

# Lecture 1: Introduction to “Computer Vision”

Harbin Institute of Technology ( Shenzhen )

Jingyong Su

[sujingyong@hit.edu.cn](mailto:sujingyong@hit.edu.cn)



哈爾濱工業大學(深圳)  
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Today's agenda

- What is computer vision
- Course overview



# A picture is worth a thousand words



Gali Tibbon / AFP



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# A word is worth a thousand pictures



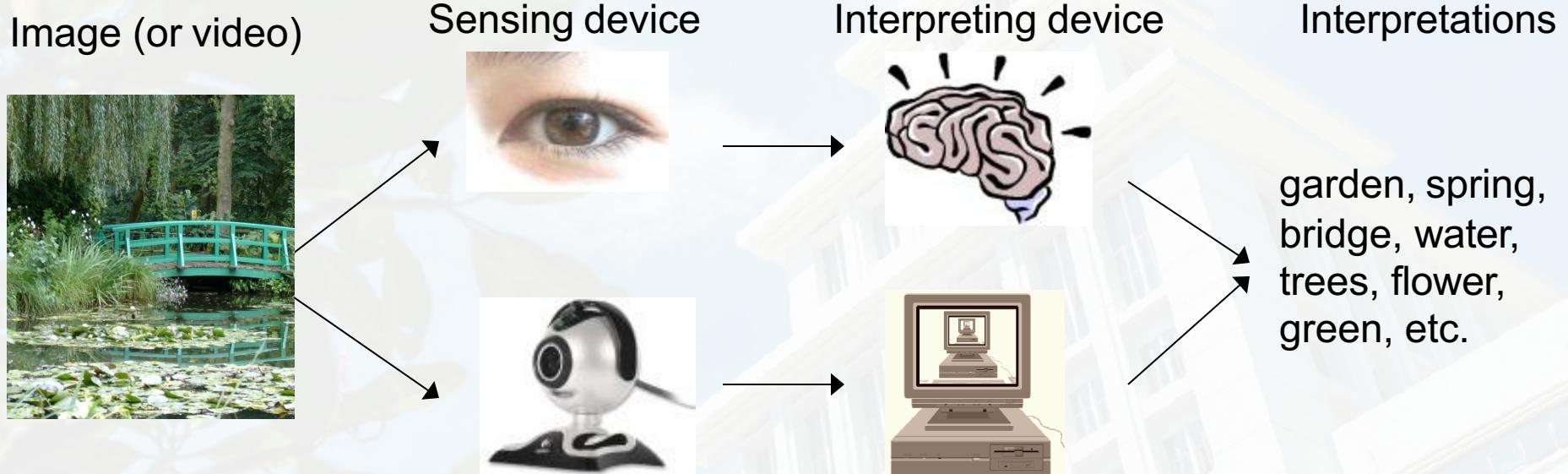
## A HUNT



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

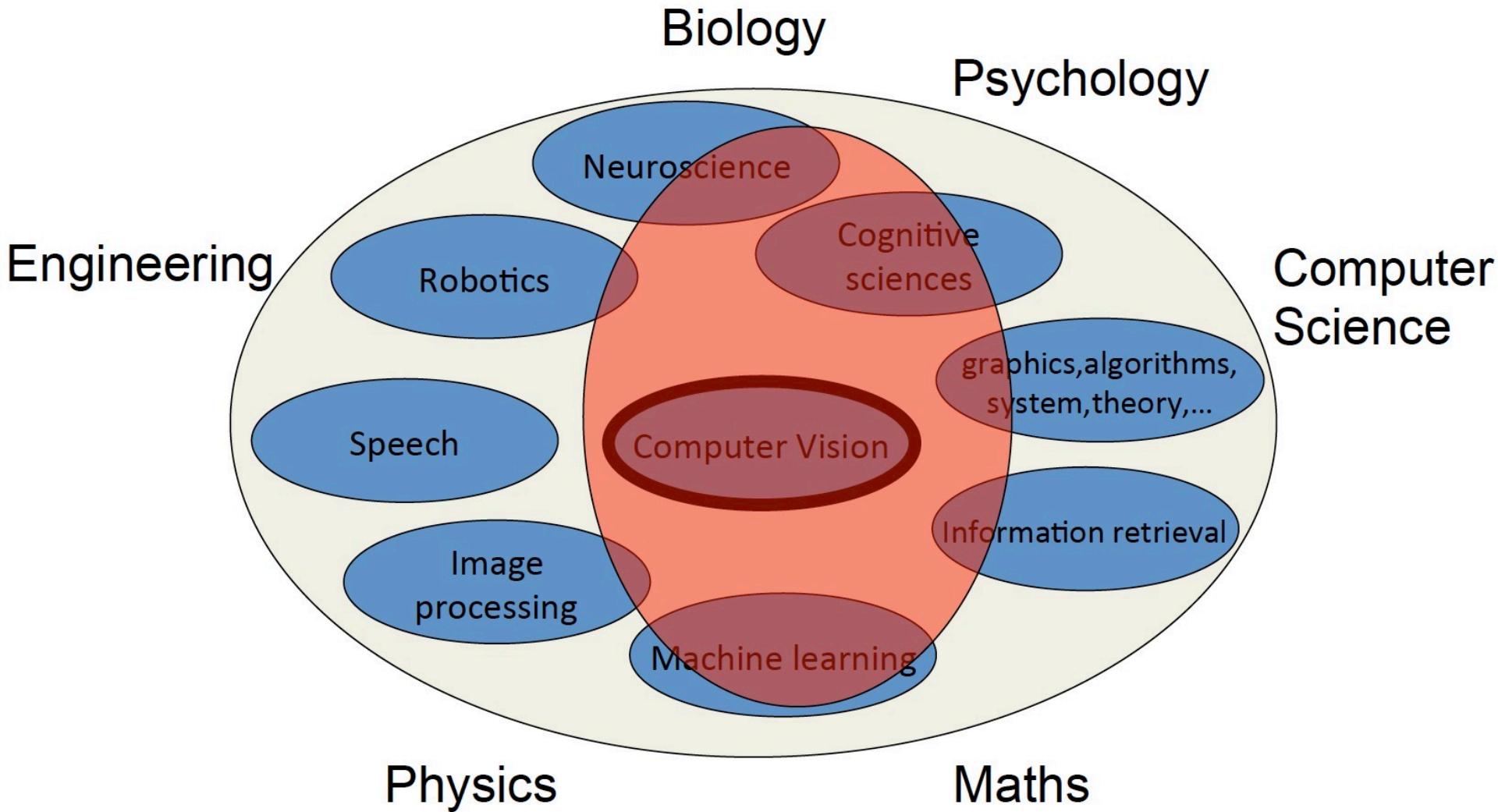
# What is (computer) vision?



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# What is related to?



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# The goal of computer vision

- To bridge the gap between pixels and “meaning”



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

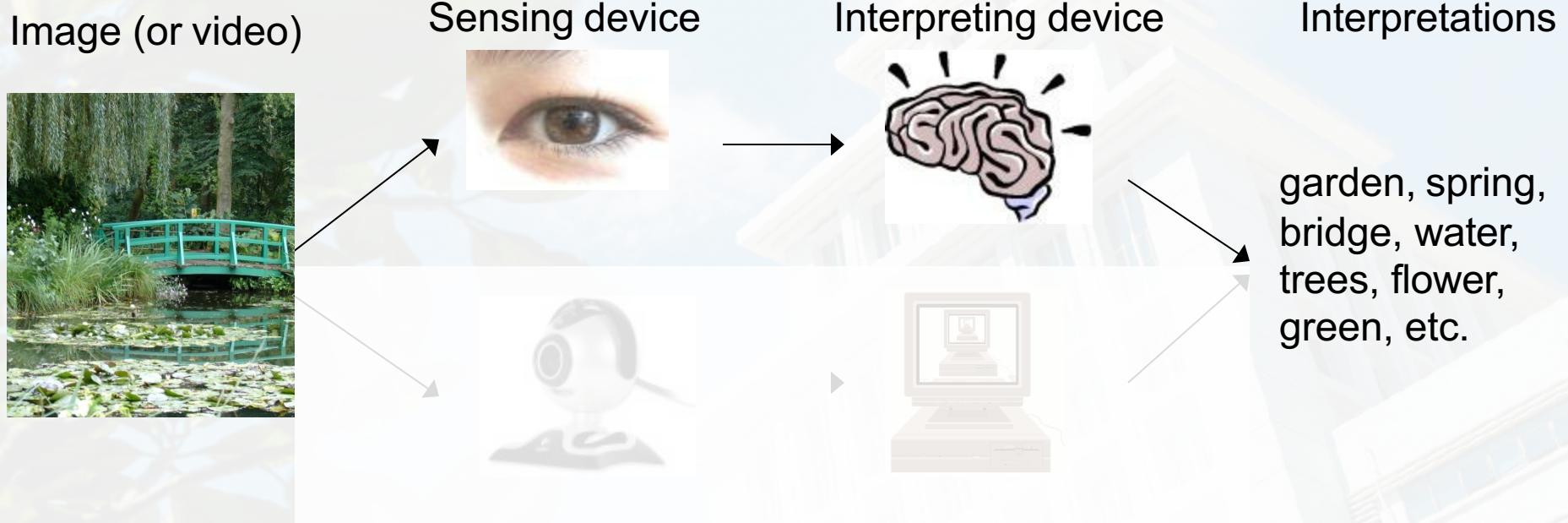
What a computer sees



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

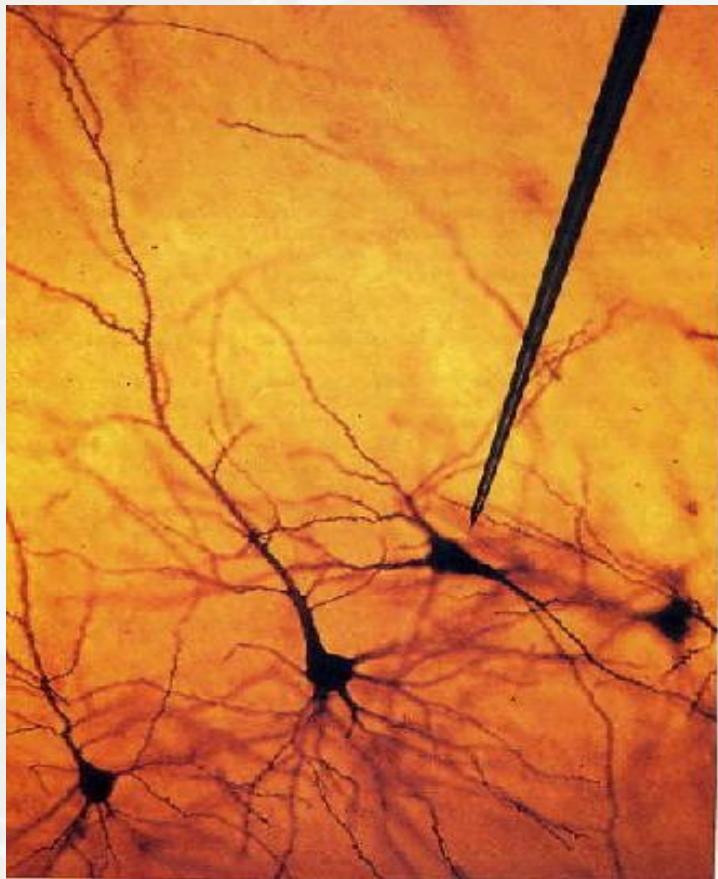
# What is (computer) vision?



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# 1981: Nobel Prize in medicine



Hubel & Wiesel at Harvard



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Human vision is superbly efficient

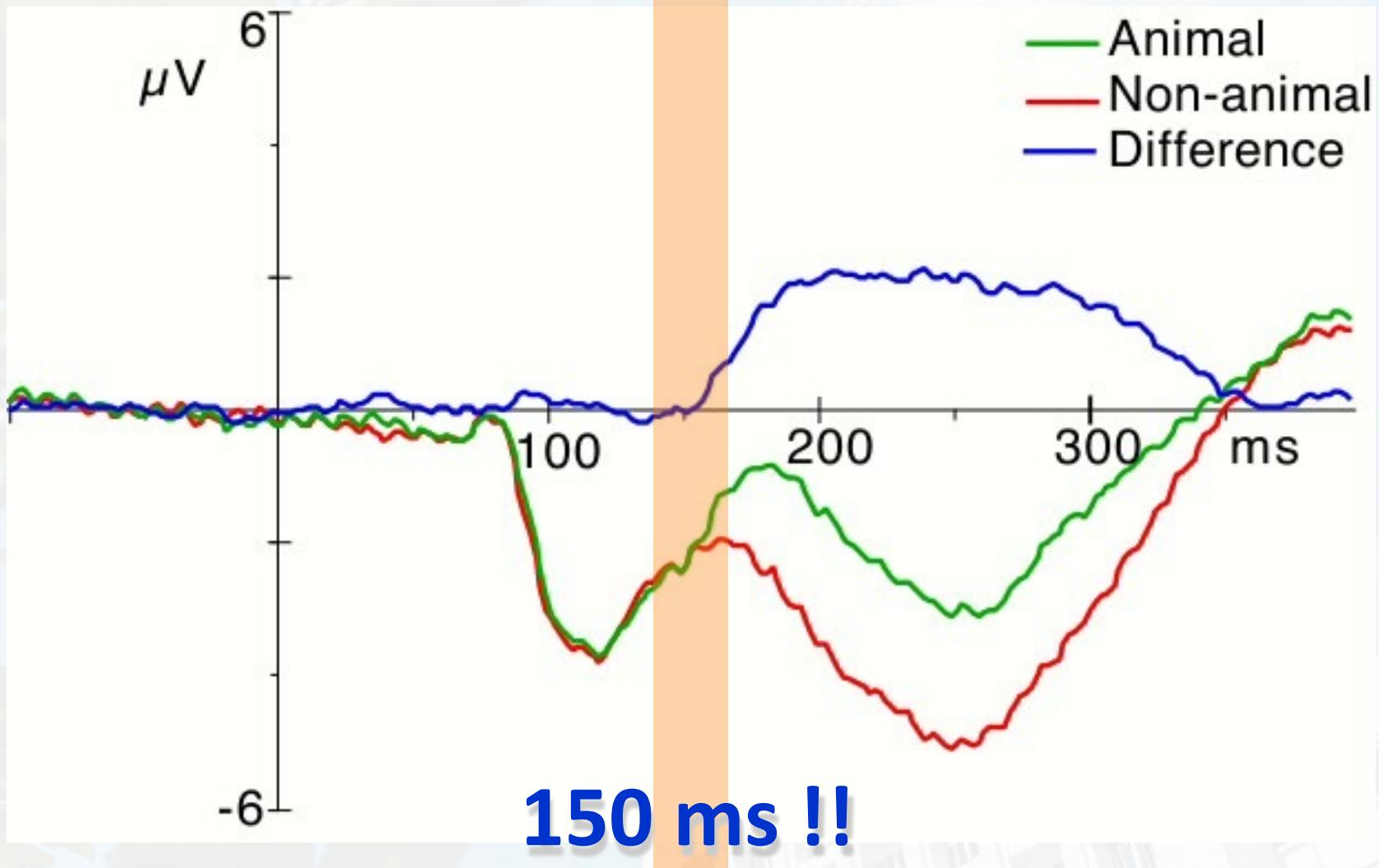


Thorpe, et al. *Nature*, 1996



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN



Thorpe, et al. *Nature*, 1996



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Fooling humans

However, this speed is obtained at the price of some drawbacks.

- Changing small irrelevant parts of an image such as water reflection or background can go unnoticed because the human brain focuses on the important parts of an image.
  
- If the signal is very close to the background, it can be difficult to detect and segment the relevant part of the image.



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Change blindness



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Change blindness



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Segmentation

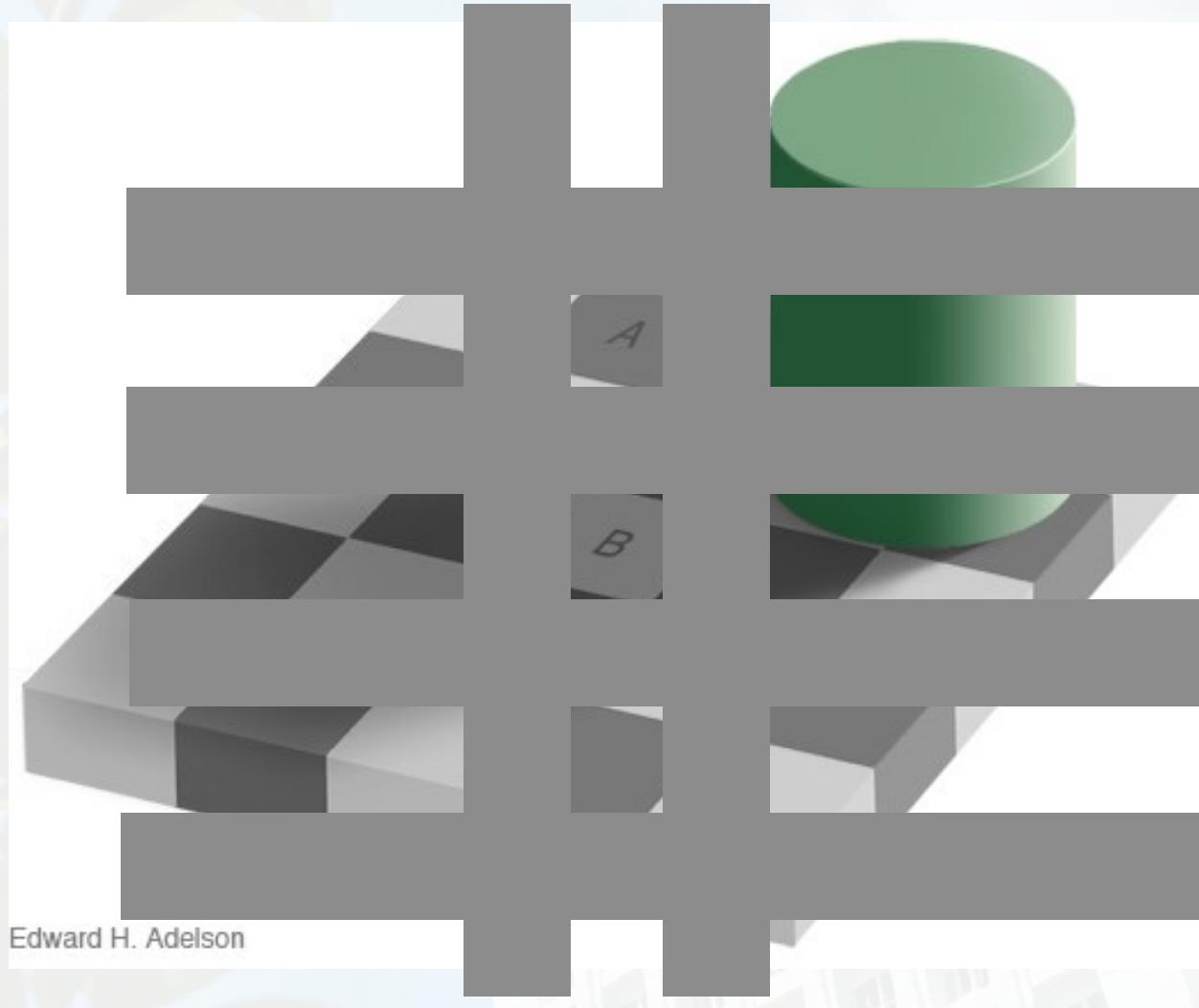
Dalmatian  
dog



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Perception



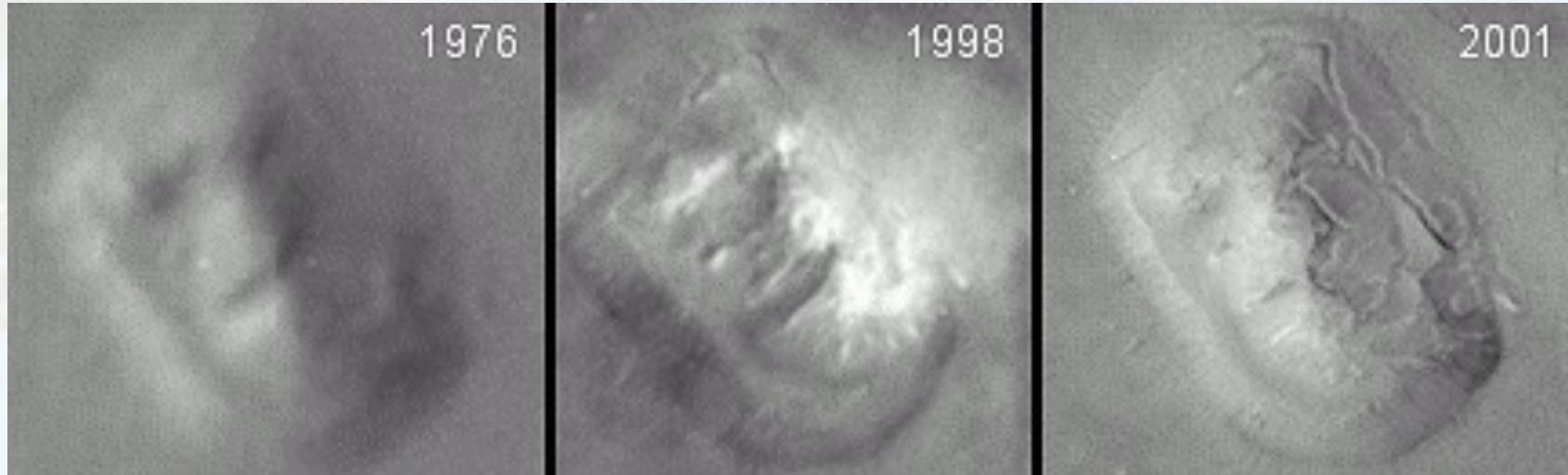
Edward H. Adelson



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Perception



Face on Mars



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Gestalt



哈爾濱工業大學(深圳)

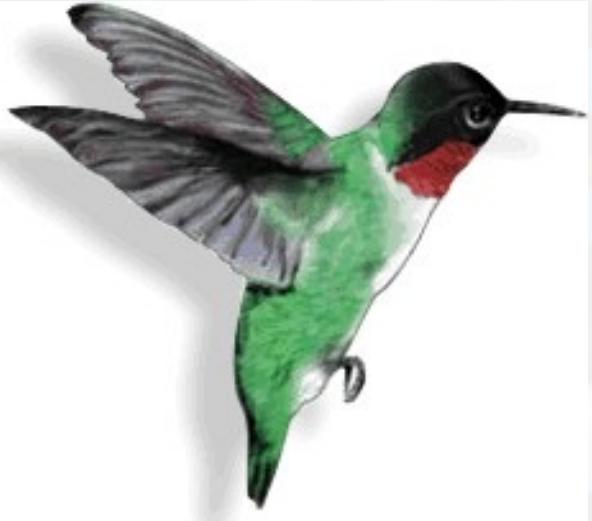
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Gestalt



哈爾濱工業大學(深圳)

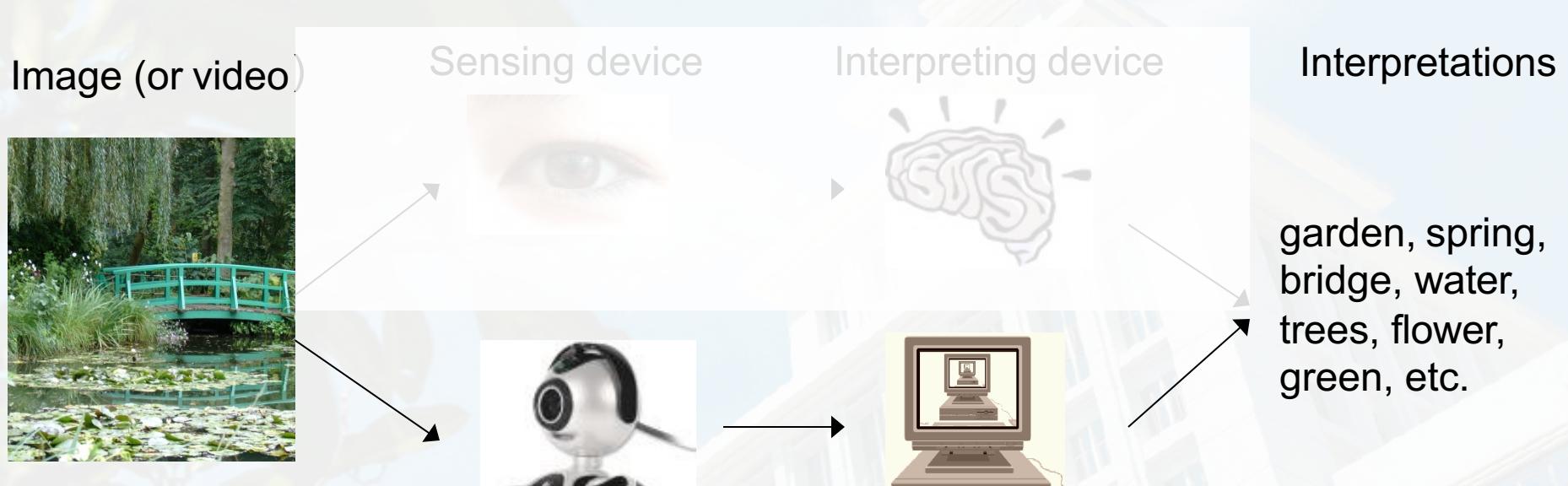
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# What is (computer) vision?



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# The goal of computer vision

- To bridge the gap between pixels and “meaning”



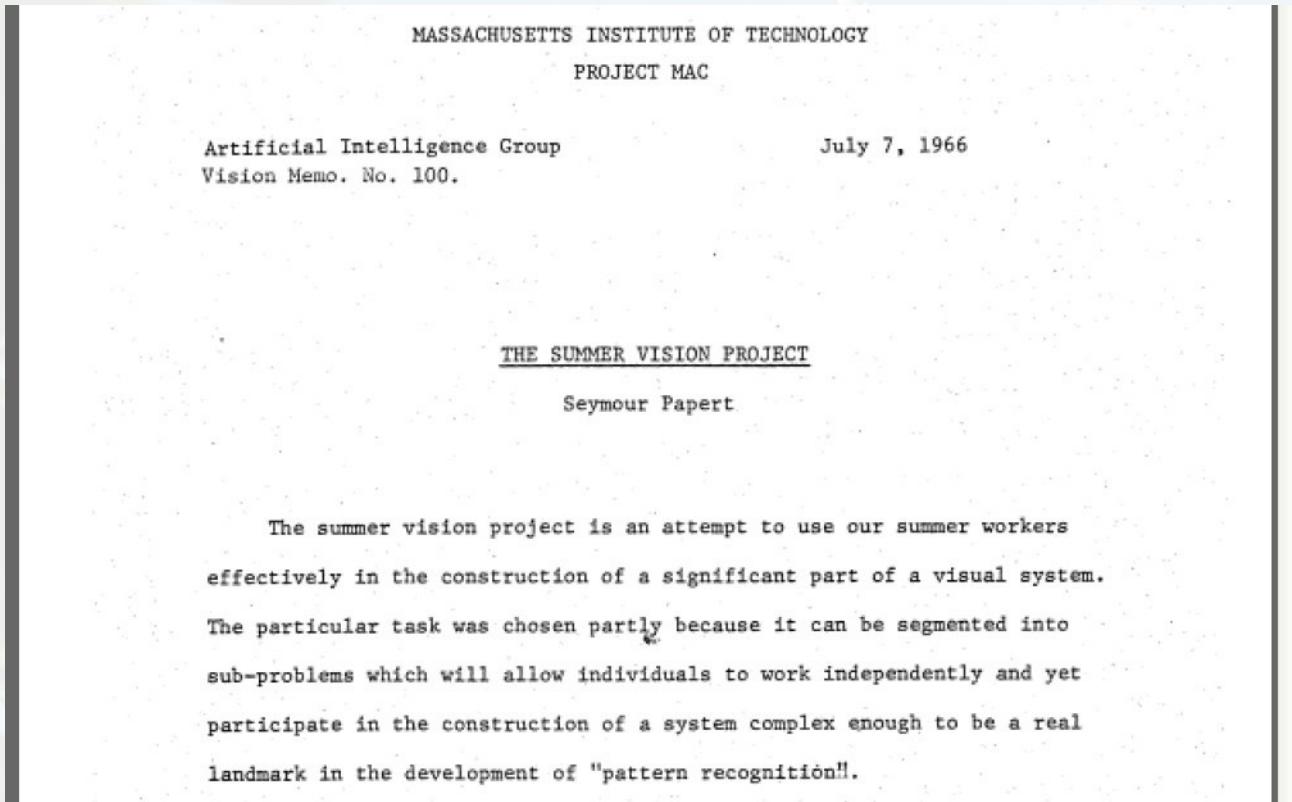
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Origins of computer vision: an MIT undergraduate summer project



哈爾濱工業大學(深圳)

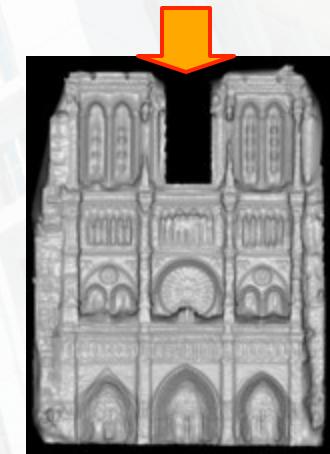
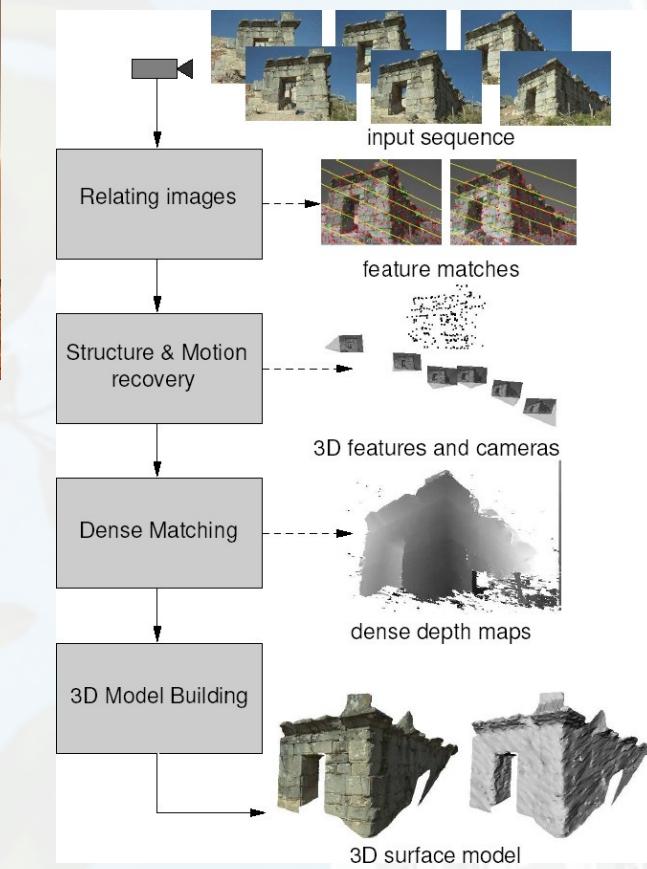
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# What kind of information can we extract from an image?

- Measurement
- Semantic information



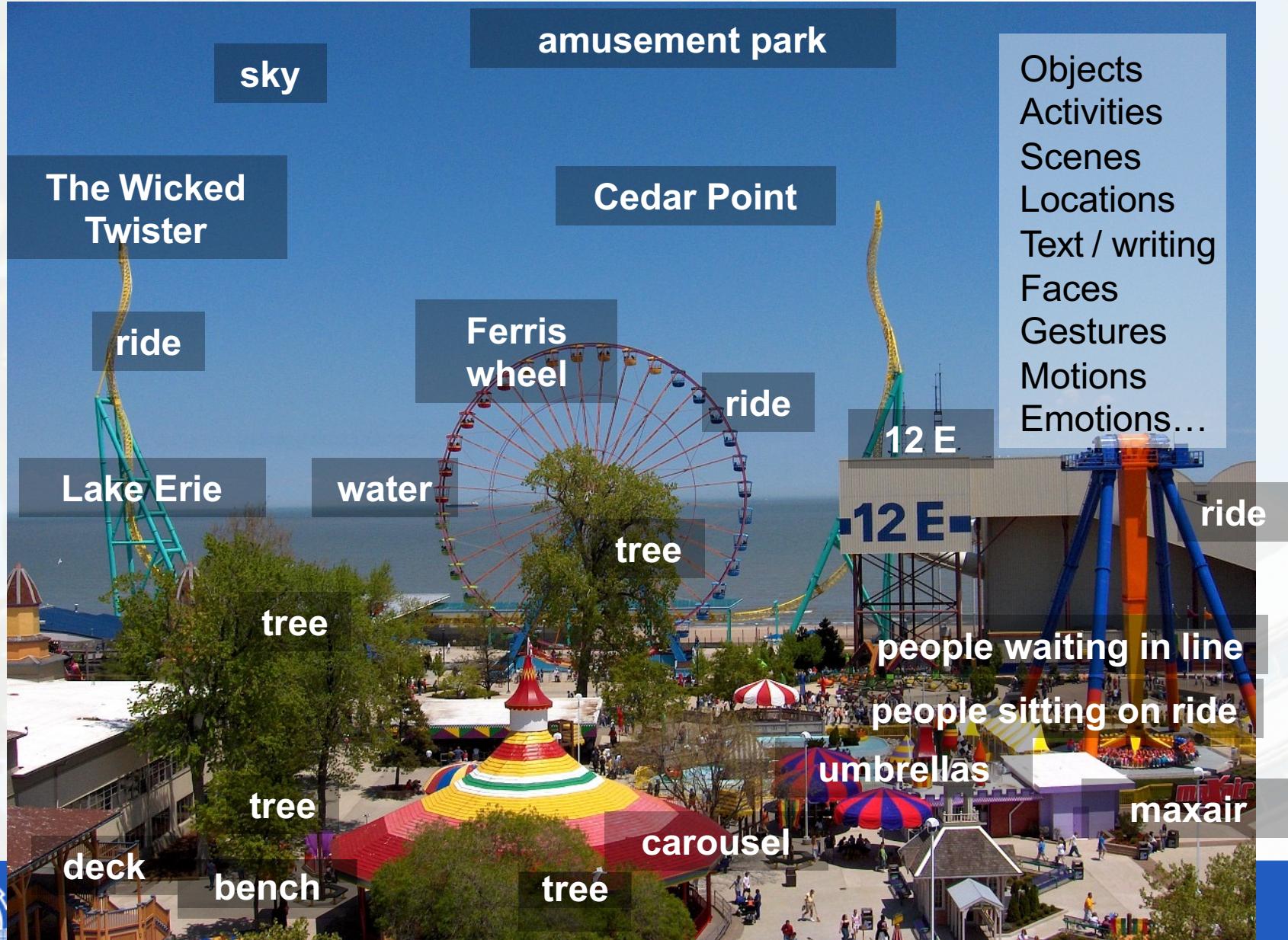
# Vision as a measurement device



哈爾濱工業大學(深圳)

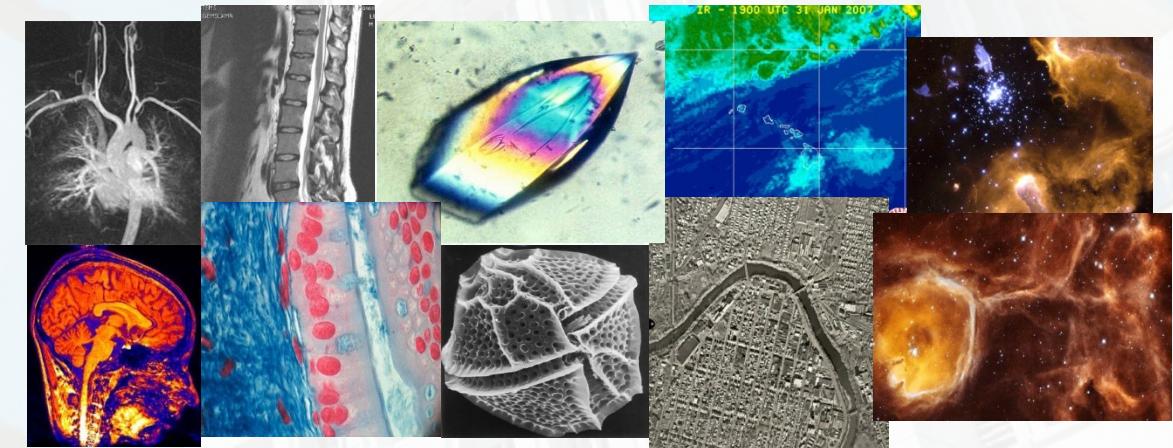
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Vision as a source of semantic information



# Why study computer vision?

- Vision is useful: Images and video are everywhere!



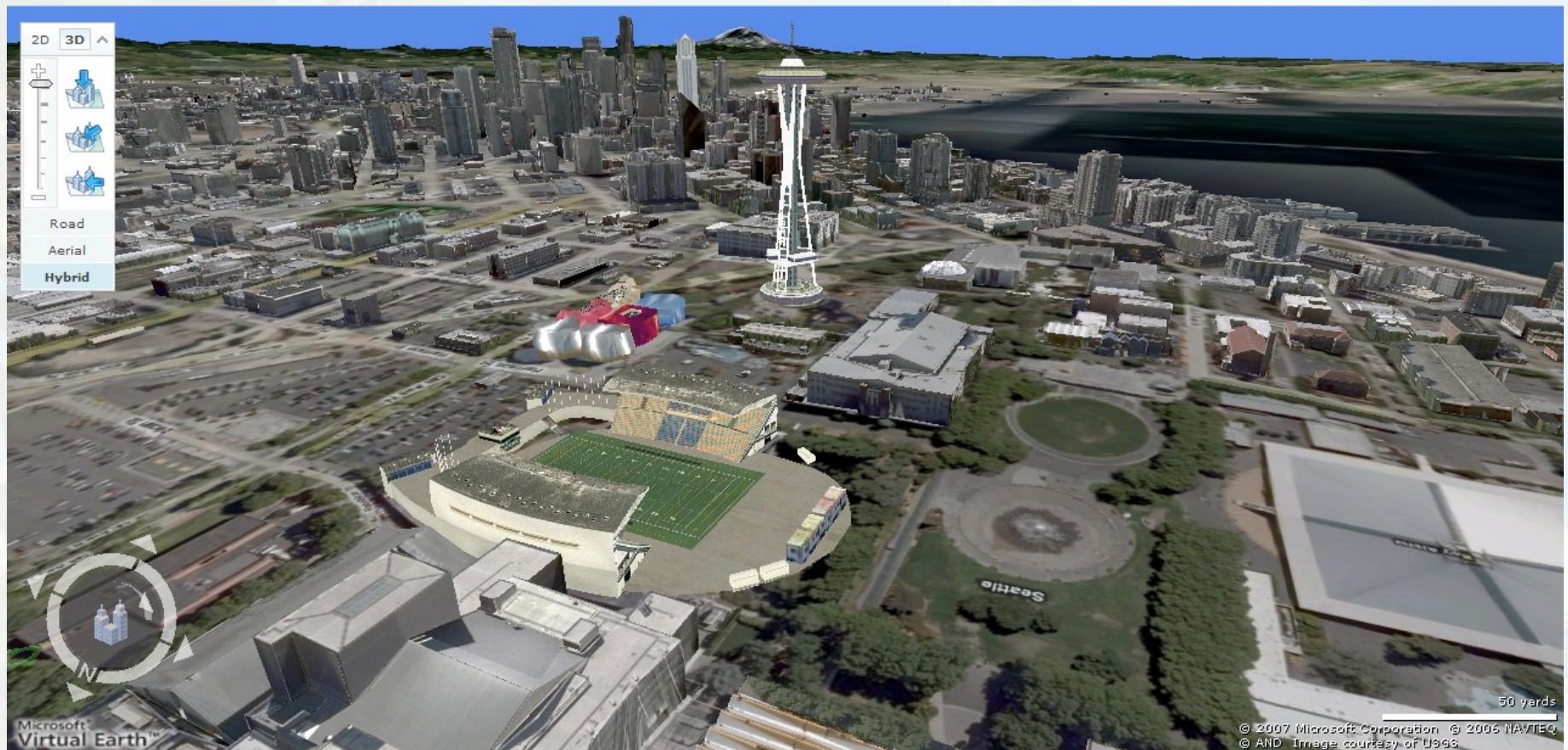
# Special effects: shape and motion capture



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# 3D urban modeling



Bing maps, Google Streetview



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# 3D urban modeling: Microsoh Photosynth



[hip://labs.live.com/photosynth/](http://labs.live.com/photosynth/)



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Face detection



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



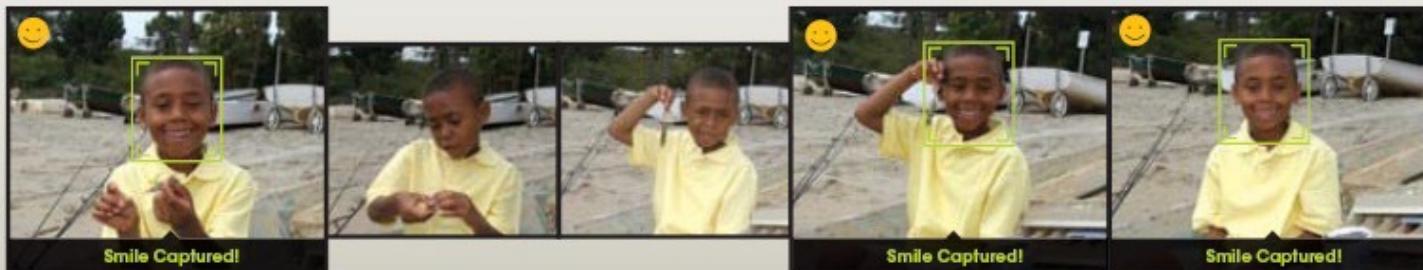
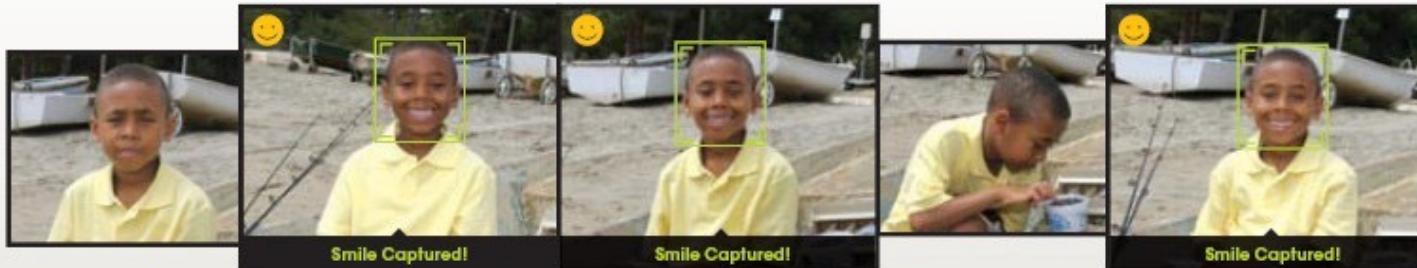
哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# 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.



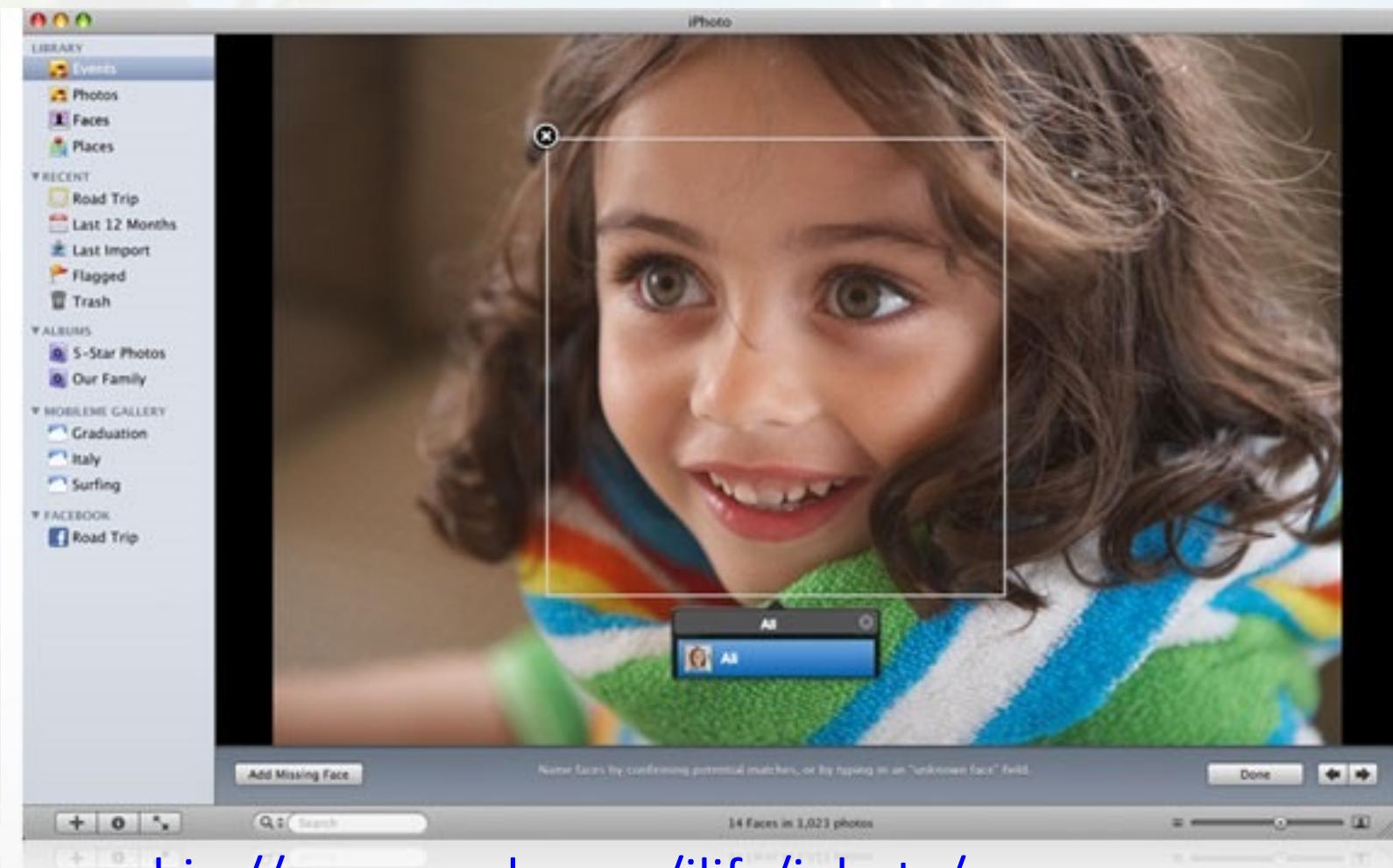
[Sony Cyber-shot® T70 Digital Still Camera](#)



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Face recognition: Apple iPhoto software



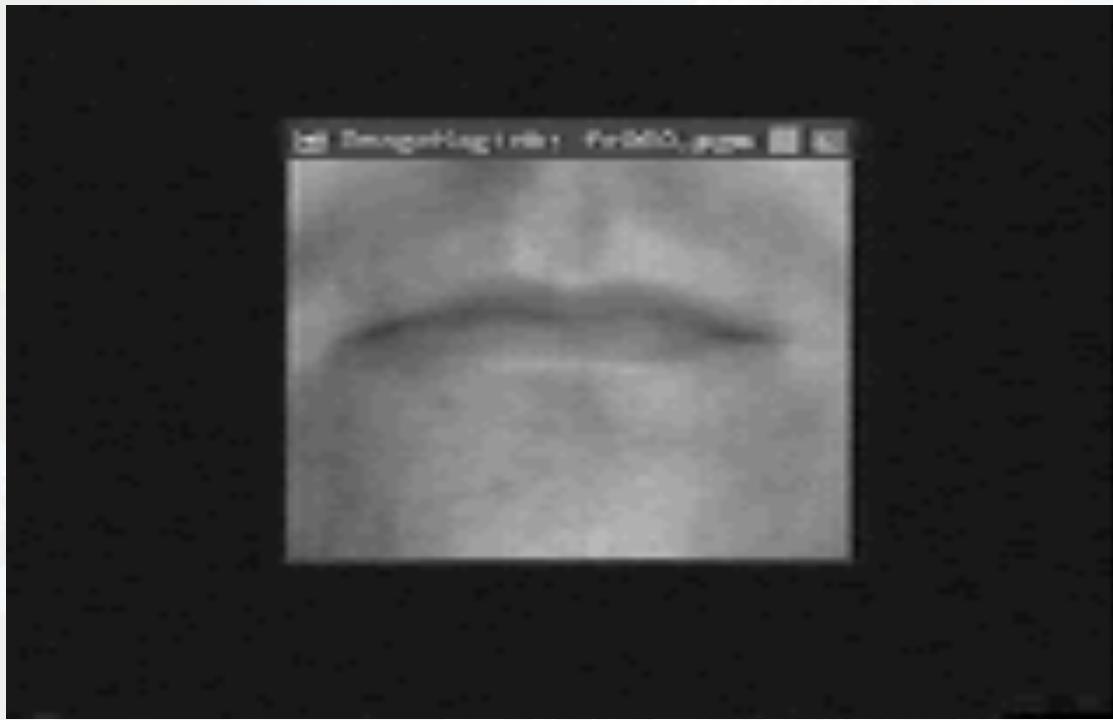
<http://www.apple.com/ilife/iphoto/>



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

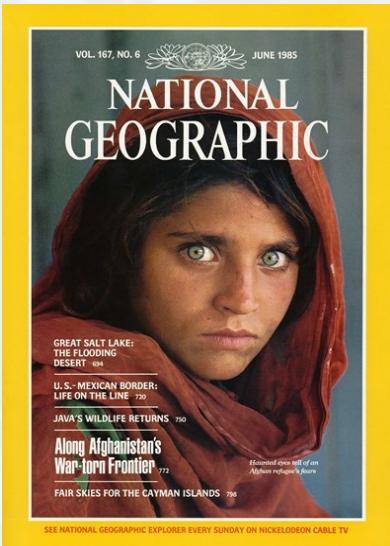
# Lipreading



哈爾濱工業大學(深圳)

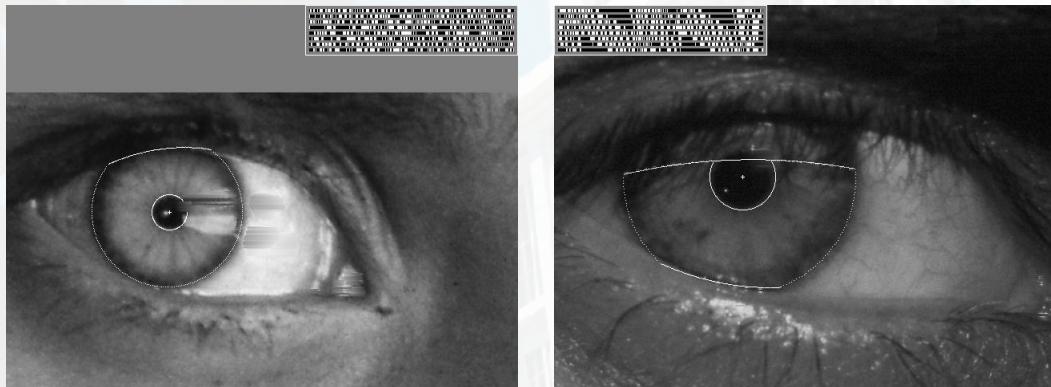
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Biometrics



1984 vs 2002

How the Afghan Girl was Identified by Her Iris Patterns



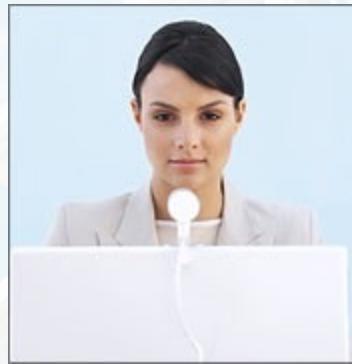
哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Biometrics



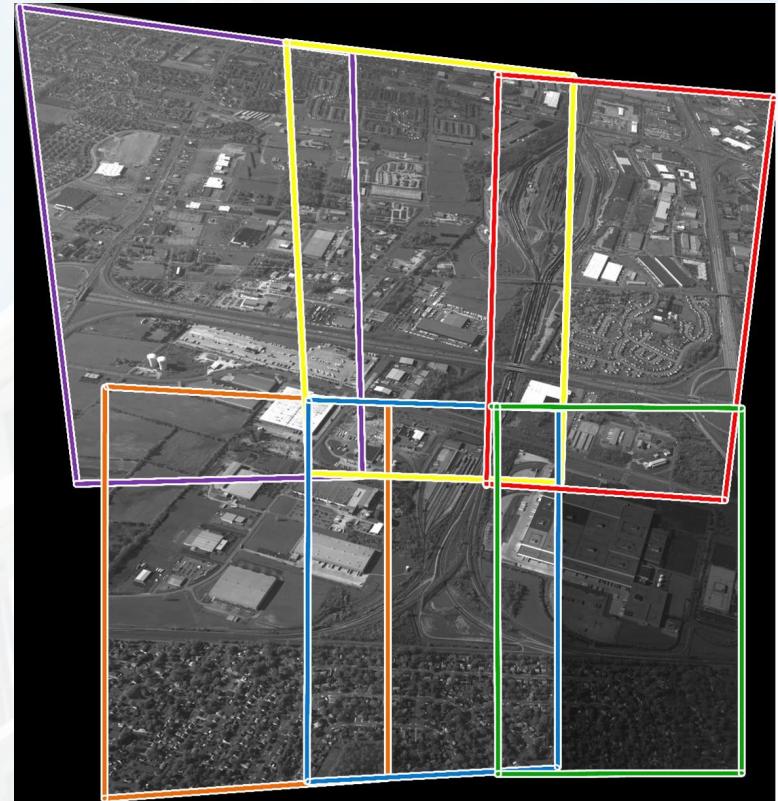
Fingerprint scanners on  
many new laptops,  
other devices



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



# Wide area surveillance



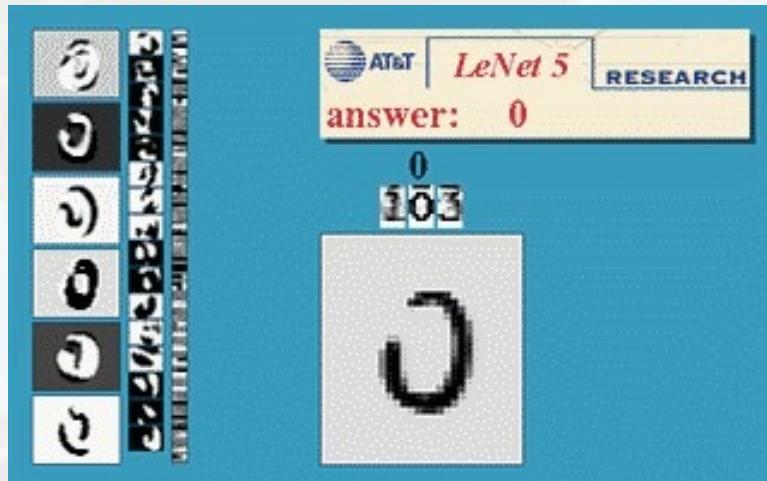
哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs



License plate readers

[http://en.wikipedia.org/wiki/Automatic\\_number\\_plate\\_recognition](http://en.wikipedia.org/wiki/Automatic_number_plate_recognition)

Source: S. Seitz



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Toys and Robots



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Mobile visual search: Google Goggles

## Google Goggles in Action

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



[Landmark](#)



[Book](#)



[Contact Info.](#)



[Artwork](#)



[Places](#)



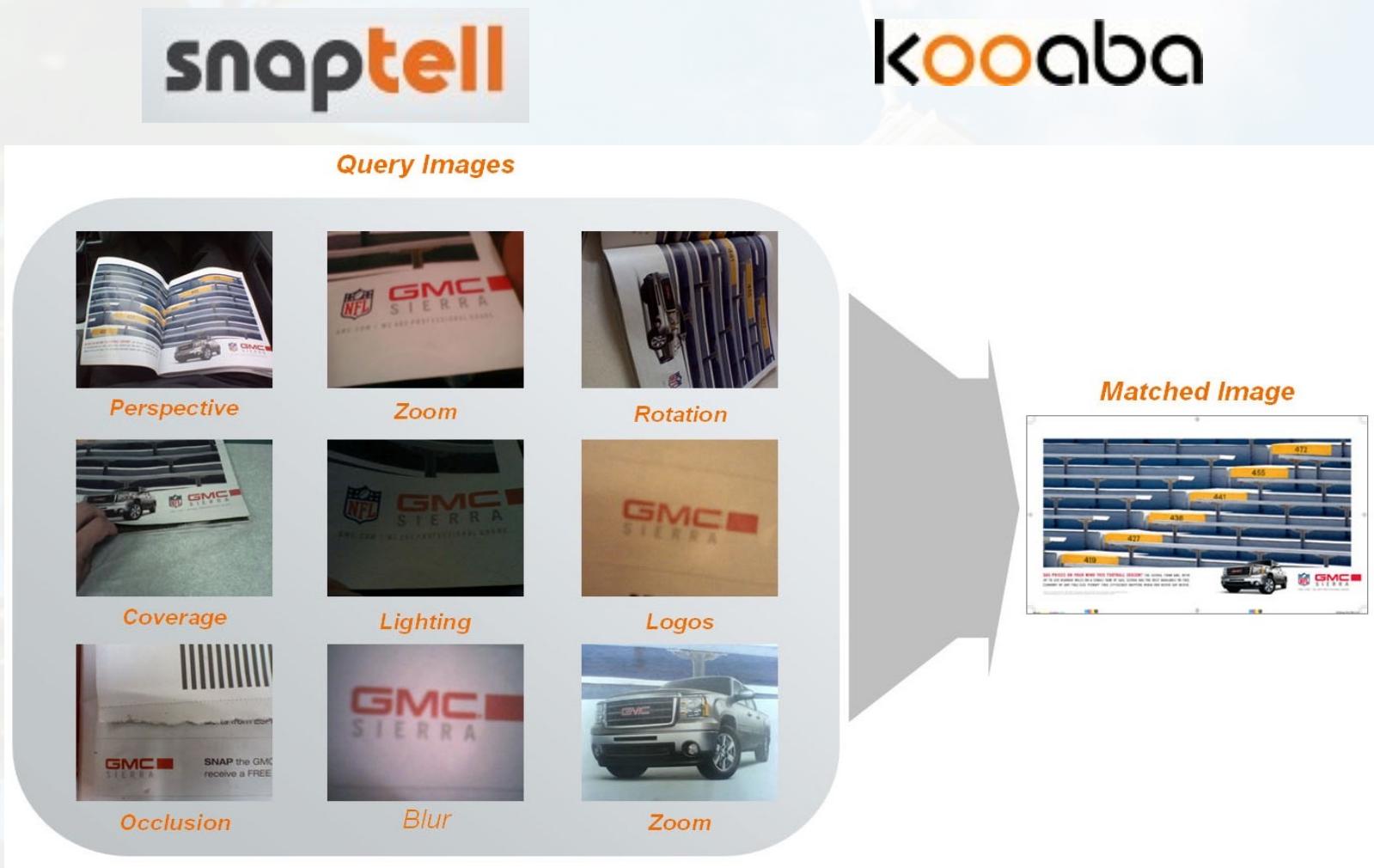
[Wine](#)



[Logo](#)



# Mobile visual search: iPhone Apps



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Automotive safety

▶▶ manufacturer products      consumer products ◀◀

## Our Vision. Your Safety.

rear looking camera      forward looking camera      side looking camera

### News

- ▶ Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System
- ▶ Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end

> all news

› **EyeQ** Vision on a Chip

› **Vision Applications**

Road, Vehicle, Pedestrian Protection and more

› **AWS** Advance Warning System

› read more      › read more      › read more

### Events

- › Mobileye at Equip Auto, Paris, France
- › Mobileye at SEMA, Las Vegas, NV

› read more

- Mobileye: Vision systems in high-end BMW, GM, Volvo models
  - “In mid 2010 Mobileye will launch a world's first application of full emergency braking for collision mitigation for pedestrians where vision is the key technology for detecting pedestrians.”



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Vision in supermarkets



## LaneHawk by EvolutionRobotics

“A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it...”



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Vision--based interaction (and games)



Microsoft's Kinect



Sony EyeToy



Assistive technologies



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Vision for robotics, space



[NASA'S Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

## Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “[Computer Vision on Mars](#)” by Maihies et al.



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Today's agenda

- What is computer vision
- Course overview



# Contacting instructor and TA

Recommended but not required:

Computer Vision: Algorithms and Applications, 2nd ed.  
Richard Szeliski

Instructor: Jingyong Su @ L1501

QQ Group: 338475061



TAs: Yijun Liu, Tong Shao, Jiaxun Dun, Zhenyu Tang



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Topics (16 lectures)

1. Introduction (2)
2. Pixels and Features (2)
3. Regions of Images and Segmentation (2)
4. Object Recognition and Face Detection(2)
5. Neural Networks (2)
6. Recurrent Neural Networks: RNN, LSTM (2)
7. Attention and Transformers (1)
8. Generative Modeling (1)
9. Visual Foundation Model (1)
10. Data-Centric Artificial Intelligence (1)



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Your Grades

- Homework Assignment: 60% (12% each)
  - Only the “**.ipynb**” file
  - Procedures and results are required
- Final Project: 40%



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Image and Transformations



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Types of Images

Binary



Gray Scale



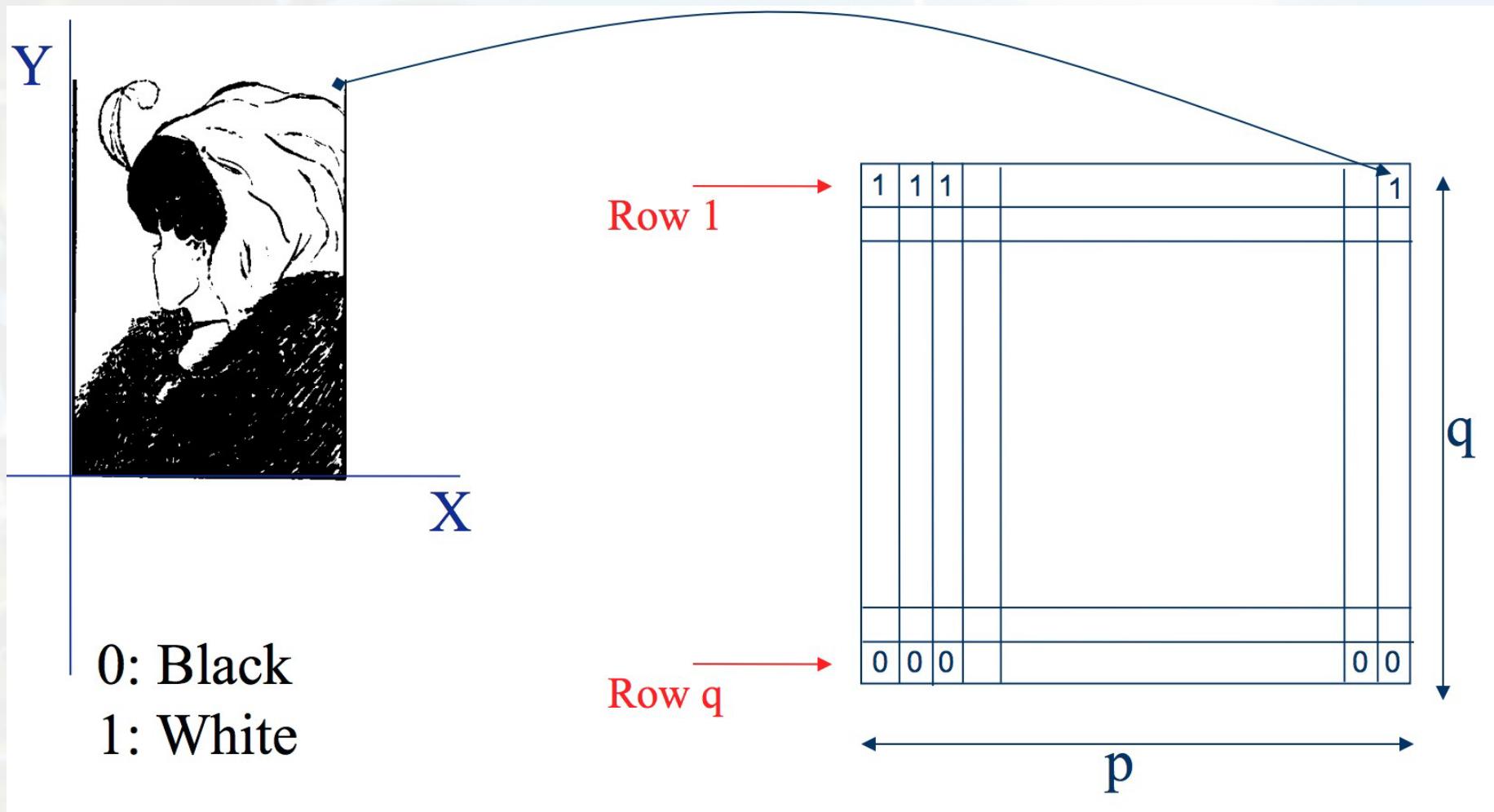
Color



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Binary image representation



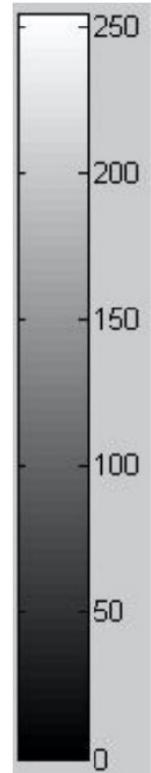
哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Grayscale image representation



10	5	9							
									100



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Color image representation



Phil Noble / AP



Phil Noble / AP



Phil Noble / AP



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Digital Images Are Sampled

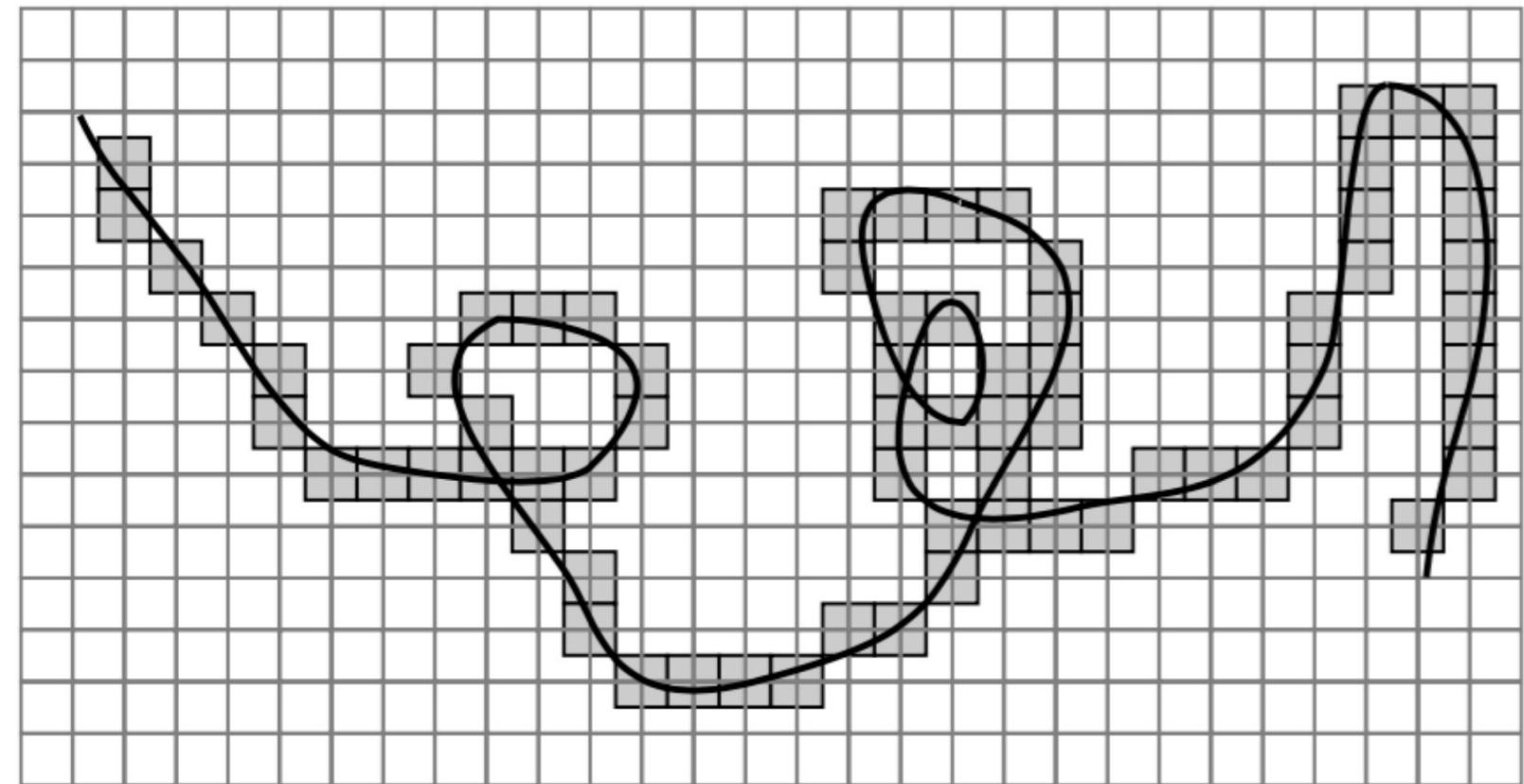
What happens when we zoom into the images we capture?



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Errors due to Sampling



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Resolution

is a **sampling** parameter, defined in dots per inch (DPI) or equivalent measures of spatial pixel density



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

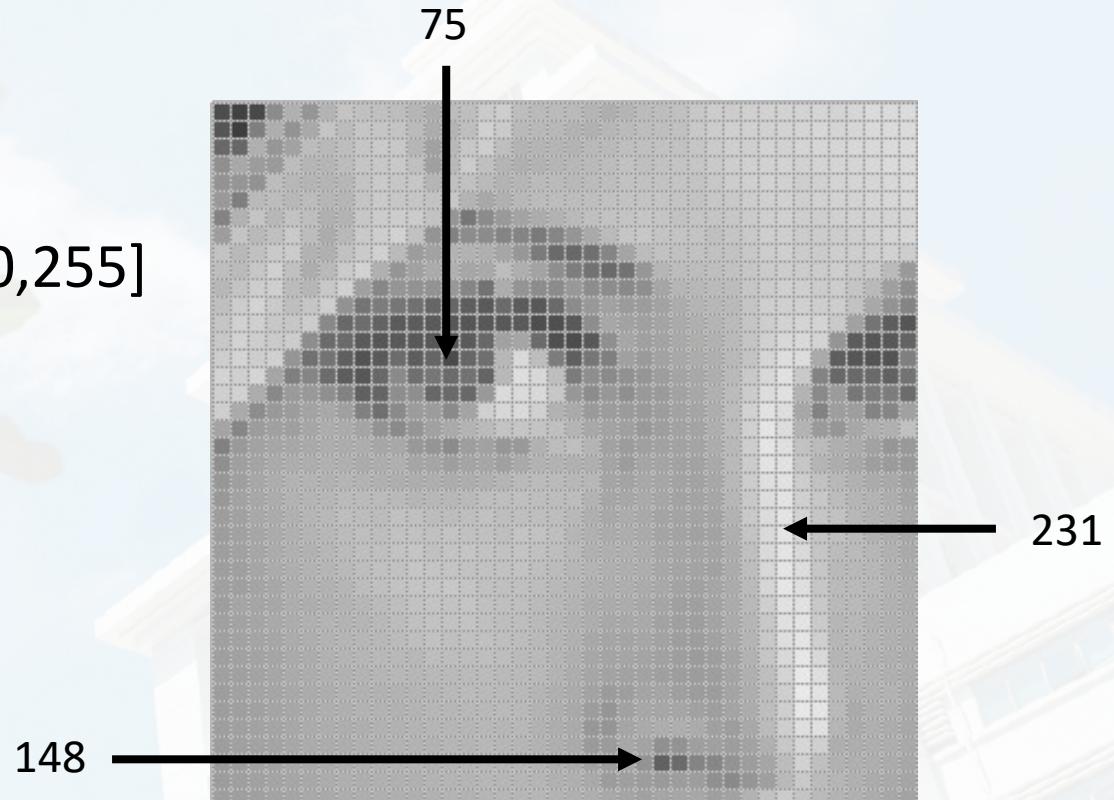
# Images are Sampled and Quantized

- An image contains discrete number of pixels

- Pixel value:

- “grayscale”

- (or “intensity”): [0,255]



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Images are Sampled and Quantized

- An image contains discrete number of pixels

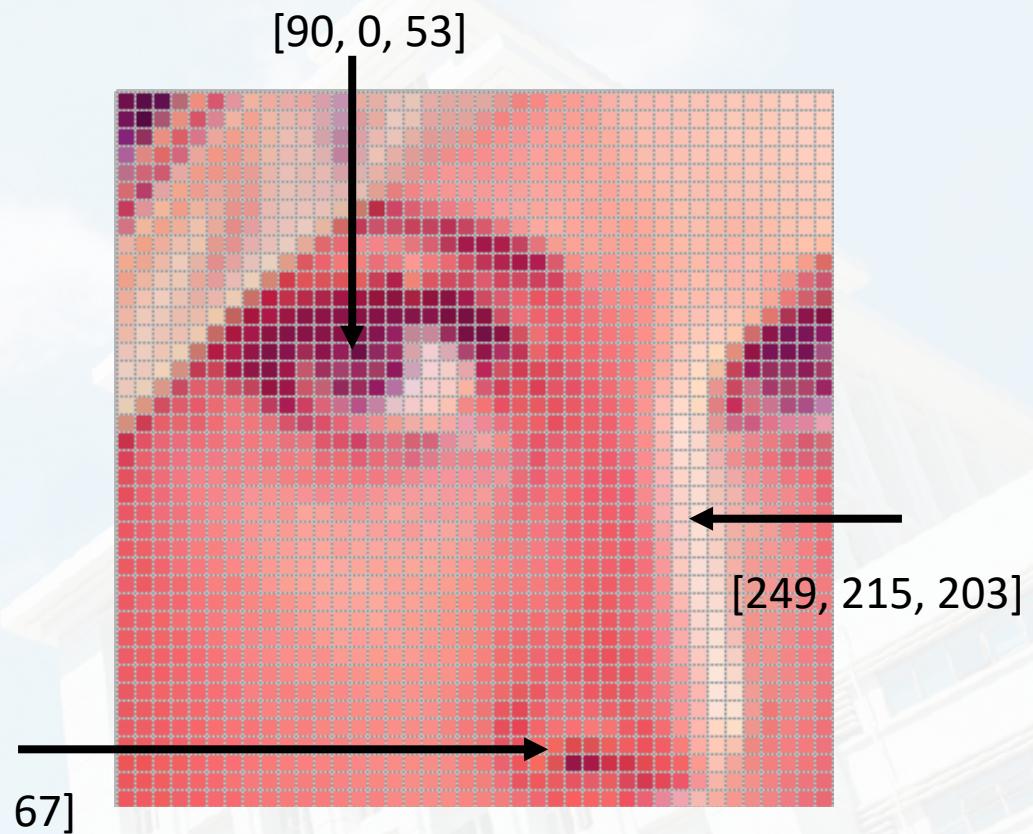
- Pixel value:

- “color”

- RGB: [R, G, B]

- Lab: [L, a, b]

- HSV: [H, S, V]

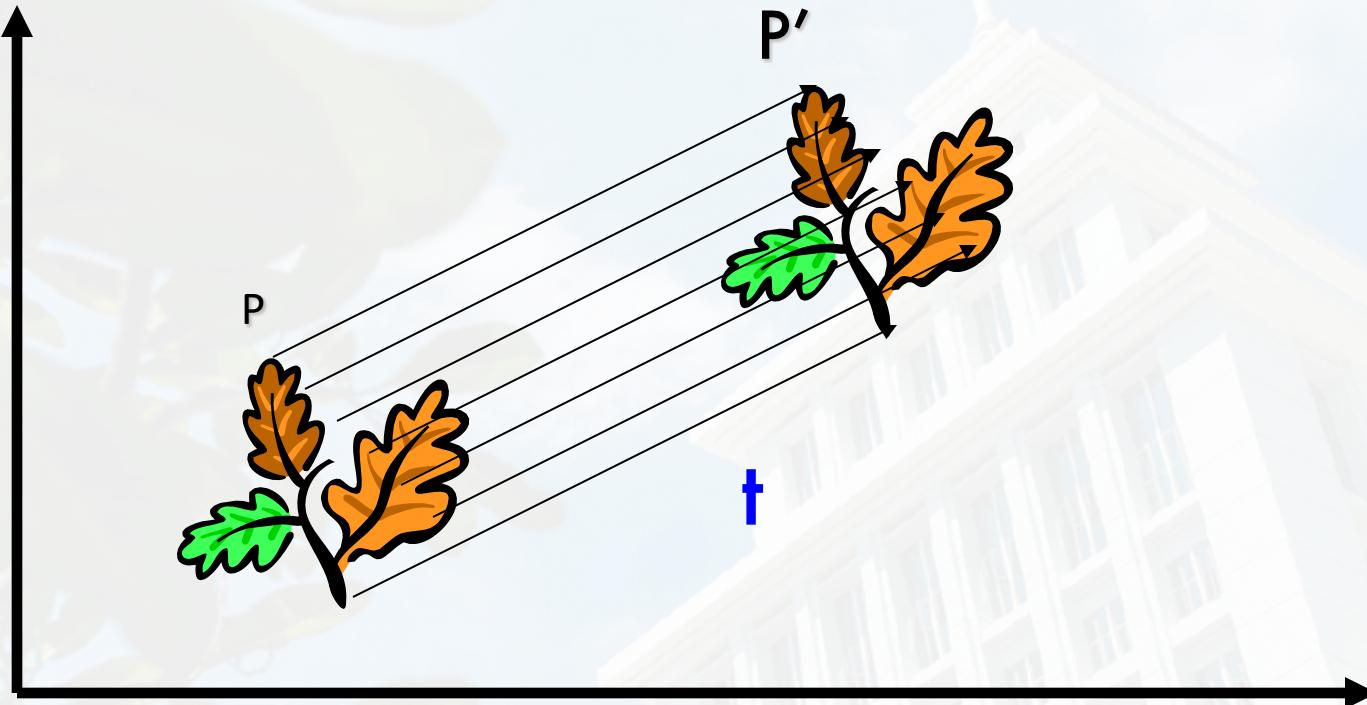


# Transformations

- ❖ Translation
- ❖ Rotation
- ❖ Scaling



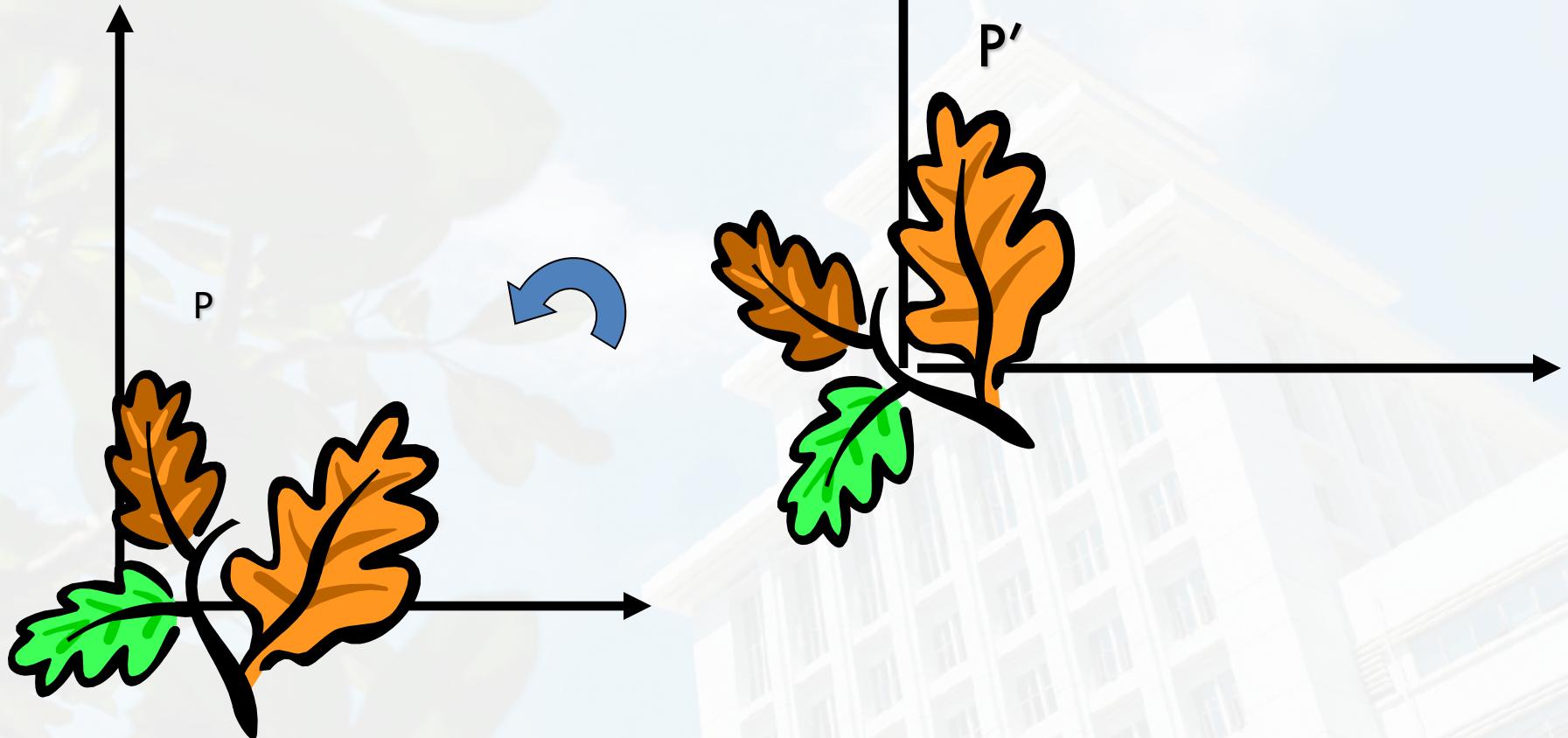
# Translation



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

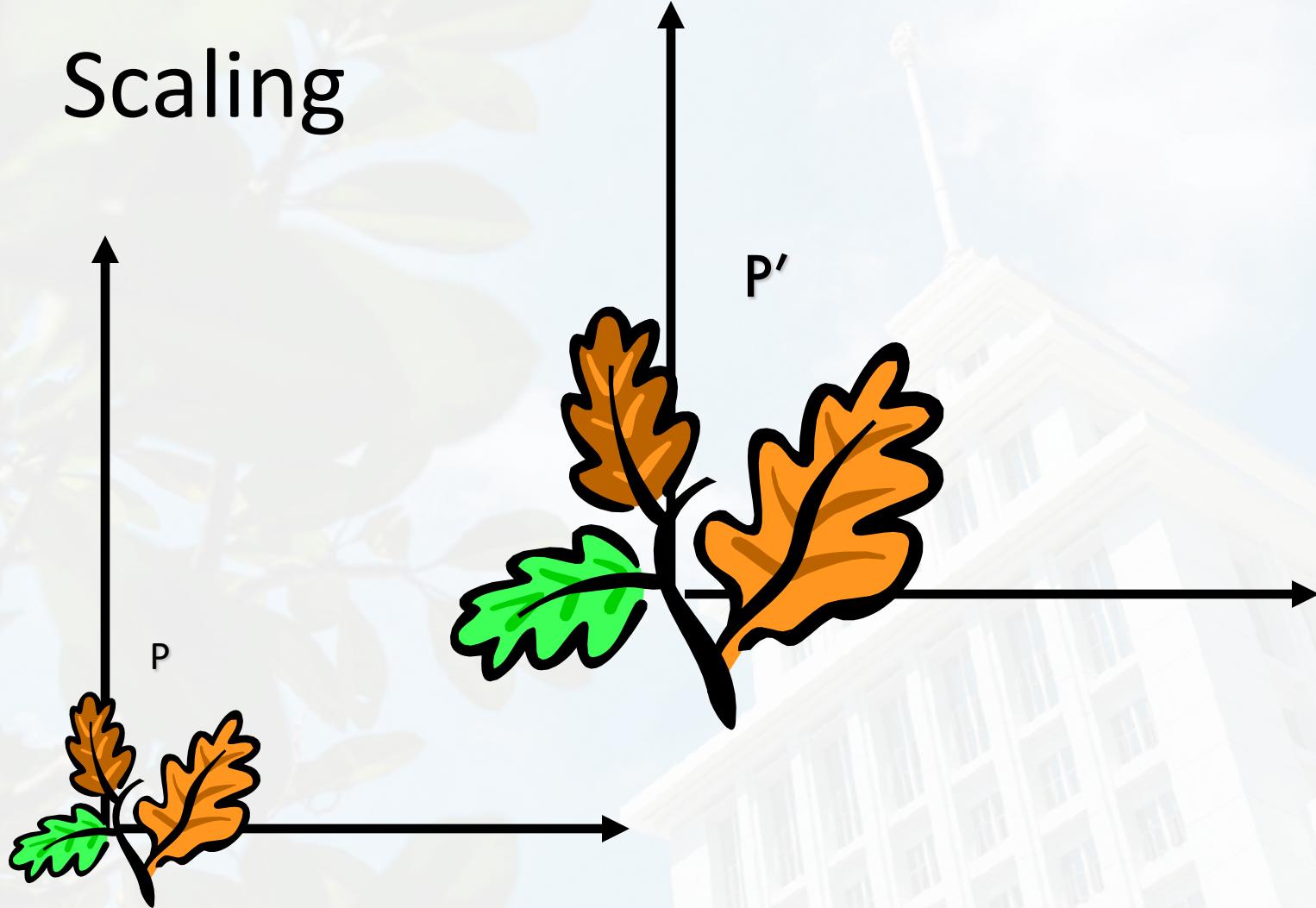
# Rotation



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# Scaling



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN