كلية الهندسة
قسم هندسة الروبوت و الأنظمة الذكية

ذكاء صنعي 2

محاضرة 6

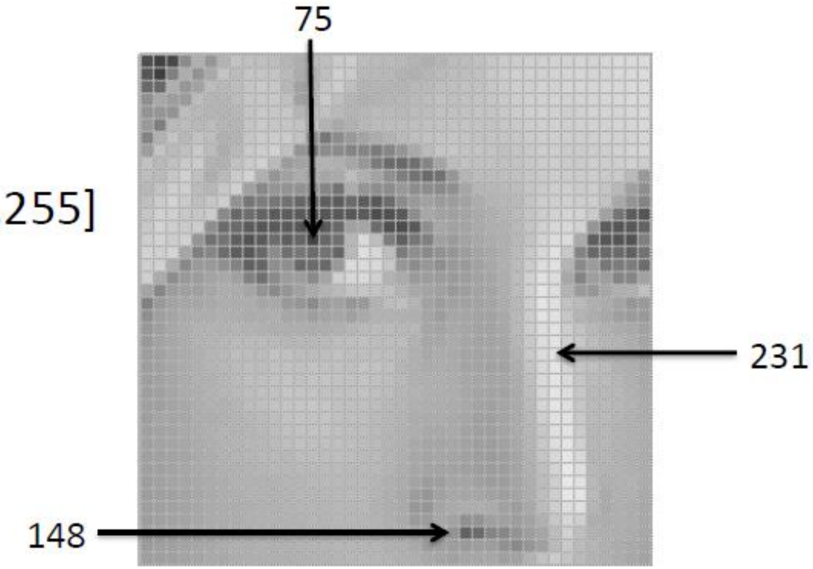
Computer Vision

د. فادي متوج

Images as functions

- An image contains discrete number of pixels
  - A simple example
  - Pixel value:
    - "grayscale"

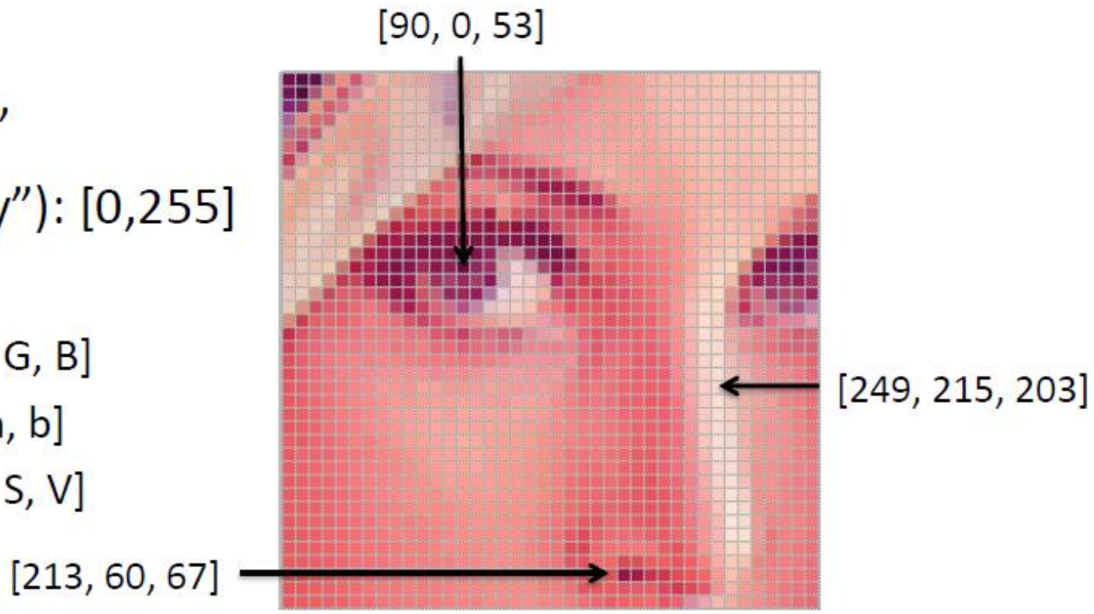
      (or "intensity"): [0,255]



Images as functions

- An image contains discrete number of pixels
  - A simple example
  - Pixel value:
    - "grayscale"

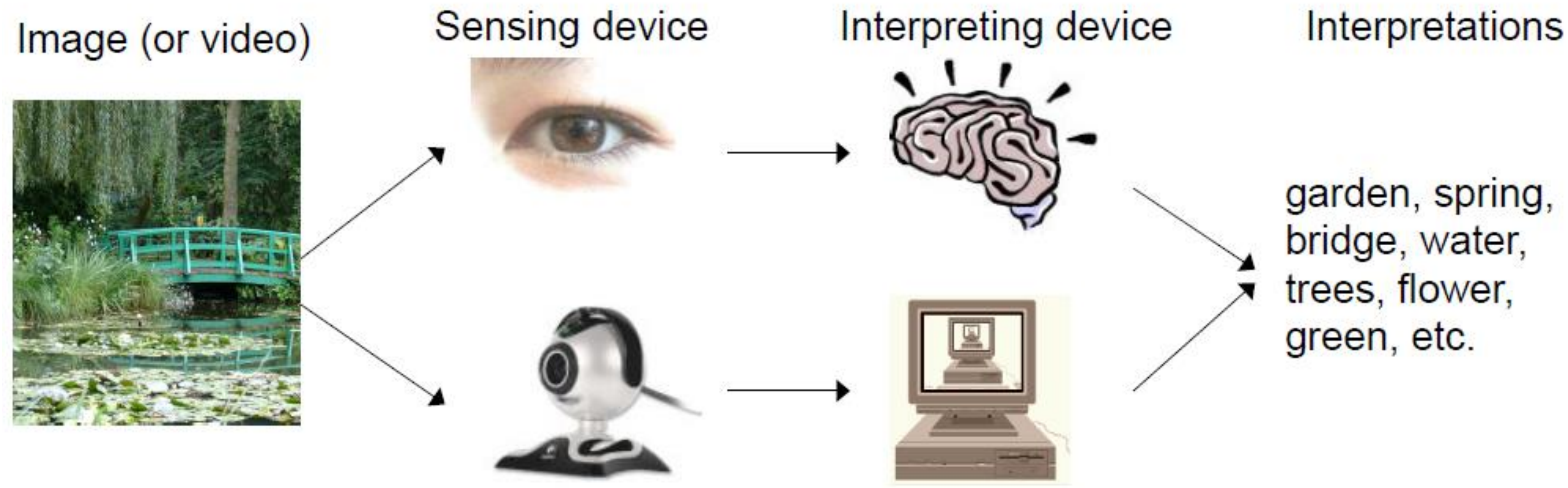
    (or "intensity"): [0,255]
    - "color"
      - RGB: [R, G, B]
      - Lab: [L, a, b]
      - HSV: [H, S, V]



Images as functions

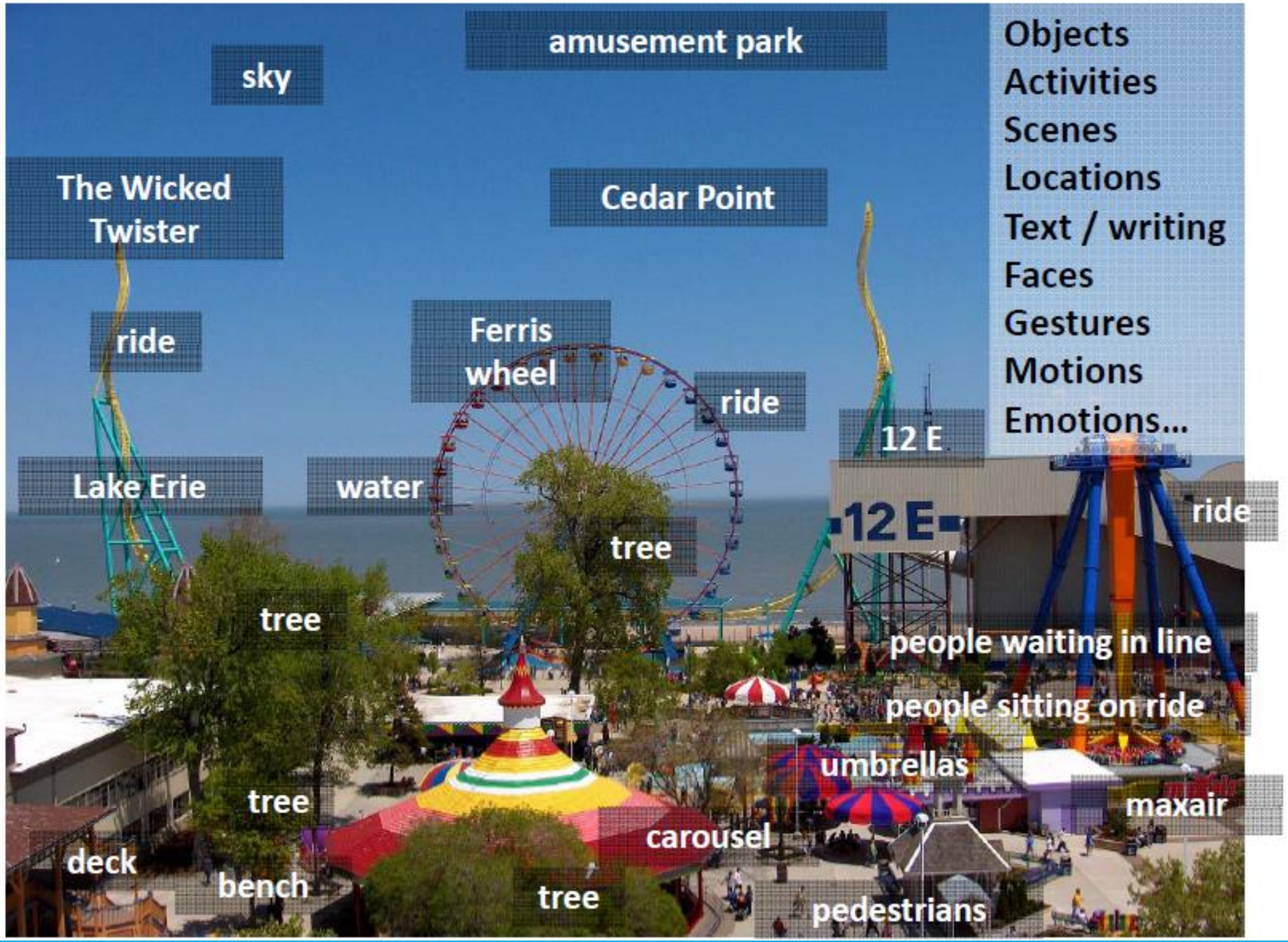
- **An Image** as a function f from R^2 to R^M:
  - $f(x, y)$ gives the **intensity** at position (x, y)

What is Computer Vision?



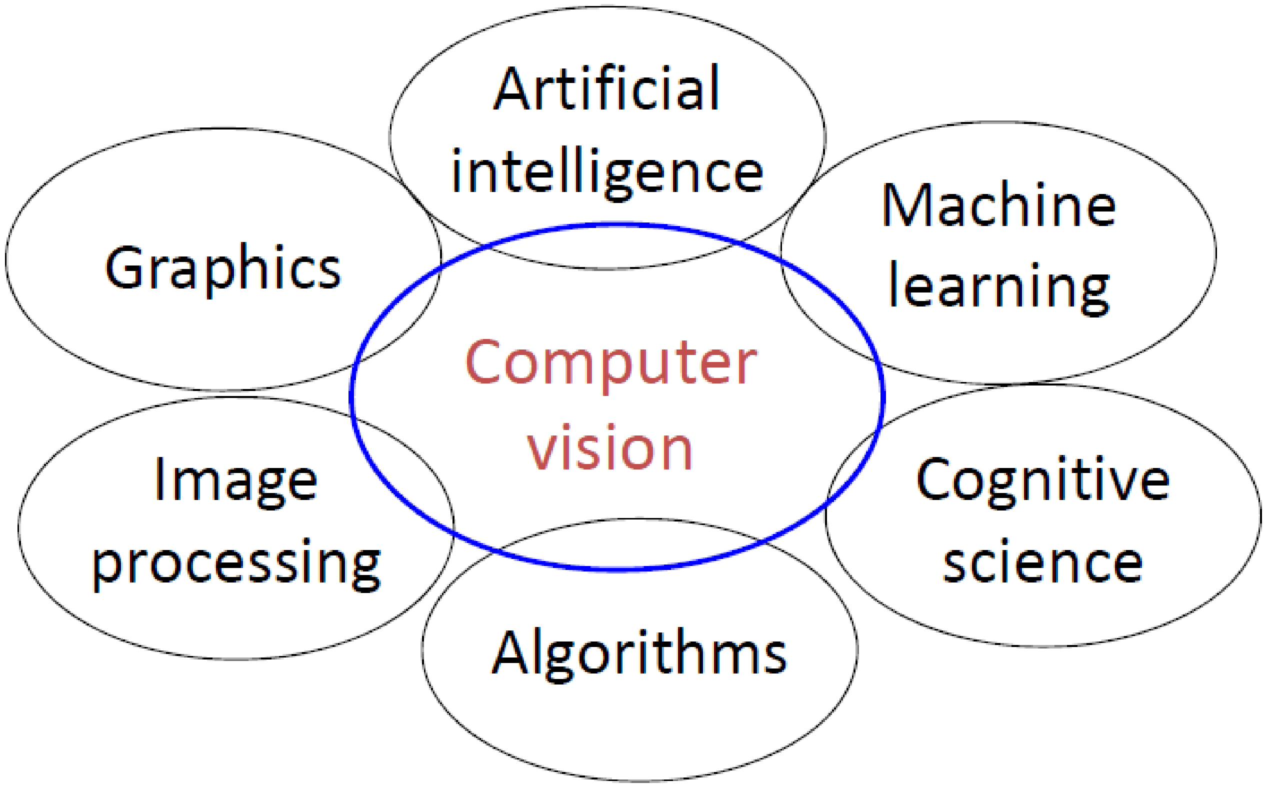
جامعة المنارة

Vision for perception, interpretation



Objects
Activities
Scenes
Locations
Text / writing
Faces
Gestures
Motions
Emotions…

Related disciplines



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

Computer Vision

- Make computers understand images and video.

What kind of scene?

Where are the cars?

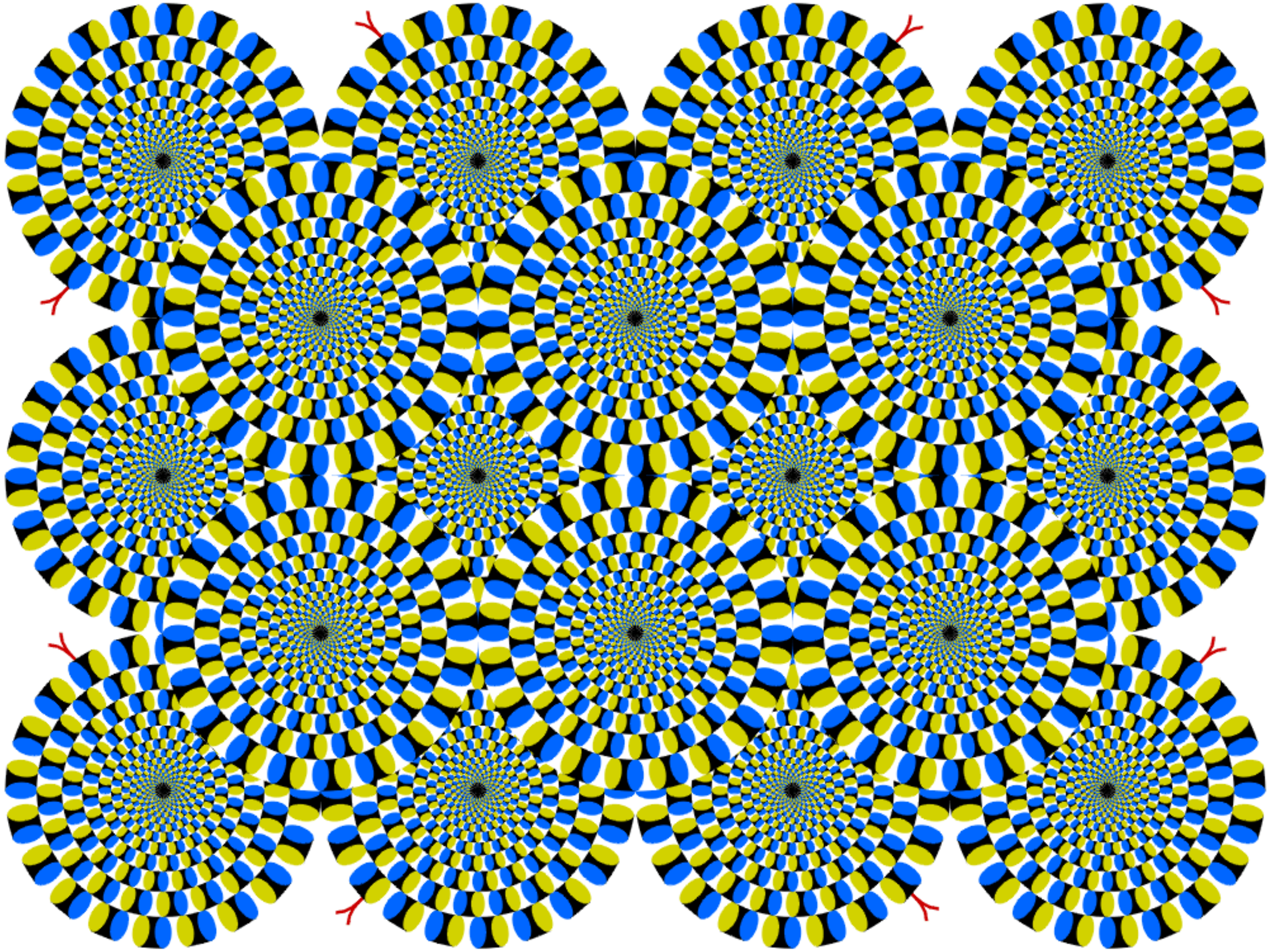
How far is the building?

…

Every picture tells a story

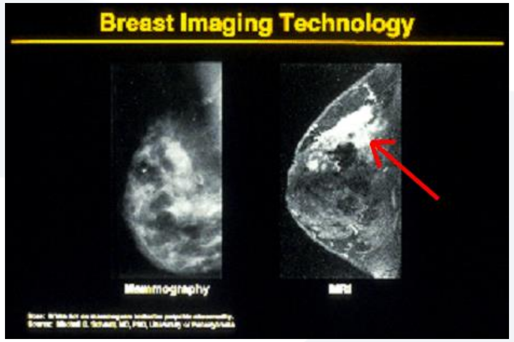
- Goal of computer vision is to write computer programs that can interpret images

Human perception has its shortcomings...



Why computer vision matters

Safety



Health

Security

Comfort

Fun

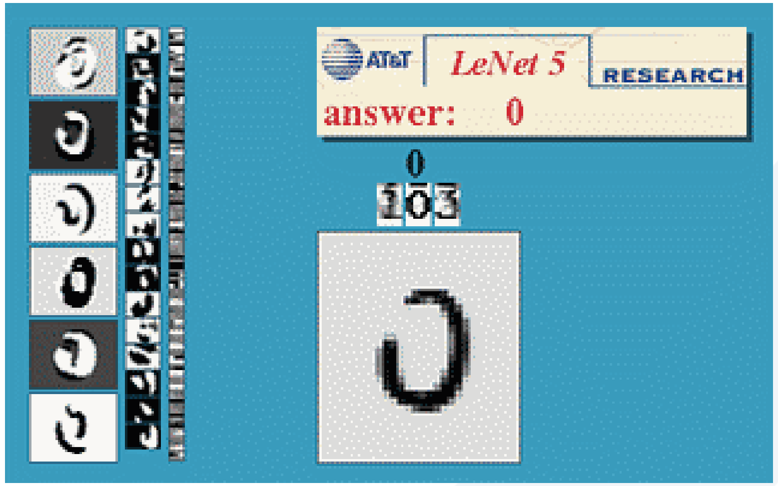
How vision is used now

- Examples of **state-of-the-art**

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
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

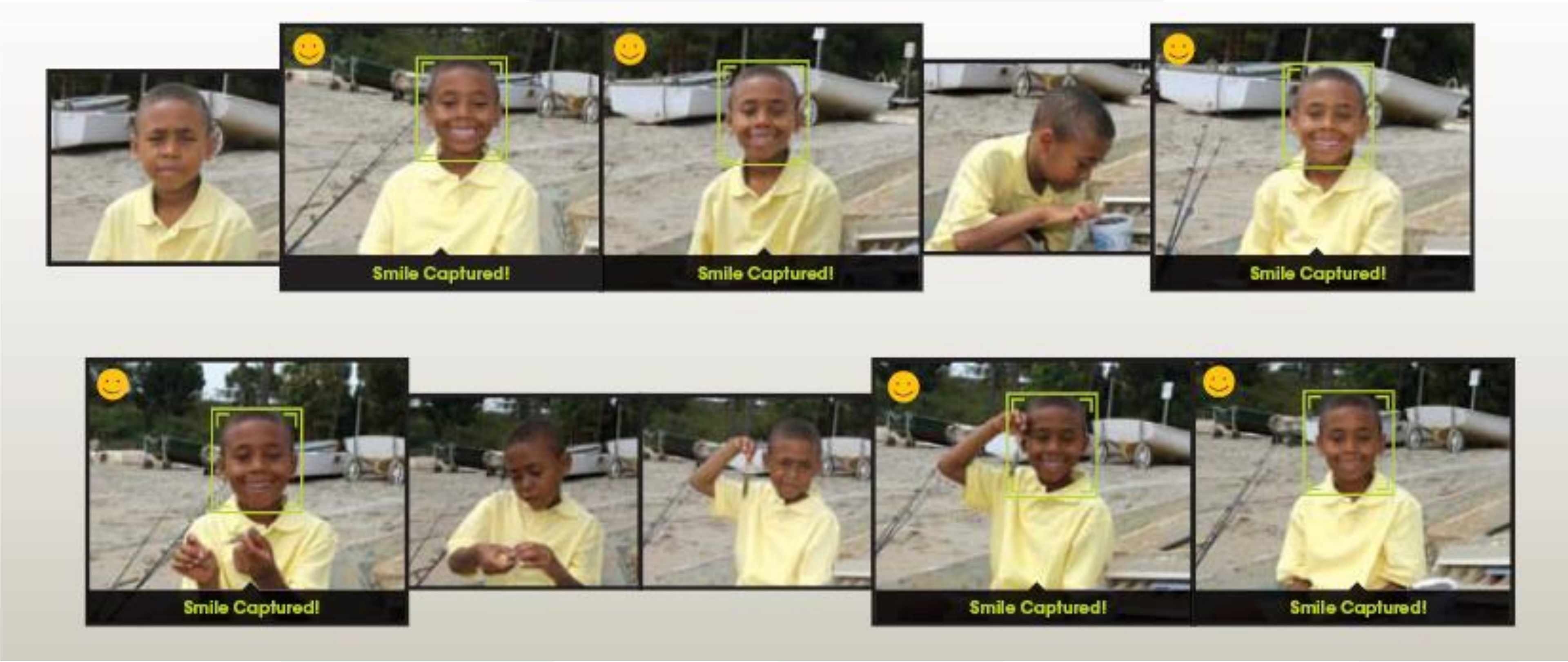
Face detection

- 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.



3D from thousands of images

Building Rome in a Day: Agarwal et al. 2009

Object recognition (in supermarkets)

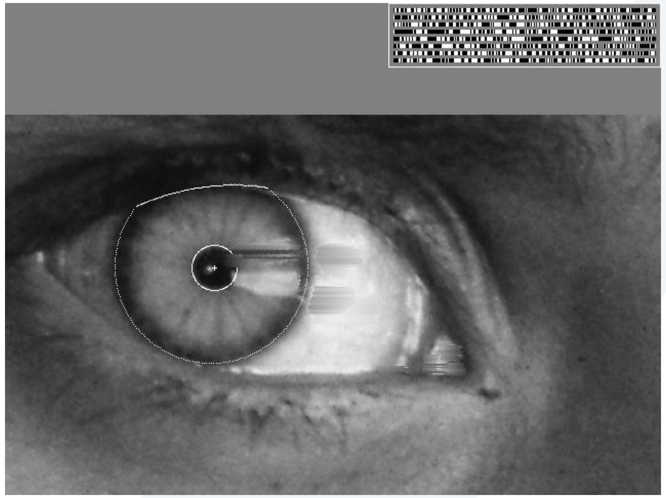
LaneHawk by EvolutionRobotics

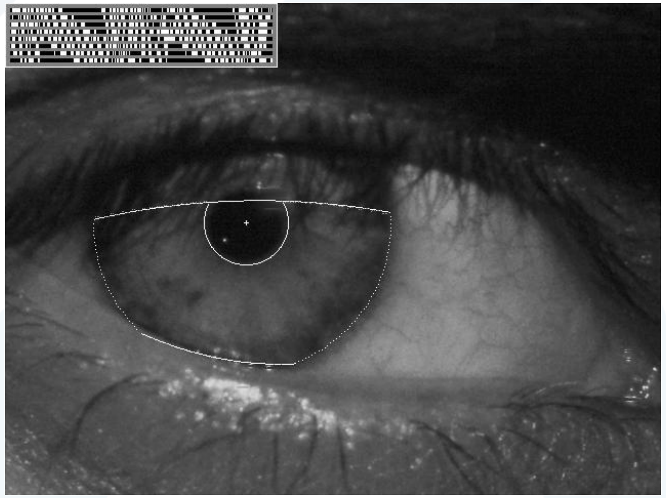
"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you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

Vision-based biometrics

"How the Afghan Girl was Identified by Her Iris Patterns"

http://www.cl.cam.ac.uk/~jgd1000/afghan.html





Login without a password…


Fingerprint scanners on many new laptops, other devices


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

Object recognition (in mobile phones)

[Google Goggles]

Every picture tells a story

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: shape capture

Pirates of the Carribean, Industrial Light and Magic

Sports

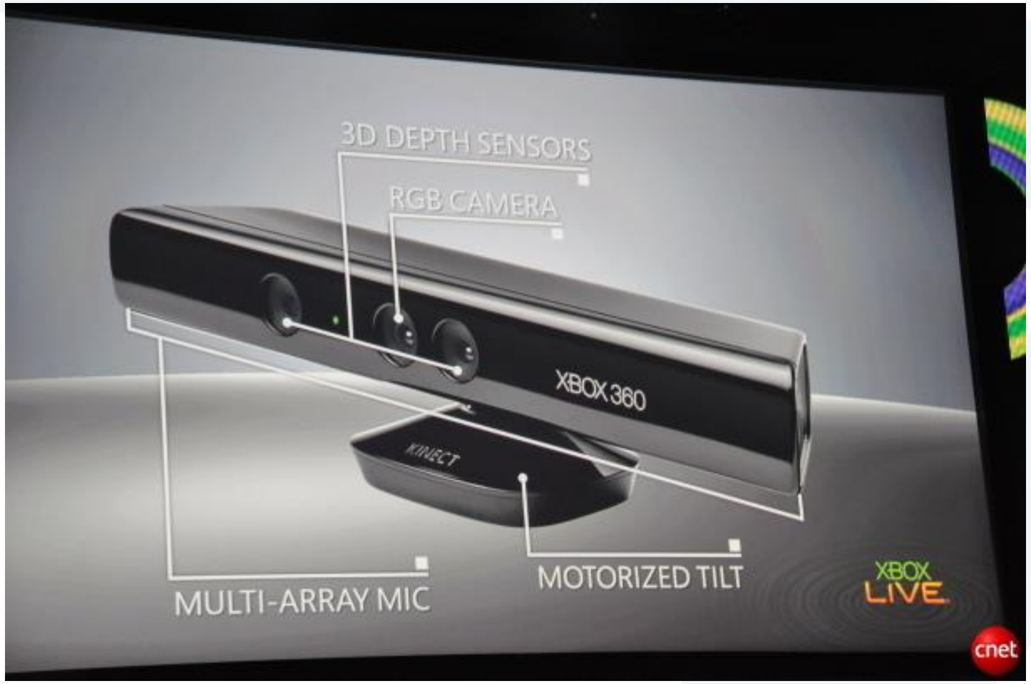
Smart cars

- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers.

Google cars

https://www.nytimes.com/2010/10/10/science/10google.html

Interactive Games: Kinect



Vision in space

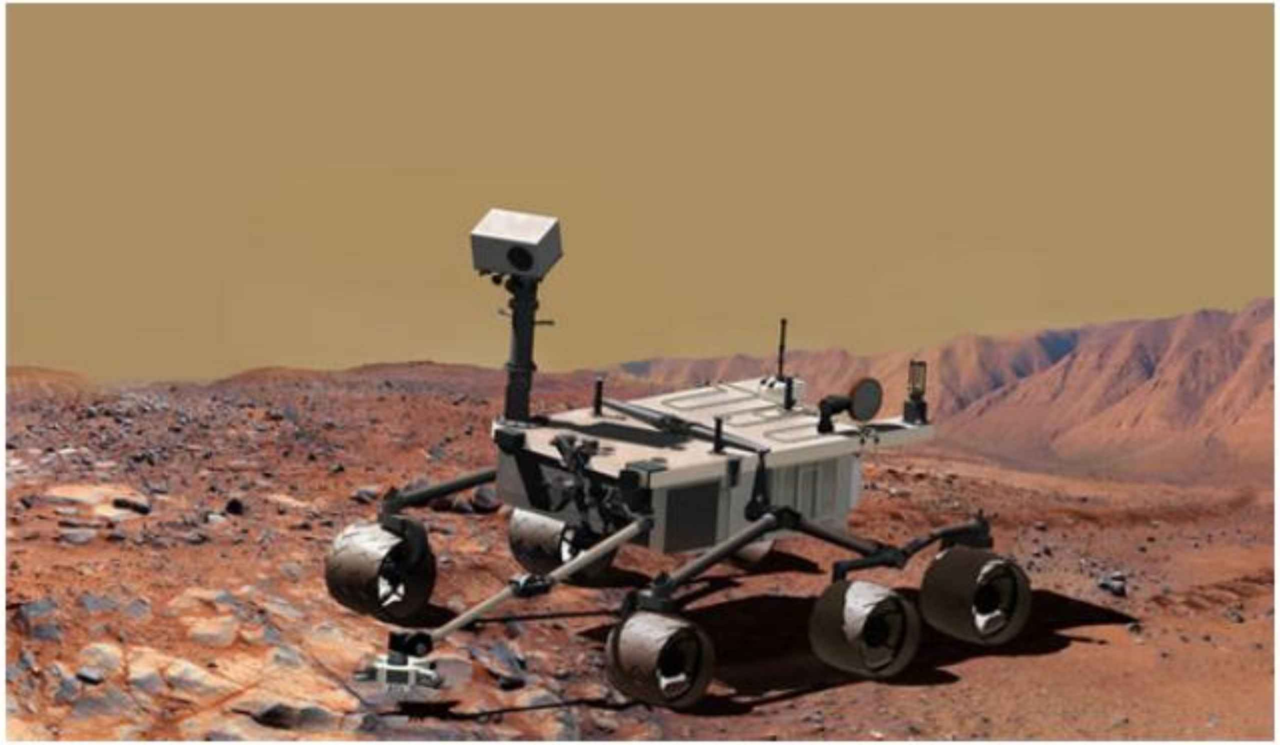
Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking

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

جـامعة
الـمَـنارة
MANARA UNIVERSITY

Opportunity
Launched 2003
Landed 2004



Curiosity
Launched 2011
Landed 2012

Industrial robots

Vision-guided robots position nut runners on wheels www.cognex.com/customersuccessstories/CustomerStoryDetail.aspx?ID=4610&langtype=1033

COGNEX

Call Cognex Sales: 855-4-COGNEX (855-426-4639) | Contact us

United States | Eng

Products Industries Applications Support How to Buy Resources Company Search Cognex

Applications

Home > Applications > Other Industries

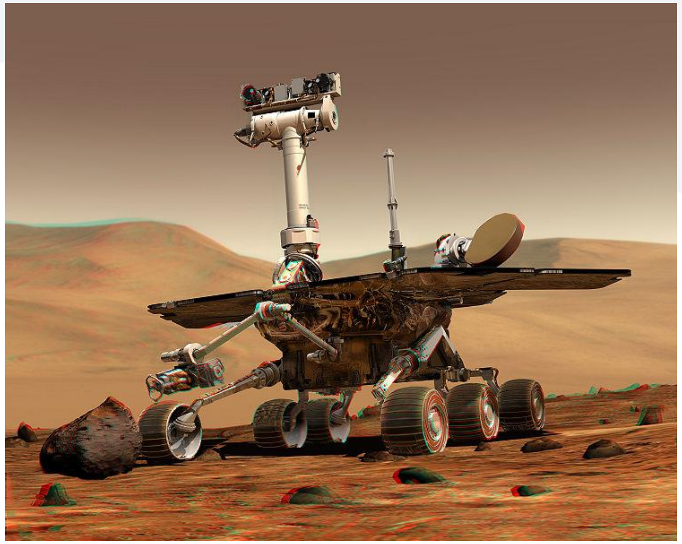
- Automotive
- Consumer Products
- Electronics
- Food & Beverage
- Logistics
- Medical Devices
- Pharmaceuticals
- Packaging
- Solar
- Web and Surface Inspection
- Other Industries

Vision Guided Robot Key to First-Known Automated Lug Nut Fastening Application

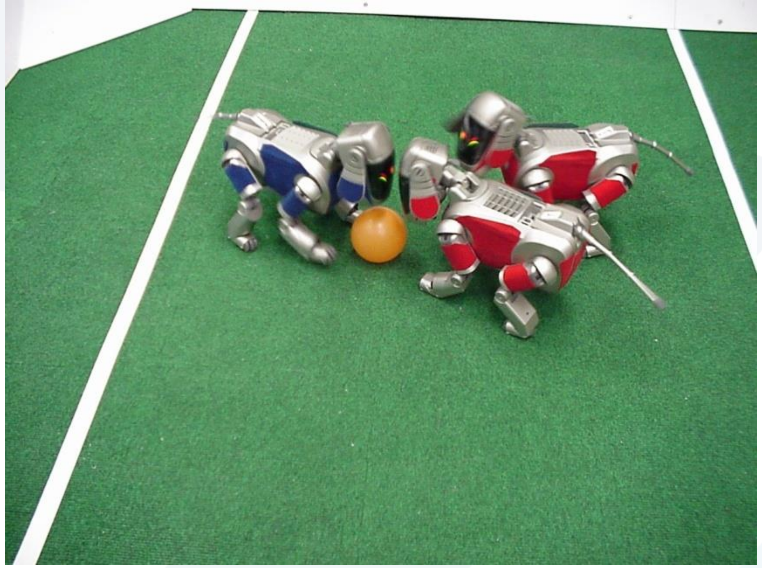
Running down and torquing the lug nuts that hold the wheel to the hub is seemingly one of the simpler aspects of building an automobile, but it has proven one of the most difficult to automate. This is a difficult manual job, as well, because of the size and weight of the nutrunner and the need to tighten the nuts on two wheels in approximately 40 seconds. If the position of the lug nuts is known, a robot can easily position the nutrunner to deliver the needed torque. The problem is that typically the vehicle is only roughly positioned by a conveyor and the wheels themselves are free to rotate, tilt, and turn. Therefore an ordinary blind robot would never be able to find the nuts.

Radix Controls, Oldcastle, Ontario, successfully automated this application by using a vision system to determine the position of the wheel including its fore and aft and side-to-side positions and three

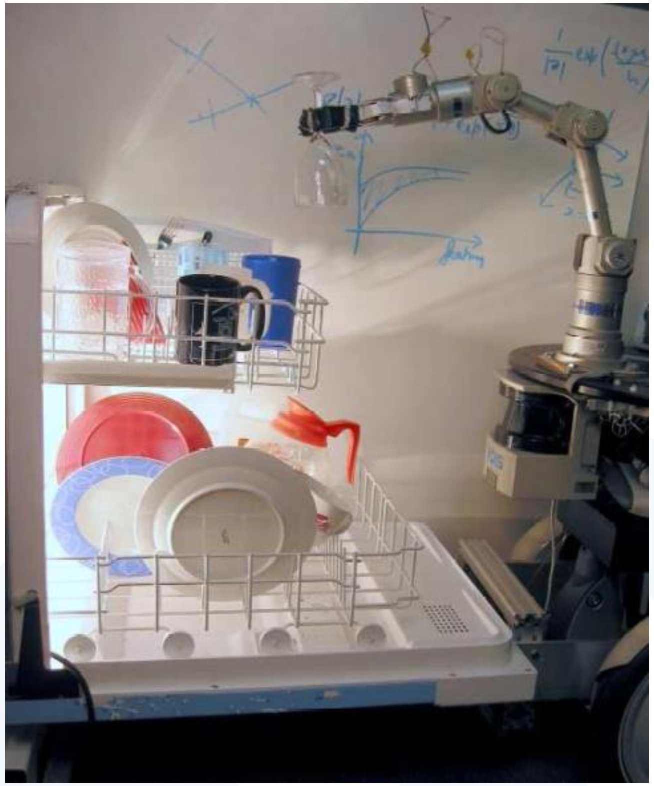
Mobile robots



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



http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

stair.stanford.edu

STAIR: STanford Artificial Intelligence Robot

Artificial Intelligence Laboratory, Computer Science Department, Stanford University

Home | People | Multimedia | Papers | Data | Sponsors | Contact

Since its birth in 1956, the AI dream has been to build systems that exhibit broad-spectrum competence and intelligence. In the STAIR (STanford AI Robot) project, we are building a robot that can navigate home and office environments, pick up and interact with objects and tools, and intelligently converse with and help people in these environments.

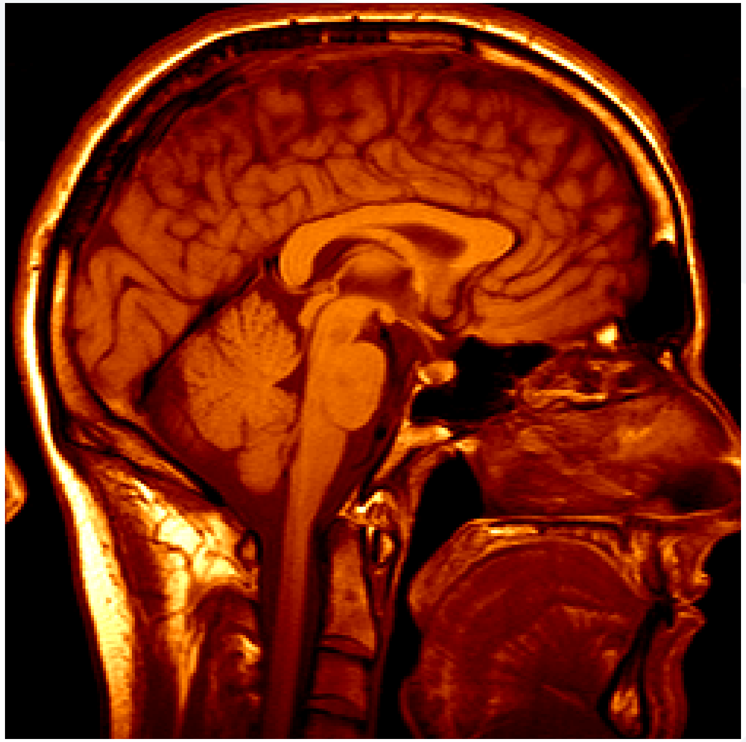
Our single robot platform will integrate methods drawn from all areas of AI, including machine learning, vision, navigation, manipulation, planning, reasoning, and speech/natural language processing. This is in distinct contrast to the 30-year trend of working on fragmented AI sub-fields, and will be a vehicle for driving research towards true integrated AI.

Over the long term, we envision a single robot that can perform tasks such as:

- Fetch or deliver items around the home or office.
-
- Tidy up a room, including picking up and throwing away trash, and using the dishwasher.
- Prepare meals using a normal kitchen.
- Use tools to assemble a bookshelf.

A robot capable of these tasks will *revolutionize* home and office automation, and have important applications ranging from home assistants to elderly care. However, carrying out such tasks will require significant advances in integrating learning, manipulation, perception, spoken dialog, and reasoning.

Medical imaging



3D imaging
MRI, CT

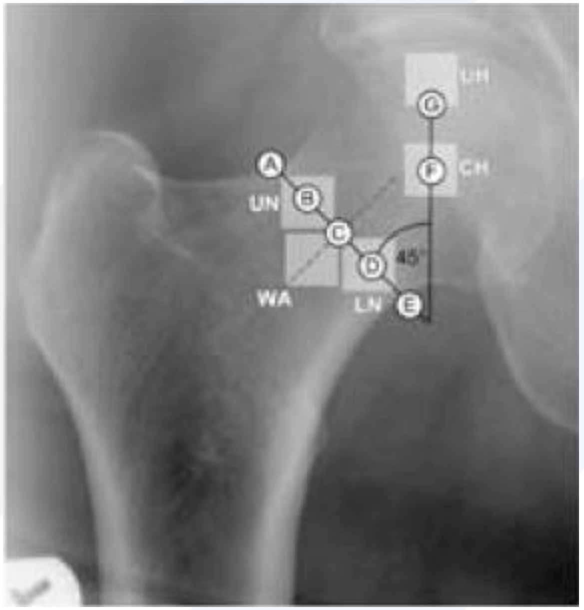
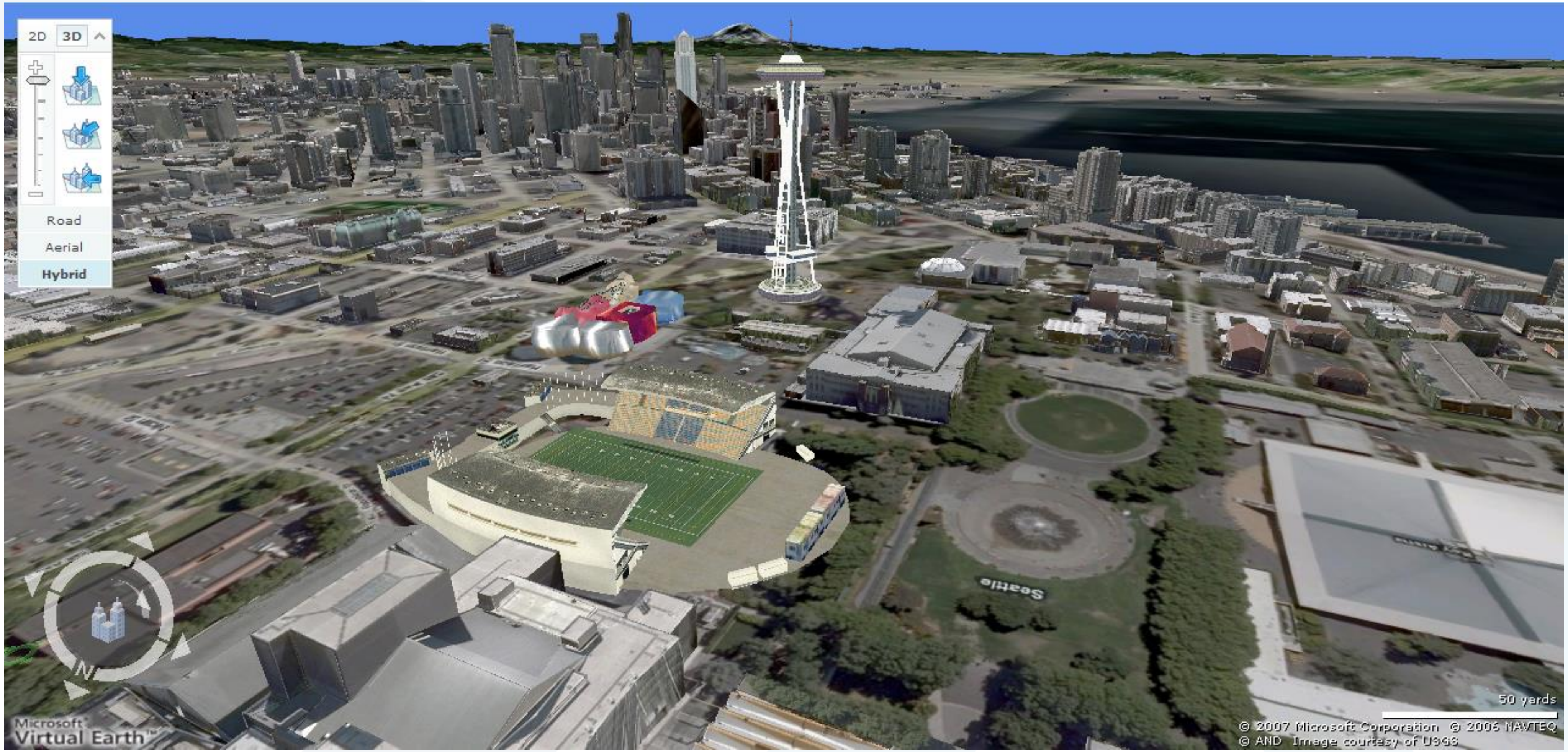


Image guided surgery
[Grimson et al., MIT]

Earth viewers (3D modeling)



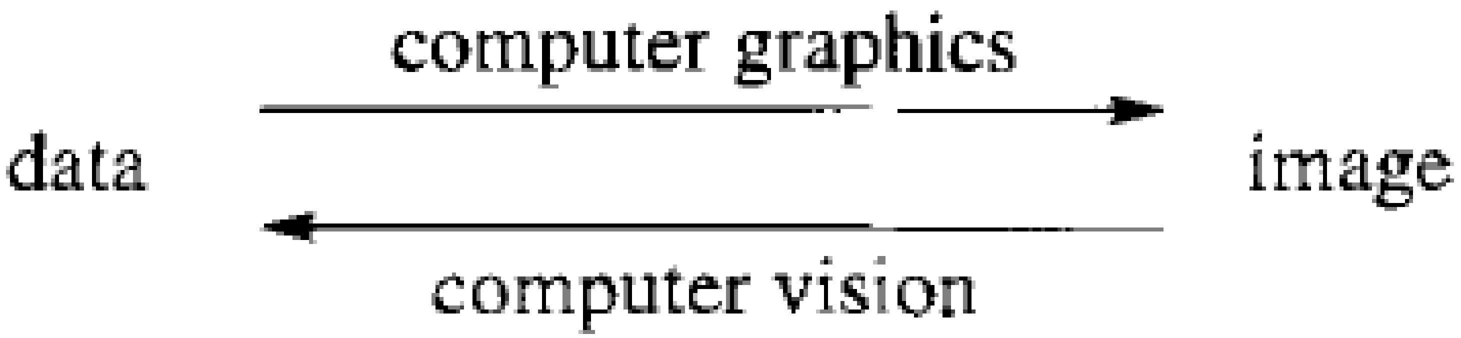
Textbook

Computer Vision: Algorithms and Applications

© 2010 Richard Szeliski, Microsoft Research

http://szeliski.org/Book/

Computer graphics versus computer vision



Every picture tells a story

- In **computer graphics** data are used to create pictures, for example, the case of computer-aided design (CAD).
- In **computer vision** one starts with a picture and attempts to abstract data from it.
- Usually experts in computer vision are also experts in computer graphics and use this expertise to create graphic user interfaces (GUIs) which make their vision products easier to use and more attractive.

2-D vision and 3-D stereo vision

- Almost all digital images are two-dimensional.
- In some applications it is necessary to reconstruct the three-dimensional scene from the image to obtain information in three dimensional coordinates.
- ***Stereo-imaging*** allows powerful 3-D imaging, as our own 3-D binocular vision illustrates.
- Stereo imaging involves **two cameras** with known vision and geometric properties arranged so that their 2-D pixels can be correlated. Various **algorithms** are then used to recognize objects and reconstruct some 3-D data.

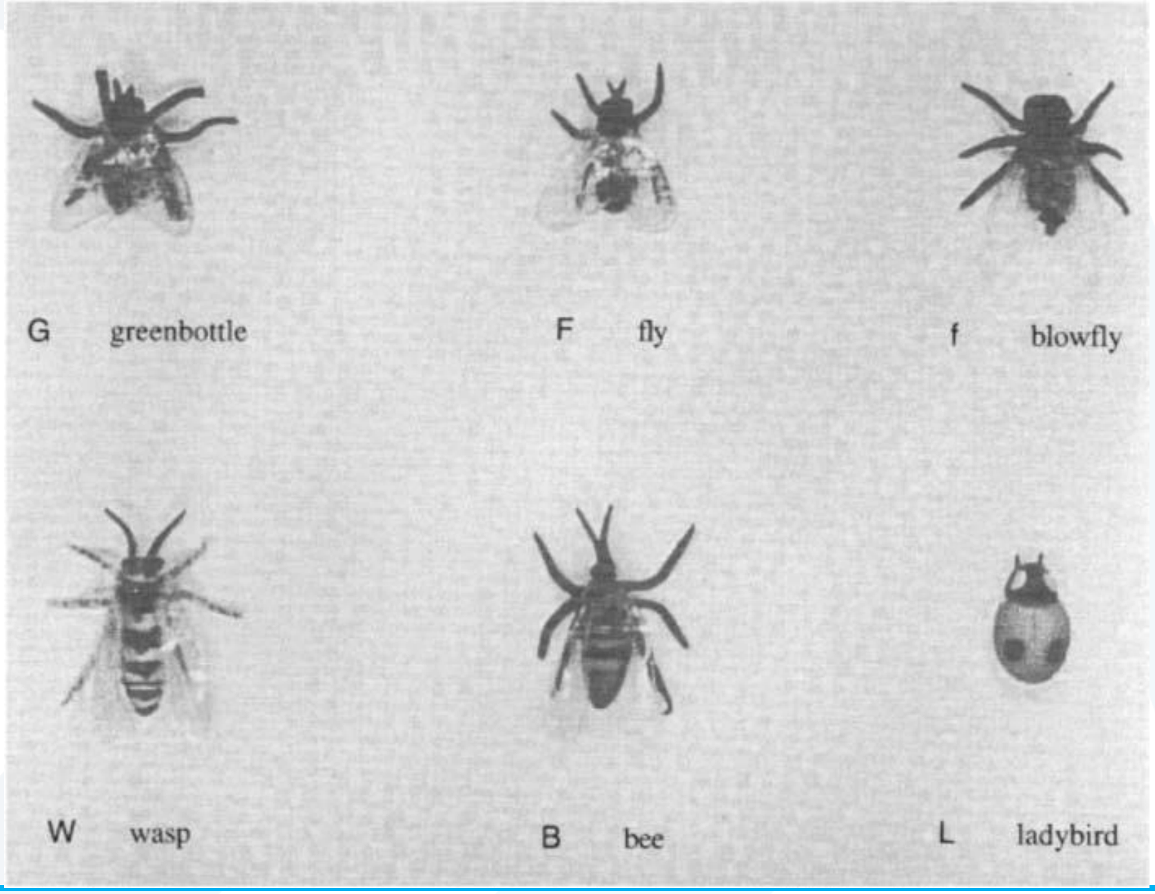
Object Recognition

- In computer vision one must be clear which data are actually required in any particular case.
- Once this is known the engineer can choose the least expensive or most effective way of delivering that information (assuming it is possible).
- The simplest information one can demand of a vision system is ***whether an object is present or not***.

Detecting insects in a digital image using neural networks

- In some cases the cost of **erroneous recognition** may be unacceptably high, and the vision system must be more discriminating than the simple infra-red sensor.
- To illustrate this point consider a machine whose purpose is to kill some undesirable insects by ultraviolet irradiation but not to kill other benign insects.

- Let us suppose that the insects of interest are the following, as illustrated in Figure



- How can we begin the job of recognizing these objects in an image?
- What information can we use?
- Of many possibilities, we might immediately think of characteristics such as colour, shape, size, movement pattern, speed, and so on.
- As always in computer vision, these things are easy to say but much more difficult to pin down in an explicit representation.

- For example, how would we represent the concept of insect shape within a computer?
- It can be done, but, of the various possibilities, **colour** is one of the easiest characteristics to represent.
- One of the first problems is distinguishing the objects of interest (**the insects**) from objects of no interest (**the background**).

- As Figure shows, the background can be very complex, and in the case of flowers it can move around considerably.

The first problem in computer vision: discriminating objects of interest from" the background clutter

- Thus the problem of deciding if the image contains a bee is compounded by the problem of knowing what else the image contains.
- A simple solution to the problem of ***background clutter*** is to constrain the system so that the background is fixed and simple.
- So suppose that the machine will have a platform on which the insects will walk, and suppose that this platform will be of a constant light blue colour which allows the pixels to be classified as either background or insect with reasonable fidelity.
- In practice, of course, some pixels are misclassified so that in the following experiments some background pixels have their data included in the 'insect' statistics, while some genuine insect pixels have got lost in the background.

- Having abstracted a set of 'insect pixels' from the image, suppose that the greyscale values from a colour camera can be used to classify the pixels into one of the following colours:

Red	Green
Yellow	Ochre
Black	White

- So that each insect will have its pixels assigned to these six classes.

PIXEL FREQUENCIES BY COLOUR FOR SIX PAIRS OF INSECTS

Object	Red	Green	Yellow	Ochre	Black	White	Unclassified
Ladybird 1	12872	554	423	291	9107	223	145
	13009	483	243	314	8728	364	403
Ladybird 2	18579	842	228	28	16879	917	283
	18402	938	229	25	16880	969	218
Fly 1	2775	3383	1845	294	16566	368	1966
	2831	3417	1878	334	16675	285	1482
Fly 2	1399	2322	3875	375	13271	3445	5924
	1414	2378	3797	384	13398	3316	5685
Bee 1	3669	4821	1281	2576	21829	693	779
	3594	4542	1336	2762	21442	632	762
Bee 2	3695	3037	535	2021	18198	344	266
	3623	3173	621	2102	18086	333	256
Wasp 1	2314	3590	12753	1686	8386	2846	2330
	2269	3834	12675	1608	8540	2724	2155
Wasp 2	3097	5034	15124	3626	9599	4406	845
	2756	3814	15018	3159	8936	6701	1556
Greenbottle 1	1678	6797	1207	106	15737	1402	1891
	1691	6734	1152	98	15831	1371	1356
Greenbottle 2	842	7931	3192	63	14925	3287	3113
	873	7900	3183	57	14708	3254	3142
Blowfly 1	1801	4033	432	231	24722	24	1185
	2161	3994	412	240	25271	27	1143
Blowfly 2	3958	3578	212	153	29592	11	458

- The ladybirds have the highest numbers of red pixels.
- The wasps have the highest numbers of yellow pixels.
- The greenbottles have the highest numbers of green pixels, and so on.
- However, not all the insects can be classified by having predominance in one colour.
- For example, honey bees are mostly black but have a yellowy-brown 'ochre' colour. Even though this colour characterizes the bees, their ochre count is less than that of the wasps whose bright yellow becomes this ochre colour in certain lights and shadows.

- Thus every insect is represented by six numbers: the number of its pixels classified as red, the number classified as green, the number classified as yellow, the number classified as ochre, the number classified as black, and the number classified as white.

- For example, for the first ladybird the numbers can be arranged as a sequence, or *vector:*

red	green	yellow	ochre	black	white
(12872,	554,	423,	291,	9107,	223)

- while for the second ladybird the statistics are:

red	green	yellow	ochre	black	white
(18579,	842,	228,	28,	16879,	917)

- Having established a pre-processing procedure which maps the various insects into this colour/frequency space, the computer vision task can be completed in a number of ways.
- This case is particularly well suited to the application of **neural networks**.
- The **training data** for the network are the vectors of colour frequencies as inputs, and the insect classes as outputs.
- So we might use a six-input and six-output network with six nodes in the hidden layer.
- The pattern recognition was very successful in correctly assigning new insects to their class.

- In this application of neural networks we exploit some of their **useful features**:

1. no two insects have exactly the same pixel colour frequencies and so the generalization of the network to 'similar' data is essential.
2. ability to cope with redundancy in the data. It happens that there is very little useful information in the frequency of white pixels. This is because the wings of the insects, although transparent in some lights, are highly reflective and can produce quite large 'white' responses in rather a random way.

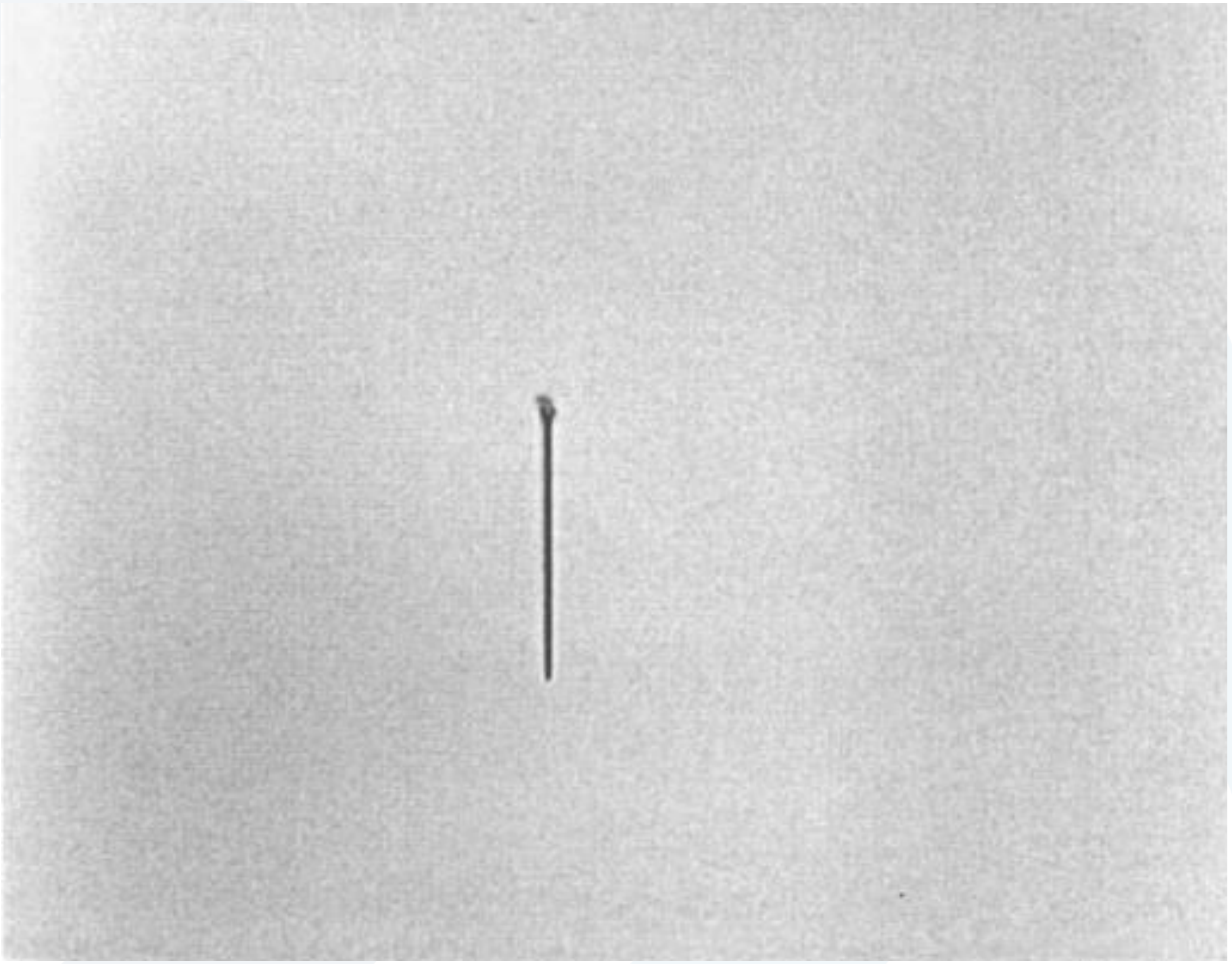
- Once the insect has been recognized - or more precisely, classified - the system can take whatever action is appropriate according to its specification.

Measuring the diameter of a pin using sub-pixel edge detection

- The kind of object recognition in the last section is of the 'yes/no', 'it-is-here-or not‘ kind.
- In many applications of computer vision much more than this is required.
- Apart from knowing that an image contains an object, more information about that object may be needed.

172	169	173	168	172	171	175	172	172	169	172	166	166	166	166	167	168	169	166	169	175	171	166	169	168	168	168	171	165	172	171
170	176	173	162	166	170	176	170	168	165	173	172	172	164	170	177	174	171	172	170	166	171	173	174	170	177	169	172	170	172	169
177	169	166	167	176	165	170	167	168	169	171	166	170	171	168	169	173	171	170	171	165	168	167	168	165	173	176	180	169	173	175
166	169	170	161	168	171	171	168	169	163	171	168	166	163	168	169	170	170	165	163	165	166	166	170	166	171	166	169	166	174	171
172	167	170	166	170	168	170	165	170	168	168	165	168	163	165	167	168	169	172	170	164	168	168	169	164	176	175	172	166	168	165
170	169	169	167	168	172	177	166	166	171	177	169	166	165	168	169	178	173	166	163	164	166	163	169	170	171	167	172	166	172	173
171	170	166	166	172	168	171	170	176	171	172	170	174	170	171	174	172	168	166	170	170	169	162	163	168	174	166	168	164	170	168
169	164	163	166	169	167	175	166	165	166	172	167	170	166	164	165	168	171	170	166	164	167	162	164	165	172	165	168	166	170	169
176	171	172	166	172	170	174	169	166	166	172	170	172	168	161	154	160	171	172	173	170	168	151	160	163	168	166	170	168	171	176
174	167	167	165	171	170	174	169	172	169	172	176	172	154	146	142	142	156	174	176	169	170	163	166	169	174	161	169	171	174	170
176	166	168	169	172	166	175	168	167	171	183	179	160	133	115	112	114	131	160	173	169	170	167	168	170	174	172	175	171	174	170
174	170	170	165	167	163	170	166	168	168	175	173	165	124	78	71	83	114	144	163	165	161	161	167	170	175	169	171	167	172	170
169	163	167	162	167	164	170	165	164	166	170	167	156	114	59	45	65	100	139	162	159	163	165	165	167	174	169	171	169	171	166
172	168	169	167	173	168	172	166	167	167	175	166	165	134	66	1	69	130	150	158	165	170	167	168	171	176	173	172	167	173	175
170	167	172	166	174	169	175	172	167	163	176	172	174	150	77	0	74	150	165	163	171	174	168	166	167	173	167	169	171	173	174
171	168	169	166	173	171	178	173	166	166	177	174	173	159	88	0	81	151	166	159	164	172	166	168	166	172	167	166	164	170	168
178	172	172	168	167	163	173	170	170	169	178	181	183	162	88	2	53	152	170	162	167	172	167	165	168	173	170	172	167	169	168
169	170	176	171	177	174	175	171	169	168	177	175	183	171	102	2	51	155	176	166	164	172	173	172	167	175	175	173	171	178	174
174	170	172	168	173	168	177	172	165	170	181	175	177	170	95	1	58	154	172	164	166	172	167	166	165	170	172	173	167	172	173
167	164	167	165	174	169	175	172	170	169	178	174	180	168	99	4	33	150	172	159	162	168	169	170	170	175	176	178	169	167	170
176	170	169	168	176	169	176	171	165	169	178	173	177	166	97	2	43	151	176	164	164	173	173	173	174	176	171	170	167	170	168
172	166	174	171	180	175	181	175	168	168	177	177	183	169	106	0	38	152	175	184	168	178	174	173	171	174	171	173	169	175	172
173	168	167	171	177	168	175	171	171	168	178	173	178	165	98	1	39	150	177	163	166	173	175	173	171	178	171	173	171	175	171
173	167	174	168	170	170	176	166	164	170	176	177	182	168	111	1	27	147	175	159	158	170	171	167	165	175	173	174	171	177	172
171	166	168	168	173	164	173	169	167	169	178	177	183	167	105	3	29	145	173	162	165	173	169	166	163	171	169	171	169	173	172
175	171	173	172	174	168	176	172	168	164	173	175	177	173	112	2	29	141	170	161	160	168	169	167	168	176	171	173	170	176	174
175	171	173	171	178	169	172	174	170	167	175	176	180	170	120	7	18	141	176	165	164	172	168	168	165	173	171	175	171	176	174
171	166	168	167	172	165	175	173	170	171	177	175	178	173	116	0	19	140	172	159	158	170	168	168	167	173	169	170	167	170	166
169	171	172	163	170	168	171	169	169	167	173	176	179	172	125	2	12	137	176	160	157	168	167	168	165	175	167	171	173	172	167

- Inspection of the greyscale numbers in Figure shows there is a marked vertical column of low values in the centre of the image.
- This suggests there is a **dark object** against a lighter background.



- conventional attempts at edge detection have not been very successful because they depend on thresholding.
- Figure shows the results of a new method of edge detection .
- Using this method, the geometric 'edges' of the pin have been detected to **sub-pixel** accuracy approaching one-tenth of a pixel.

172	169	173	168	172	171	175	172	172	169	172	166	166	166	166	167	168	169	166	169	175	171	166	169	168	168	166	171	165	172	171
170	176	173	162	166	170	176	170	168	165	173	172	172	164	170	177	174	171	172	170	166	171	173	174	170	177	169	172	170	172	169
177	169	166	167	176	165	170	167	168	169	171	166	170	171	168	169	173	171	170	171	165	168	167	168	165	173	176	180	169	173	175
166	169	170	161	168	171	171	168	169	163	171	168	166	163	168	169	170	170	165	163	165	166	166	170	166	171	166	169	166	174	171
172	167	170	166	170	168	170	165	170	168	168	165	168	163	165	167	168	169	172	170	164	168	168	169	164	176	175	172	166	168	165
170	169	169	167	168	172	177	166	166	171	177	169	166	165	168	169	178	173	166	163	164	166	163	169	170	171	167	172	166	172	173
171	170	166	166	172	168	171	170	176	171	172	170	174	170	171	174	172	168	166	170	170	169	162	163	168	174	166	169	164	170	168
169	164	163	166	169	167	175	166	165	168	172	167	170	166	164	165	168	171	170	166	164	167	162	164	165	172	165	168	166	170	169
176	171	172	166	172	170	174	169	166	166	172	170	172	168	161	154	160	171	172	173	170	168	161	160	163	168	166	170	168	171	176
174	167	167	165	171	170	174	169	172	169	172	176	172	154	146	142	142	156	174	176	169	170	163	166	169	174	161	169	171	174	170
176	166	168	169	172	166	175	168	167	171	183	179	160	133	115	112	114	131	160	173	169	170	167	168	170	174	172	175	171	174	170
174	170	170	165	167	163	170	166	168	168	175	173	165	124	78	71	83	114	144	163	165	161	161	167	170	175	169	171	167	172	170
169	163	167	163	167	164	170	165	164	166	170	167	156	114	59	45	63	100	139	162	159	163	165	165	167	174	169	171	169	171	166
172	168	169	167	173	168	172	166	167	167	175	168	165	134	66	1	69	130	150	158	165	170	167	168	171	176	173	172	167	173	175
170	167	172	166	174	169	175	172	167	163	176	172	174	150	77	0	74	150	165	163	171	174	168	166	167	173	167	169	171	173	174
171	168	169	166	173	171	178	173	166	166	177	174	173	159	88	0	61	151	168	159	164	172	166	168	166	172	167	166	164	170	168
178	172	172	168	167	163	173	170	170	169	176	181	183	162	88	2	53	152	170	162	167	172	167	165	166	173	170	172	167	169	168
169	170	176	171	177	174	175	171	169	168	177	175	183	171	102	2	51	155	176	166	164	172	173	172	167	175	175	173	171	178	174
174	170	172	168	173	168	177	172	166	170	181	175	177	170	95	1	58	154	172	164	166	172	167	166	165	170	172	173	167	172	173
167	164	167	165	174	169	175	172	170	169	178	174	180	168	99	4	33	150	172	159	162	168	169	170	170	175	176	178	169	167	170
176	170	169	168	176	169	176	171	165	169	178	173	177	166	97	2	43	151	176	164	164	173	173	173	174	176	171	170	167	170	168
172	166	174	171	180	175	181	175	168	168	177	177	183	169	106	0	38	152	175	164	168	178	174	173	171	174	171	173	169	175	172
173	168	167	171	177	168	175	171	171						98	1	39	150						173	171	178	171	173	171	175	171
173	167	174	168	170	170	176	166	164						111	1	27	147						167	165	175	173	174	171	177	172
171	166	168	168	173	164	173	169	167						105	3	29	145						166	163	171	169	171	169	173	172
175	171	173	172	174	168	176	172	168	164	173	175	177	173	112	2	29	141	170	161	160	168	169	167	168	176	171	173	170	176	174
175	171	173	171	178	169	172	174	170	167	175	176	180	170	120	7	18	141	176	165	164	172	168	168	165	173	171	175	171	178	174
171	166	168	167	172	165	175	173	170	171	177	175	178	173	116	0	19	140	172	159	158	170	168	168	167	173	169	170	167	170	166

edges detected
to sub-pixel
accuracy

edges detected
to sub-pixel
accuracy

- By this method the measurement is that the pin is **2.5 pixels** across.
- At 100 pixels per inch this means that the diameter measurement of the pin is 0.025 inches.