ESE 26
vision par ordinateur

Sabrina Bouatmane
Département d'eléctronique
Université de jijel

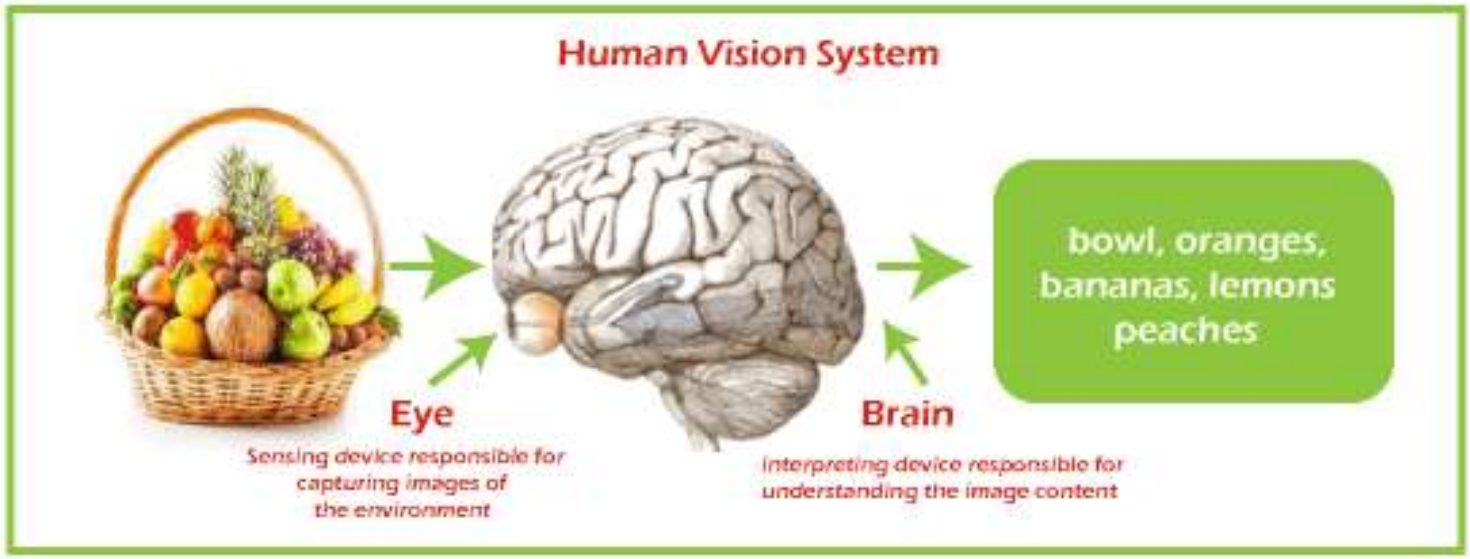
C’est quoi la vision artificielle?

- La vision par ordinateur consiste à utiliser des techniques informatiques pour extraire des informations significatives à partir d'images ou de vidéos. Cela permet à des systèmes automatisés de "voir", comprendre, et réagir à leur environnement visuel.

- Quel genre de scène?
- Où sont les voitures ?
- À quelle distance se trouve le bâtiment ?

Système de vision artificielle

a) Acquisition d'Images
caméras, scanners, vidéos, drones, etc.
Les images sont ensuite converties en données numériques (pixels) que l'ordinateur peut traiter.
b) traitement de l'Image
Filtrage, Segmentation, détection de contours,Points caracteristiques, descripteurs
c) Reconnaissance et Classification
Algorithmes d'apprentissage automatique.
CNN: Une architecture couramment utilisée pour traiter les images, en imitant le fonctionnement du cortex visuel humain.



- **Objectifs de la vision artificielle**
- Comprendre et interpréter des images.
- Automatiser des tâches visuelles qui nécessitent normalement une intervention humaine.

La vision est vraiment difficile

- Permettre l'interaction machine-environnement basée sur la vision.
- La vision est une caractéristique étonnante de l'intelligence
- La vision occupe plus de place chez l'humain que dans toute autre chose

Pourquoi la vision par ordinateur est-elle difficile ?

The state of Computer Vision and AI: we are really, really far away.
Oct 22, 2012

- Où cette photo a-t-elle été prise ?
- Combien de personnes y a-t-il ?
- Que font-elles
- Sur quel objet se tient la personne à gauche ?
- Pourquoi cette photo est-elle drôle ?

Defis

Éclairage

position

Désordre

Occlusions

Apparence intra-classe

Point de vue

Defis : variation Intra-classe

Histoire de la VA

Marvin Minsky, MIT
Turing award, 1969

“En 1966, Minsky a embauché un étudiant de première année de licence et lui a assigné un problème à résoudre pendant l'été :connecter une caméra à un ordinateur et faire en sorte que la machine décrive ce qu'elle voit.”

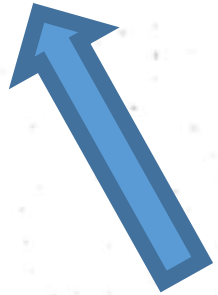
Crevier 1993, pg. 88

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966



Half a century later,
we're still working on it.

THE SUMMER VISION PROJECT

Seymour Papert

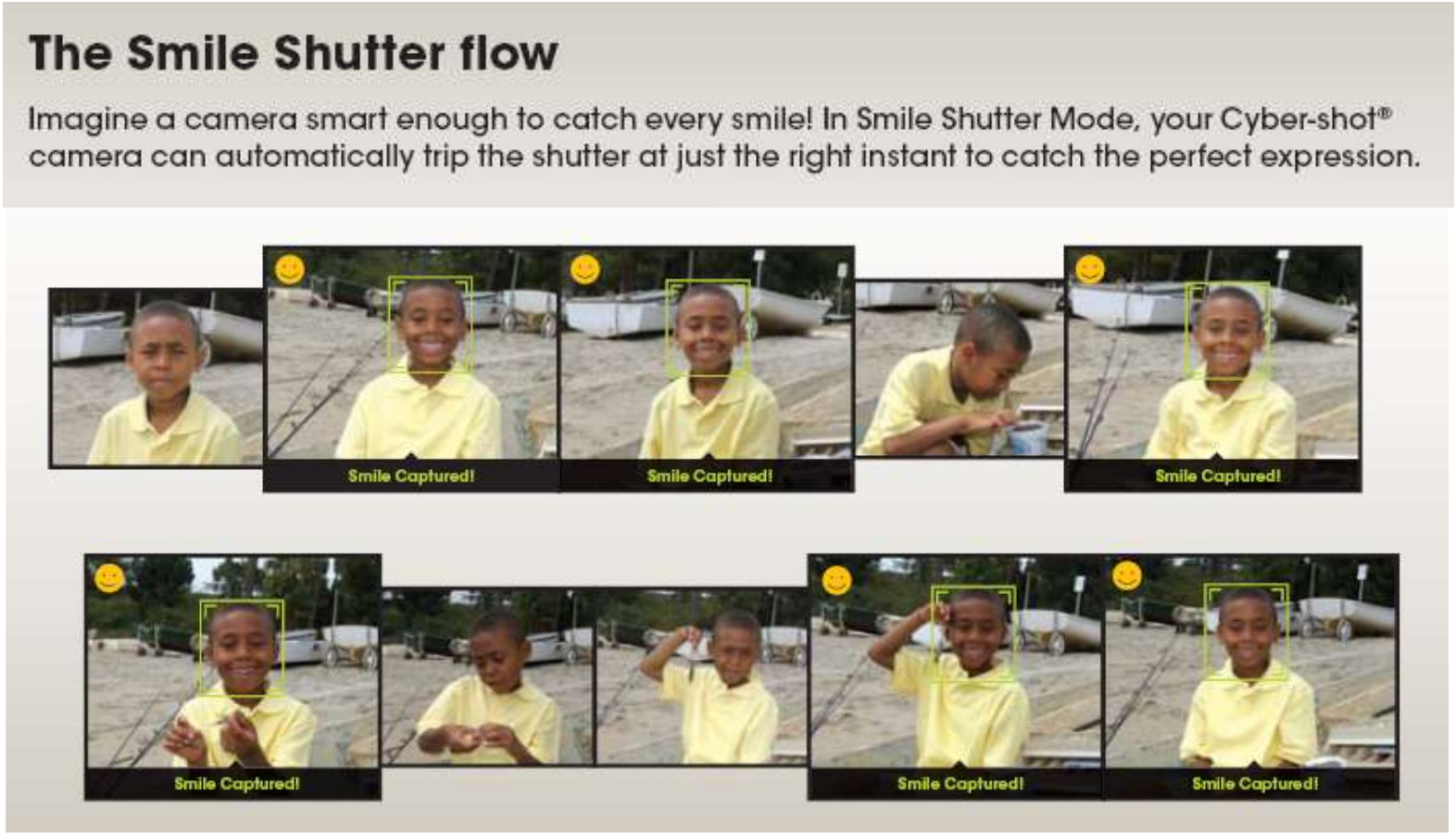
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the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

Application: detection de visage (Face detection)

Face recognition

Facebook face auto-tagging

Smile Detection

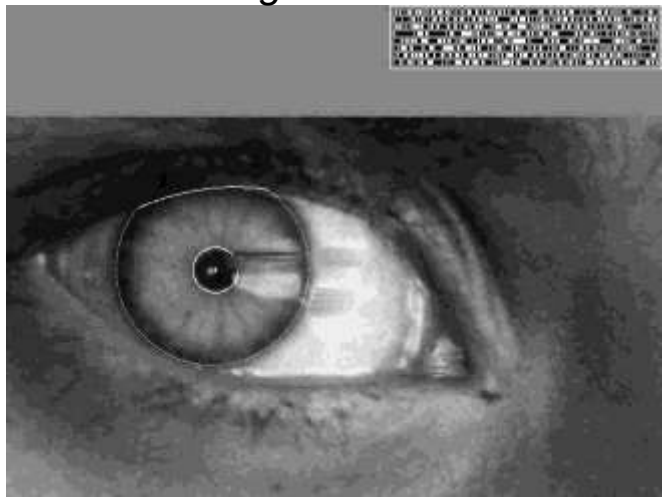


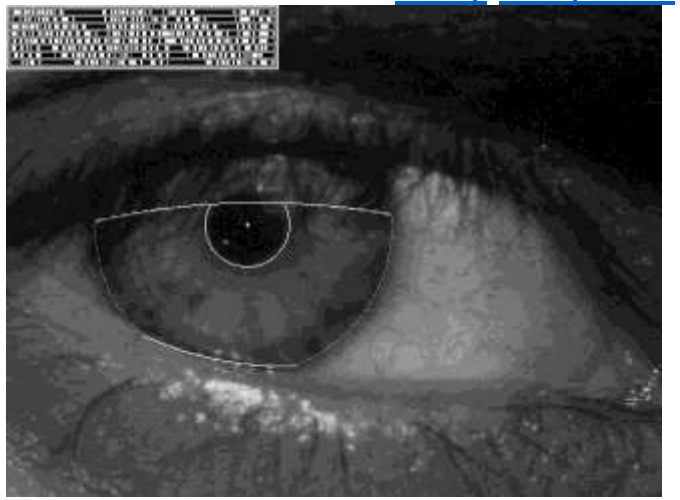
Sony Cyber-shot® T70 Digital Still Camera

Slide credit: Steve Seitz

biometrie

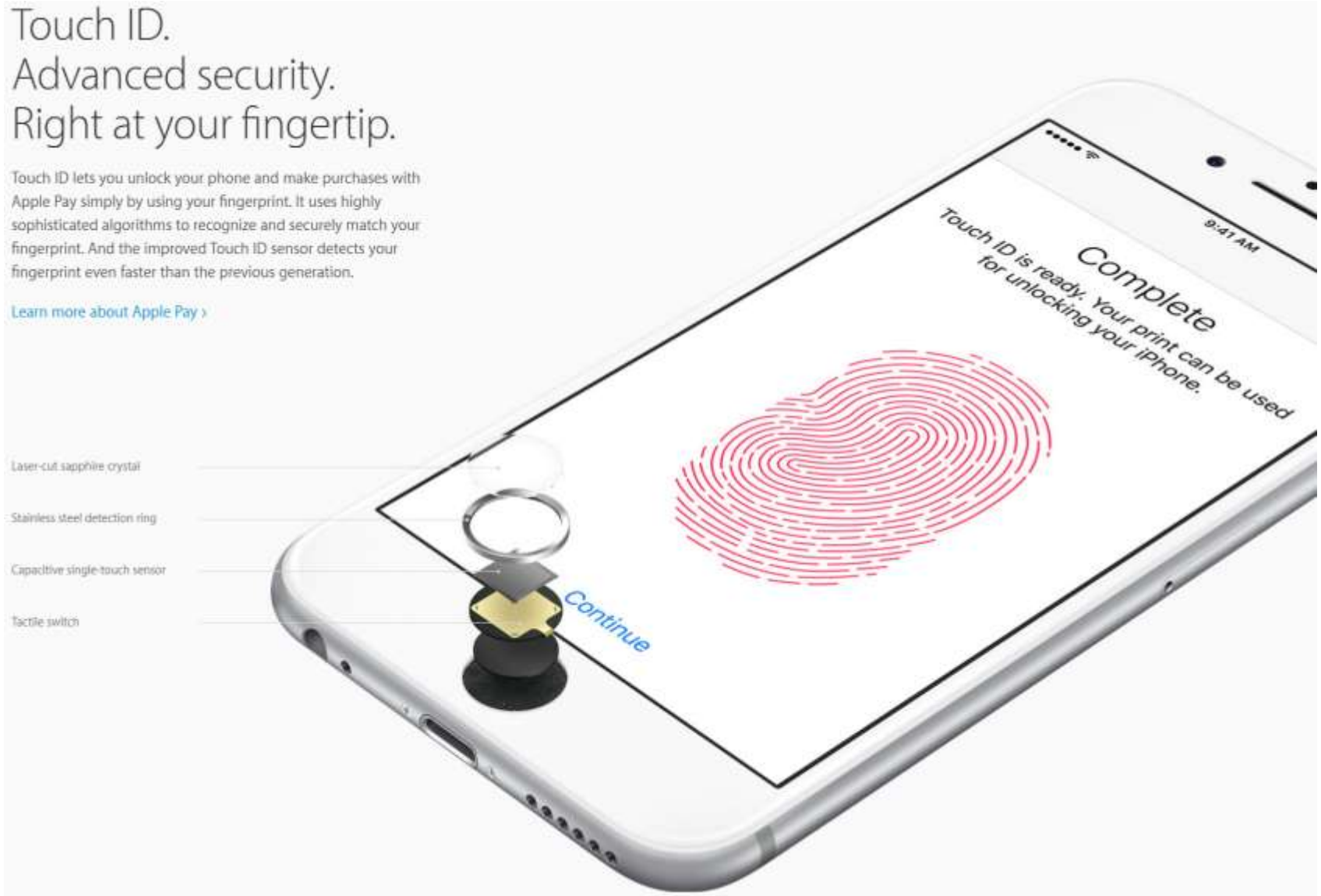
"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story wikipedia





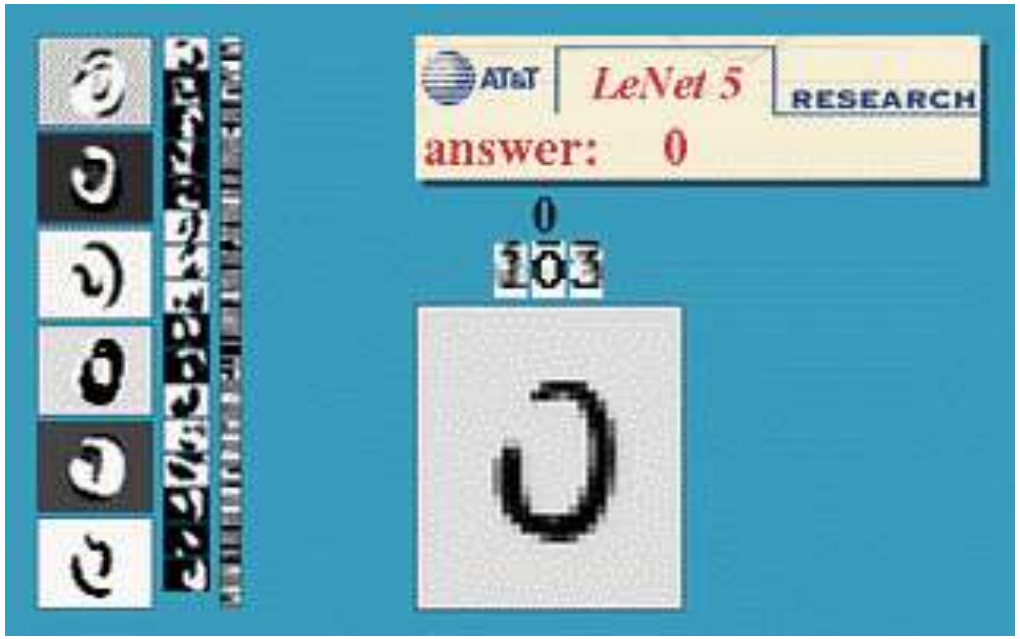
Slide credit: Steve Seitz

Vision-based Biometrics



Optical Character Recognition (OCR)

- Technologie permettant de convertir des documents numérisés en texte

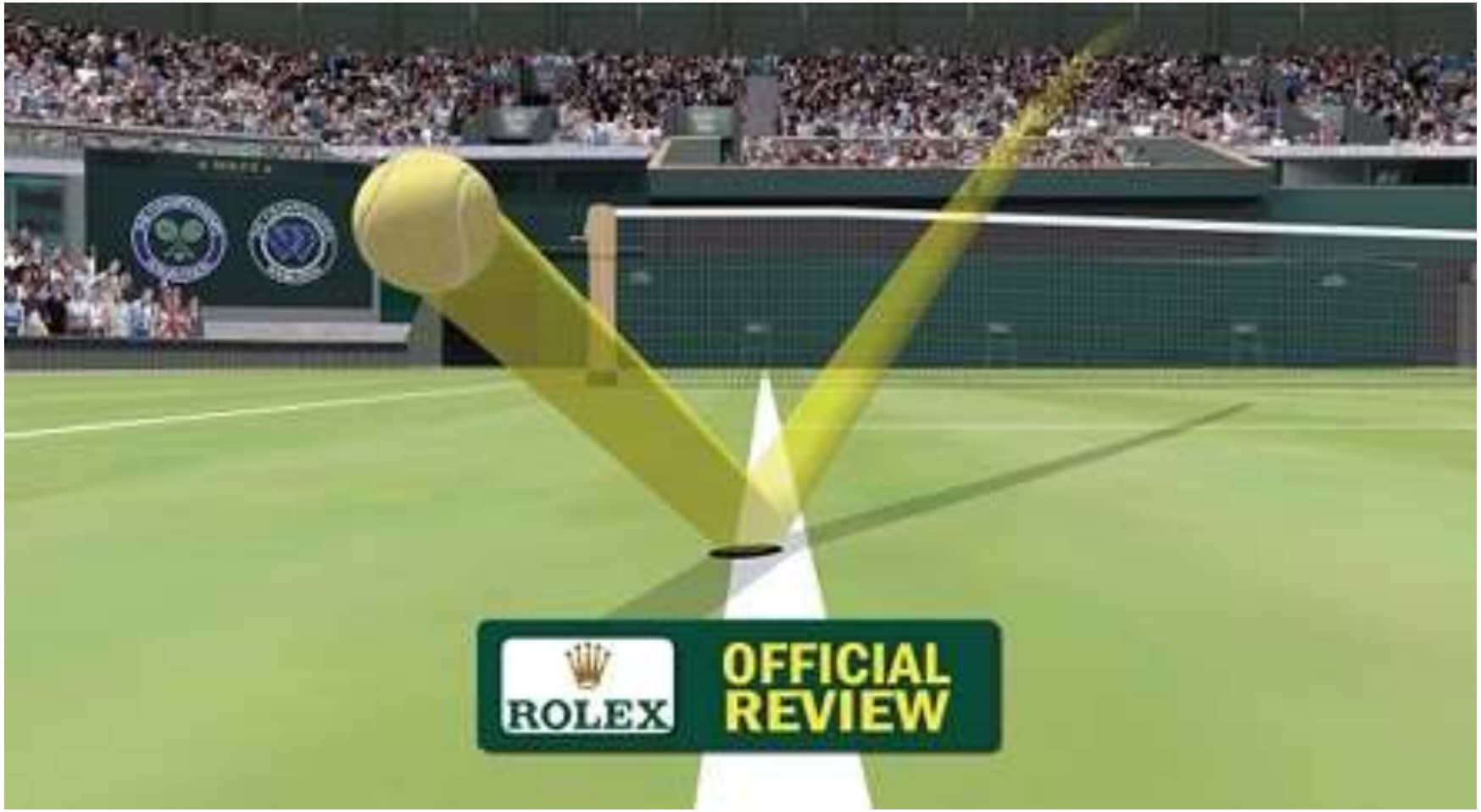


Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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

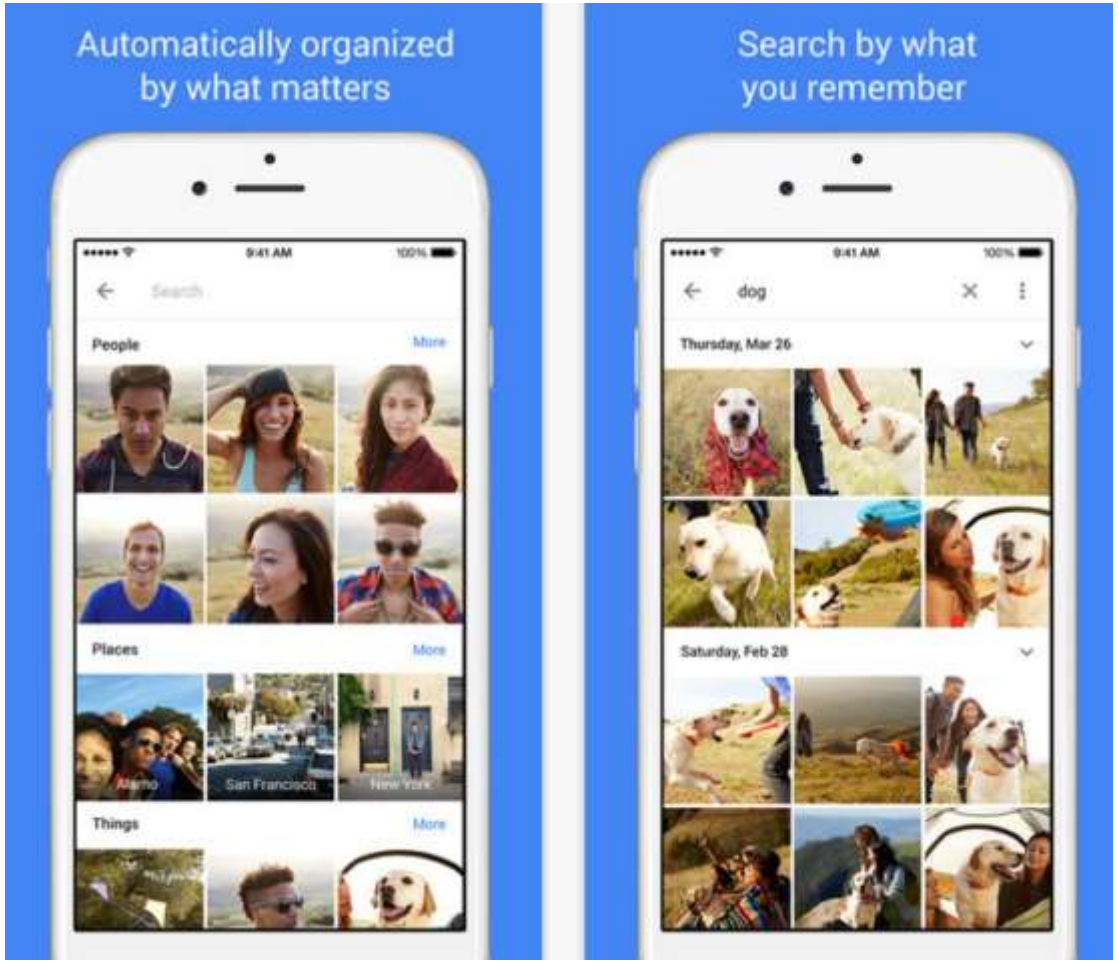
Slide credit: Steve Seitz

Computer vision in sports



Hawk-Eye: helping/improving referee decisions

Visual recognition for photo organization



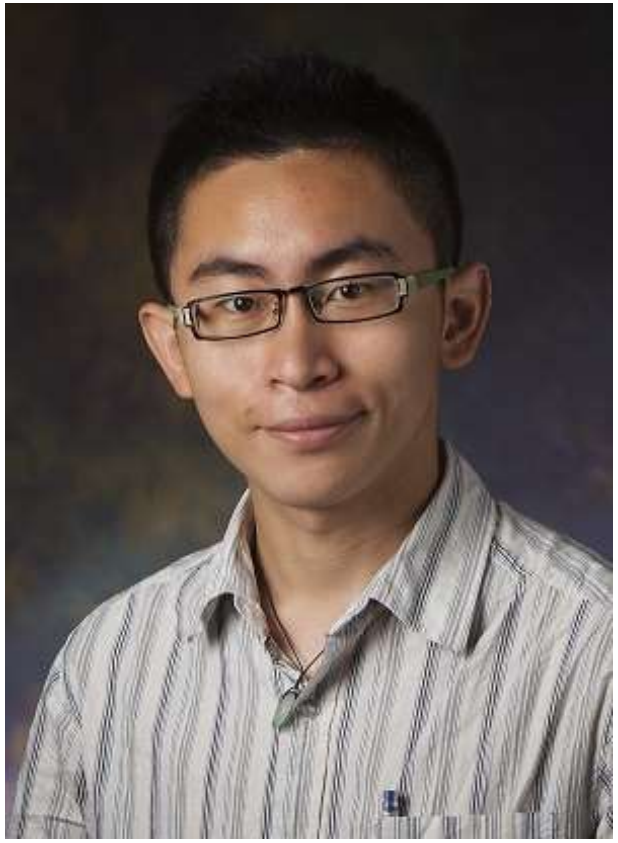
Google photo

3D from thousands of images

[Furukawa et al. CVPR 2010]

Style transfer

Source image (**Style**)



Target image (**Content**)

Output (deepart)

A Neural Algorithm of Artistic Style [Gatys et al. 2015]

Special effects: Shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Slide credit: Steve Seitz

Vision in 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

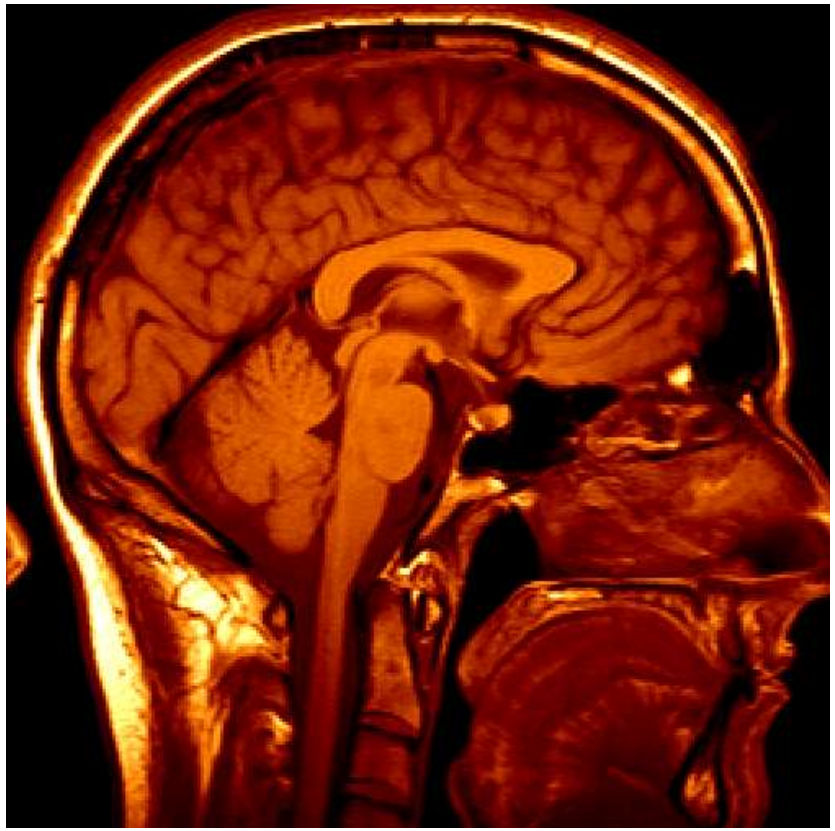
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “Computer Vision on Mars” by Matthies et al.

Industrial robots

Vision-guided robots position nut runners on wheels

http://www.automationworld.com/computer-vision-opportunity-or-threat

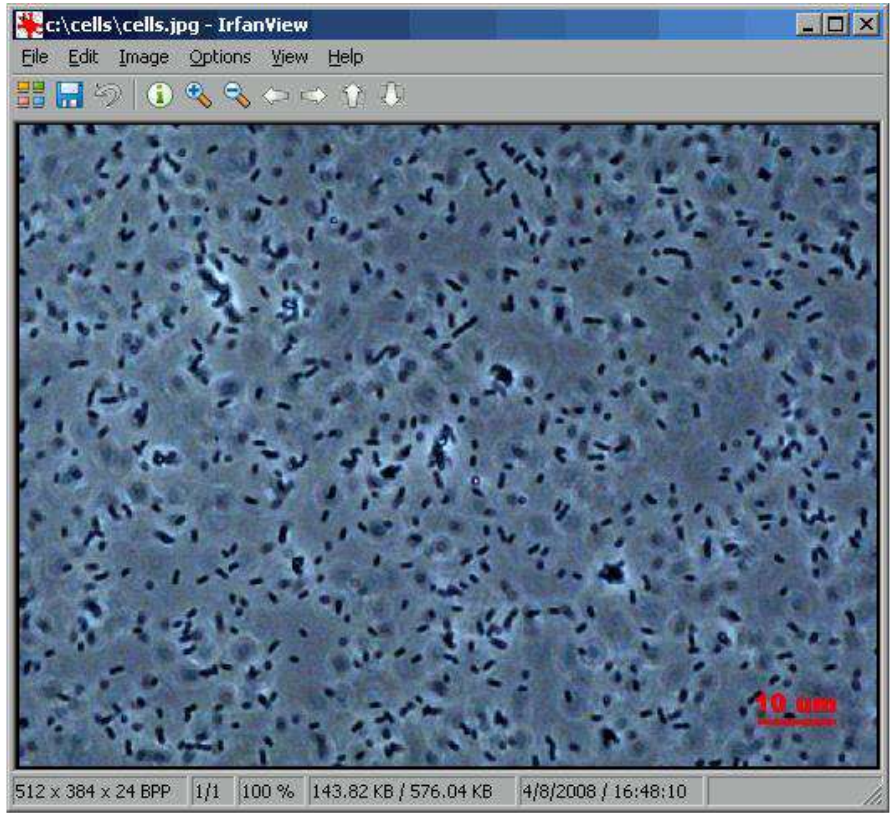
Medical imaging



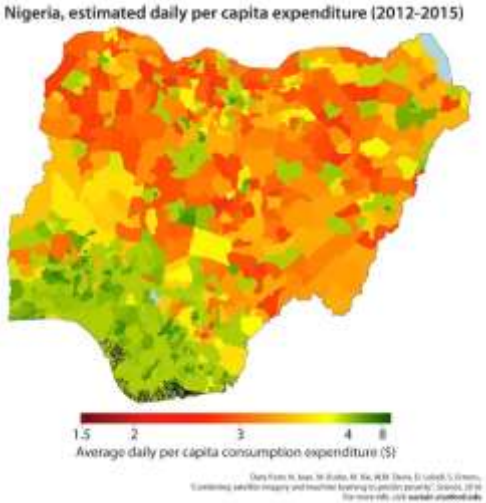
3D imaging
MRI, CT

Image guided surgery

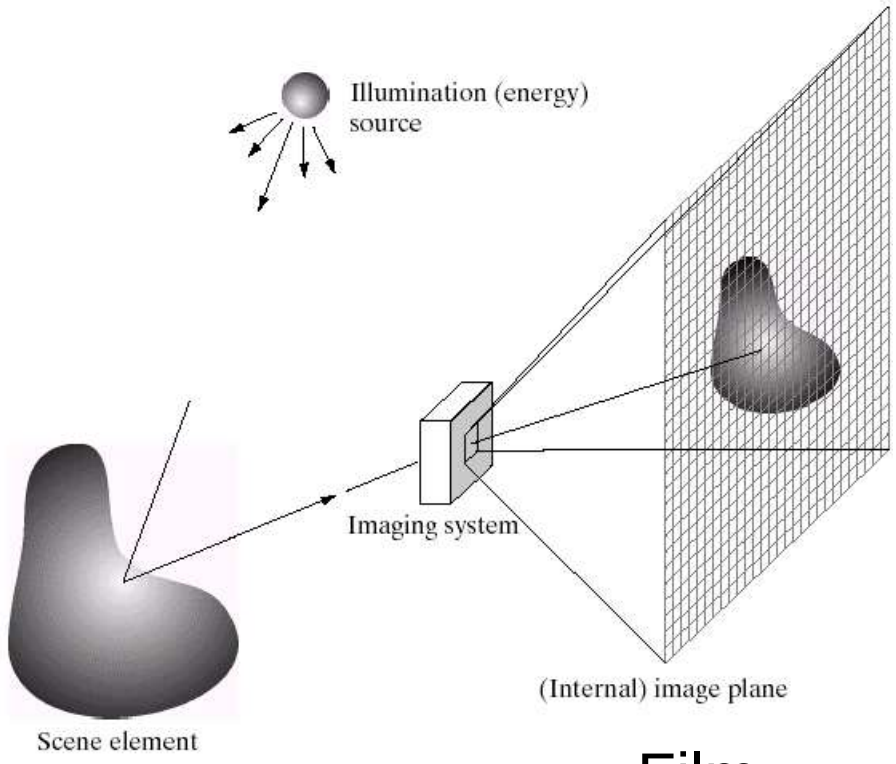
Computer vision for the mass



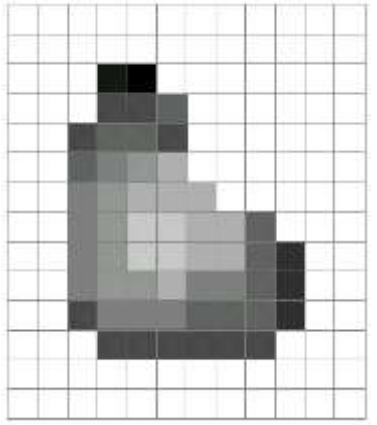
Counting cells



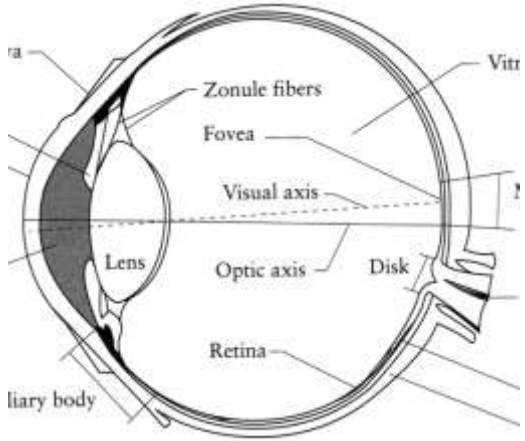
Formation d’image



Film

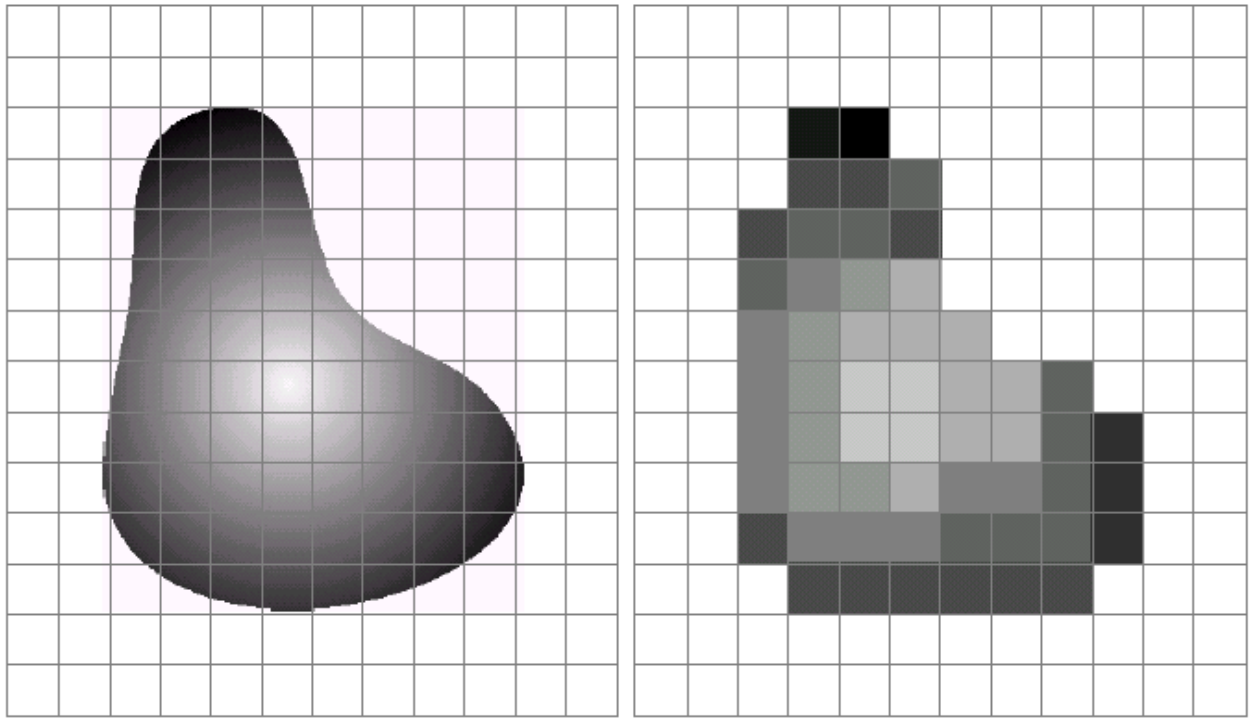


Camera numerique



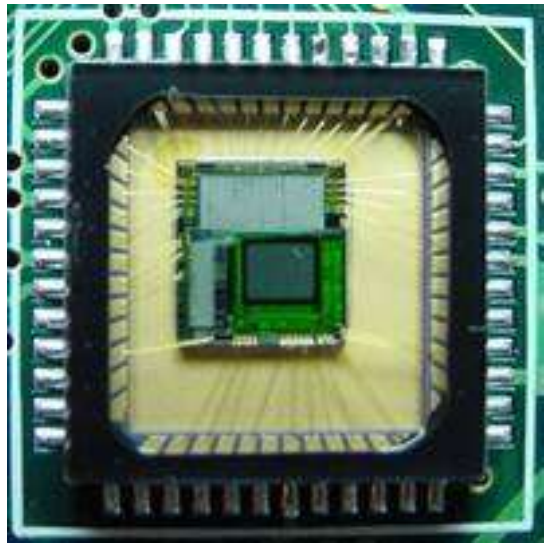
L’oeil

Formation d’image

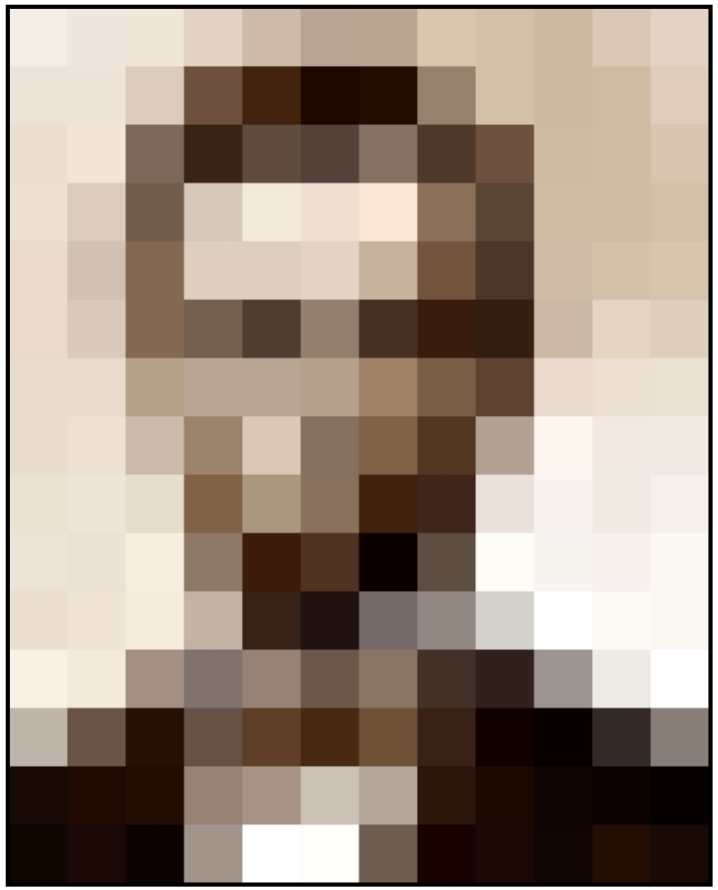


a b

FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.



CMOS sensor

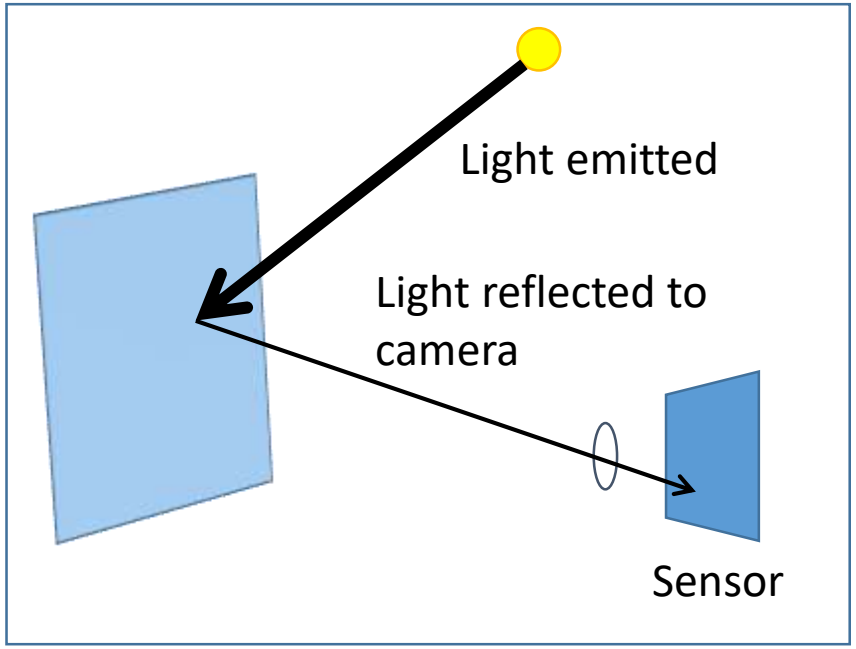


243	239	240	225	206	185	188	218	211	206	216	225
242	239	218	110	67	31	34	152	213	206	208	221
243	242	123	58	94	82	132	77	108	208	208	215
235	217	115	212	243	236	247	139	91	209	208	211
233	208	131	222	219	226	196	114	74	208	213	214
232	217	131	116	77	150	69	56	52	201	228	223
232	232	182	186	184	179	159	123	93	232	235	235
232	236	201	154	216	133	129	81	175	252	241	240
235	238	230	128	172	138	65	63	234	249	241	245
237	236	247	143	59	78	[illegible]	94	255	248	247	251
234	237	245	193	55	33	115	144	213	255	253	251
248	245	161	128	149	109	138	65	47	156	239	255
190	107	39	102	94	73	114	58	[illegible]	[illegible]	51	137
23	32	33	148	168	203	179	43	27	17	12	[illegible]
17	26	19	160	255	255	109	22	26	19	35	24

Comment un pixel obtient-il sa valeur

Facteurs principaux

- Intensité et direction de l'éclairage
- Géométrie de la surface
- Matériau de la surface
- Surfaces proches
- Gain/exposition de la caméra



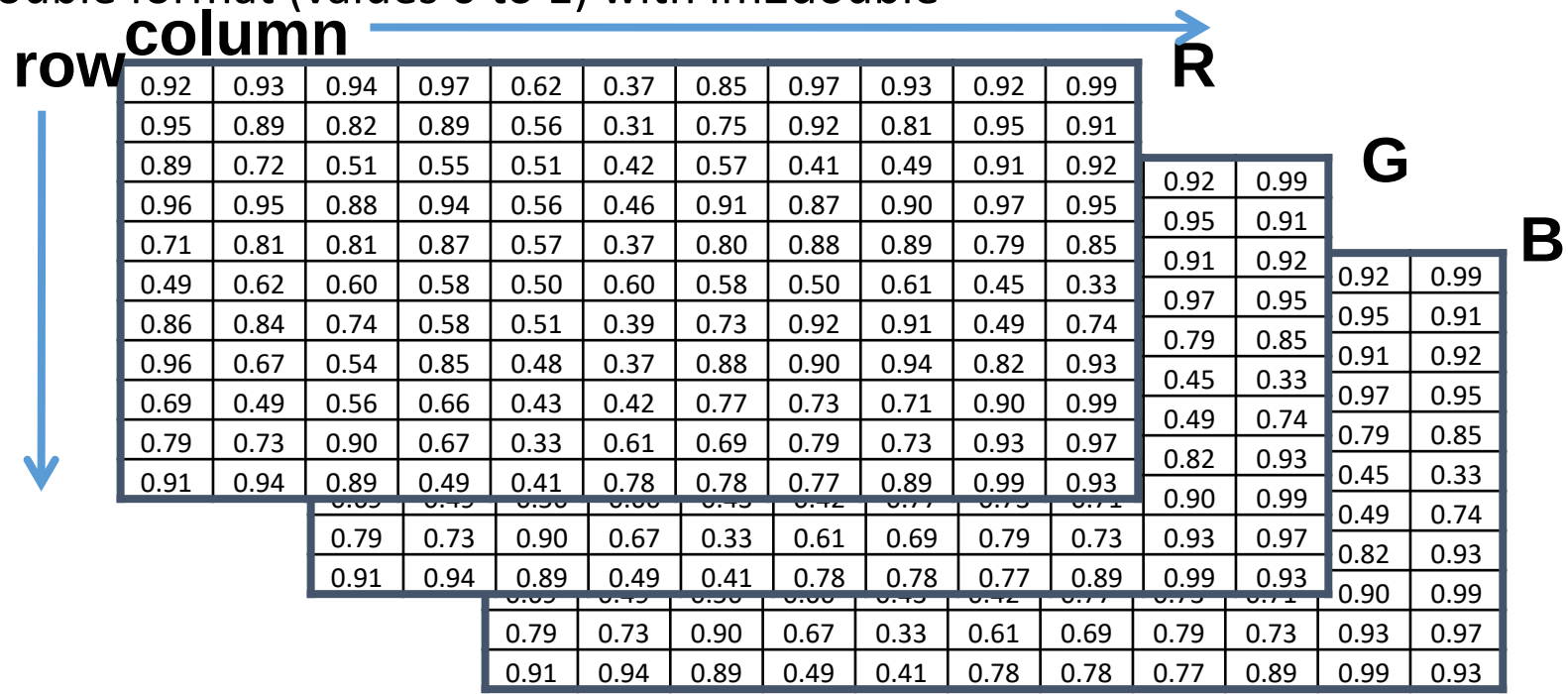
Que nous disent les différences d'intensité sur la forme ?

- Modifications de la normale de surface
- Texture
- Proximité
- Indentations et bosses
- Rainures et plis

Photos Koenderink slides on image texture and the flow of light

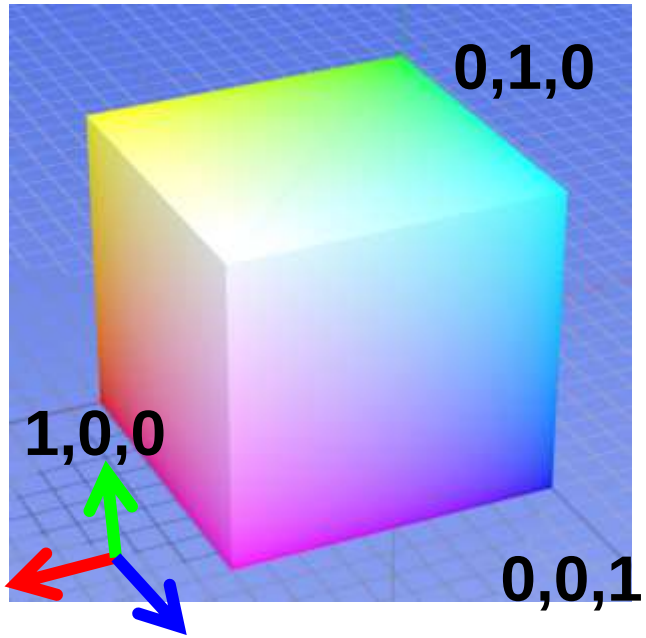
Images in Matlab

- Images represented as a matrix
- Suppose we have a NxM RGB image called “im”
  - im(1,1,1) = top-left pixel value in R-channel
  - im(y, x, b) = y pixels down, x pixels to right in the b^{th} channel
  - im(N, M, 3) = bottom-right pixel in B-channel
- imread(filename) returns a uint8 image (values 0 to 255)
  - Convert to double format (values 0 to 1) with im2double



Color spaces: RGB

Default color space



RGB cube

- Easy for devices
- But not perceptual
- Where do the grays live?
- Where is hue and saturation?

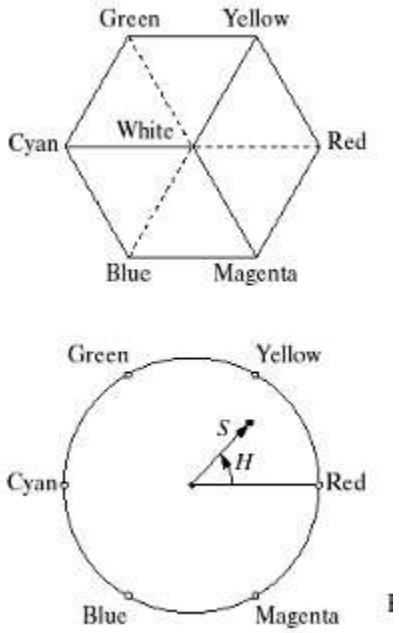
R
(G=0,B=0)

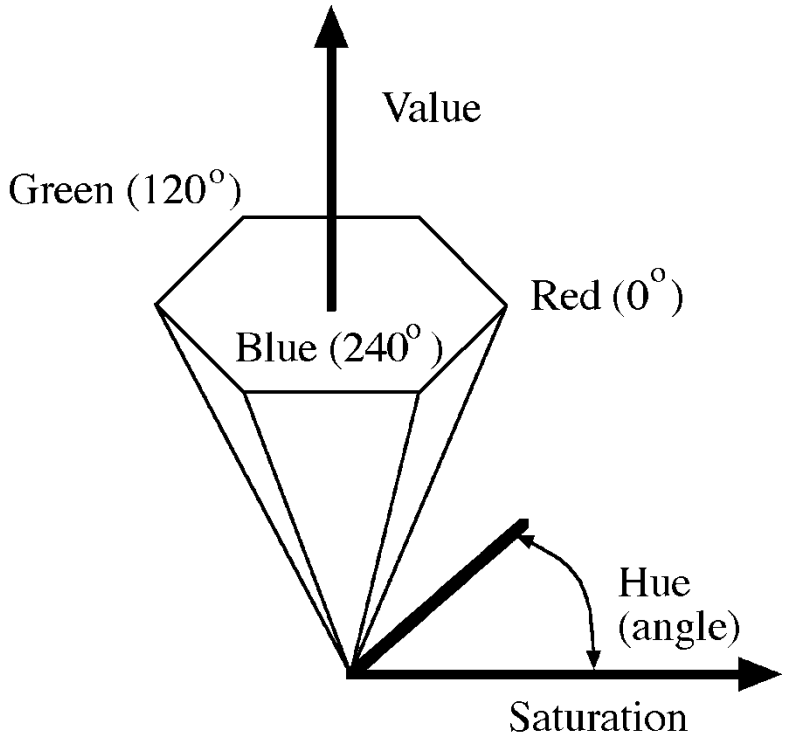
G
(R=0,B=0)

B
(R=0,G=0)

Image from: http://en.wikipedia.org/wiki/File:RGB_color_solid_cube.png

HSV



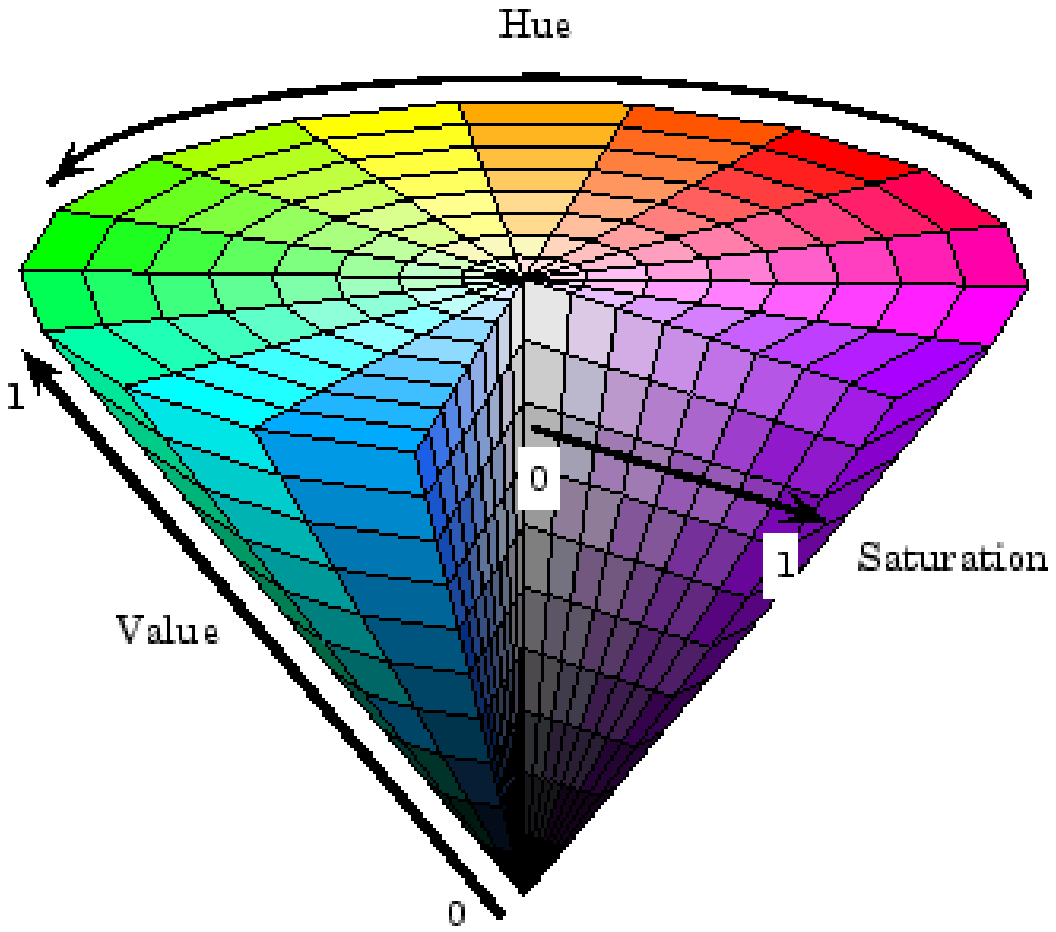


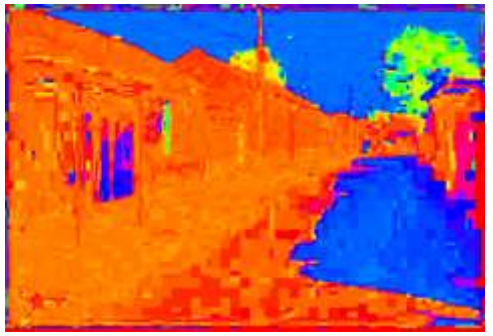
- Hue, Saturation, Value (Intensity)
  - RGB cube on its vertex
- Decouples the three components (a bit)
- Use rgb2hsv() and hsv2rgb() in Matlab

Slide by Steve Seitz

Color spaces: HSV

Intuitive color space





H

(S=1,V=1)

S

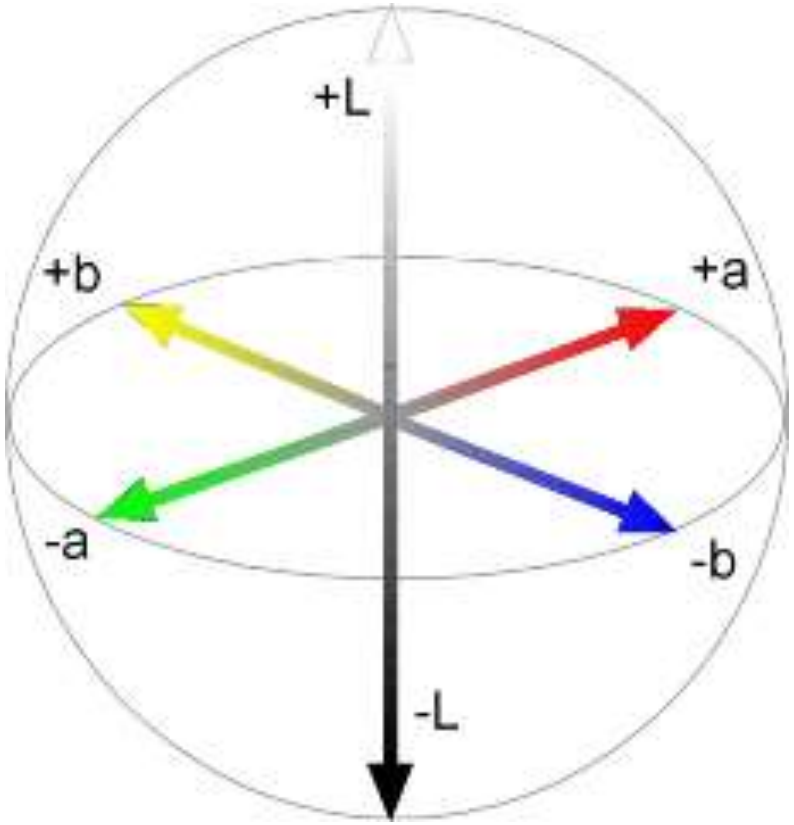
(H=1,V=1)

V

(H=1,S=0)

Color spaces: L*a*b*

"Perceptually uniform" color space



L
(a=0,b=0)

a
(L=65,b=0)

b
(L=65,a=0)