

Visión

**Sesión 1: Introducción a la
visión artificial**

Departamento CCIA
<http://www.rvg.ua.es/master/vision/2011/>

Hoy

- Introducción a la asignatura
- Introducción a la Visión Artificial
- Prácticas: Python, Scipy, Numpy, OpenCV

Introducción a la asignatura

Temario

Teoría

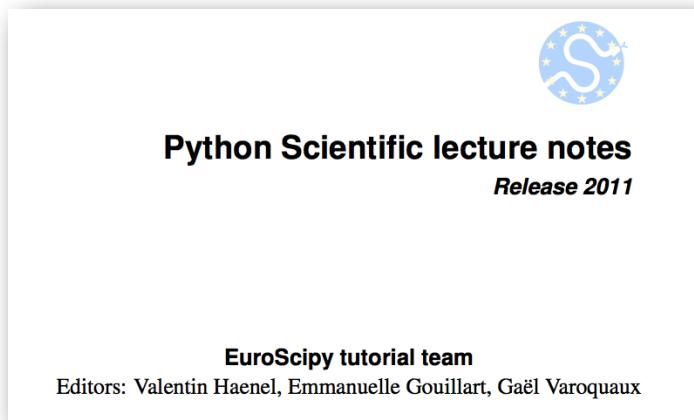
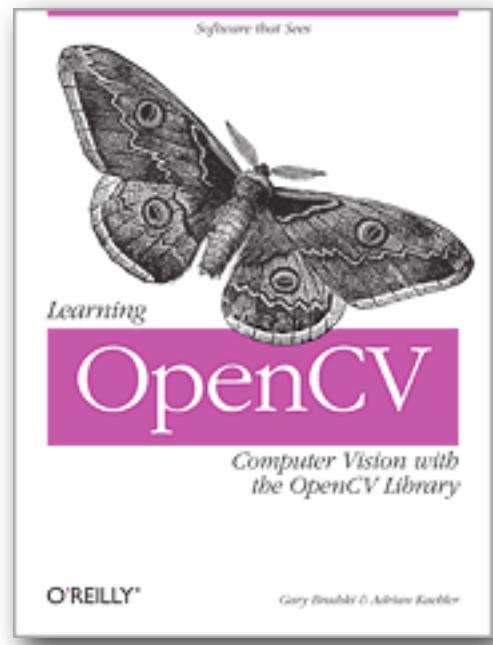
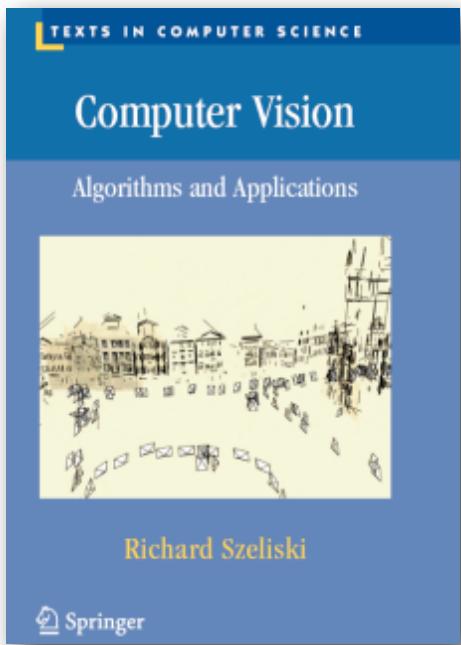
1. Introducción a la Visión Artificial y a sus aplicaciones
2. Clasificación de imágenes mediante histogramas y características
3. Extracción y emparejamiento de contornos
4. Tracking de objetos
5. Visión 3D: modelo y calibración de cámara, cámaras estéreo
6. Matching y reconstrucción 3D a partir de nubes de puntos

Prácticas

1. Desarrollo de software para visión artificial con OpenCV, Scipy y Python
2. Visión 3D con Point Cloud Library (PCL)

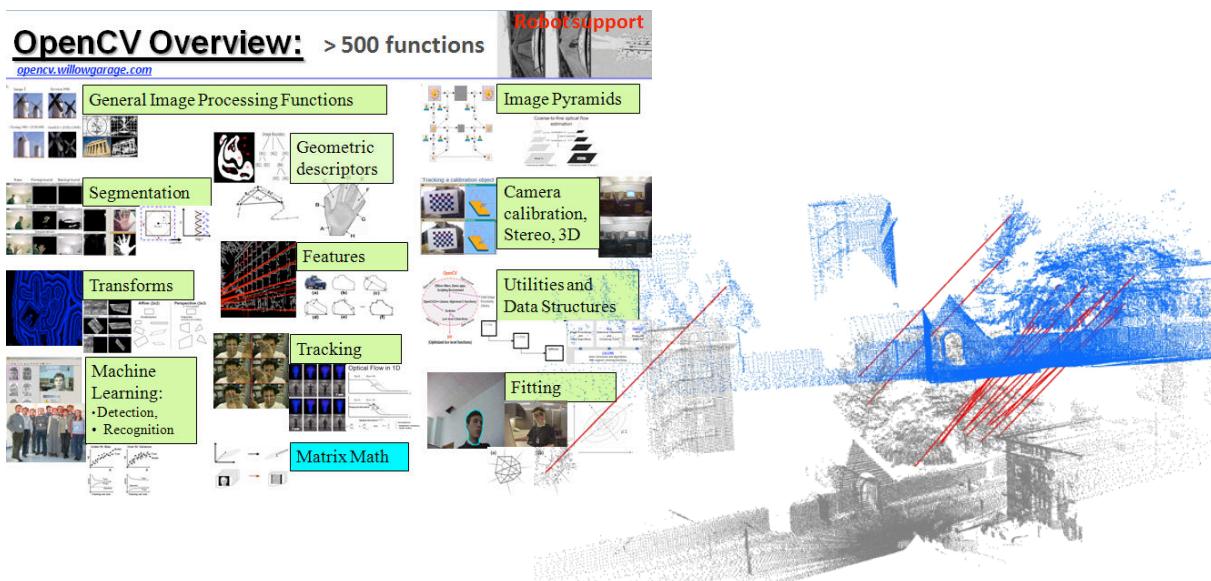
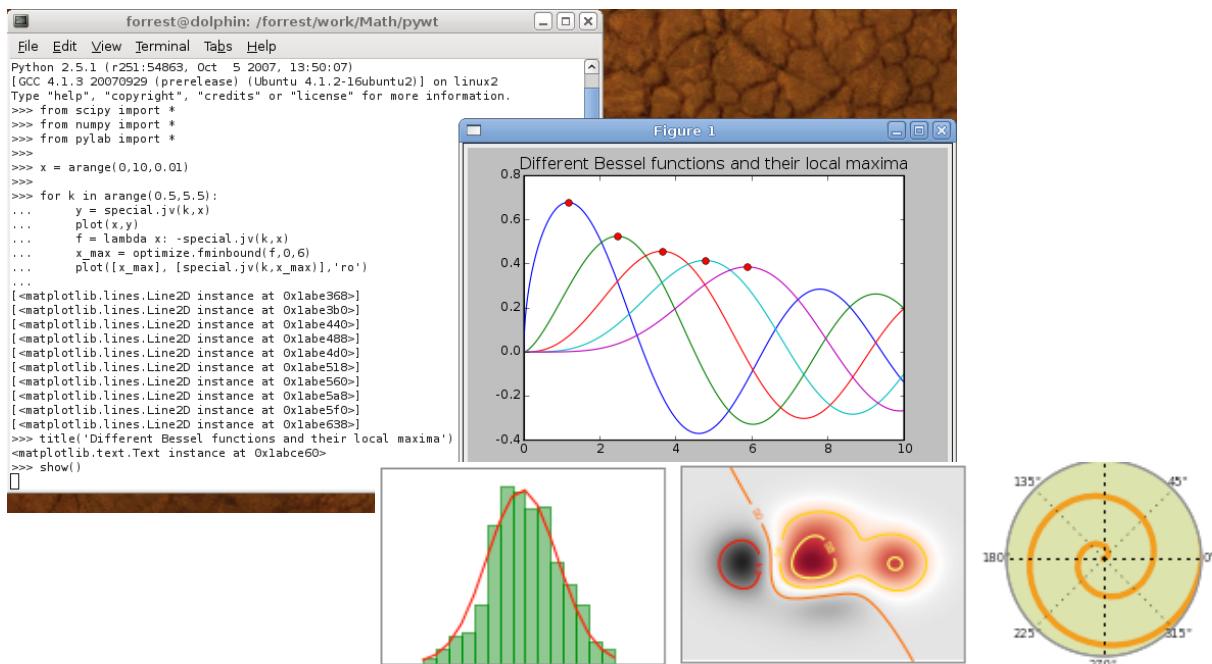
Bibliografía

- [Computer Vision: Algorithms and Applications de Szelinsky](#)
- [Learning OpenCV](#)
- [Python Scientific Lecture Notes](#)



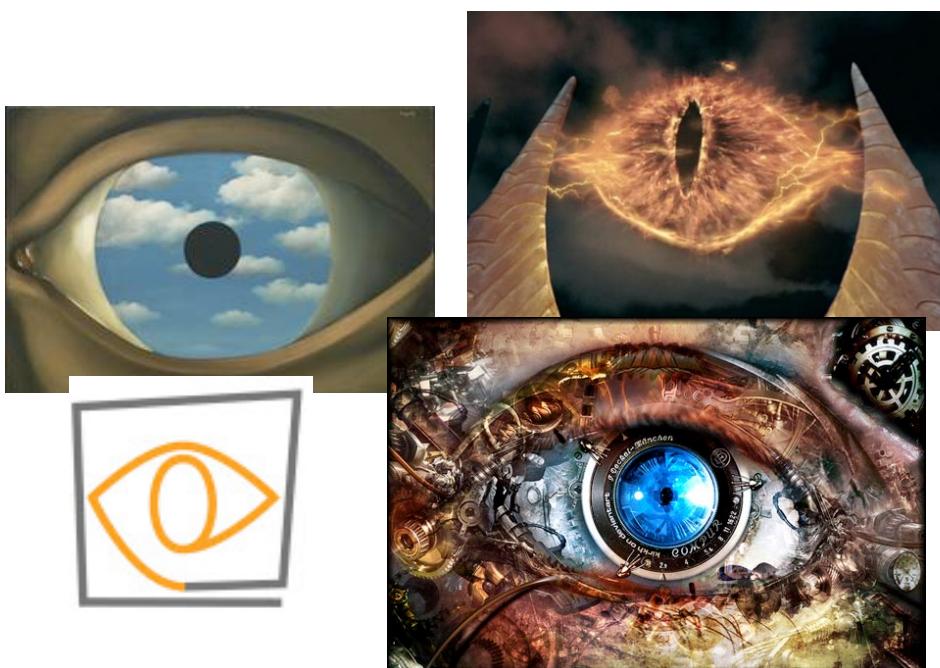
Software

- Python + Numpy/Scipy + Matplotlib
- OpenCV (interfaz Python)
- PCL (Point Cloud Library)



Transparencias del curso

- Las transparencias del curso están basadas en los materiales disponibles en las webs de los siguientes cursos
 - [Svetlana Lazebnik](#), University of North Carolina at Chapel Hill, [Computer Vision](#)
 - [William T. Freeman](#) y [Antonio Torralba](#), MIT, [Advances in Computer Vision](#)
 - [Steven Seitz](#), y [Rick Szelinski](#), University of Washington: Computer Vision, [Winter 2006](#) y [Spring 2008](#)
 - [Derek Hoiem](#), University of Illinois, [Computer Vision](#)
 - [Michael J. Black](#), Brown University, [Introduction to Computer Vision](#)
 - [James Hays](#), Brown University, [Introduction to Computer Vision](#)



Evaluación

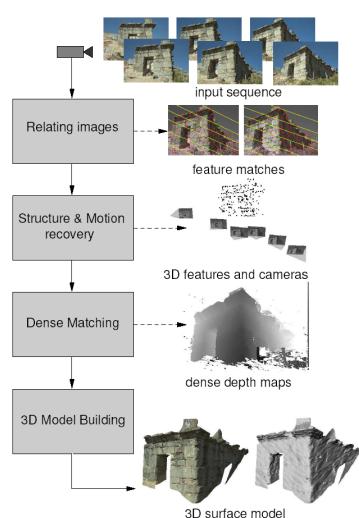
- 50% Práctica 1
- 20% Práctica 2
- 30% Presentación de un trabajo
 - Presentación de un artículo sobre alguna **aplicación práctica** que use alguna de las técnicas de la asignatura, entrando en detalles técnicos de implementación
 - Se realizará en las horas de prácticas del último día de la asignatura (miércoles 21 de diciembre)

Introducción a la Visión Artificial

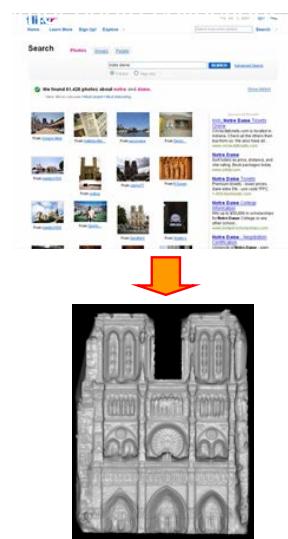
¿Qué es la visión artificial?

- El objetivo de la visión artificial es desarrollar programas que sean capaces de interpretar imágenes y vídeo, obteniendo información sobre ellas

Visión como un instrumento de medida



Pollefeys et al.



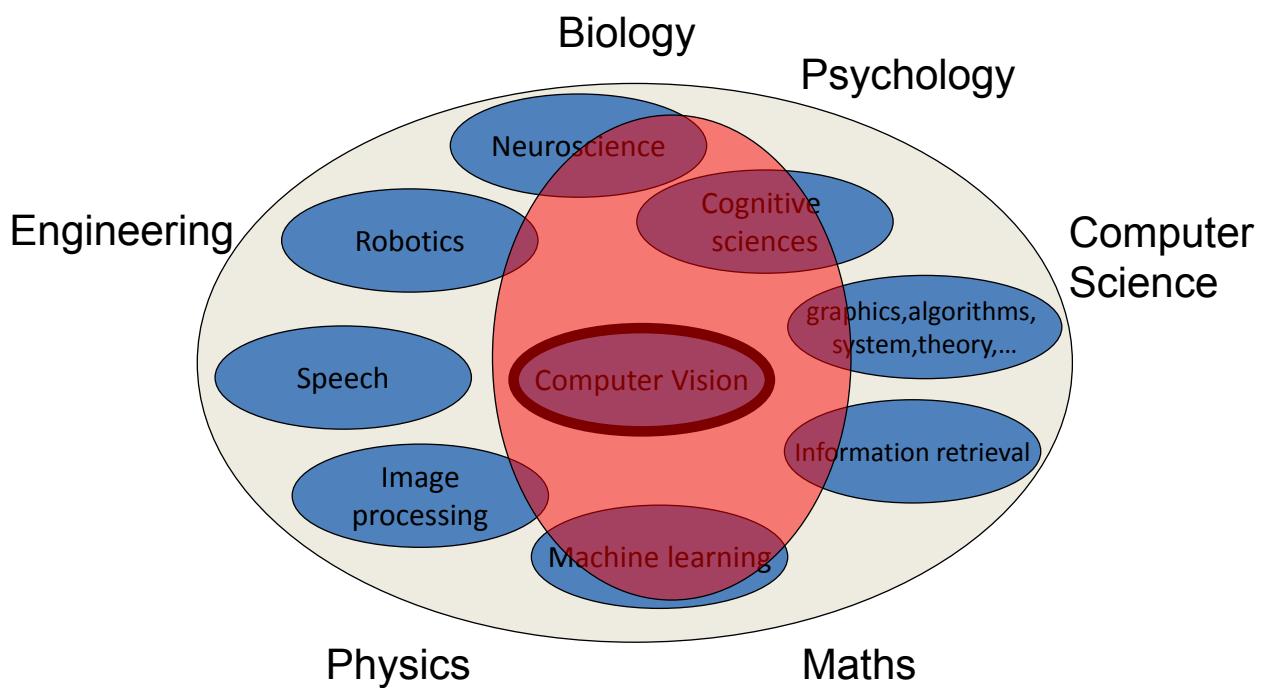
Goesele et al.

Visión como fuente de información semántica



Relaciones con otros campos

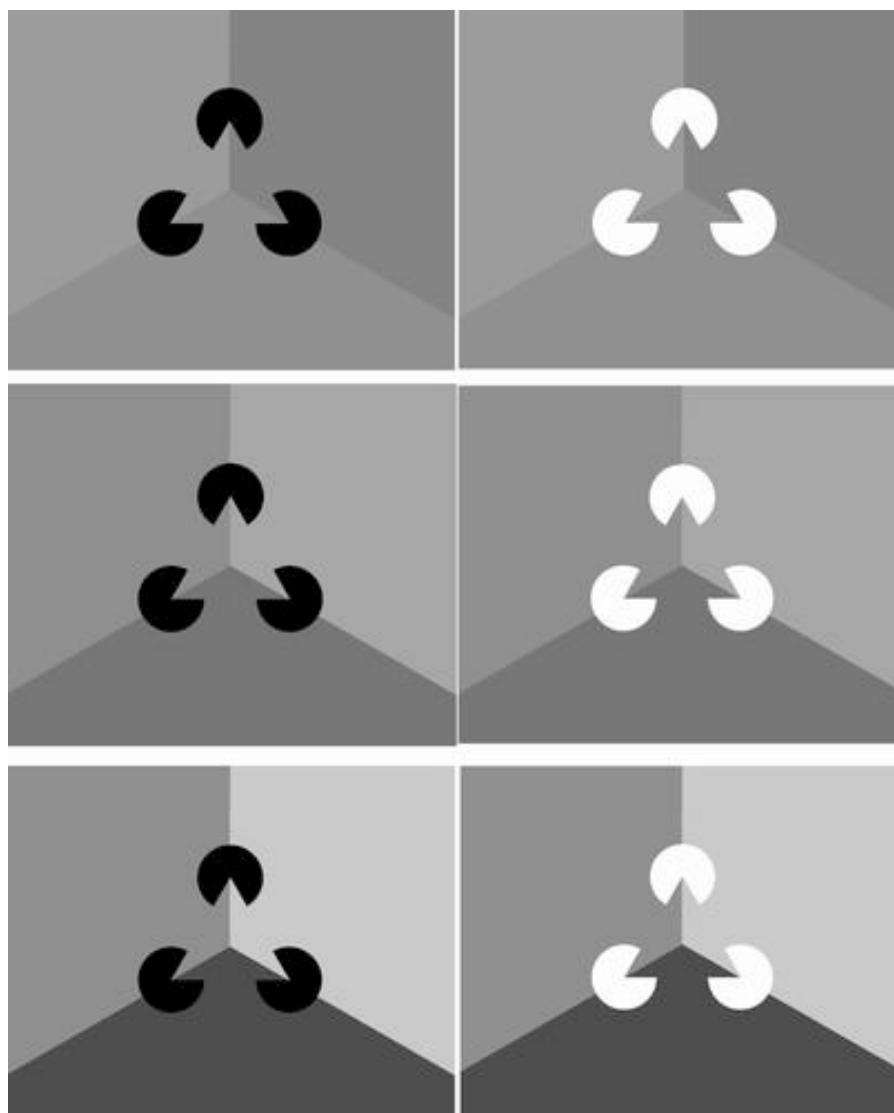
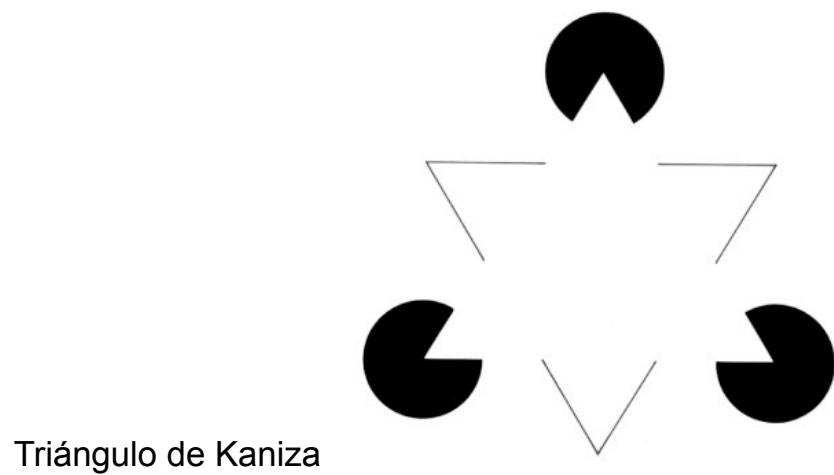
- Biología
- Psicología
- Informática
- Matemáticas
- Física
- Ingeniería



¿Por qué investigar en visión artificial?

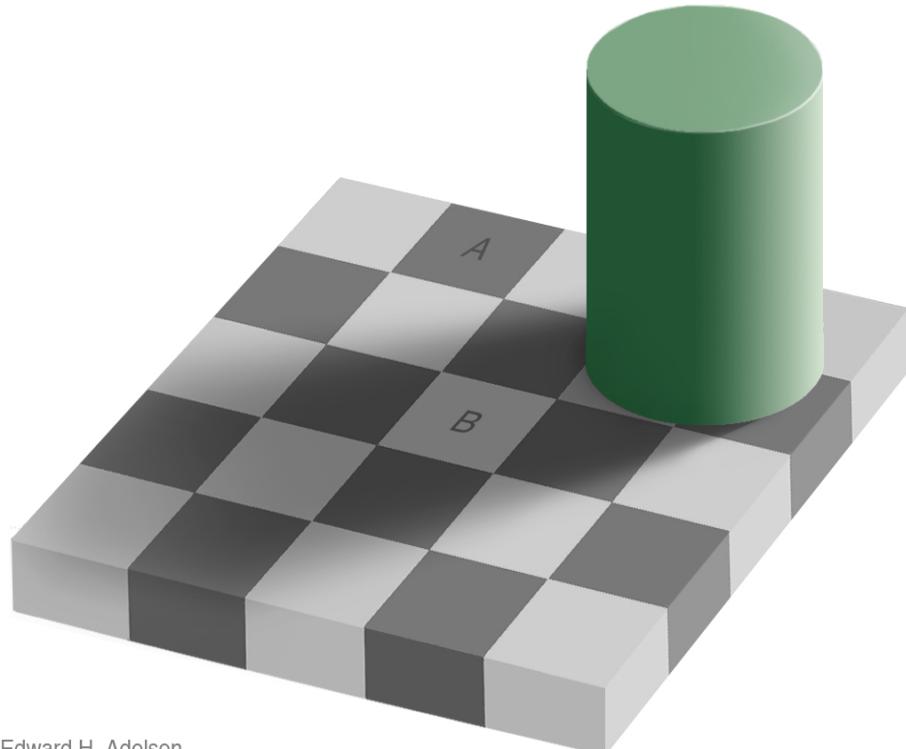
- La visión es útil
 - Aplicaciones en muchos campos de interés
 - No hace falta resolver el “problema completo” para hacer algo útil
- La visión es interesante
 - Permite distintos enfoques y técnicas
 - Resultados visuales y atractivos
- La visión es difícil
 - La mitad del cortex cerebral de los primates se utiliza para el procesamiento visual
 - La comunidad investigadora es bastante exigente para la publicación de papers en congresos y revistas

El cerebro interpreta las escenas

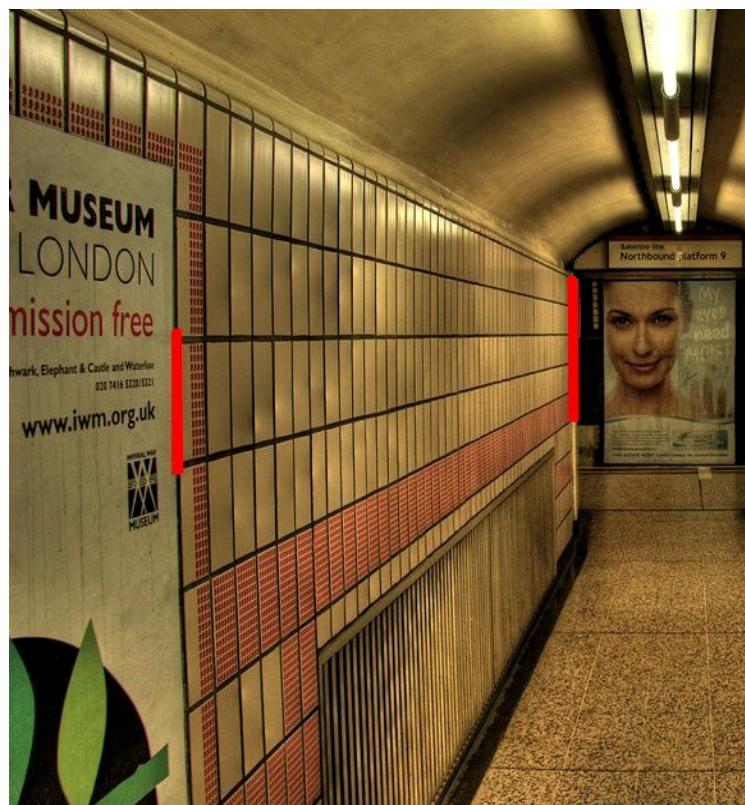


Kaniza en 3D

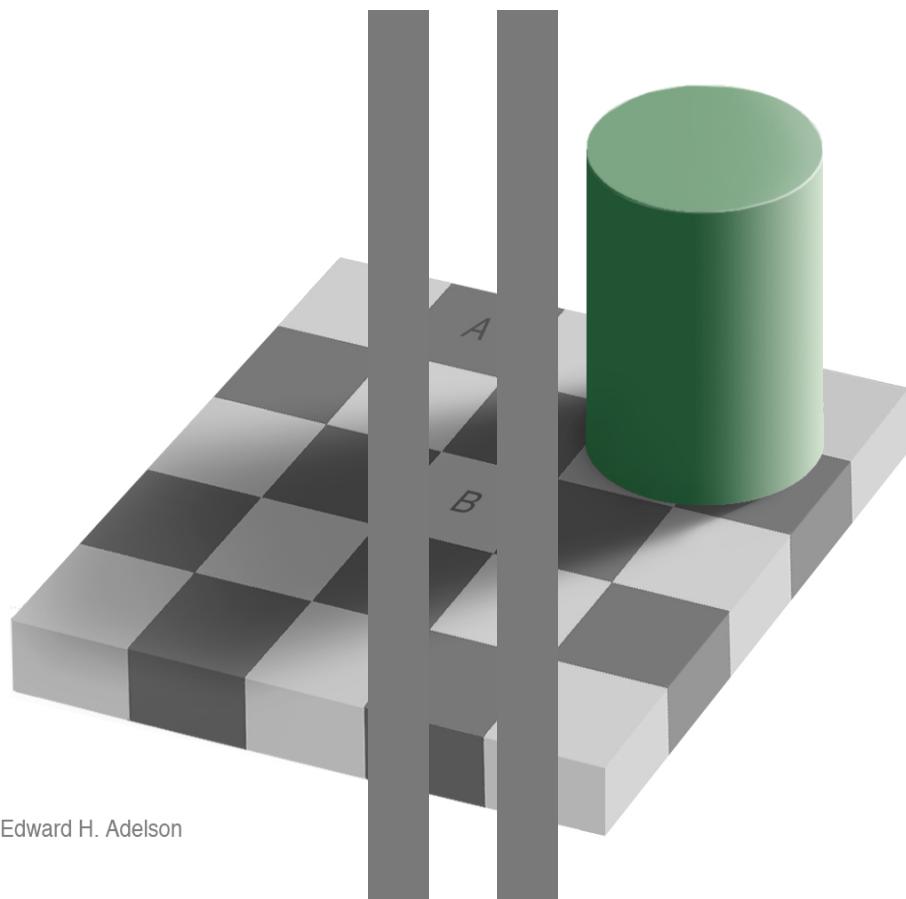
El cerebro interpreta las escenas



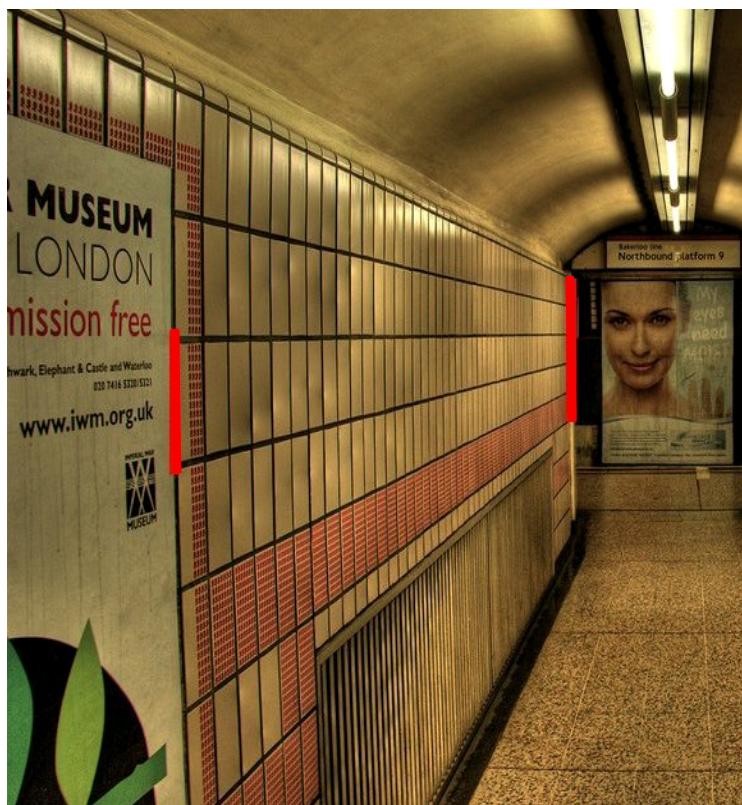
Edward H. Adelson



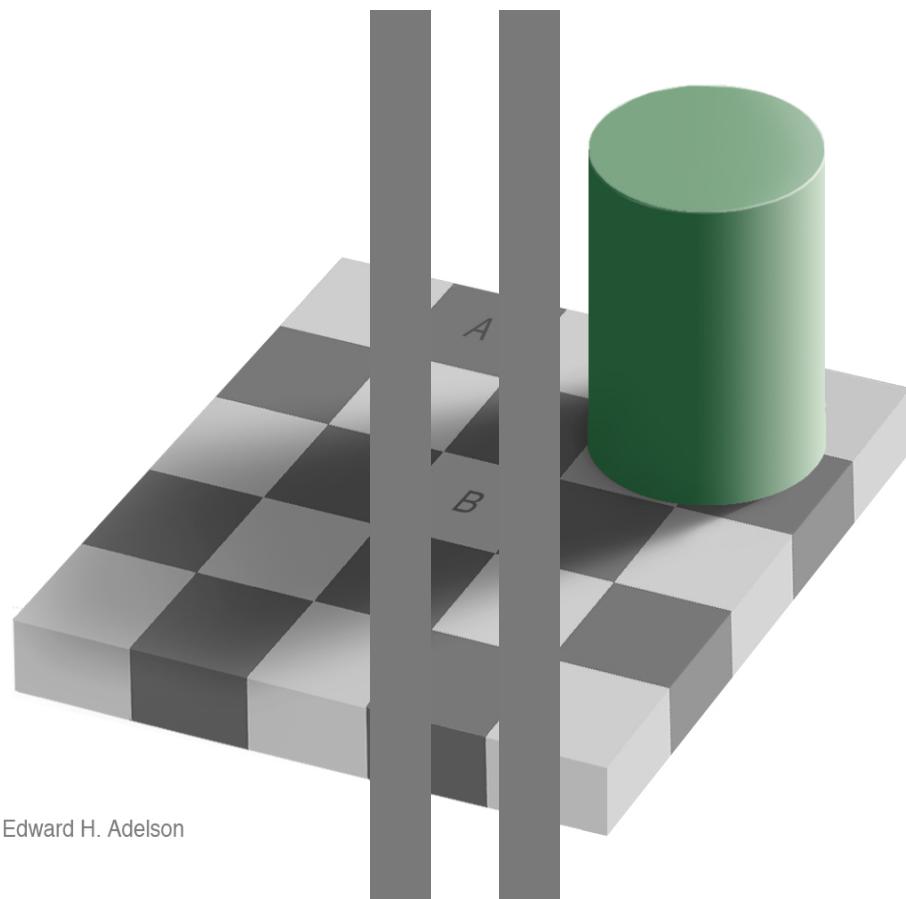
El cerebro interpreta las escenas



Edward H. Adelson



El cerebro interpreta las escenas



La percepción es ambigua

- Muchas escenas 3D diferentes pueden dar lugar a la misma escena 2D

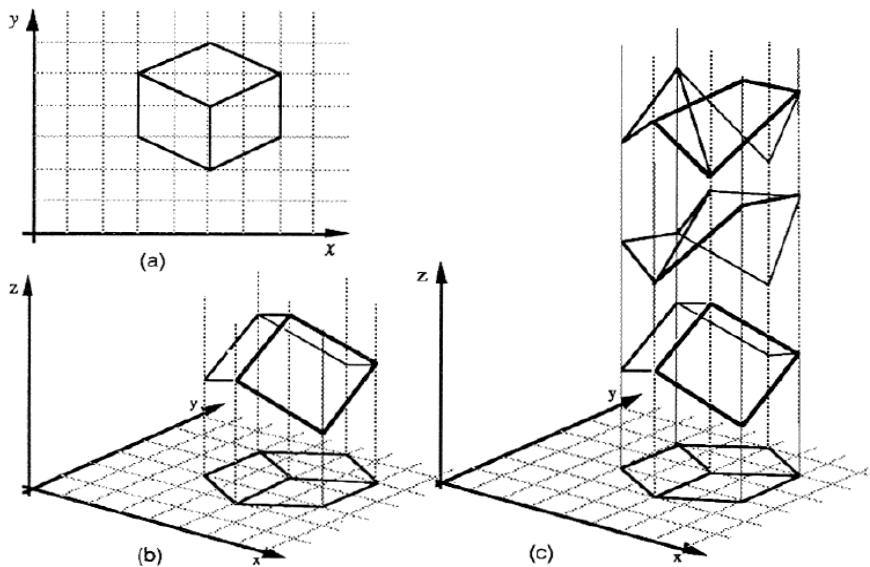
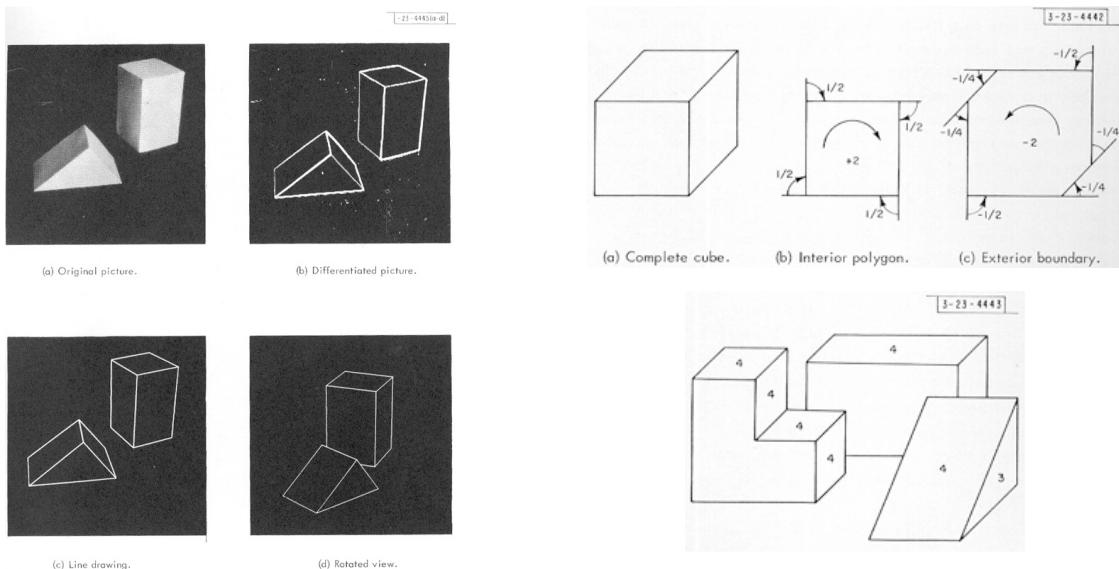


Figure 1. (a) A line drawing provides information only about the x , y coordinates of points lying along the object contours. (b) The human visual system is usually able to reconstruct an object in three dimensions given only a single 2D projection (c) Any planar line-drawing is geometrically consistent with infinitely many 3D structures.

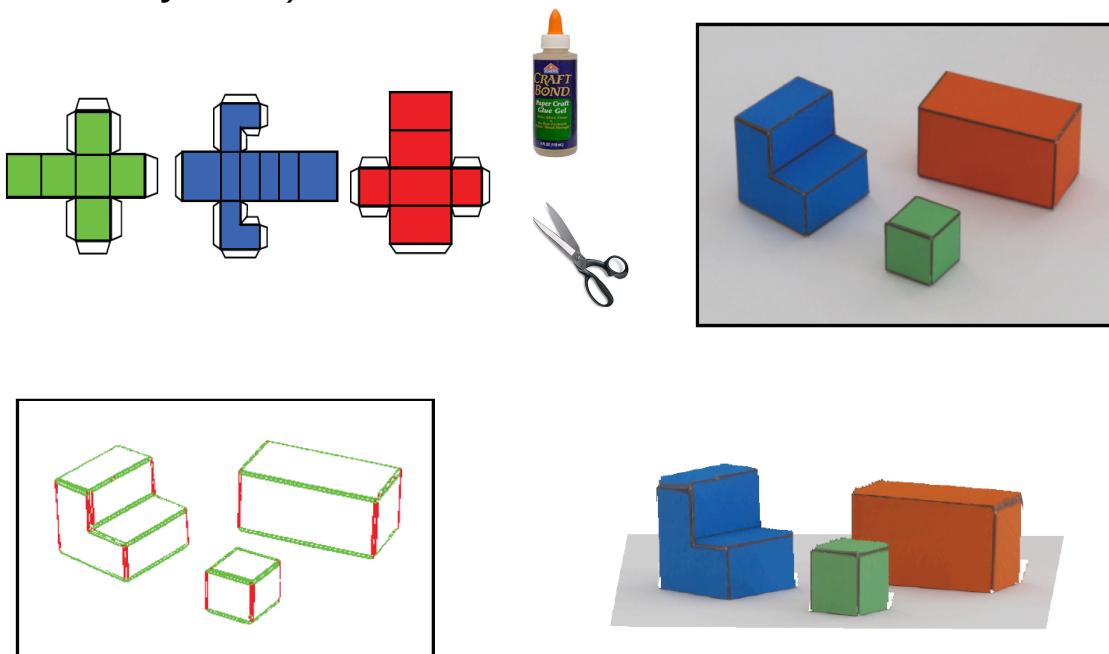
Sinha & Adelson 93

Orígenes de la visión artificial

- Lawrence G. Roberts (Premio Príncipe de Asturias 2002), [Machine Perception of Three Dimensional Solids](#), Ph.D. thesis, MIT Department of Electrical Engineering, 1963 ([PDF original](#))

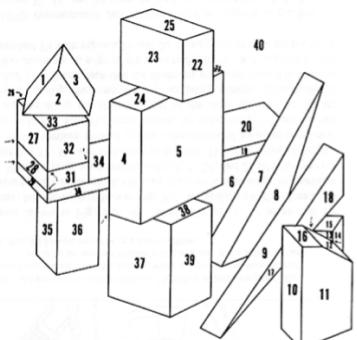


- Hecho realidad en las [notas de clase](#) del curso [6.869: Advances in Computer Vision](#) del MIT (*A Simple Vision System*)

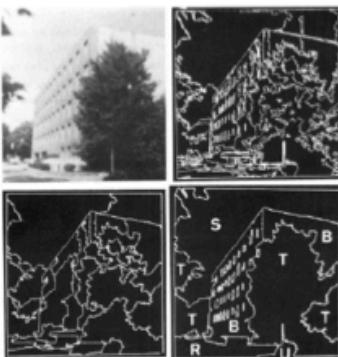


Una historia rápida de la visión artificial

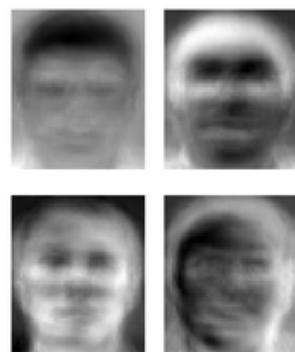
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts



Guzman '68



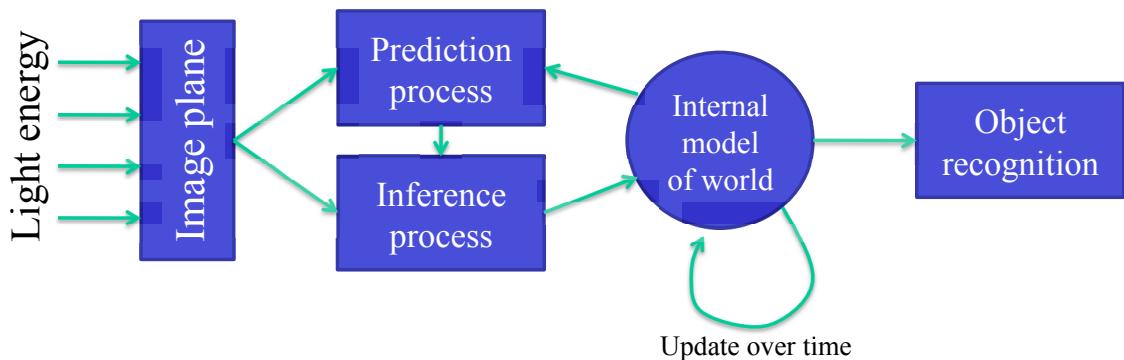
Ohta Kanade '78



Turk and Pentland '91

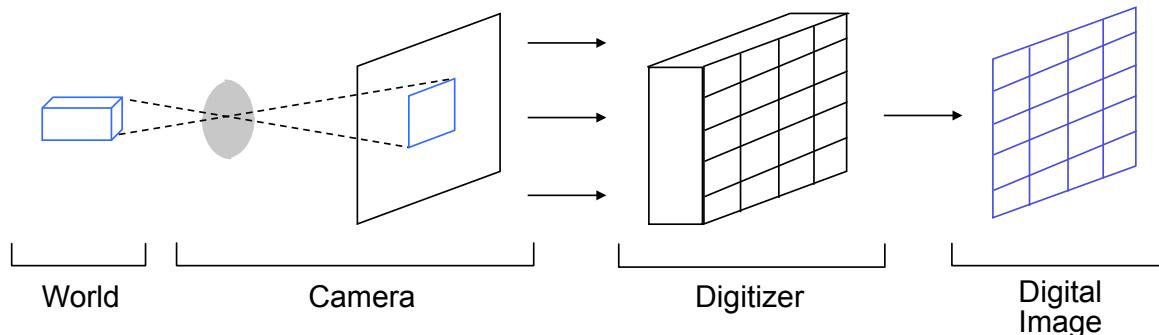
Definición más formal

- Necesitamos una definición más formal, que tenga en cuenta los aspectos físicos, computacionales y matemáticos
 - ¿Qué propiedades (pistas) del mundo visual podemos extraer o medir?
 - ¿Cómo podemos usar nuestro conocimiento (a priori) sobre el mundo para interpretarlo?



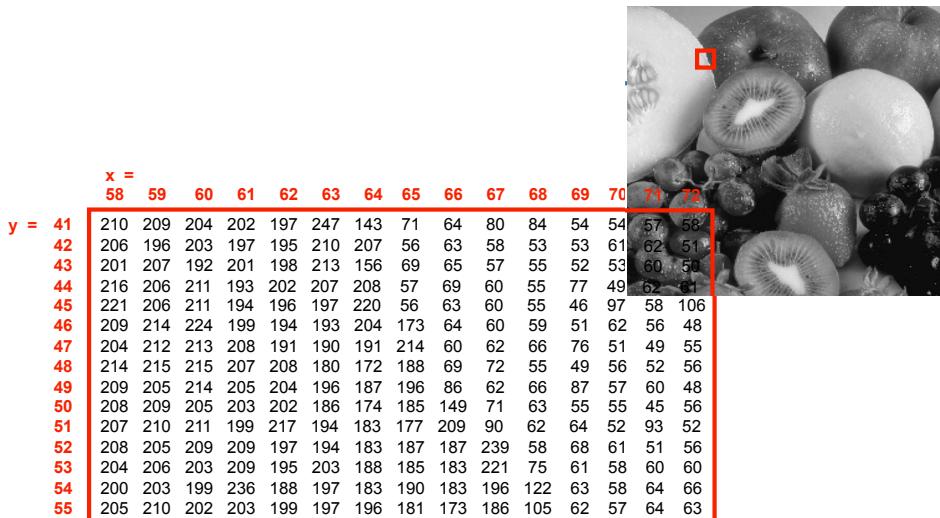
(Michael J. Black)

- Formación de la imagen

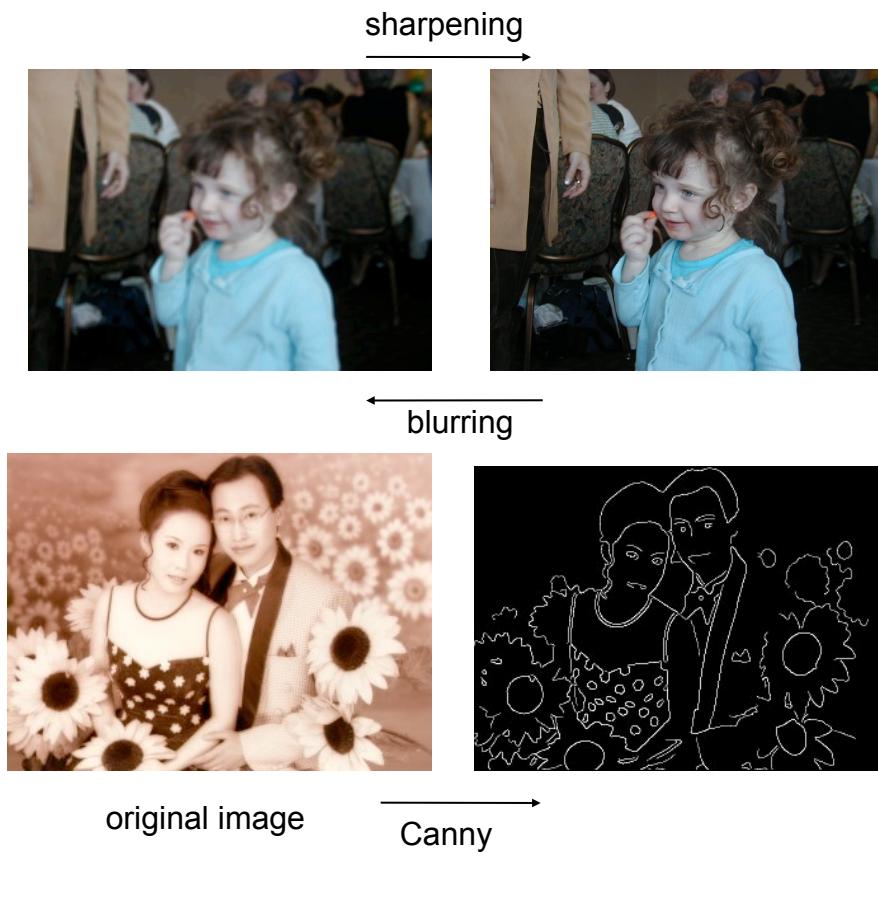


Técnicas y niveles

- Imágenes



- Bajo nivel: imágenes a imágenes



(Linda Shapiro,
Computer vision,
Washington Univ)

Técnicas y niveles

- Nivel medio: imágenes a características



edge image

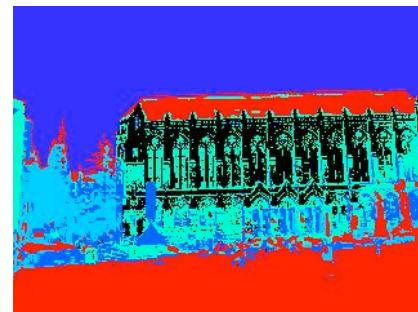


circular arcs and line segments

data
structure



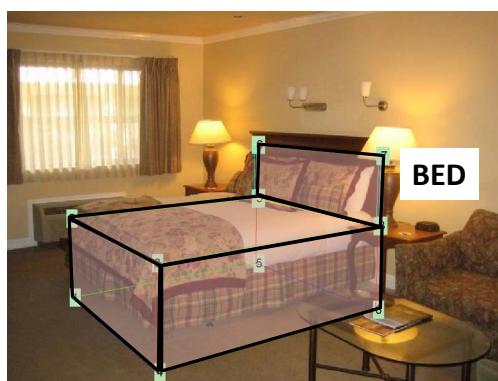
original color image



K-means
clustering

regions of homogeneous color

- Alto nivel: procesamiento de características



(Hoeim)

La visión artificial en funcionamiento

- Modelado 3D

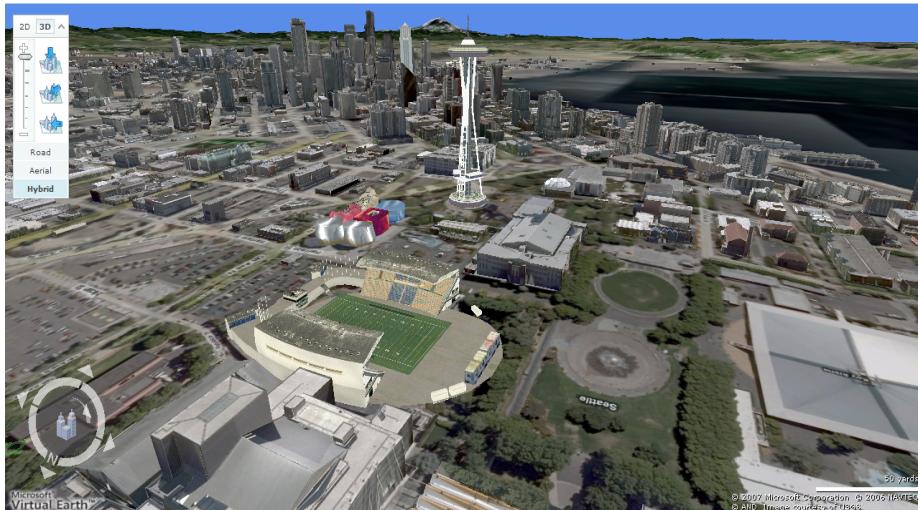


Image from Microsoft's [Virtual Earth](#)
(see also: [Google Earth](#))

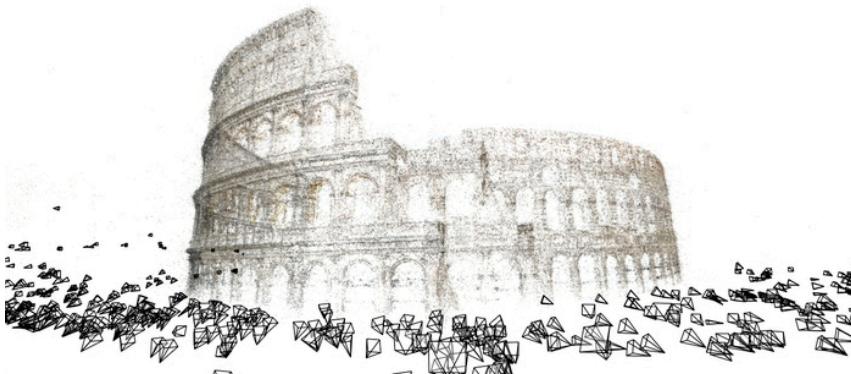
- Photosynth.net



Based on [Photo Tourism](#)
by Noah Snavely, Steve Seitz, and Rick Szeliski
<http://photosynth.net/>

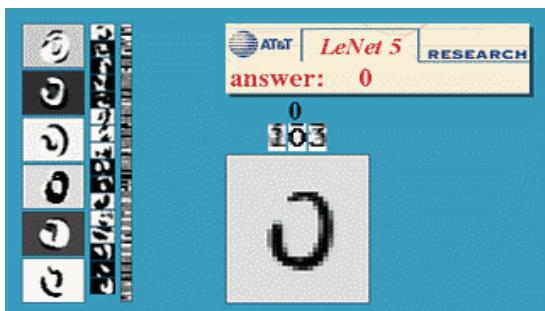
La visión artificial en funcionamiento

- 3D a partir de miles de imágenes



[Building Rome in a Day: Agarwal et al. 2009](#)

- Reconocimiento de caracteres



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



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

La visión artificial en funcionamiento

- Detección y reconocimiento de caras



Cámaras digitales



iPhoto de Apple

- Efectos especiales



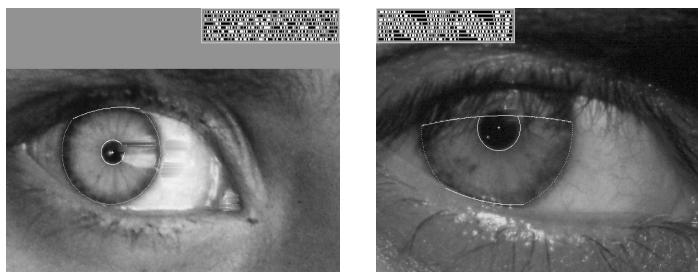
Lord of the Rings,
WETA Digital



Pirates of the Caribbean,
Industrial Light and Magic

La visión artificial en funcionamiento

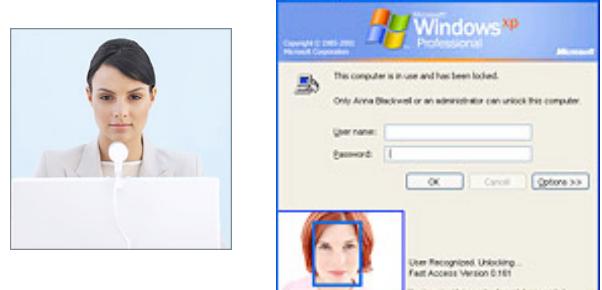
- Datos biométricos a partir de imágenes ([John Daugman, 2001 -paper sobre reconocimiento basado en el iris-](#))



- Sistemas de seguridad



Fingerprint scanners on many new laptops, other devices



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

La visión artificial en funcionamiento

- Retransmisiones deportivas



Tracking de objetos

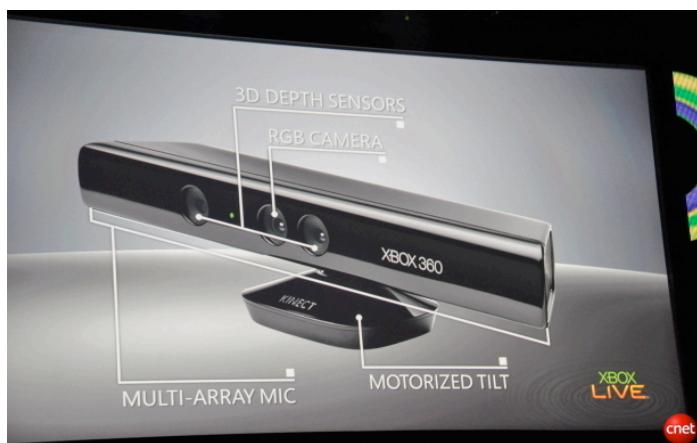


Sportvision first down line

Nice [explanation](#) on www.howstuffworks.com

<http://www.sportvision.com/video.html>

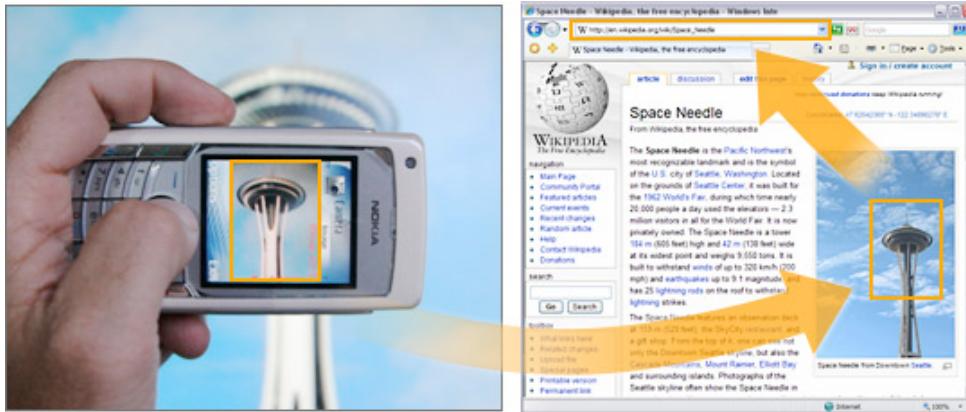
- Juegos interactivos: Kinect



- [Object Recognition](#)
- [3D](#)
- [Robot](#)

La visión artificial en funcionamiento

- Reconocimiento de lugares en dispositivos móviles



Point & Find, Nokia, Google Goggles

- Conducción autónoma (artículo del [New York Times](#))

The New York Times
October 10, 2010

Autonomous Driving

Google's modified Toyota Prius uses an array of sensors to navigate public roads without a human driver. Other components, not shown, include a GPS receiver and an inertial motion sensor.

LIDAR
A rotating sensor on the roof scans more than 200 feet in all directions to generate a precise three-dimensional map of the car's surroundings.

VIDEO CAMERA
A camera mounted near the rear-view mirror detects traffic lights and helps the car's onboard computers recognize moving obstacles like pedestrians and bicyclists.

RADAR
Four standard automotive radar sensors, three in front and one in the rear, help determine the positions of distant objects.

POSITION ESTIMATOR
A sensor mounted on the left rear wheel measures small movements made by the car and helps to accurately locate its position on the map.

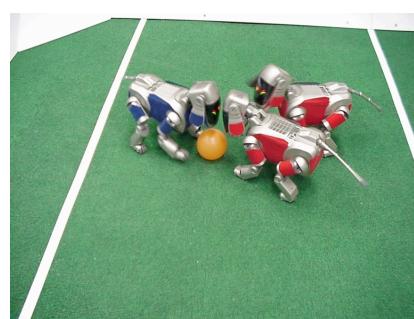
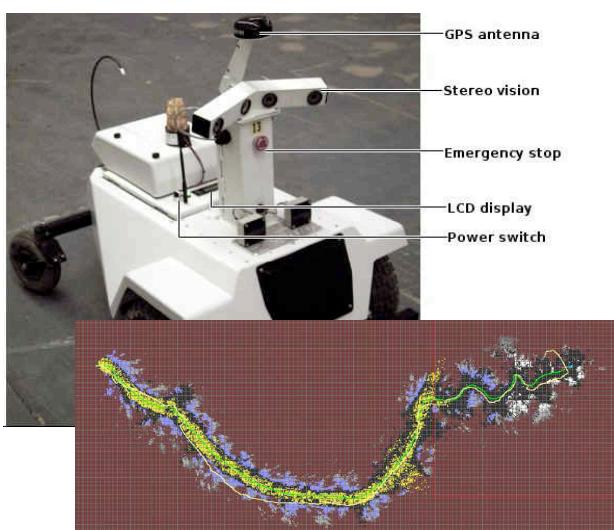
Source: Google
THE NEW YORK TIMES; PHOTOGRAPHS BY RAMIN RAHIMIAN FOR THE NEW YORK TIMES

La visión artificial en funcionamiento



[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.
- Robots móviles

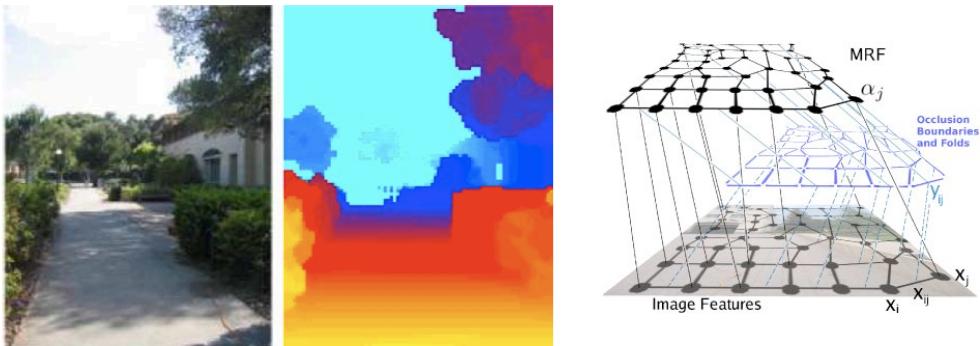


<http://www.robocup.org/>

[Kurt Konolige](#)

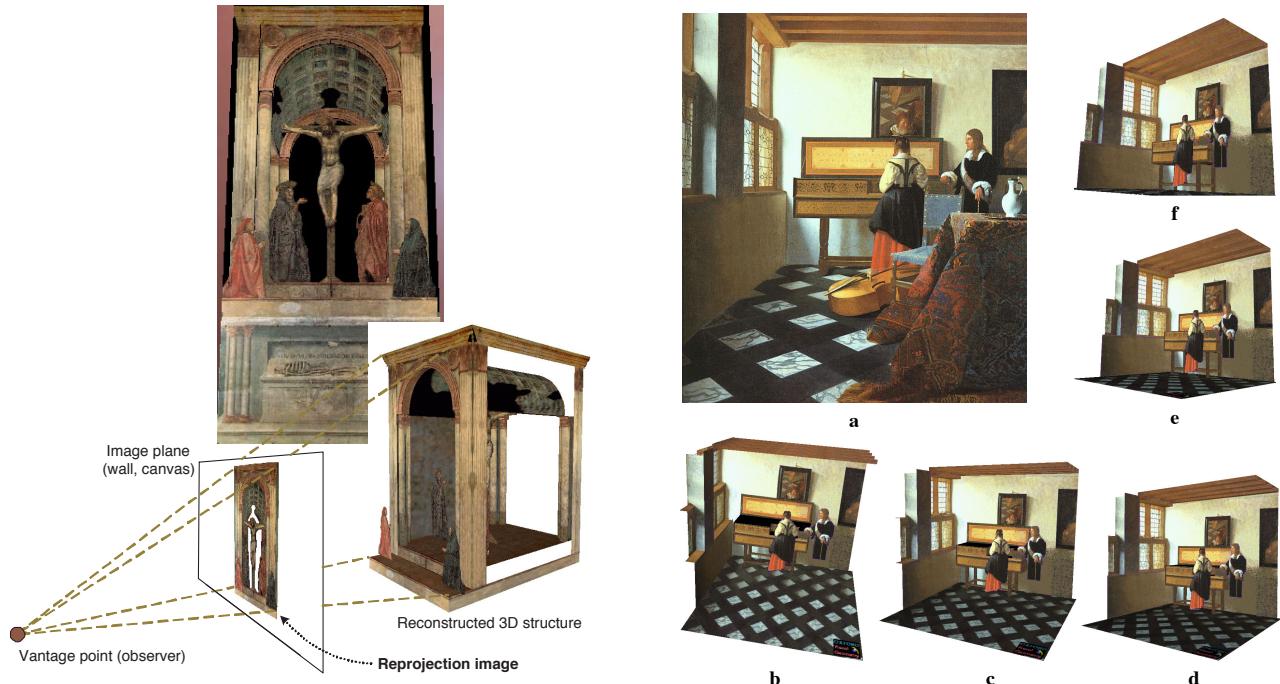
Dos ejemplos de investigación

- Profundidad a partir de una única imagen. Ashutosh Saxena, Sung H. Chung, Andrew Y. Ng, Make 3D ([IJCV Paper, 2007](#))



<http://make3d.cs.cornell.edu/>

- Bringing pictorial space to life. Antonio Criminisi, Martin Kemp, Andrew Zisserman (2002 [paper](#), [vídeo](#))



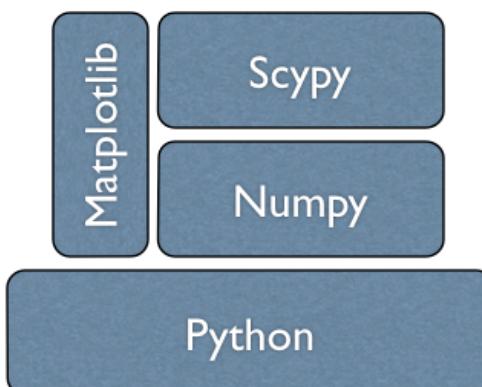
Estado del arte actual

- Hemos visto ejemplos de sistemas de los últimos años
- Vamos a ver cada vez más aplicaciones en dispositivos cada vez más variados
 - Explosión de aplicaciones en dispositivos móviles
- El área de investigación es muy dinámica
- Para saber más acerca de aplicaciones y compañías de visión:
 - [David Lowe](#) mantiene una excelente página con compañías y aplicaciones de visión: <http://www.cs.ubc.ca/spider/lowe/vision.html>

Python, Scipy/Numpy, OpenCV

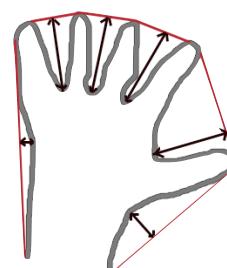
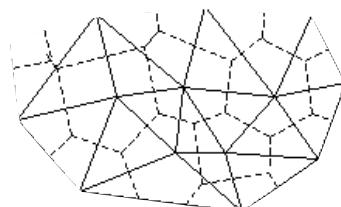
Software científico

- Problemas de Matlab
- Python: lenguaje de programación muy extendido en áreas científicas
 - Lenguaje de script
 - Interpretado
 - Débilmente tipado
 - Prototipado rápido, no es muy eficiente
- Módulos de computación científicos en Python:
 - Numpy: arrays y operaciones con matrices
 - Scipy: librerías de alto nivel
 - Matplotlib: visualización, gráficas
 - Open Source y multi-plataforma
 - Librerías bastante extendidas y probadas en la comunidad científica
- Referencias
 - Documentación de Numpy/Scipy
 - Blog



OpenCV

- Historia
 - Librería desarrollada inicialmente por Intel
 - Mantenido en la actualidad por la empresa de robótica [Willow Garage](#)
- Características principales
 - Proporciona una librería open source bien testeada y optimizada para procesamiento de imagen, visión y operaciones geométricas
 - Escrita en C, asegurándose que la implementación es rápida y portable
 - Compilado para múltiples plataformas, incluyendo plataformas embebidas y GPUs
 - Wrappers para distintos lenguajes, incluyendo python
- Referencias
 - [Wiki de OpenCV](#): Wiki de OpenCV
 - [OpenCV 2.1 Python Reference](#): Manual de referencia de la interfaz Python en WillowGarage



Referencias

- [Computer Vision: Algorithms and Applications de Szelinsky](#): cap. 1
- [Learning OpenCV](#): cap. 1 y 2
- [Python Scientific Lecture Notes](#): cap. 1