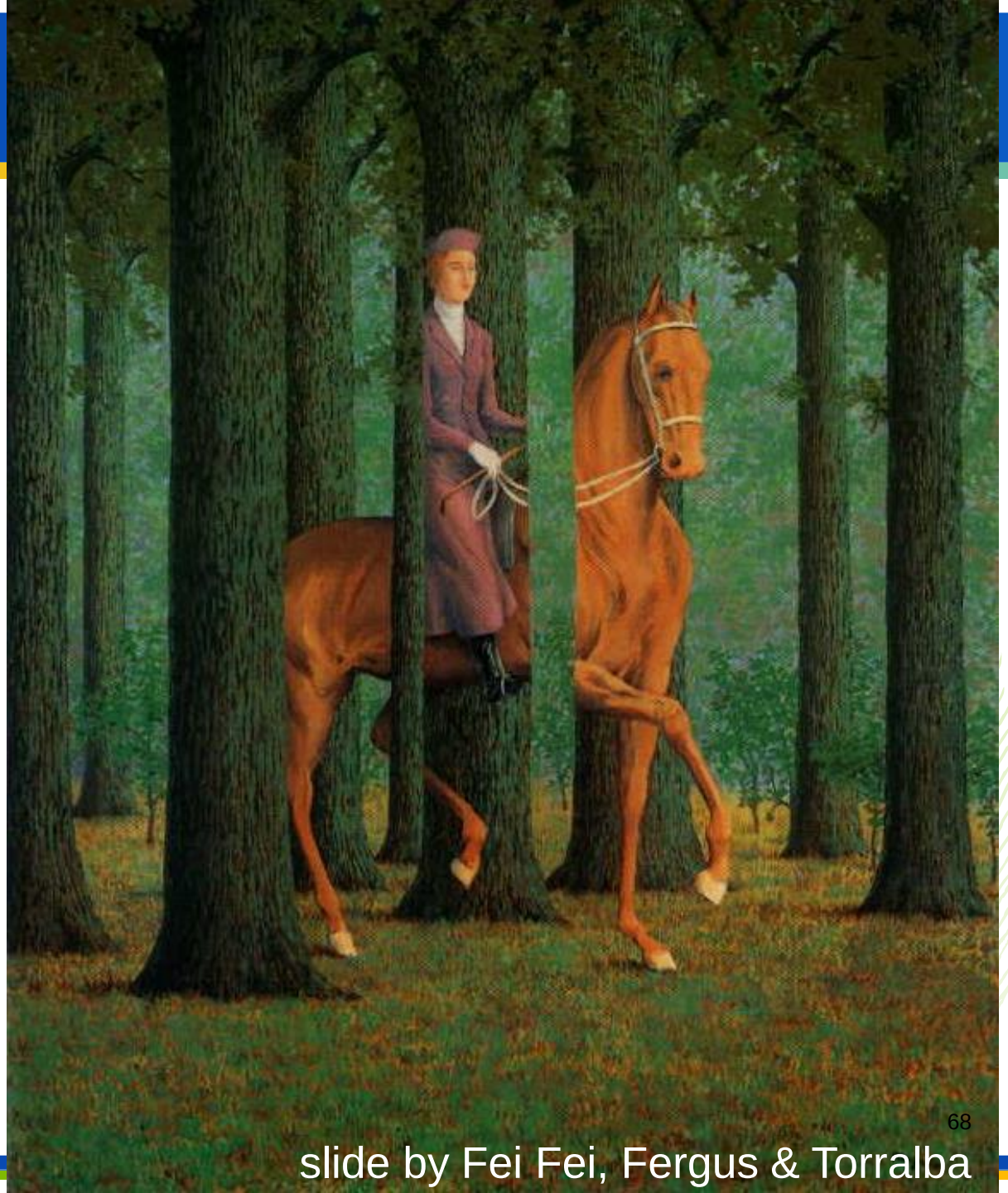
Michelangelo 1475-1564

slide by Fei Fei, Fergus & Torralba

Challenges 2: illumination

slide credit: S. Ullman

Challenges 3: occlusion



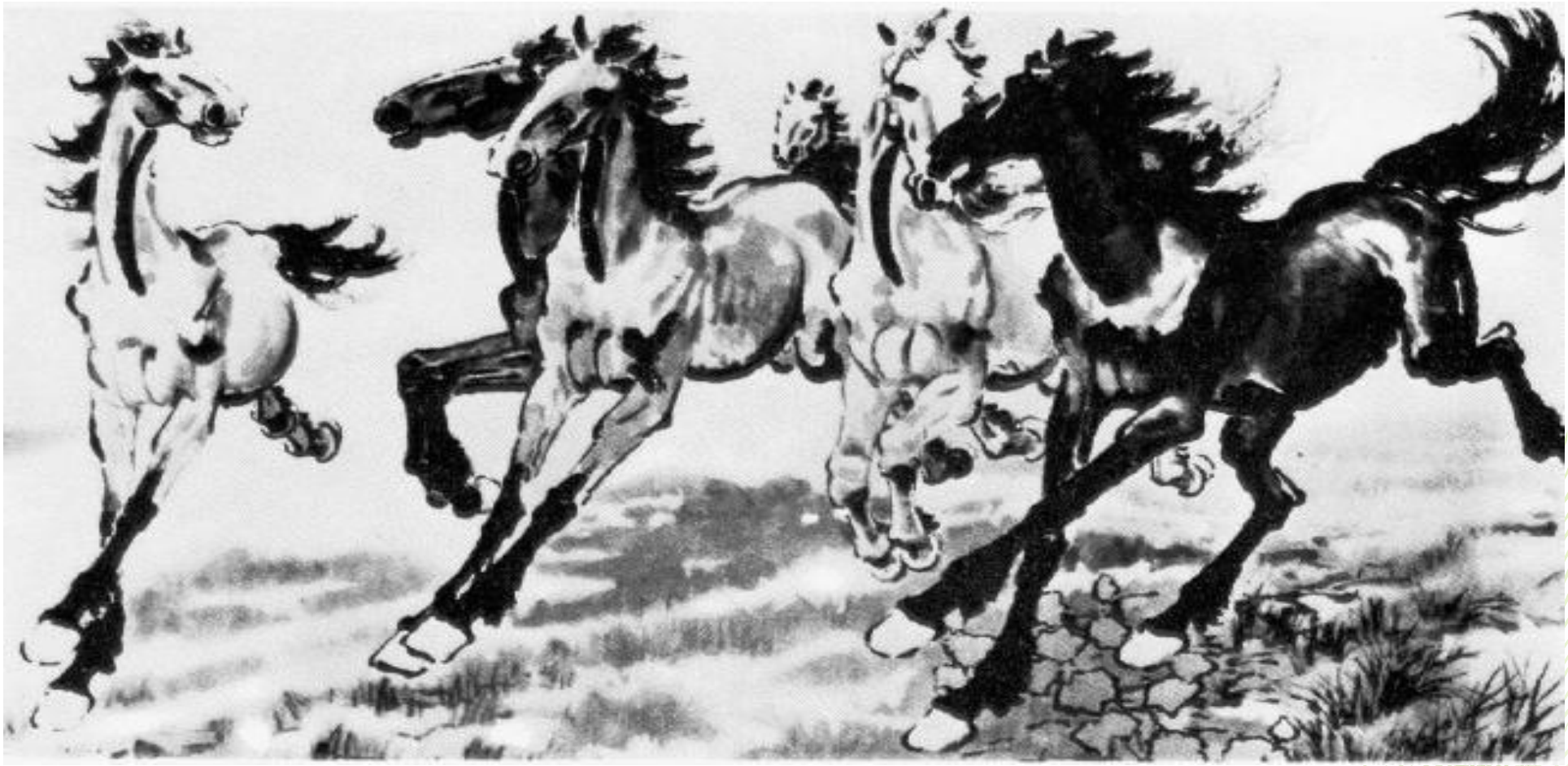
Magritte, 1957

slide by Fei Fei, Fergus & Torralba

Challenges 4: scale

slide by Fei Fei, Fergus & Torralba

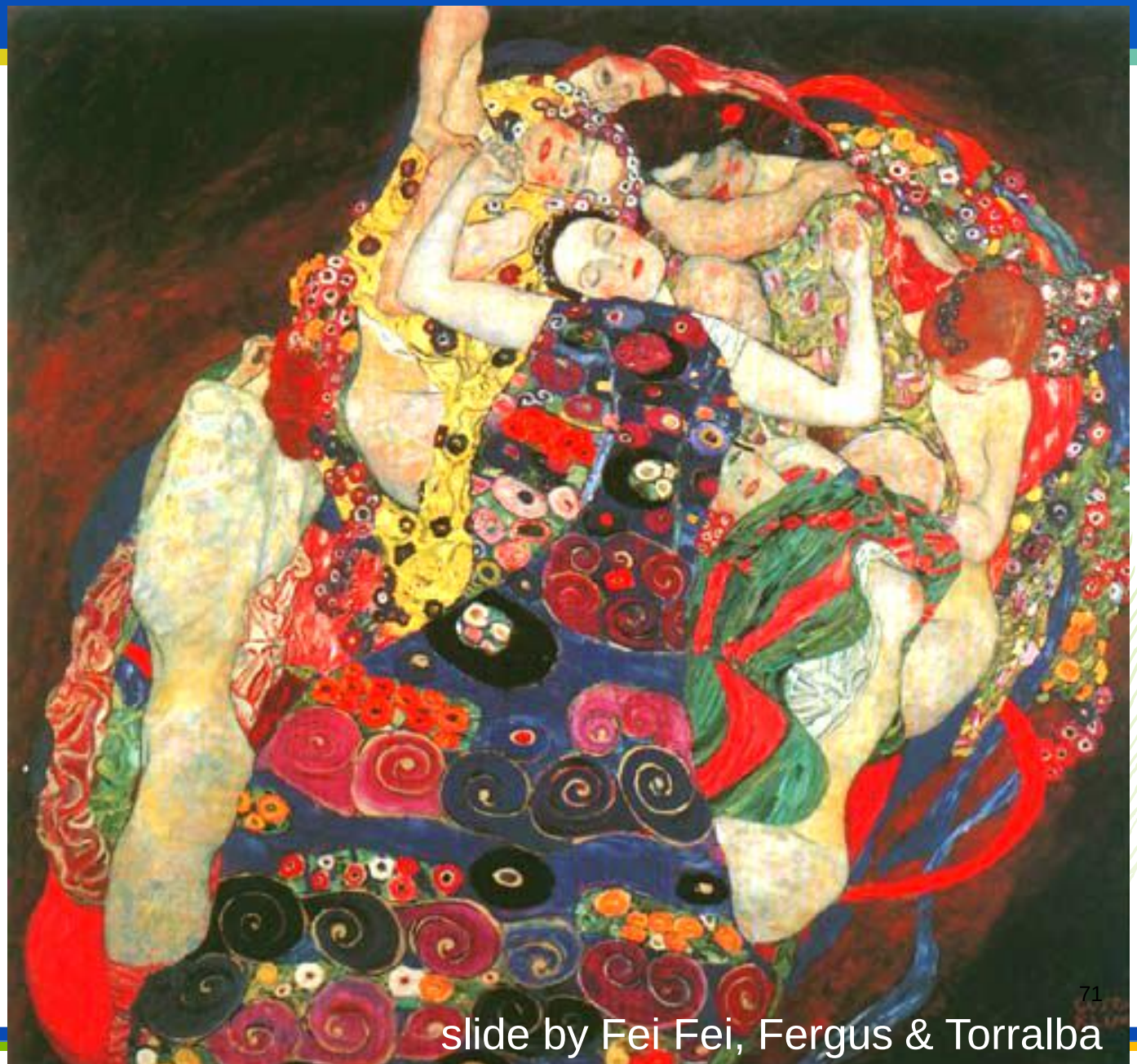
Challenges 5: deformation



slide by Fei Fei, Fergus & Torralba

Xu, Beihong 1943

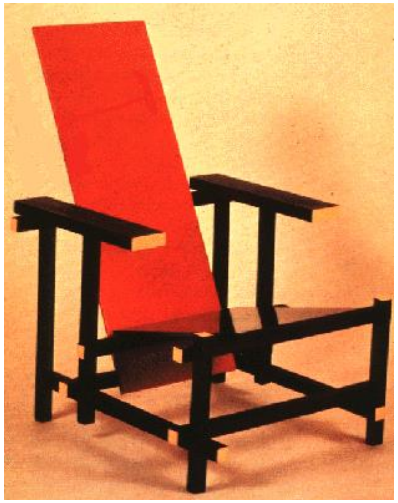
Challenges 6: background clutter



Klimt, 1913

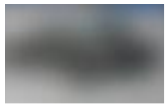
slide by Fei Fei, Fergus & Torralba

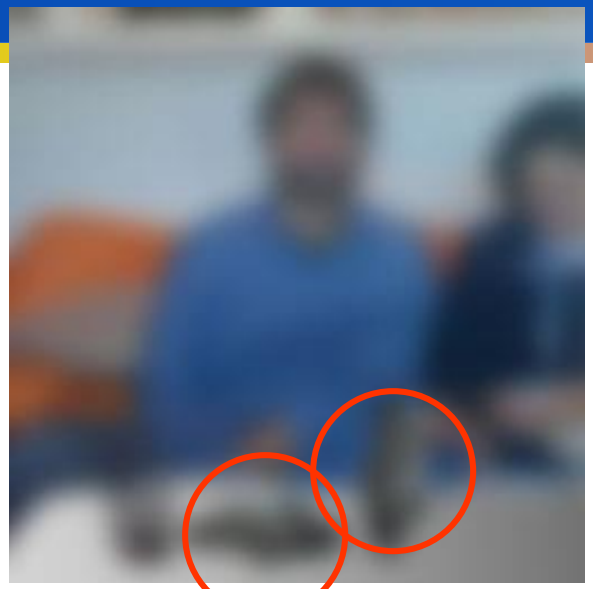
Challenges 7: object intra-class variation

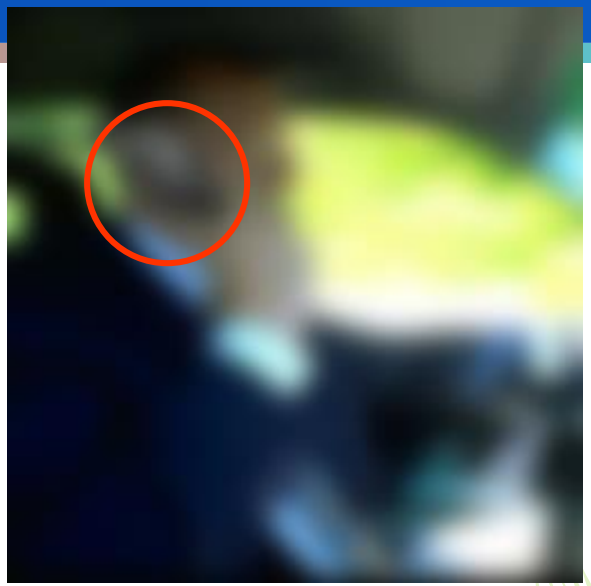


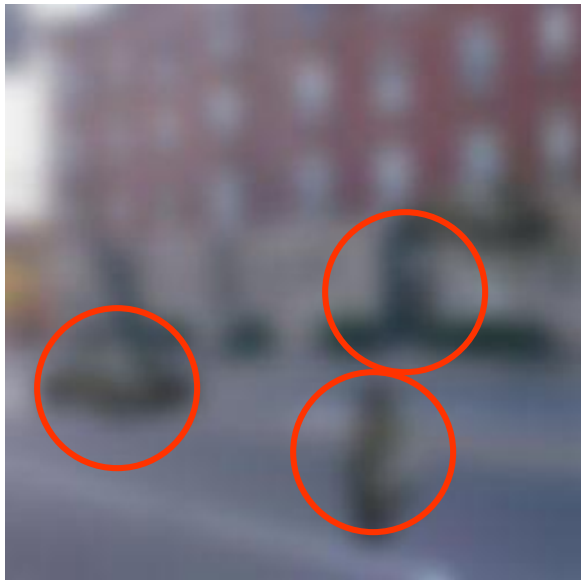
slide by Fei-Fei, Fergus & Torralba

Challenges 8: local ambiguity





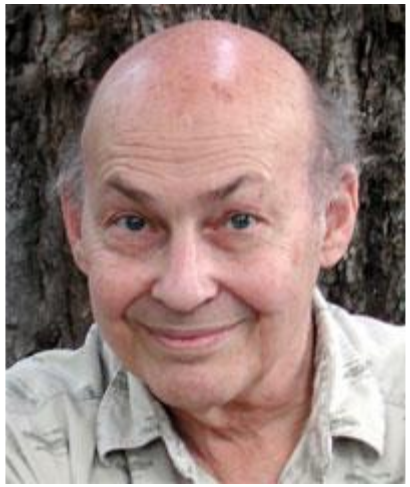




slide by Fei-Fei, Fergus & Torralba

Challenges 9: the world behind the image

How hard is computer vision?



Marvin Minsky, MIT
Turing award,1969

“In 1966, Minsky hired a first-year undergraduate student and assigned him a problem to solve over the summer: connect a television camera to a computer and get the machine to describe what it sees.”

Crevier 1993, pg. 88

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

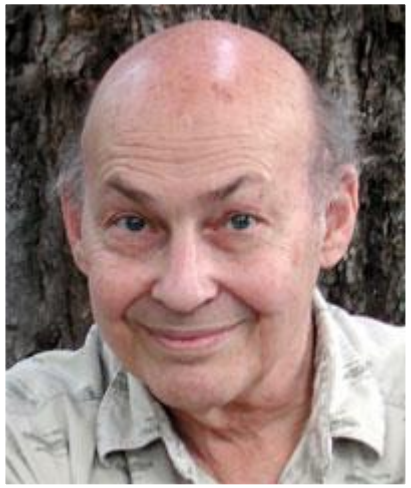
Artificial Intelligence Group — July 7, 1966

Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

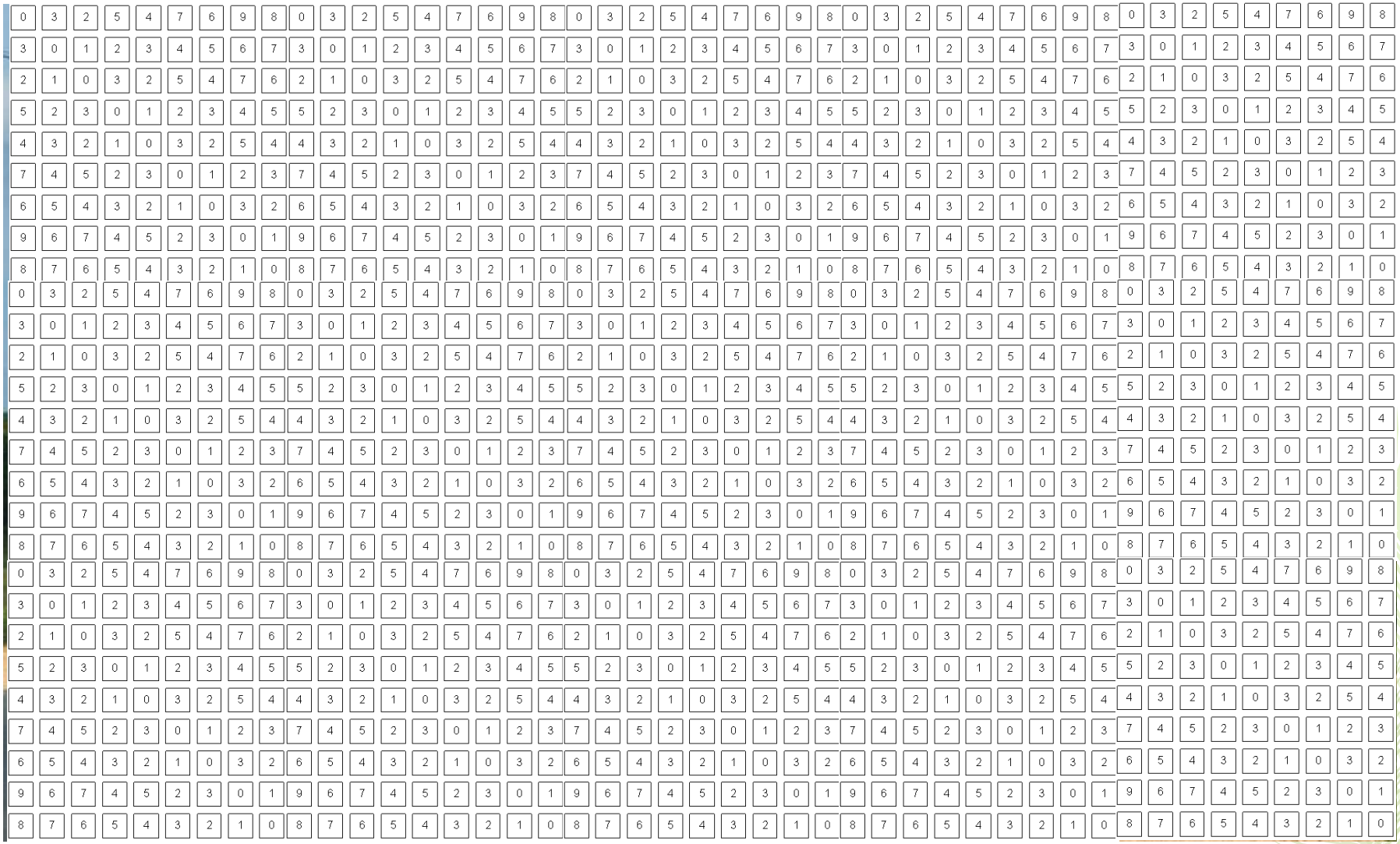
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

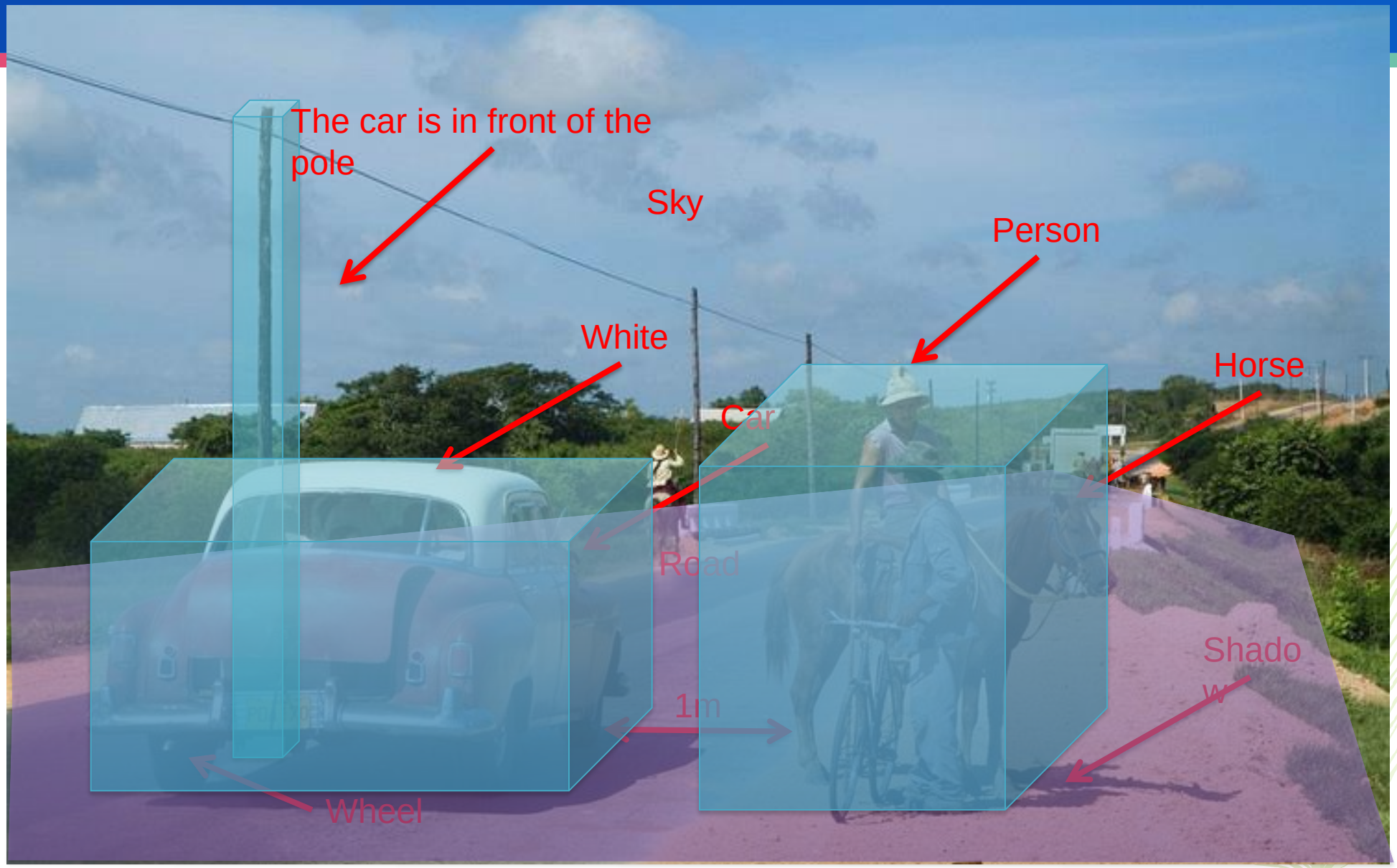


Marvin Minsky, MIT
Turing award,1969

Gerald Sussman, MIT
(the undergraduate)

"You'll notice that Sussman never worked in vision again!" – Berthold Horn



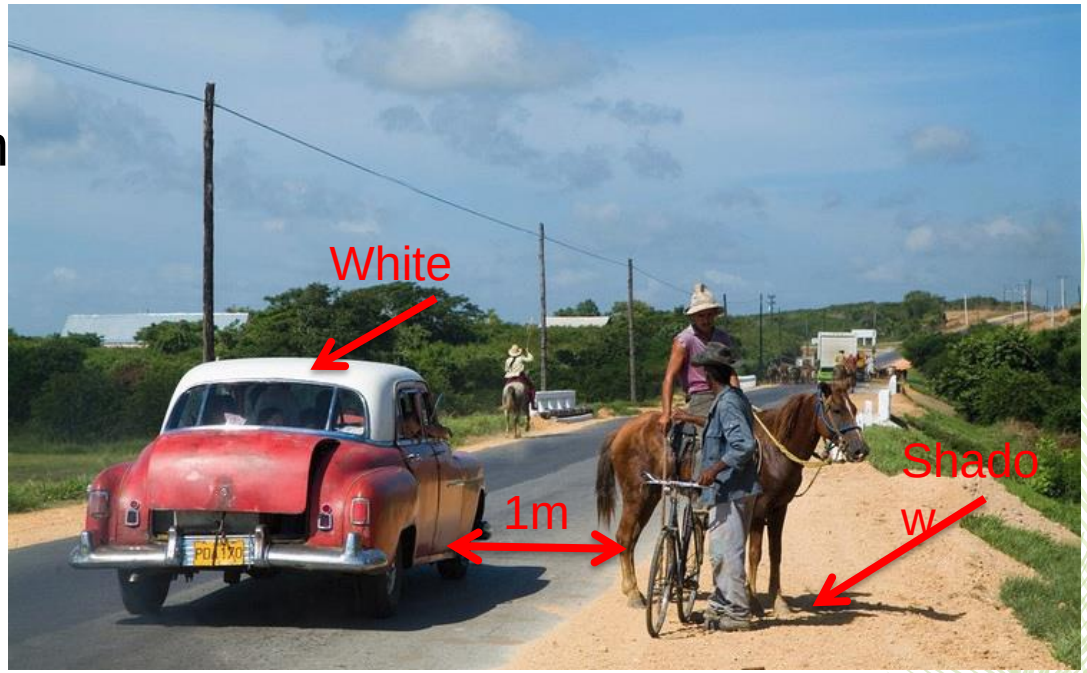

The car is in front of the pole
Sky
Person
White
Horse
Car
Road
Shado
w
1m
Wheel

Computer Vision

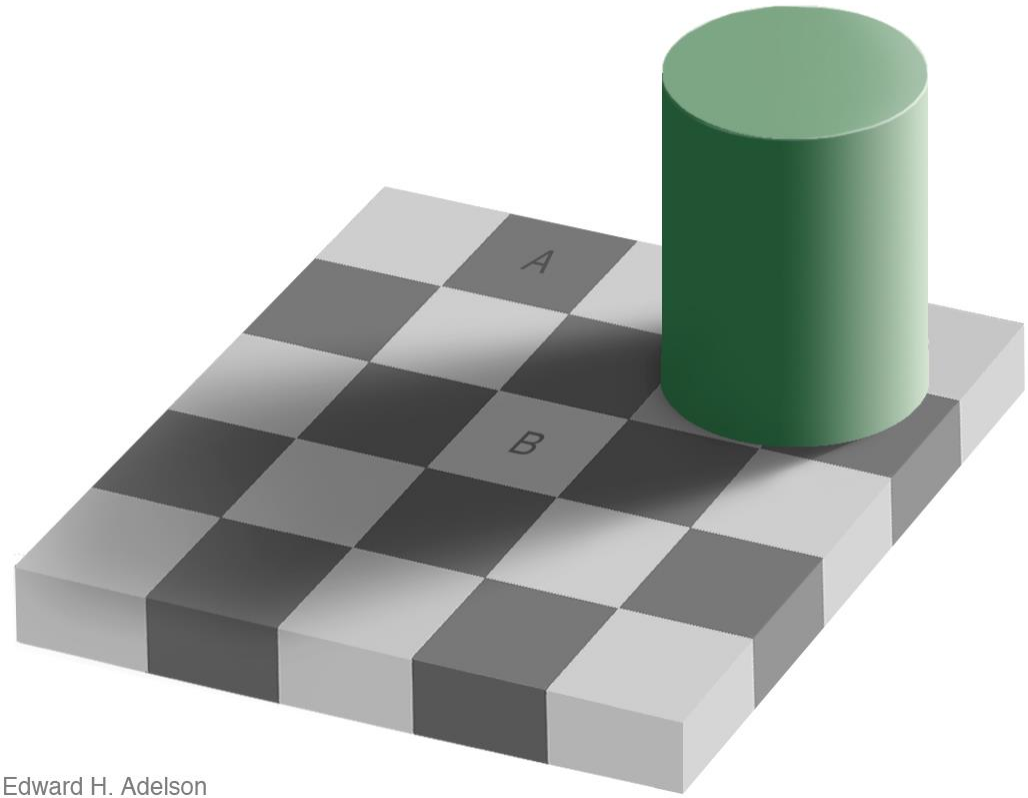
- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings

Computer Vision

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

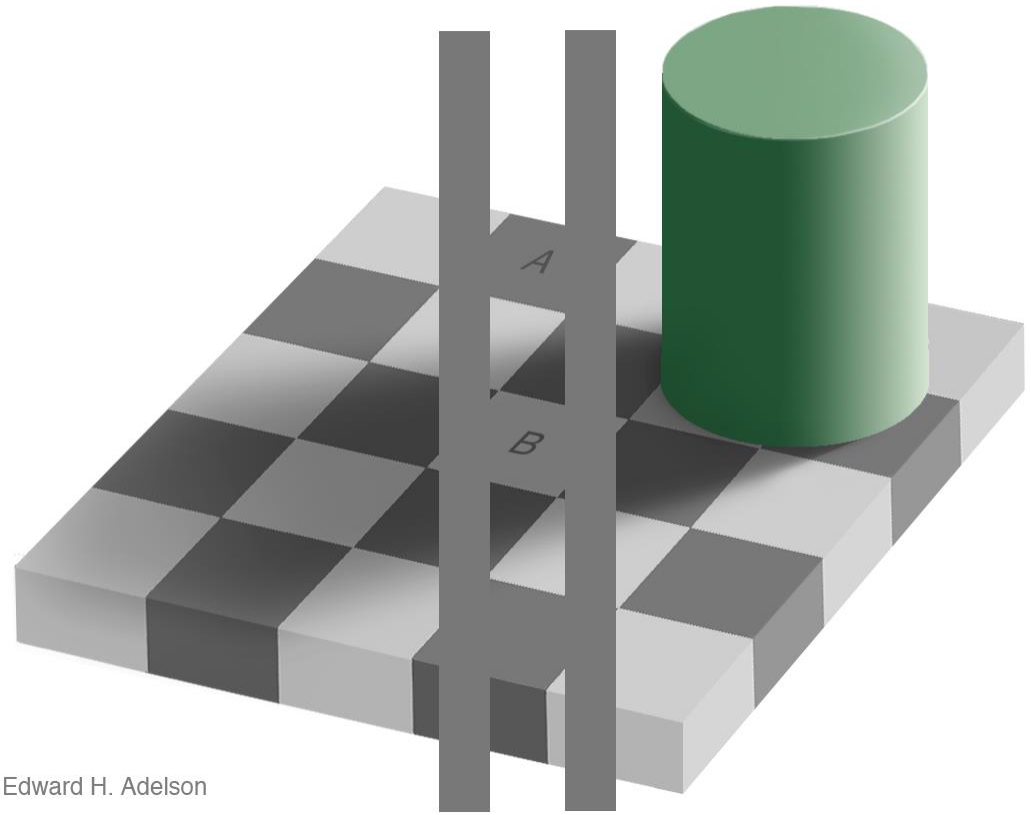


Measurement



Edward H. Adelson

Measurement

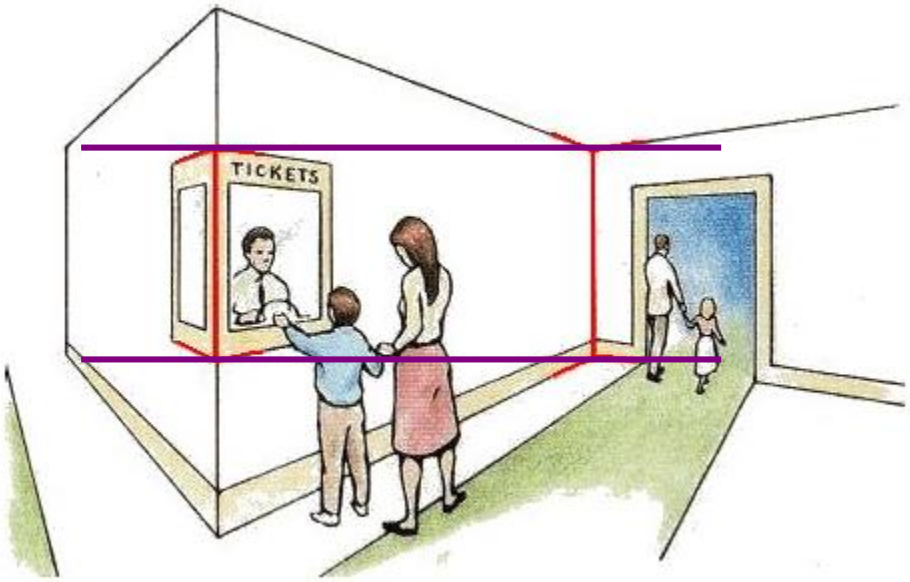


Edward H. Adelson

http://www.newworl

Slide Credit: Alyosha

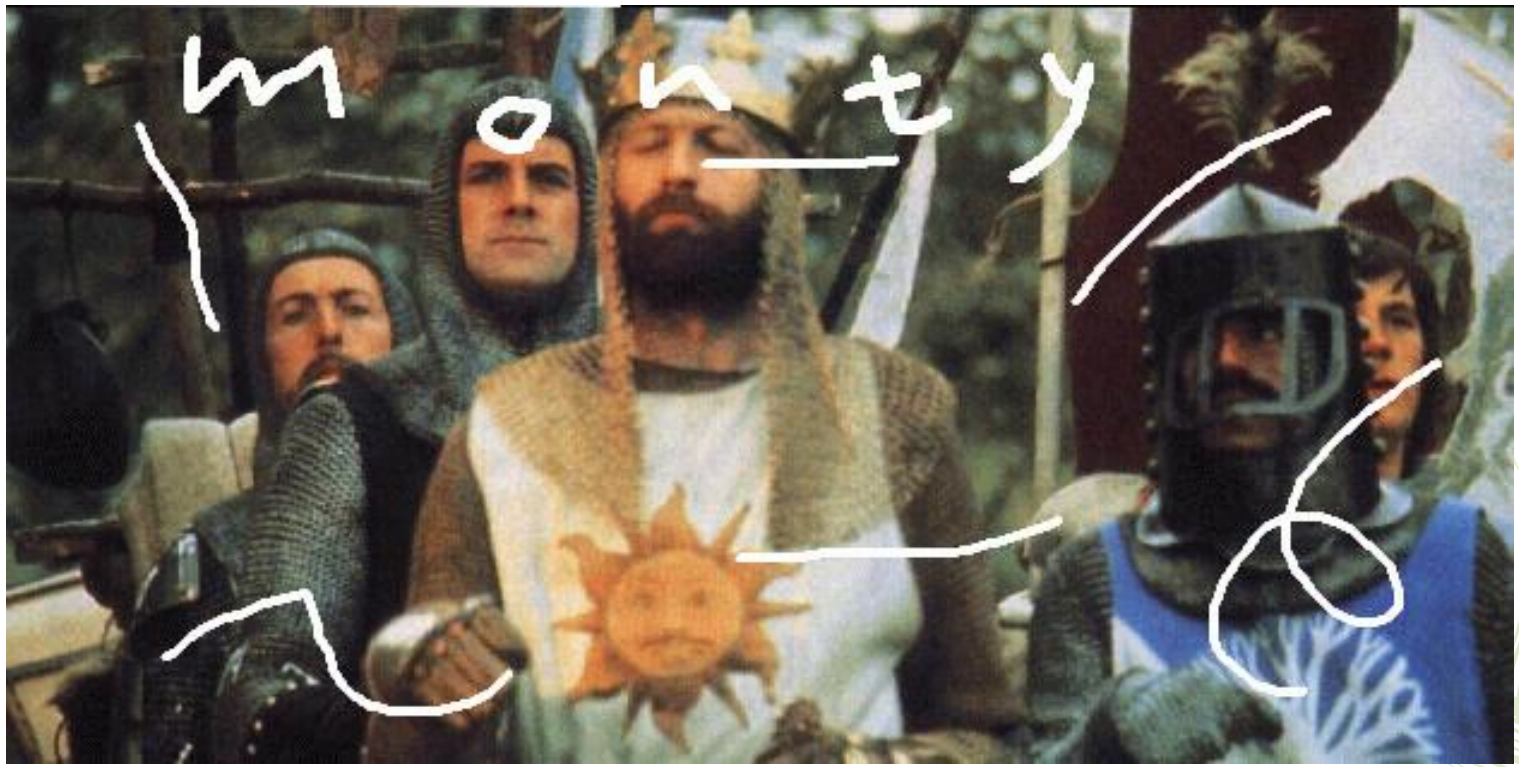
Measurement



http://www.michaelbach.de/ot/sze_muelue/index.html

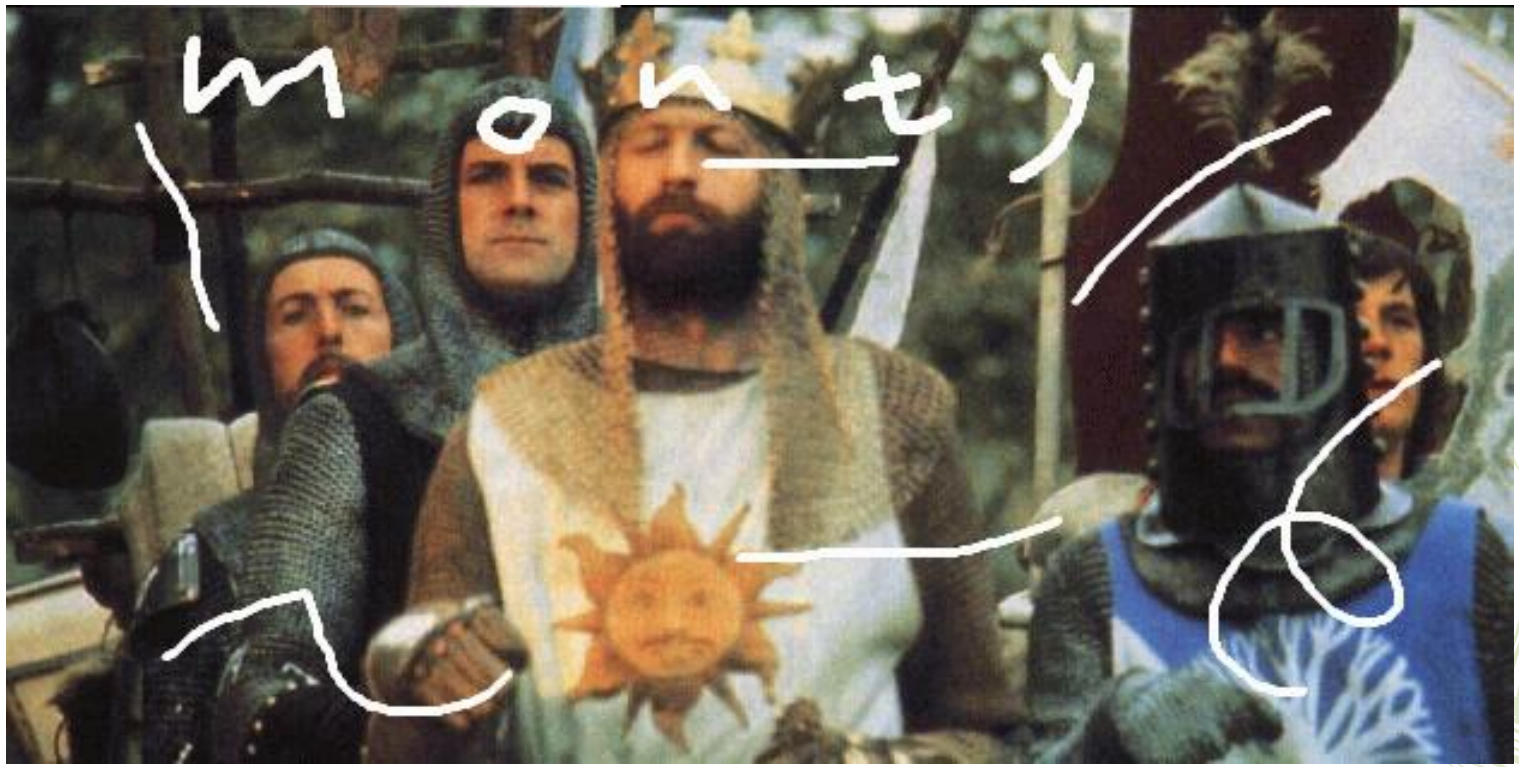
Slide Credit: Alyosha

Image Enhancement



http://www.iua.upf.es/~mbertalmio//restoration.html

Image Enhancement

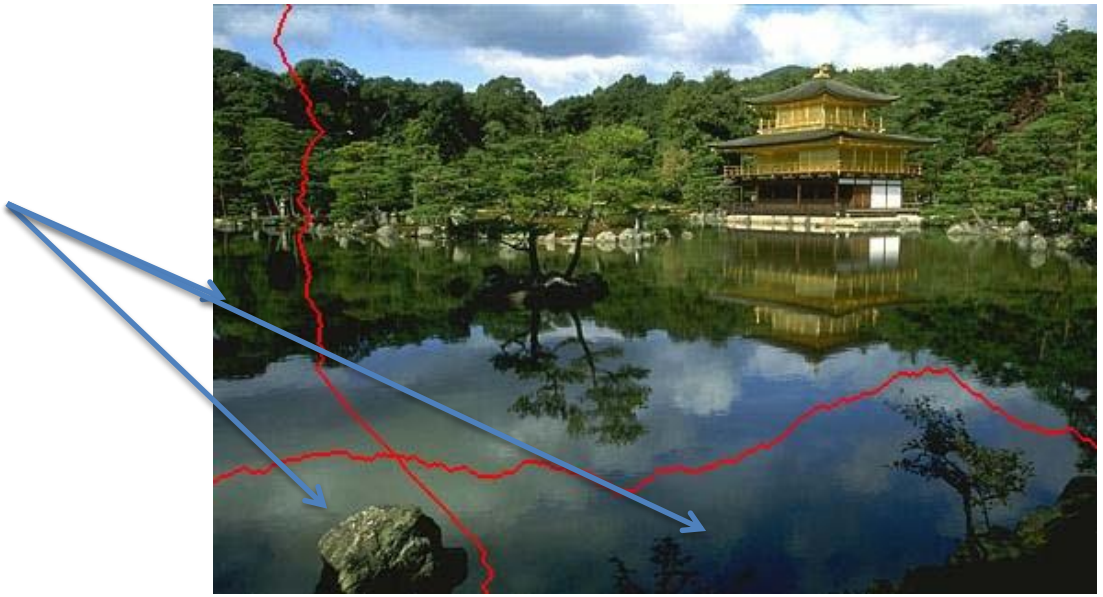


http://www.iua.upf.es/~mbertalmio//restoration.html

Image Enhancement

http://www.iua.upf.es/~mbertalmio//restoration.html

Seam Carving



[Shai & Avidan, SIGGRAPH 2007]

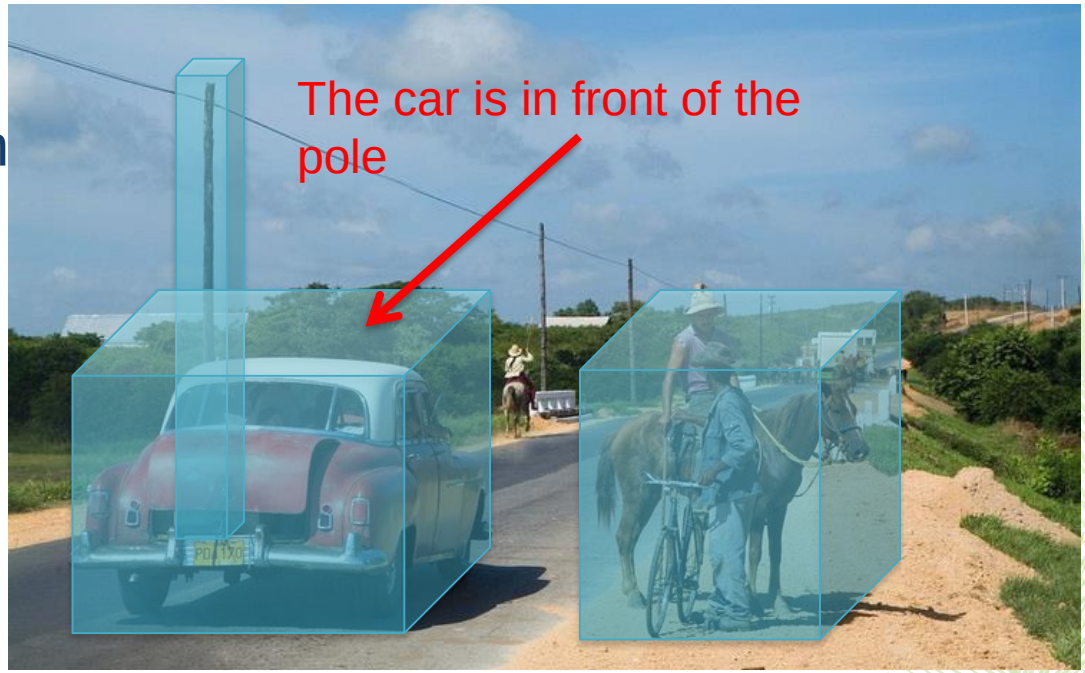
Traditional resizing uses and stretches the whole image.

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

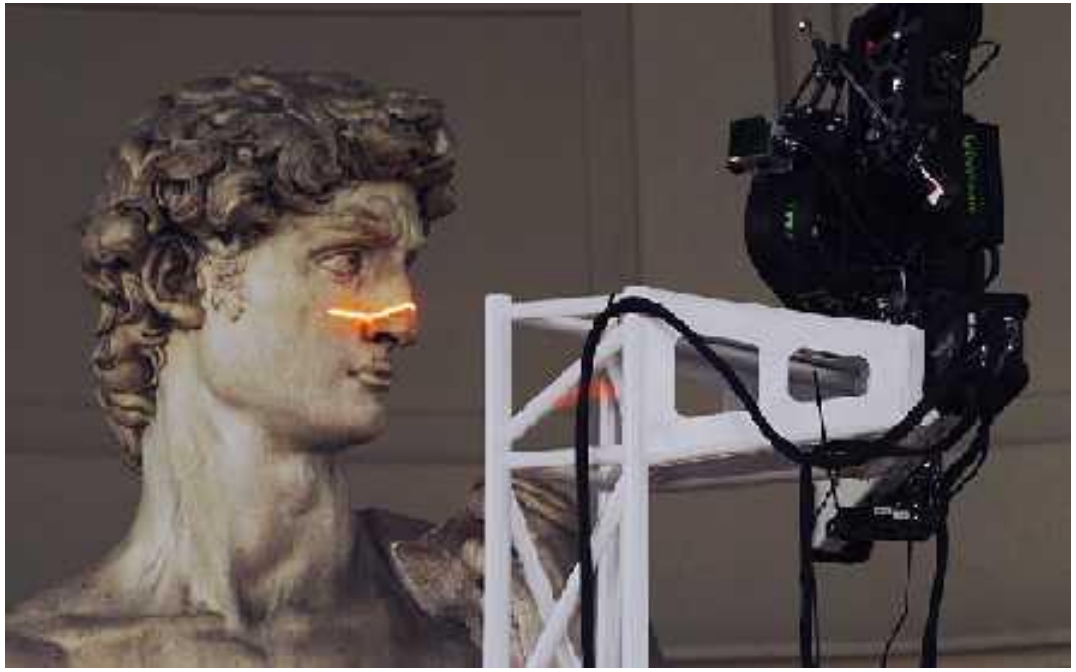
[Shai & Avidan, SIGGRAPH 2007]

Computer Vision

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

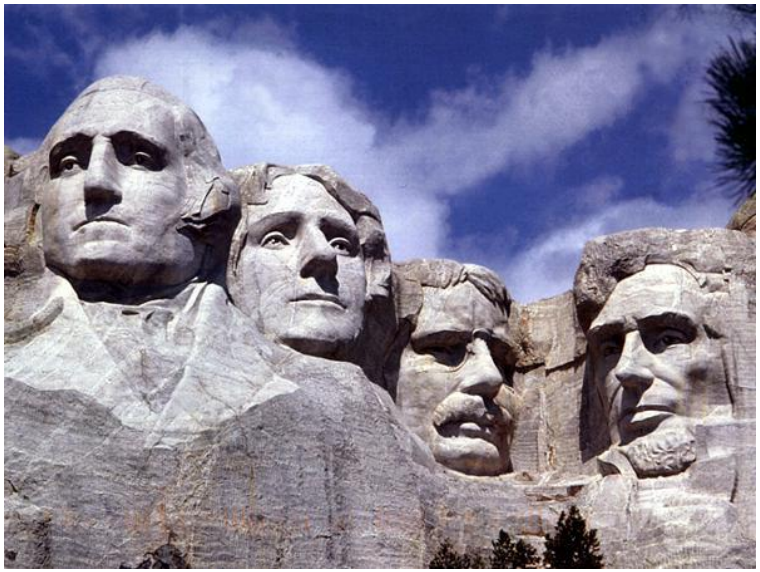


Applications: 3D Scanning

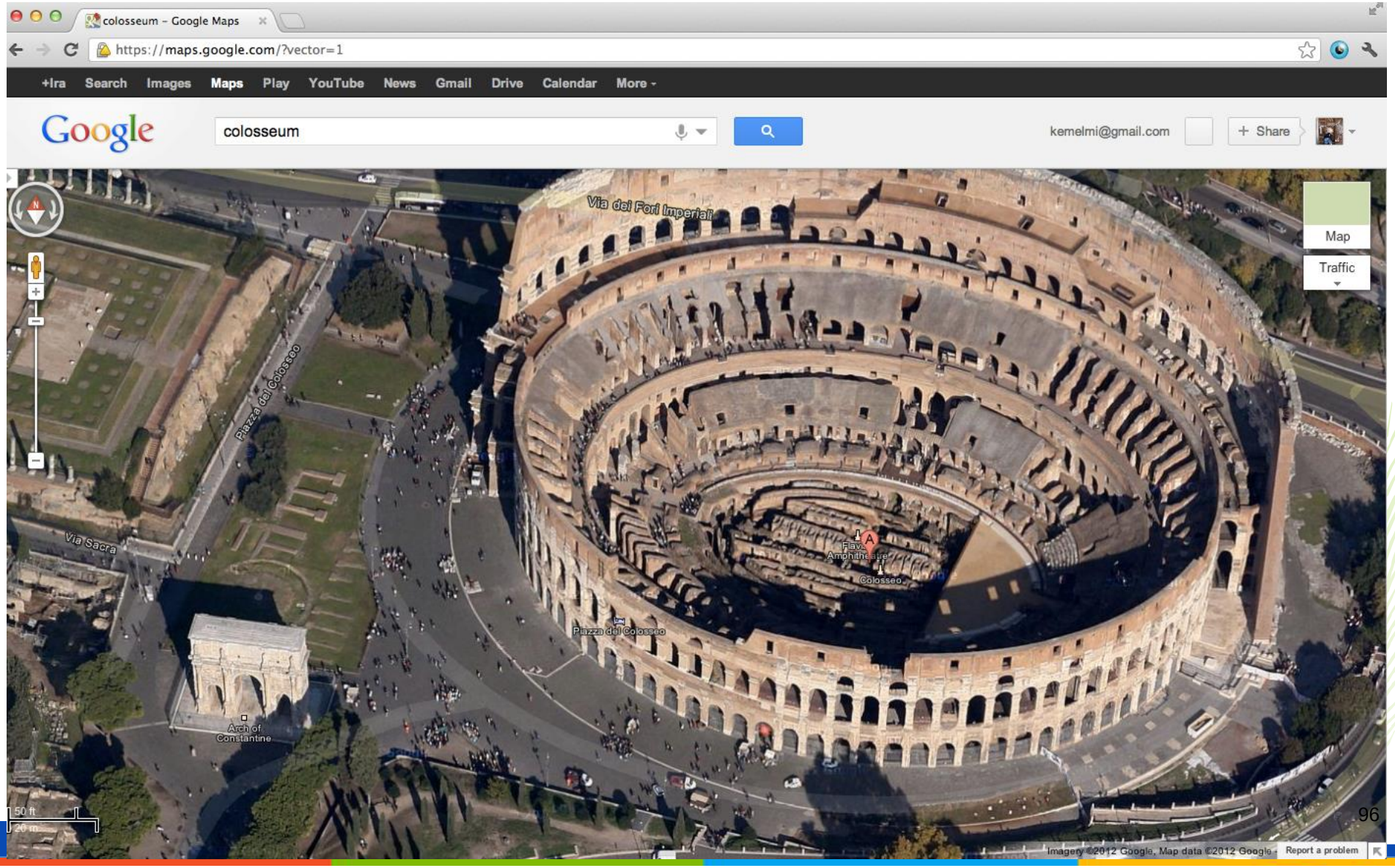


The Digital Michelangelo Project

Brian Curless



Google´s 3D Maps
Structure estimation from tourist photos

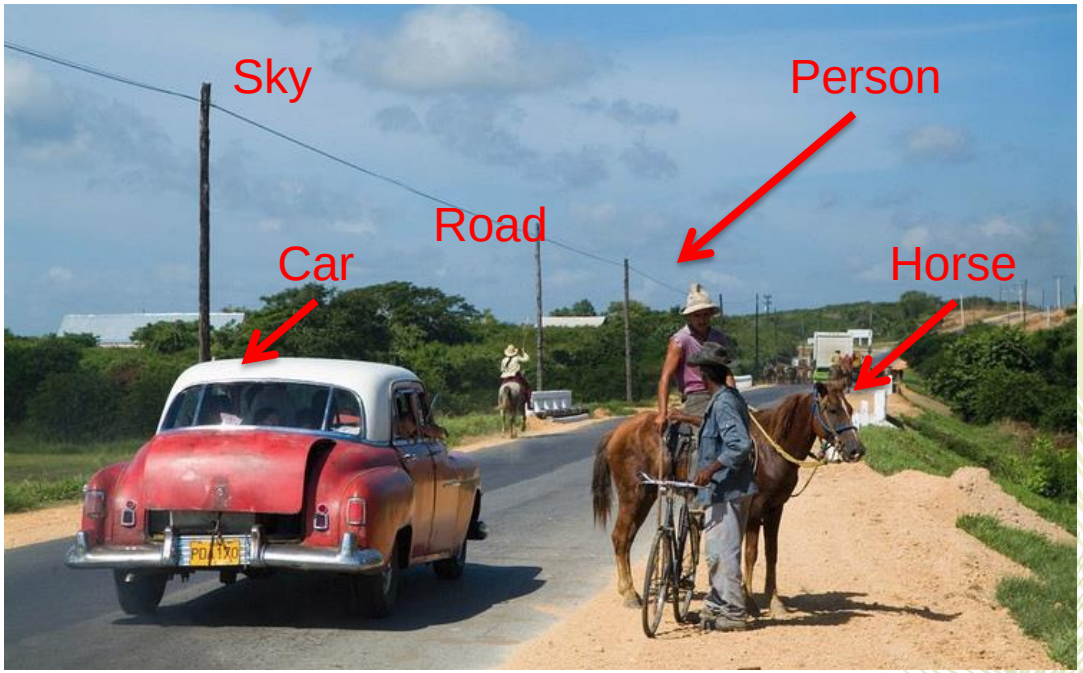


Apple's 3D maps

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

Computer Vision

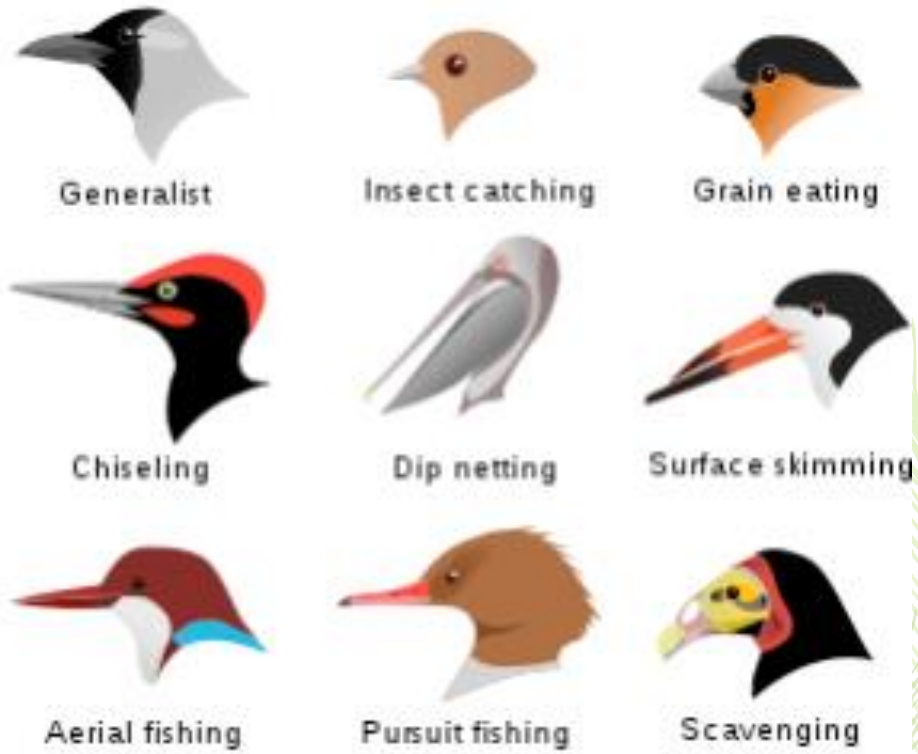
- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings
  - Pose estimation



Computer vision in other scientific fields

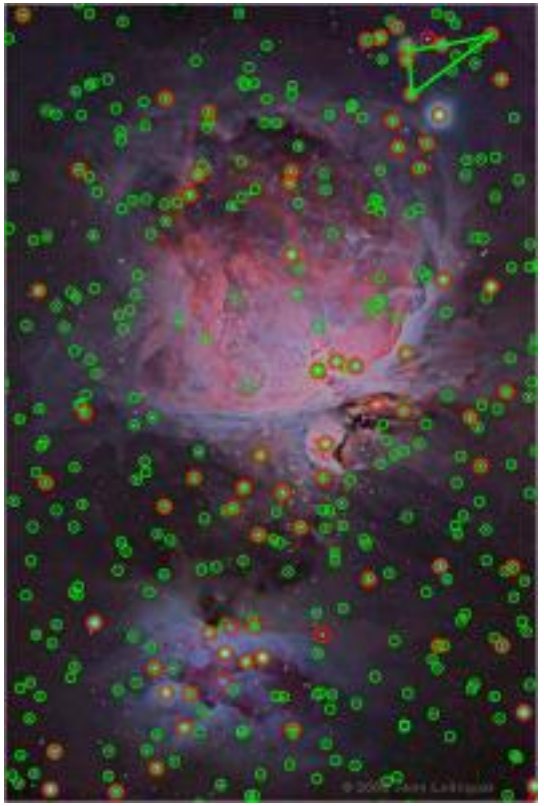
Computer vision research in biology

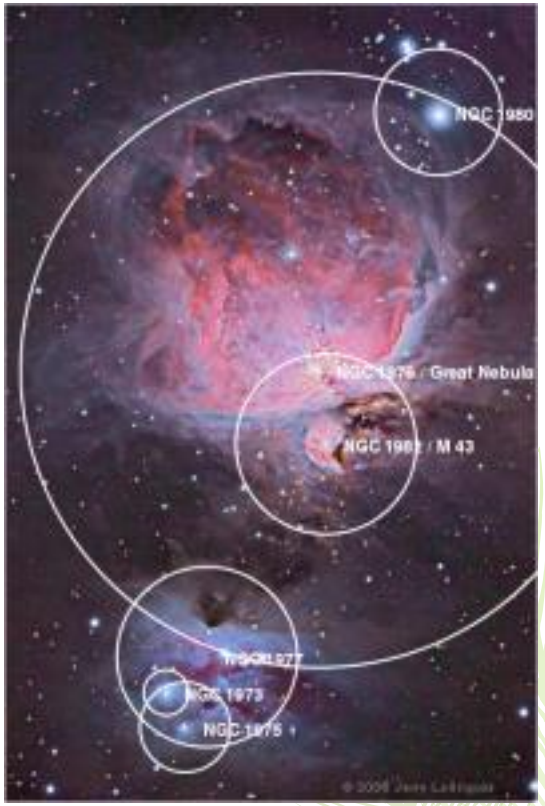
http://leafsnap.com/



http://www.vision.caltech.edu/visipedia/

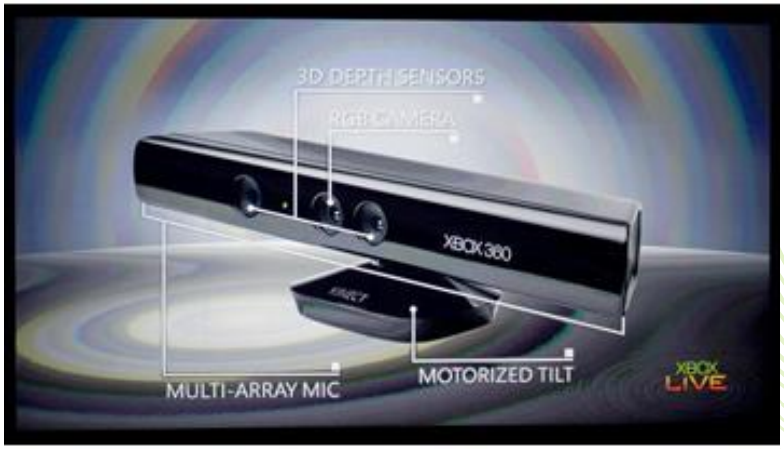
Computer vision in cosmology





http://astrometry.net/

Vision-based interaction: Xbox Kinect

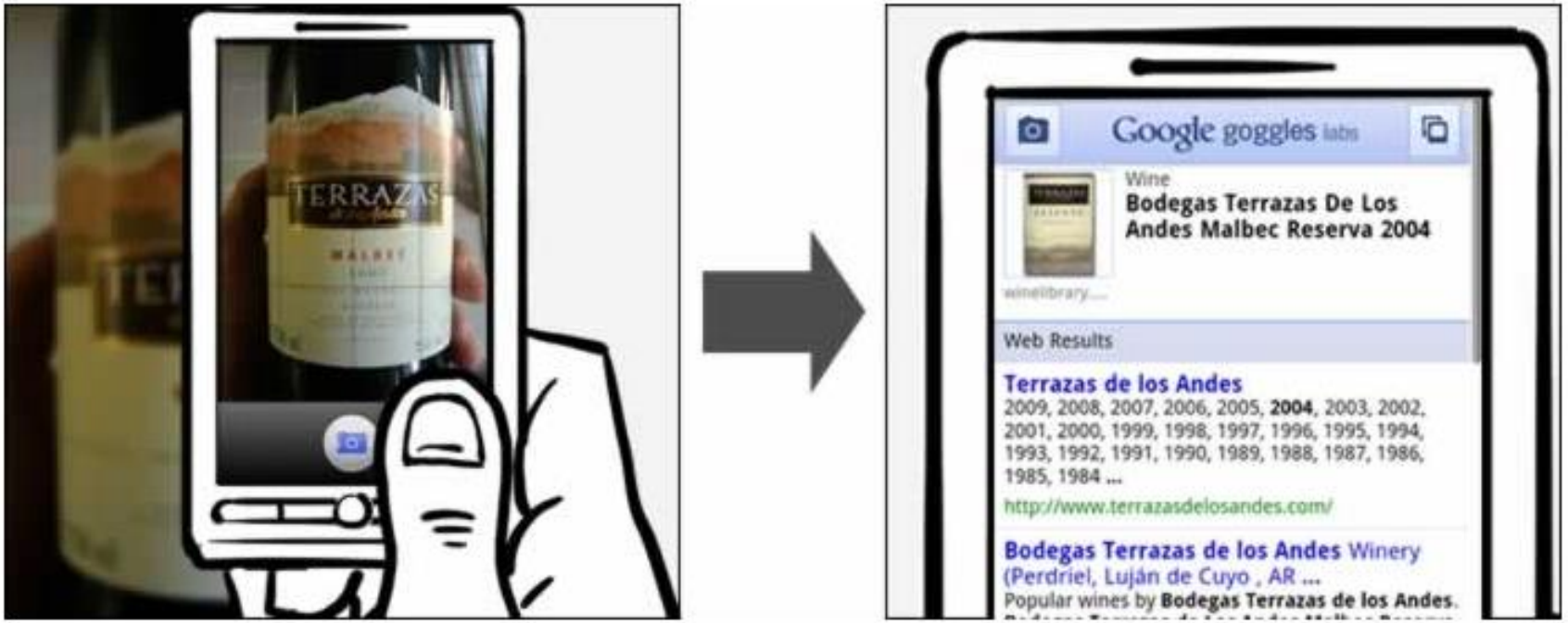


Mobile visual search: Google Goggles

Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.

Google goggles labs



Object recognition (in supermarkets)

LaneHawk by EvolutionRobotics

Safety

BBC NEWS

Watch One-Minute World News

News Front Page

Africa
Americas
Asia-Pacific
Europe
Middle East
South Asia
UK
England
Northern Ireland
Scotland
Wales
UK Politics
Education
Magazine
Business
Health
Science &

Last Updated: Wednesday, 31 August 2005, 05:44 GMT 06:44 UK

E-mail this to a friend Printable version

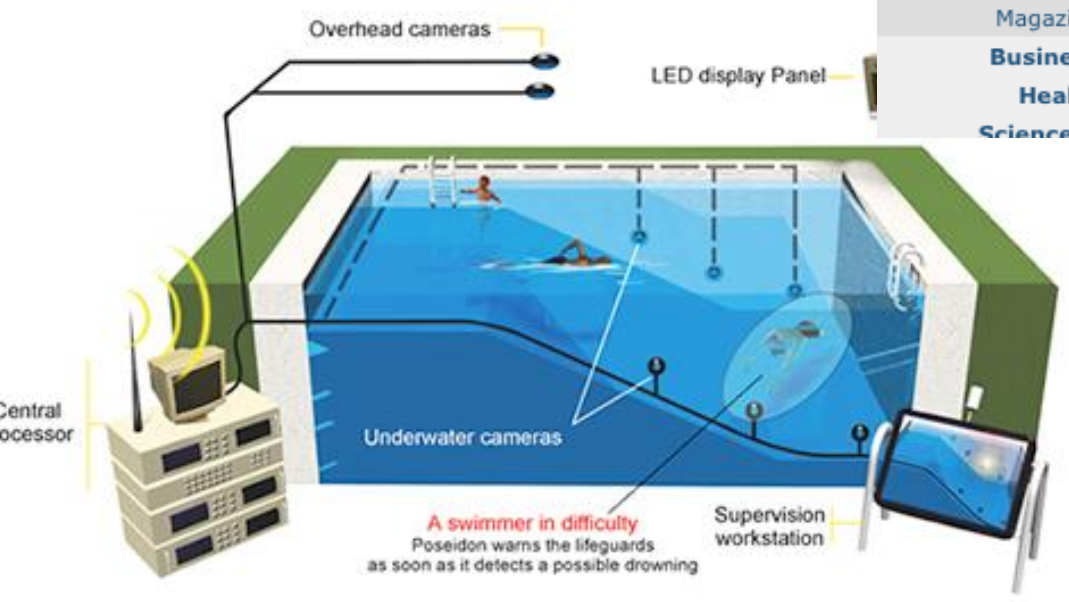
Computer alert for drowning girl

A 10-year-old girl has been saved from drowning by a computer system designed to raise the alarm when swimmers get into difficulties.

VIDEO Watch the rescue

The girl, from Rochdale, was at the deep end of the pool in Bangor, north Wales, when she sank to the bottom.

The £65,000 system, called Poseidon, detected her on the pool floor and sounded the alarm. A lifeguard pulled her out and she recovered in hospital.



Security

Local

Cameras help confirm Scott suicide ruling

Friday, December 04, 2009

TAGS: local, paul meincke

Comment Now Email Print Report a typo

Paul Meincke
More: Bio, News Team

December 4, 2009 (CHICAGO) (WLS) -- Chicago police have closed the case in the death of Chicago School Board President Michael Scott.

Police Supt. Jody Weis says investigators used police cameras in the city to trace Scott's last steps in the hours before his body was found in November.

Scott's death has been ruled a suicide. The medical examiner's office concluded --not long after Scott's body was found -- that he had committed suicide. Police did not dispute the finding but wanted to pursue all the investigative leads they could. They say they have done that and have now reached the same conclusion.

Share this Story

Recommend Be the first of your friends to recommend this.

Tweet 0 +1 Recommend this on Google

News Headlines Video abcNEWS

- 2 suspects arrested in volleyball star's murder 47 min ago
- BP Gas Recall: BP finds, fixes source of bad gas
- Teachers union, board resume negotiating
- Back to School
- 5 injured in South Side shooting 49 min ago
- Pastor: Stacy Peterson said she lied for Drew

Automotive safety

- Mobileye: Vision systems in high-end BMW, GM, Volvo models
  - Pedestrian collision warning
  - Forward collision warning
  - Lane departure warning
  - Headway monitoring and warning

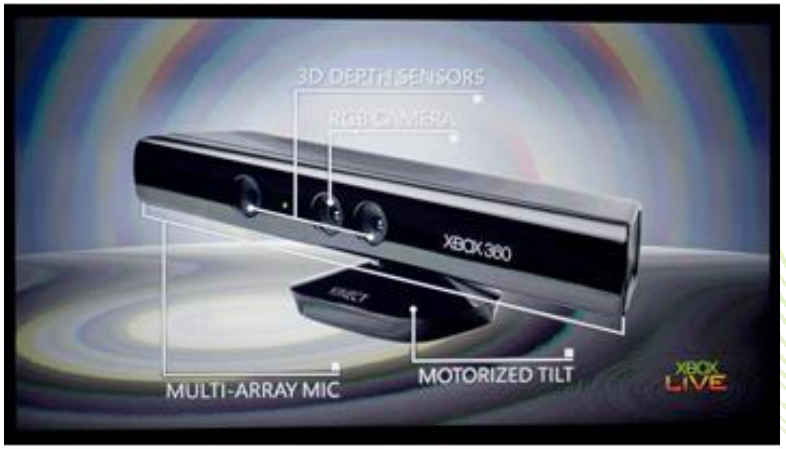
Google cars

"Google Cars Drive Themselves, in Traffic" *The New York Times*

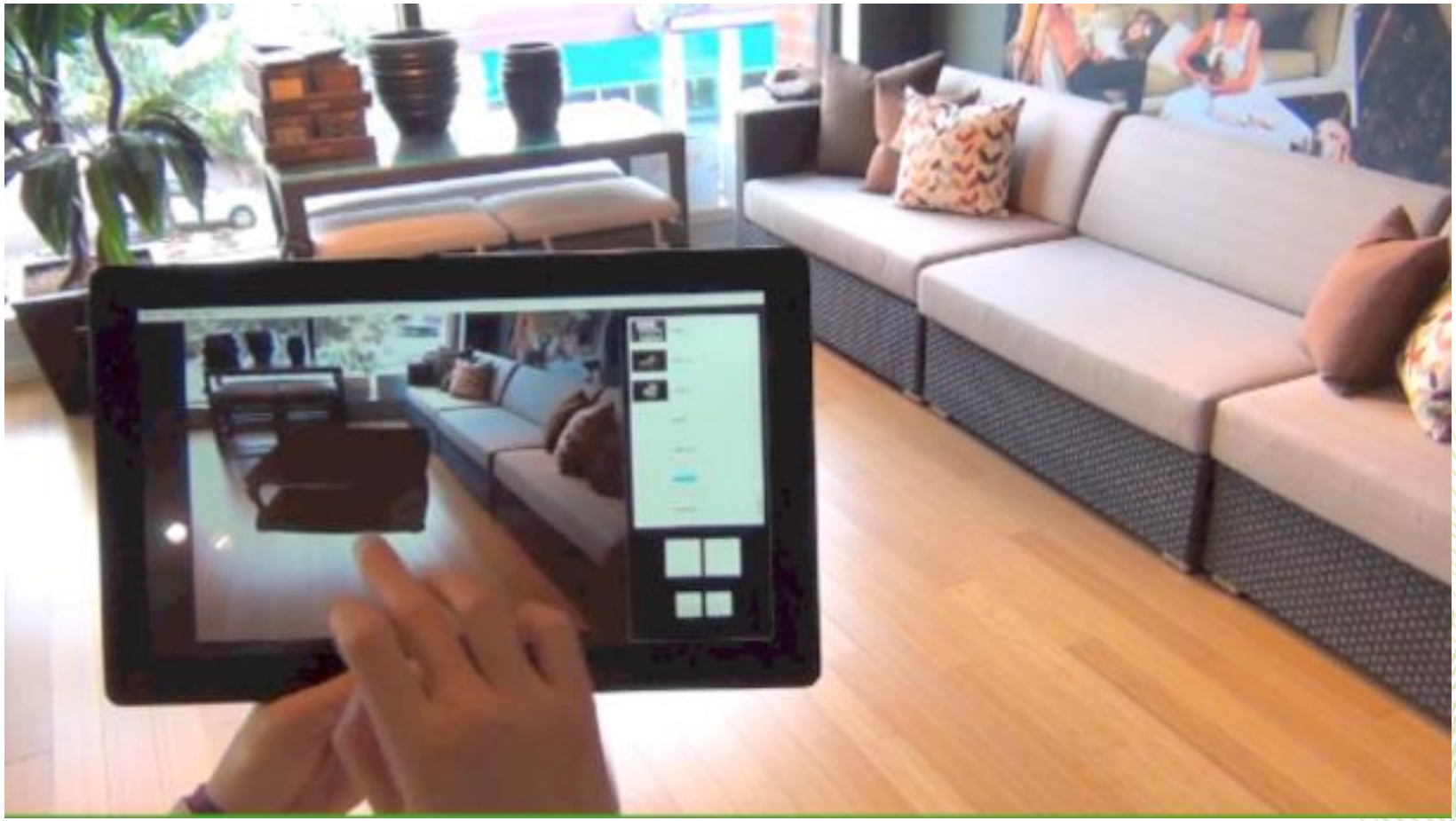
"Nevada state law paves the way for driverless cars" *Financial Post*

"Human error blamed after Google's driverless car sparks five-vehicle crash". *The Star*

Vision-based interaction: Xbox Kinect



Augmented reality, consumer products



http://nconnex.com/wp/

Special effects: shape and motion capture

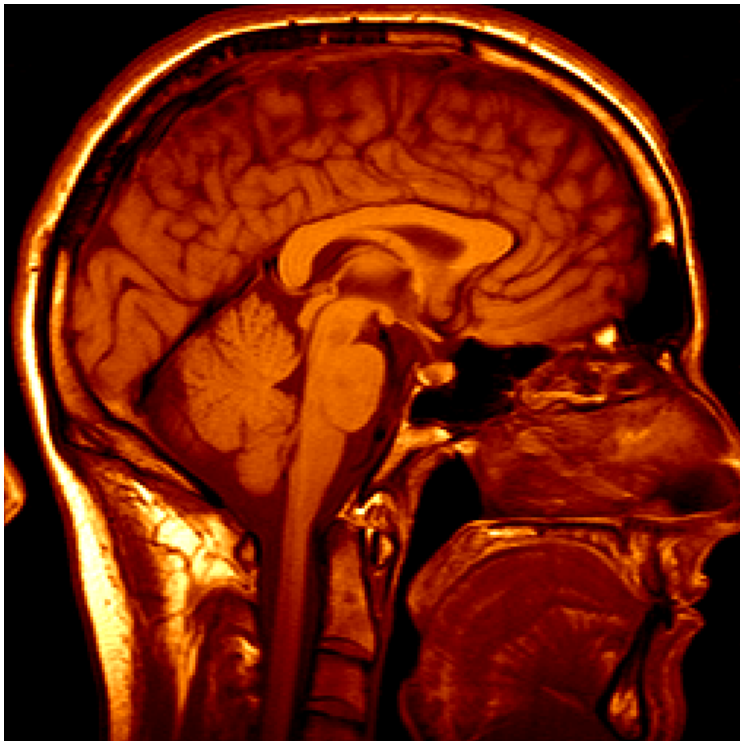
Source: S. Seitz

Vision for robotics, space exploration

NASA'S Mars Exploration Rover Spirit

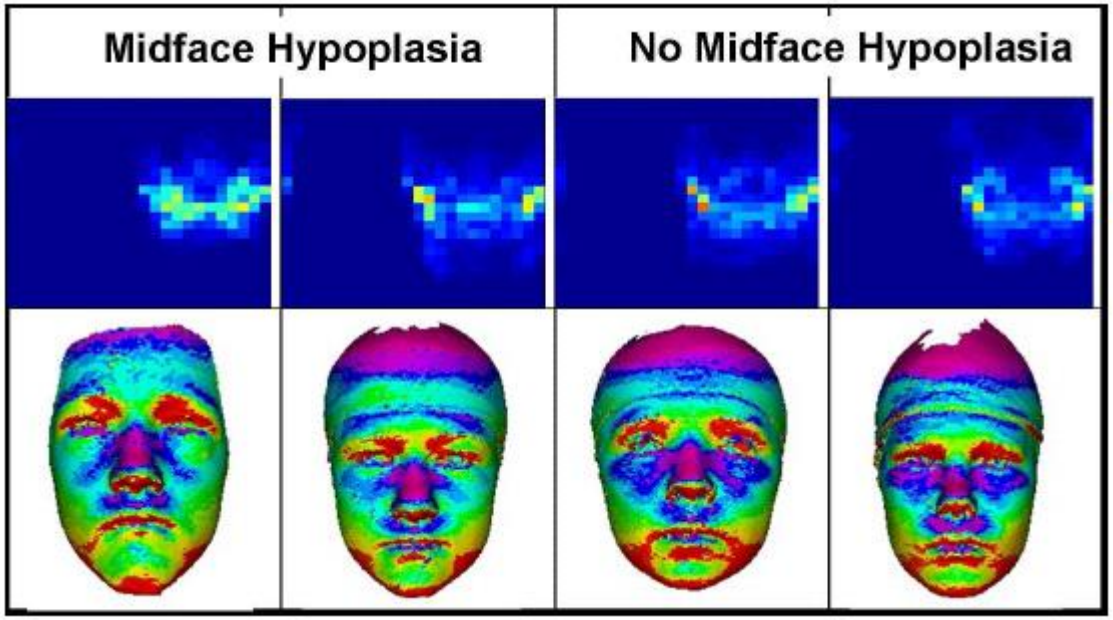
Computer Vision on Mars

Medical imaging



Grimson et al., MIT

Classification of 22q11.2DS



- Treat 2D azimuth-elevation angle histogram as feature vector

	8×8	16×16	24×24	32 × 32	Experts' median
Whole 2D hist	0.651	0.569	0.79	0.684	0.68