



# Computer Vision

Petia Ivanova Radeva

[petia.ivanova@ub.edu](mailto:petia.ivanova@ub.edu)

2022-2023

You can find these slides in Campus Virtual

## Theory

**Petia Ivanova Radeva - Wednesday, 15:00-17:00h, B7**  
[petia.ivanova@ub.edu](mailto:petia.ivanova@ub.edu)

---

## Laboratory

Petia Radeva – grup F, Wednesday, 17:00-19:00, IF  
Eloy García – grup A, Wednesday, 17:00-19:00, ID  
Petia Radeva – grup B, Thursday, 17:00-19:00, ID

**Slides, laboratories, deliveries, forum and announcements**

# Today:

- 1. What is Artificial Vision?**
  - 2. A little bit of history**
  - 3. Main problems of Artificial Vision**
  - 4. Difficulties of the Artificial Vision**
  - 5. Applications**
  - 6. Subject organization**
- 
- **Bibliography:**
    - Szeliski, CV: A&A, Ch 1.0 (Introduction)

# Can you give an example of Computer Vision applications?



**Pilot.ai**

**Location:** Palo Alto, CA

**Employees:** 30

**Differentiator:**  
End-to-end  
Powerhouse,  
both Edge and  
Cloud

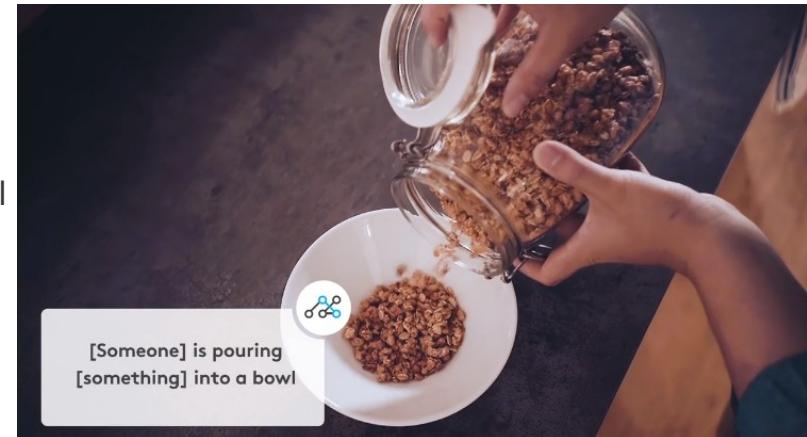
**20 billion neurons**

**Location:** Berlin, Germany

**Employees:** 25

**Differentiator:** Understanding human physical expression

**5 Breakthrough Computer Vision Startups to Watch in 2019**



Martes, 15 de Septiembre de 2021  
ESSEGRI | DIARIO DIGITAL | SUSCRIPCIONES | CLUB SEGRE | EBOLOGA

ÚLTIMAS NOTICIAS | LLEIDA | COMARCAS | MÁS NOTICIAS | SERVICIOS | OPINIÓN | LLEIDA TV | CÍRCLES

¿Quieres cambiar de casa? La tenemos al mejor precio | Ver viviendas >>> | Solvia

Panorama | TECNOLOGÍA LABORAL

## Microsoft abre en Barcelona un centro de investigación

AGENCIAS | BARCELONA Actualizada 14/09/2021 a las 08:15

Especializado en aplicaciones de Inteligencia Artificial || Contratará inicialmente a 30 ingenieros y científicos de datos

LEVERAGE A.I. BUSINESS

VENTE A ENDESA Y LLÉVATE 120€ POR CONTRATAR LA LUZ Y EL GAS.

DESCUBRE MÁS | Ver condiciones

METROPOLI

Ver contenido en catalán | [Ver contenido en catalán](#)

Interior de un Apple Store con el logo de la compañía / BCCL

## Apple compra una startup barcelonesa por 42 millones de euros

La compañía californiana ha adquirido Vilynx, especializada en el desarrollo de 'software'

METROPOLI

En el distrito 22@ se concentra el mayor número de empresas tecnológicas y digitales de Barcelona / EFE

## Apple convierte a Barcelona en un 'hub' en inteligencia artificial

La apuesta del gigante tecnológico norteamericano podría atraer a nuevos grandes inversores



SOLUTIONS | USE CASES | CASE STUDIES | ABOUT | BLOG

CONT

Restb.ai selected in the top 5 AI Startups by Nvidia.

October 13, 2017 in [Press Releases](#)

Nvidia highlights Restb.ai at their "GTC Europe Inception Awards"

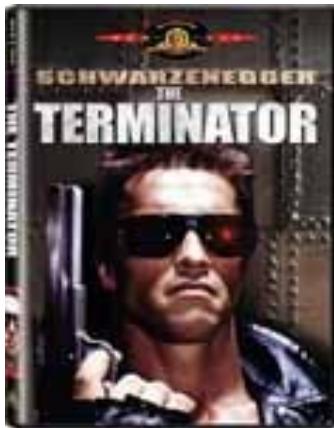
Barcelona, Spain – October 13th, 2017 – In Munich, Restb.ai was selected this week by Nvidia as one of the top 5 hottest AI startups in Europe.

The nominees were selected from among the 700 European startups already chosen to participate in [Nvidia's Inception program](#), which accelerates the development of startups involved in AI and deep learning. These are early-stage companies.



# Computer Vision and the Movie Industry

The Terminator  
(1984, USA)



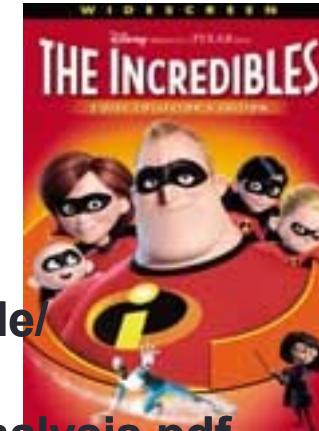
Minority report  
(2002, USA)



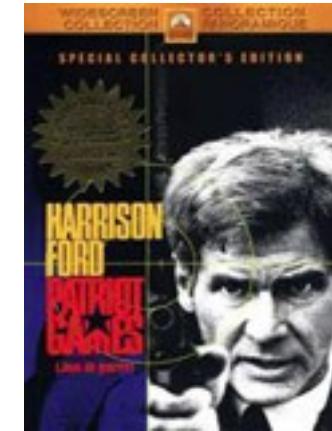
Charlie's Angels  
(2000, USA)



The Incredibles  
(2004, USA)



Patriot Games  
(1992, USA)



[https://epub.uni-regensburg.de/  
Schmit\\_EADH2021\\_MovieAnalysis.pdf](https://epub.uni-regensburg.de/Schmit_EADH2021_MovieAnalysis.pdf)

OCR object recognition, tracking and improving vision.

Gesture recognition

Iris and fingerprint recognition

Satellite images of very high resolution are used for recovering people and vehicles.

Face recognition

Person of interest  
(2007-2013, USA)



Robot and Frank(2012, USA)



*Ex Machina*  
*Blade Runner 2049*  
*A.I. Artificial Intelligence*  
*Bicentennial Man*  
*Morgan*  
*Automata*  
*Transcendence*  
*Her*  
*Moneyball*  
*The Imitation Game*

# Computer Vision and Visual Effects



**THE CHEMICAL BROTHERS**

**Wide Open**

50.571.420 visualizaciones 25 ene 2016 Best of

ChemicalBrothers: <https://goo.gl/JxwZyH>



**“Avatar” 2009**

# What is Artificial vision?

Computer vision is part of Artificial Intelligence.

It concerns with **modeling and replicating human vision** using computer software and hardware.

Formally, computer vision is a discipline that studies:

- how to perceive i.e. analyze, reconstruct and understand
- a 3D/4D scene from its 2D images or videos
- in terms of the properties of the structures and phenomena present in the scene.



Terminator 2

# Computer Vision

Make computers understand images and video.



- **What kind of scene?**
- **Where are the cars?**
- **How far is the building?**
- ...

# Vision is really hard

- Vision is an amazing feature of natural intelligence
  - More human brain devoted to vision than anything else

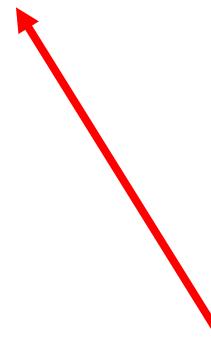


# Measurement vs. Perception



# Vision: a split personality

- What does it mean, to see? The plain man's answer (and Aristotle's, too), would be, to know what is where by looking.
- “Vision is the process of discovering from images what is present in the world, and where it is.” Marr



- Answer #1: *pixel of brightness 243 at position (124,54) ...and depth .7 meters*      **depth map**
- Answer #2: *looks like bottom edge of whiteboard showing at the top of the image*
- **Which do we want?**
  - Is the difference just a matter of scale?

# What is Computer Vision?

- Vision is about discovering from images what is present in the scene and where it is.
- In Computer Vision a **camera** (or several cameras) is linked to a **computer**.
  - The computer **interprets images** of a real scene to obtain information useful for tasks such as navigation, manipulation and recognition.

# The problem

- Want to make a computer understand images
- We know it is possible – we do it effortlessly!

- Real world scene



- Sensing device

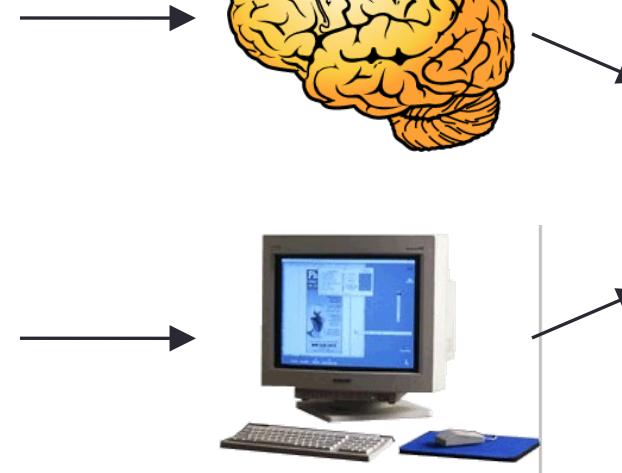


- Interpreting device



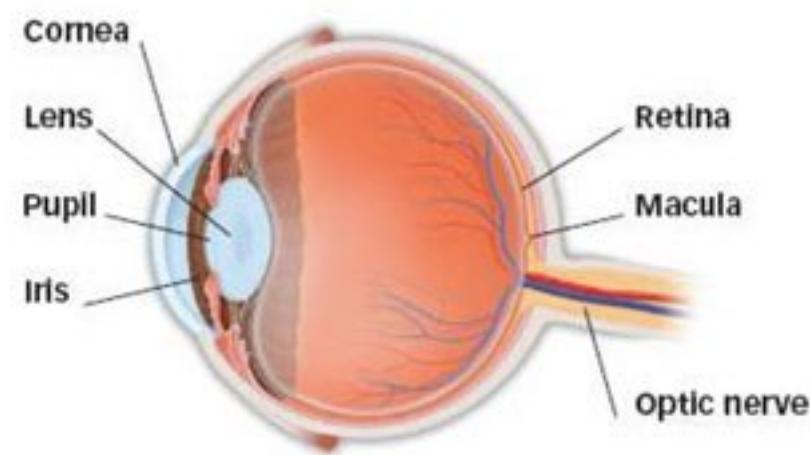
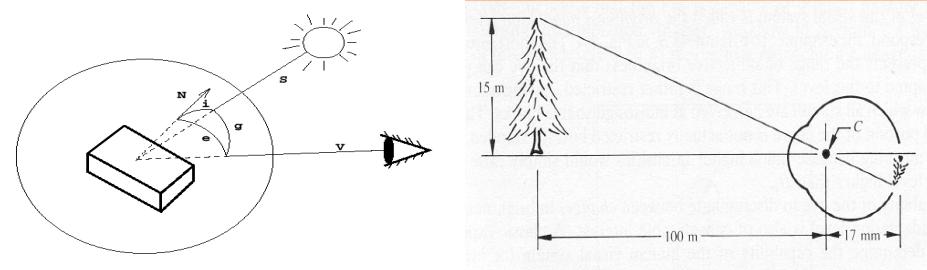
- Interpretation

- Nature
- A butterfly on rocks
- Strange color painted stones
- An artistic image



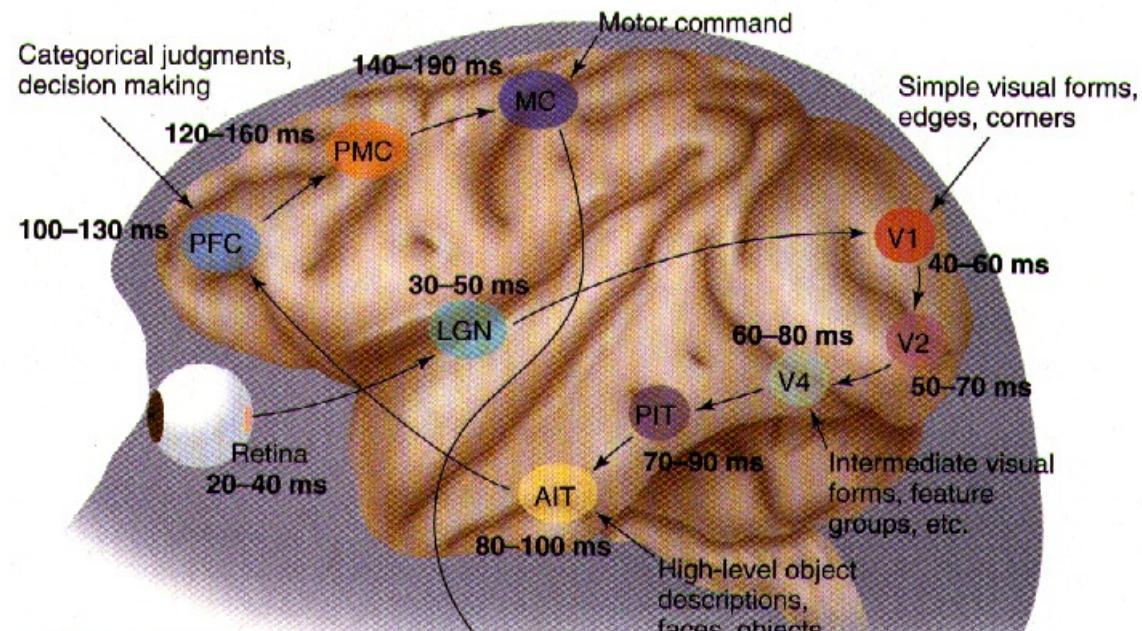
# The Human Eye

- Retina measures about  $5 \times 5$  cm and contains  $10^8$  sampling elements (rods and cones).
- The eye's spatial resolution is about  $0.01^\circ$  over a  $150^\circ$  field of view.
- Intensity resolution is about 11 bits/element, spectral range is 400–700nm.
- Temporal resolution is about 100 ms (10 Hz).
- Two eyes give a data rate of about 3 GB!



# Human visual system

- Vision is the most powerful of our own senses.



- [Thorpe et. al.]
- Around 1/3 of our brain is devoted to processing the signals from our eyes.
- The visual cortex has around  $O(10^{11})$  neurons.

# Why don't we just copy the human visual system?

- People try to, but we don't yet have a sufficient understanding of how our visual system works.
- $O(10^{11})$  neurons used in vision
- By contrast, CPUs have  $O(10^6)$  transistors (most are cache memory).
- GPUs – a great news! They have billions ( $10^9$ ) of transistors.
- Very different architectures:
  - Brain is slow, but parallel.
  - Computer is fast, but mainly serial.
- Bird vs Airplane
  - Same underlying principles.
  - Very different hardware.



# Why to study it?

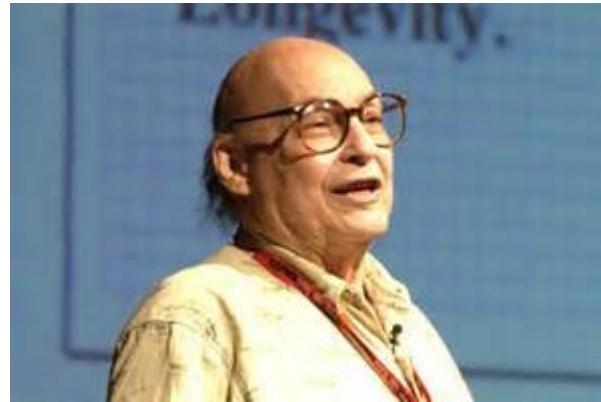
- Gain insight into how we see
  - Vision is explored extensively by neuroscientists to gain an understanding of how the brain operates
- Replicate human vision to allow a machine to see:
  - Central to that problem of Artificial Intelligence
  - Many industrial and real life human-centric applications

# Today:

1. What is Artificial Vision?
  2. A little bit of history
  3. Main problems of Artificial Vision
  4. Difficulties of the Artificial Vision
  5. Applications
  6. Subject organization
- 
- Bibliography:
    - Szeliski, CV: A&A, Ch 1.0 (Introduction)

# The Vision Story Begins...

- In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”.



- We now know that the problem is slightly more difficult than that.  
(Szeliski 2009, Computer Vision)

# Mar's Historical definition

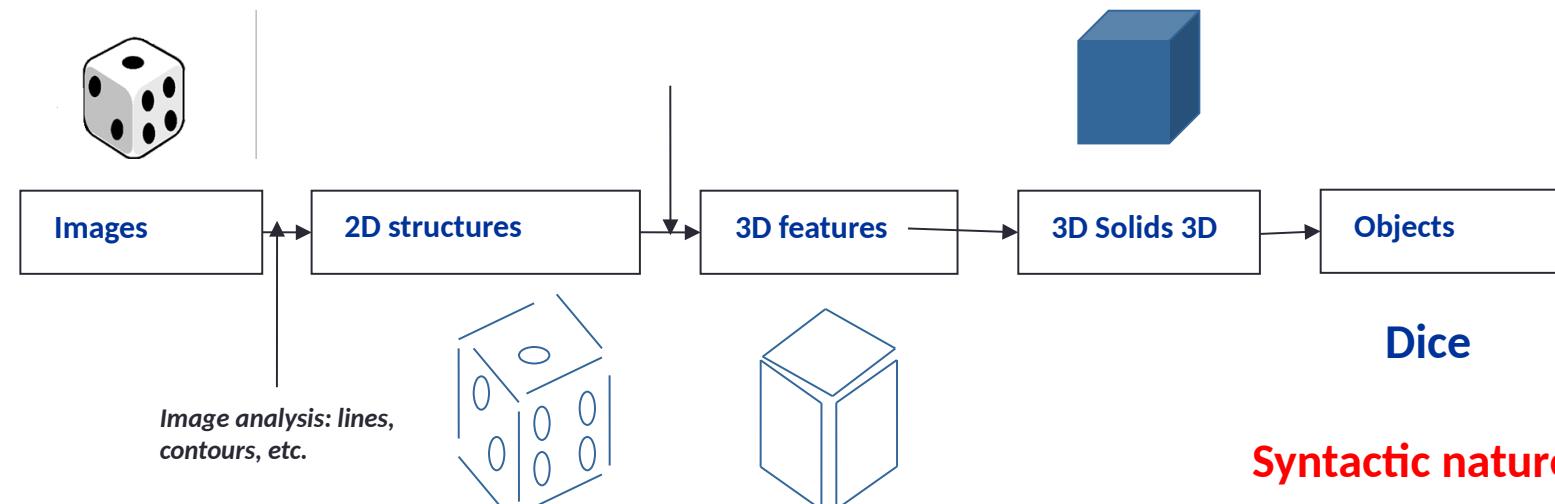
Q: How to answer:

*What is this?*

R: By a sequence of data transformations.



*Surface interference by stereo vision, motion, etc.*



# Historical definitions...

The mechanism that allows visual perception is the OPTICAL INVERSE ENGINEERING!

“Processing visual information is syntactic in nature (and consists of a set of modules that reconstruct general descriptions of the scene).”

“The semantics should be treated in a purely symbolic way as a result of the syntactic analysis.”

David Marr, Vision, Freeman, 1982.

# Vision...

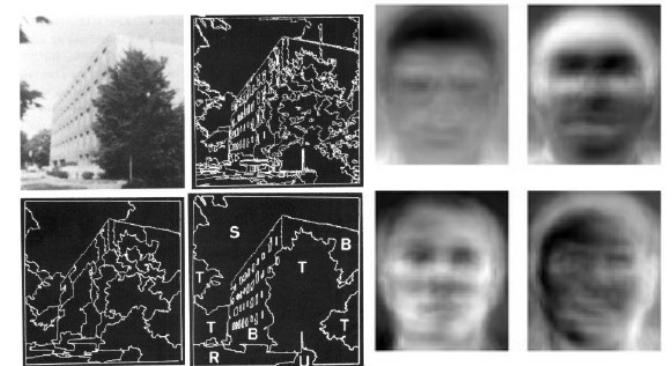
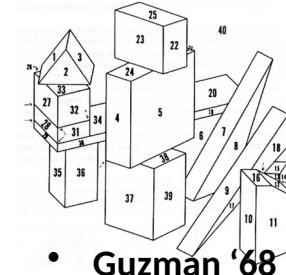
But the environment is complex...



...not accessible, continuous, dynamic and non determinist. And  
this is a **BASIC** component of the problem...

# Ridiculously brief history of computer vision

- 1960's: interpretation of synthetic worlds
- 1966: Minsky assigns computer vision as an undergrad summer project
- 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
- 2000's: Deep learning is back to stay, CNN
- 2010's: Deep learning beats humans on object recognition
- 2030's: robot uprising?



Imagenet: 15 mln of images

# Today:

1. What is Artificial Vision?
  2. A little bit of history
  3. Main problems of Artificial Vision
  4. Difficulties of the Artificial Vision
  5. Applications
  6. Subject organization
- 
- Bibliography:
    - Szeliski, CV: A&A, Ch 1.0 (Introduction)

# Each image tells a story...



- The aim of the Artificial vision is to write programs to interpret images, and perceive the story behind the image
  - Get the properties of the real world as 3D, names of persons or objects, etc.?!

## Can computers beat the men?

Yes and no (mostly not!)

- Humans are better at "hard" things
- Computers are better at "easy" things

Limit is getting thinner and thinner.

# Computer Vision: A whole series of problems

- What is in the image?
  - Object recognition problem
    - Where is it?
    - 3D spatial layout
    - Shape



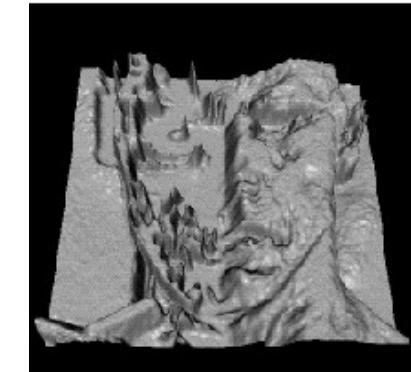
- How is the camera moving?
- What is the action?

# Computer Vision bridges 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

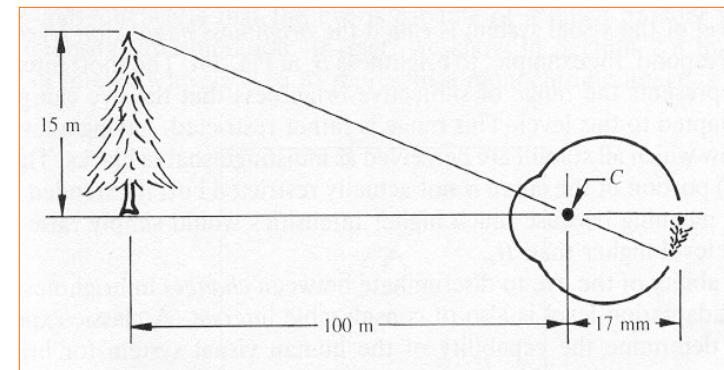
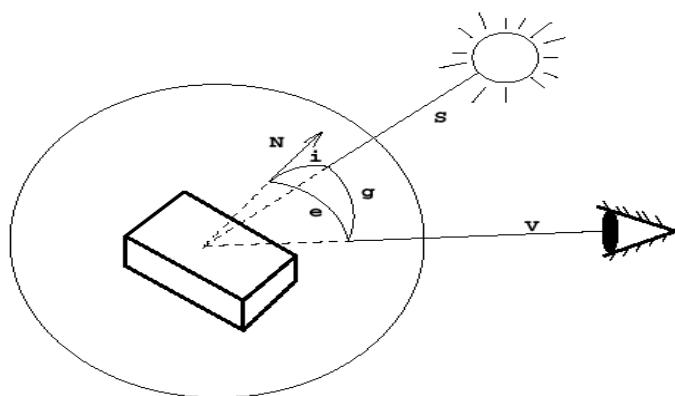


- What a computer sees

## CV problems

Formation of the image:

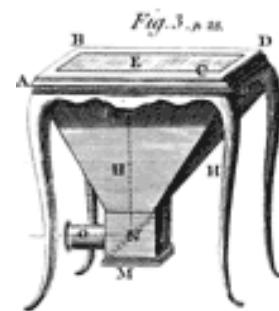
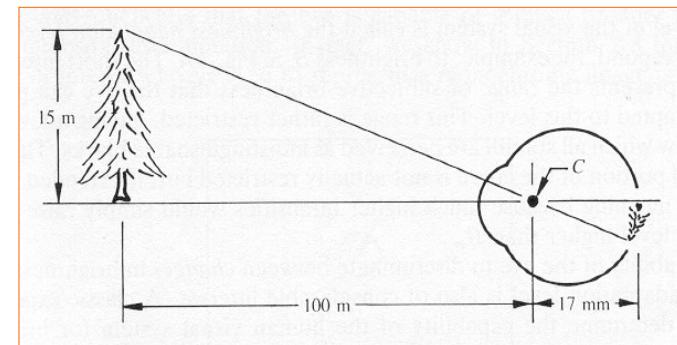
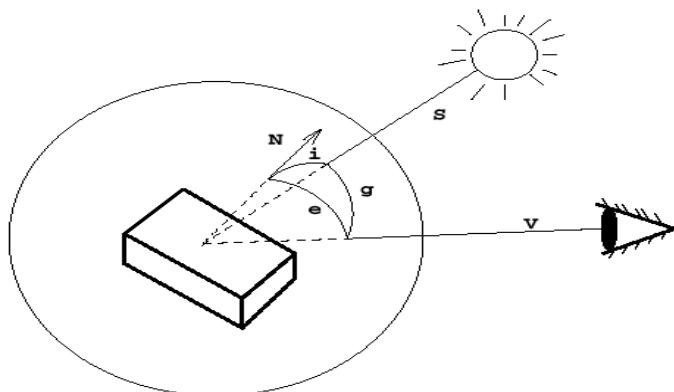
Image is a projection of the world



A point of the scene corresponds to one and only one point in the image,

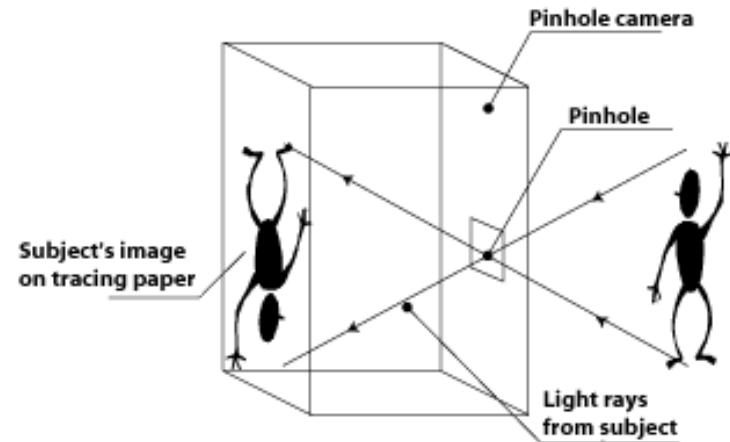
but a point of the image may correspond to infinite points of the scene (3D world).

# Image is a projection of the world



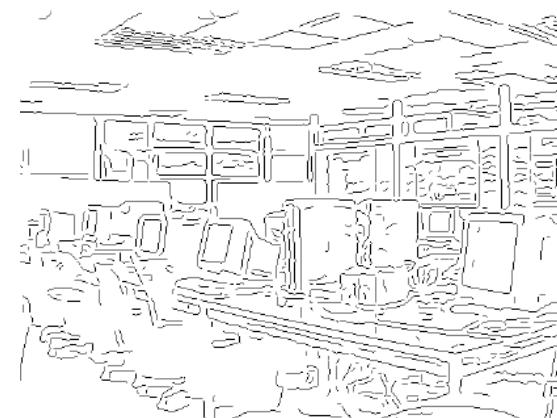
Pinhole camera

Using a pinhole camera to create an image

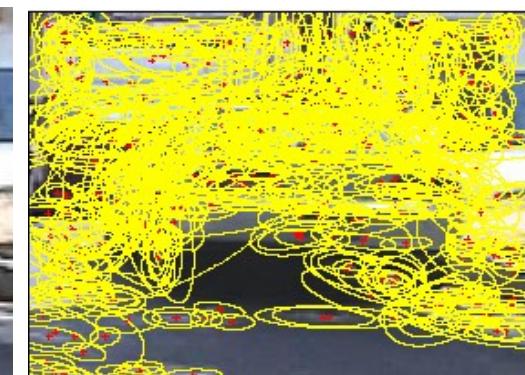


# CV problems: Low-Level Feature extraction

- Edges, corners

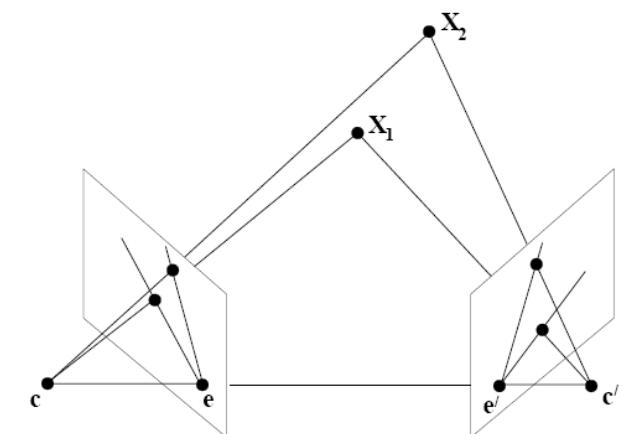
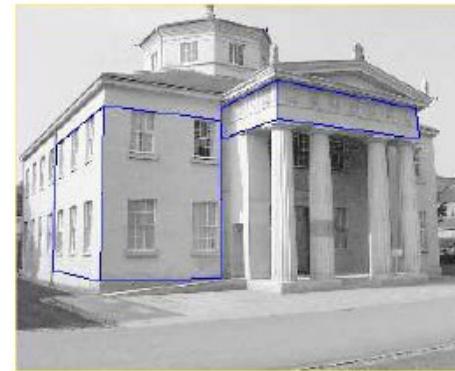
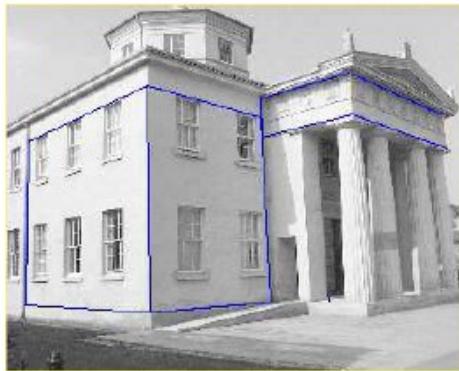


- Local regions

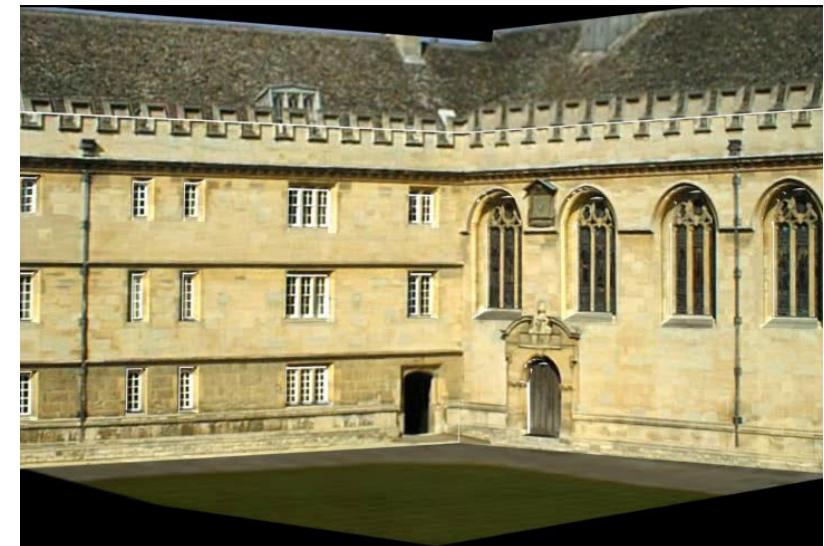
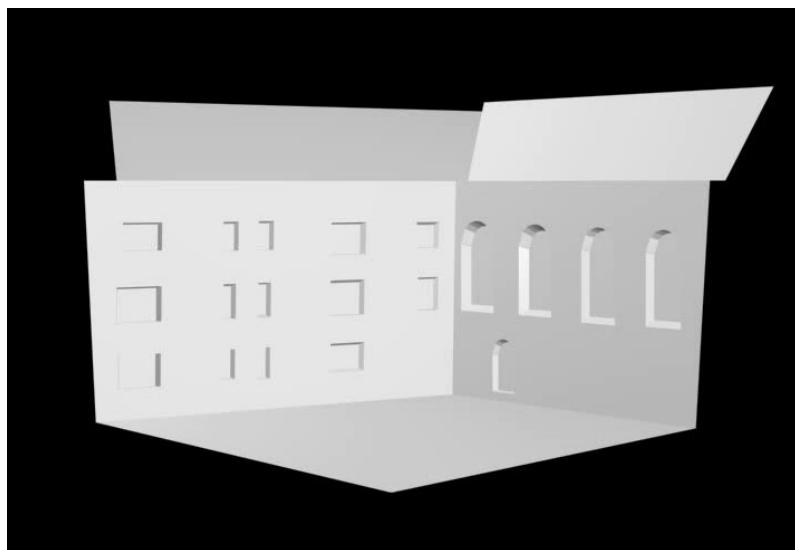
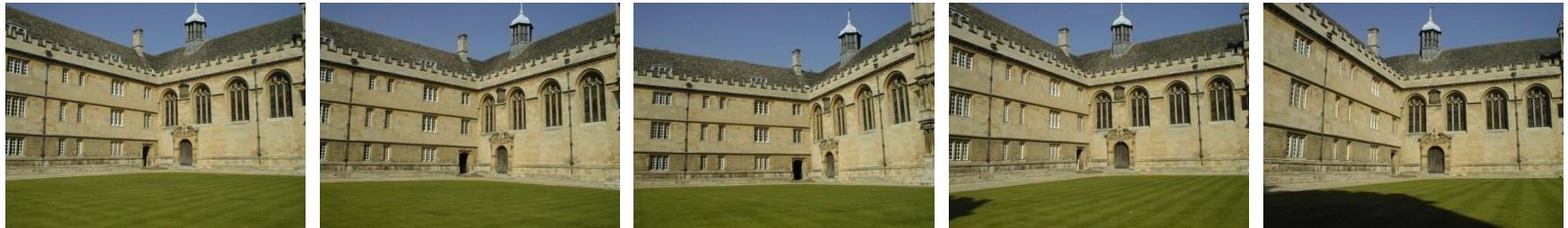


# CV problems: Stereo Vision

- By having two cameras, we can triangulate features in the left and right images to obtain the depth to the 3D point.
- Need to match features between both images:
  - Correspondence Problem

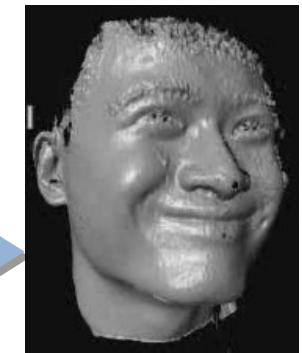
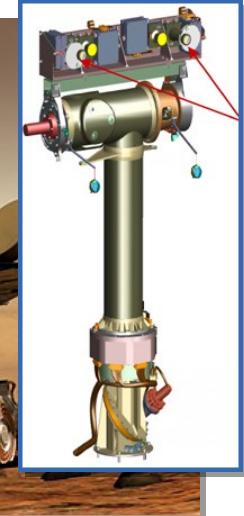
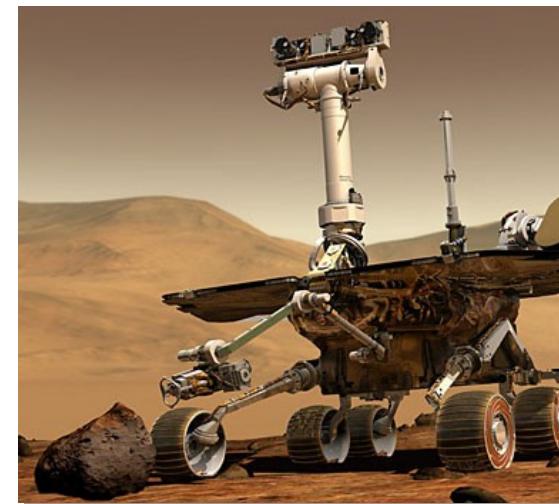


# CV problems: Geometry: 3D models of planar objects



- [Fitzgibbon et. al]
- [Zisserman et. al.]

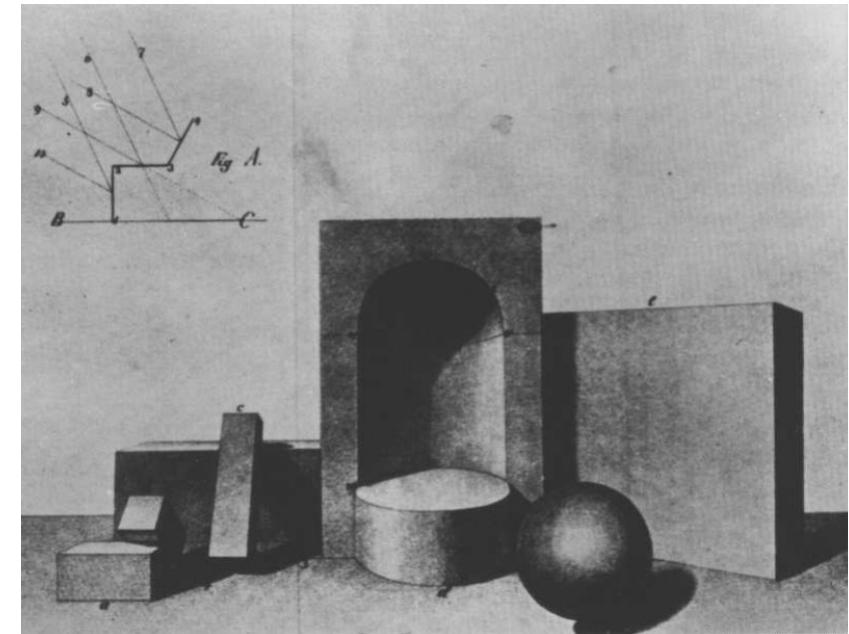
# CV problems: 3D reconstruction by depth cameras



Get the 3D shape of the objects

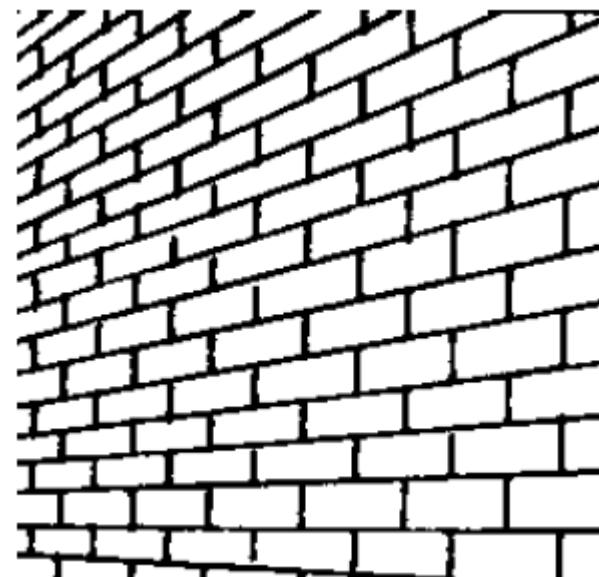
# CV problems: Shape from Shading

- Recover scene structure from shading in the image



# CV problems: Shape from Texture

- Texture provides a very strong cue for inferring surface orientation in a single image.
  - Necessary to assume homogeneous or isotropic texture.
  - Then, it is possible to infer the orientation of surfaces by analyzing how the texture statistics vary over the image.



# CV problems: Augmented reality

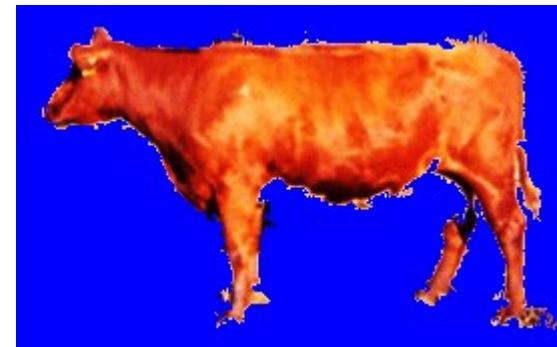


# CV problems: Segmentation

- Image



- Segmentation



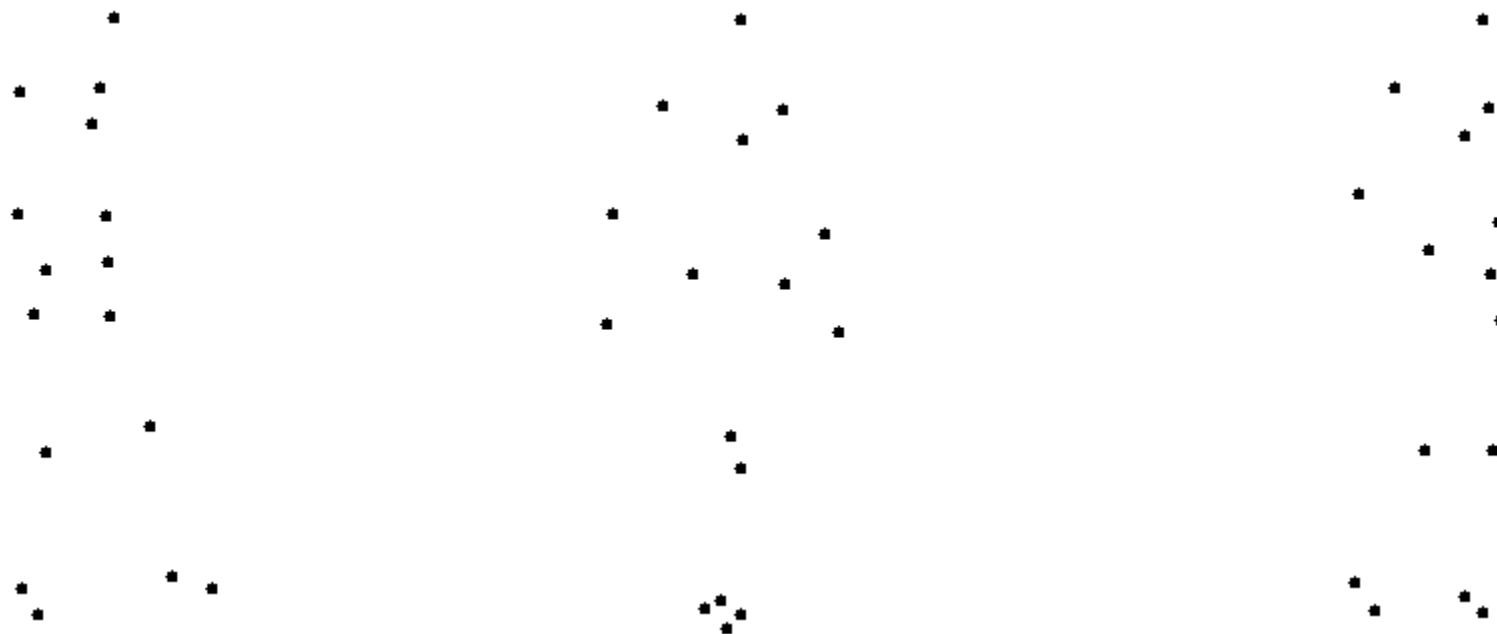
# CV problems: Scene interpretation from limited cues



# CV problems: Human motion detection



# Johansson's experiments ['70s]



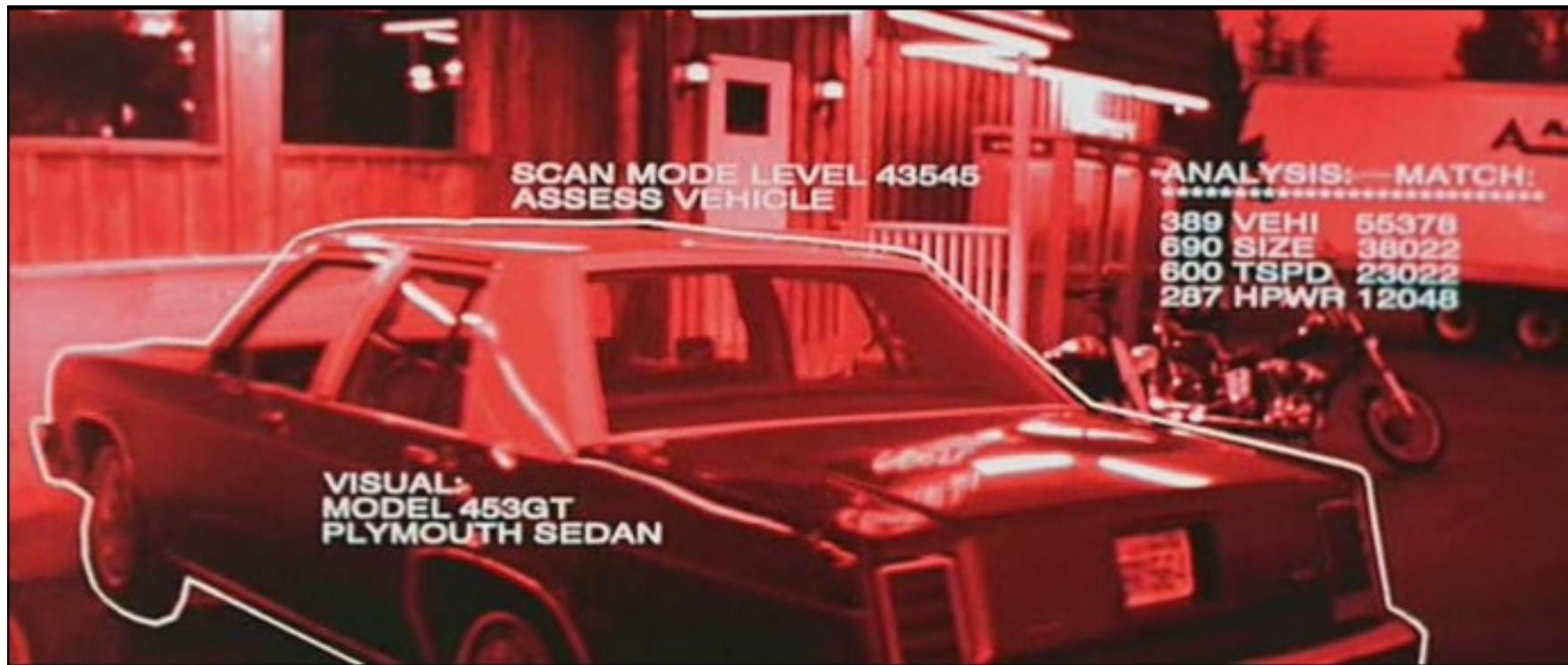
# CV problems: Object detection



# Object categorization



# Object and person recognition



# What do humans care about?



# Verification: is that a bus?



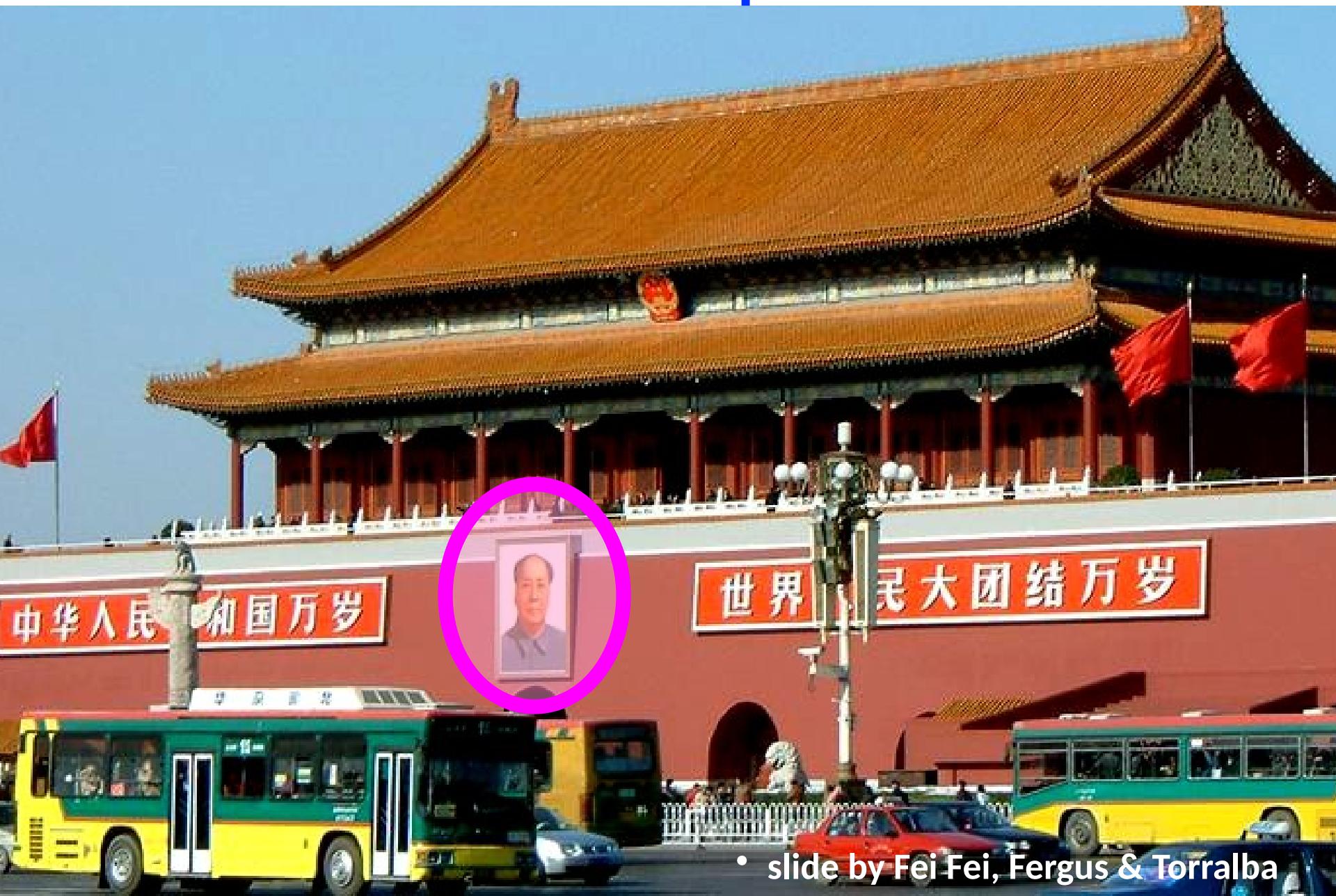
• slide by Fei Fei, Fergus & Torralba

# Detection: are there cars?



• slide by Fei Fei, Fergus & Torralba

# Identification: is that a picture of Mao?

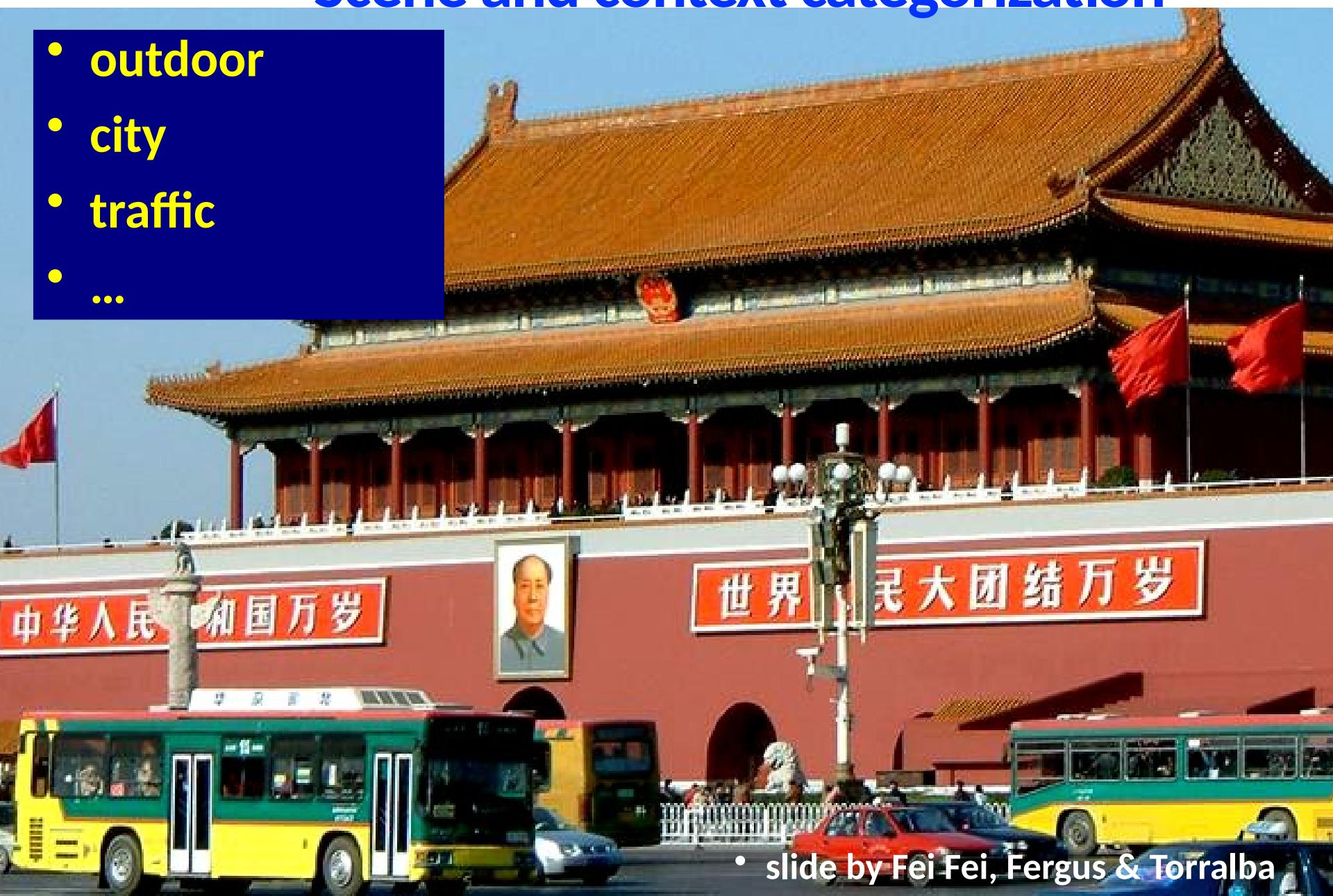


# Object categorization

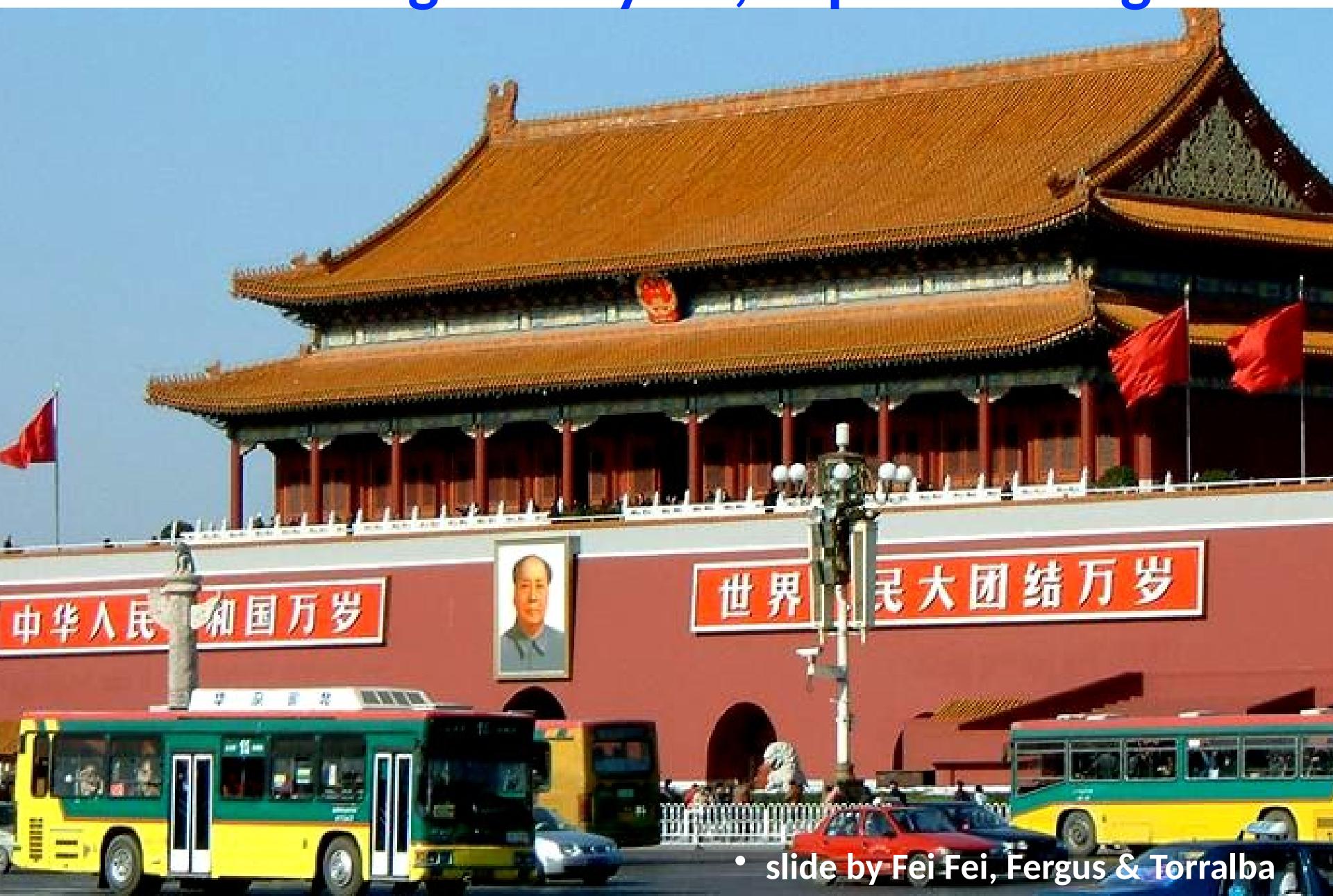


# Scene and context categorization

- outdoor
- city
- traffic
- ...



# Rough 3D layout, depth ordering



# Today:

1. What is Artificial Vision?
  2. A little bit of history
  3. Main problems of Artificial Vision
  4. Difficulties of the Artificial Vision
  5. Applications
  6. Subject organization
- 
- Bibliography:
    - Szeliski, CV: A&A, Ch 1.0 (Introduction)

# Why is it hard?



- Variation of point of view



- Illumination



- Scale

# Why is it hard?



- Intra-class variation



- Cluttered background

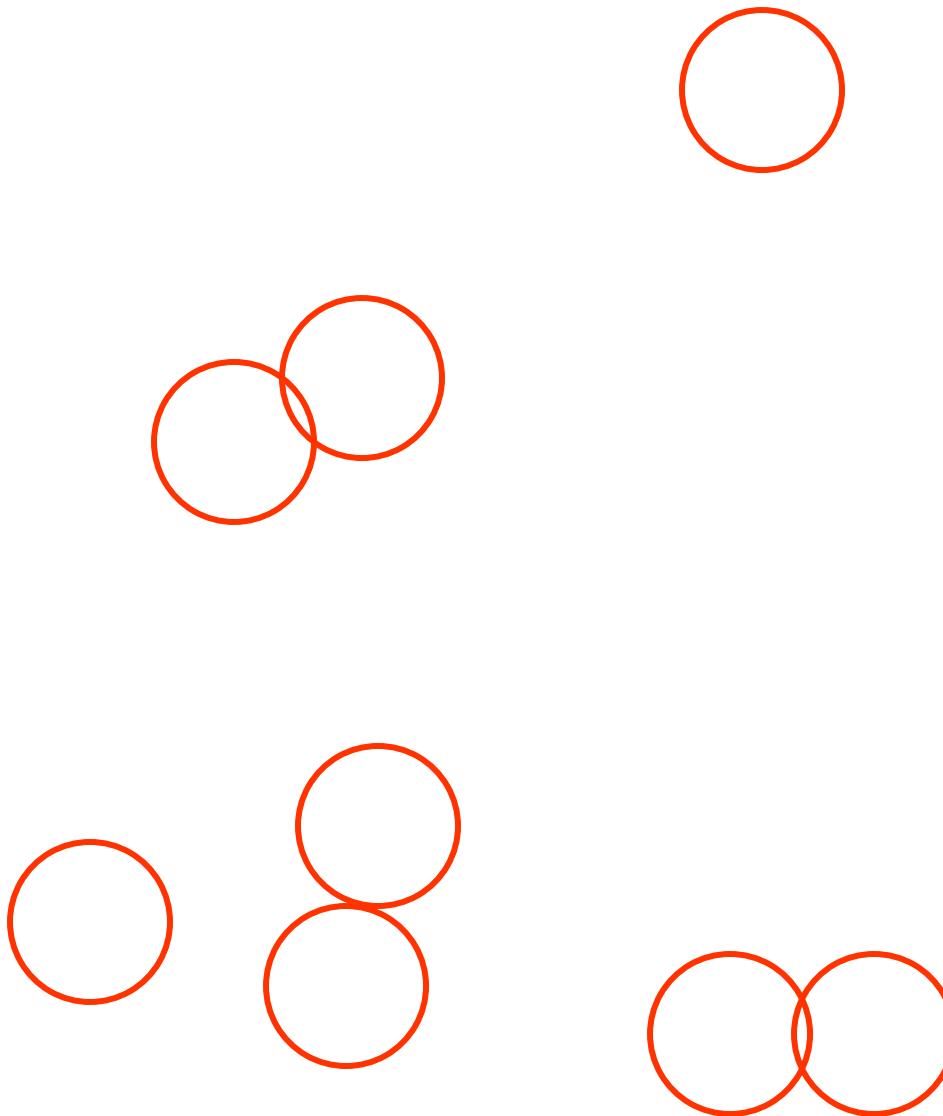


- Movement (Font: S. Lazebnik)



- Occlusion

# Why is it hard?



- Local ambiguity

# Some things know that you have eyes



- Brady, M. J., & Kersten, D. (2003). Bootstrapped learning of novel objects. *J Vis*, 3(6), 413-422

## Bottom-up data analysis

- The problem is the ambiguity of the implicit perception
  - Many 3D scenes can give the same 2D scene



Image source: F. Durand

# Is the human perception perfect?



Sinha and Poggio, Nature, 1996

- We still don't understand completely the perception in order to simulate / model / copy it

# Why is it difficult? Visual illusions

In the visual perception different factors intervene associated with aspects of the observer and his/her environment.

The knowledge acquired culturally.



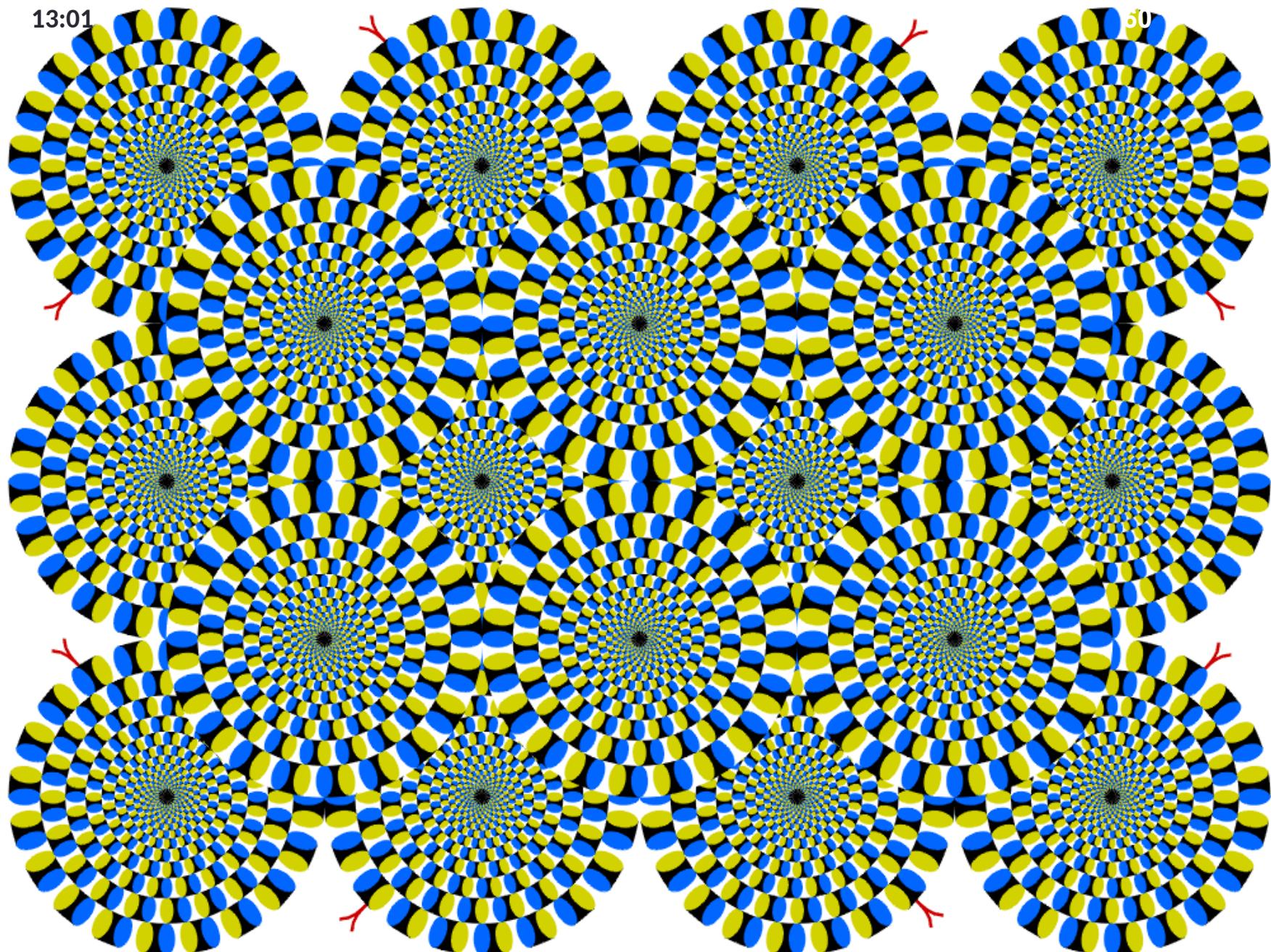
Visual memory and past experiences.

Assumptions about the world acquired for evolution.

The architecture of the human visual system.

13:01

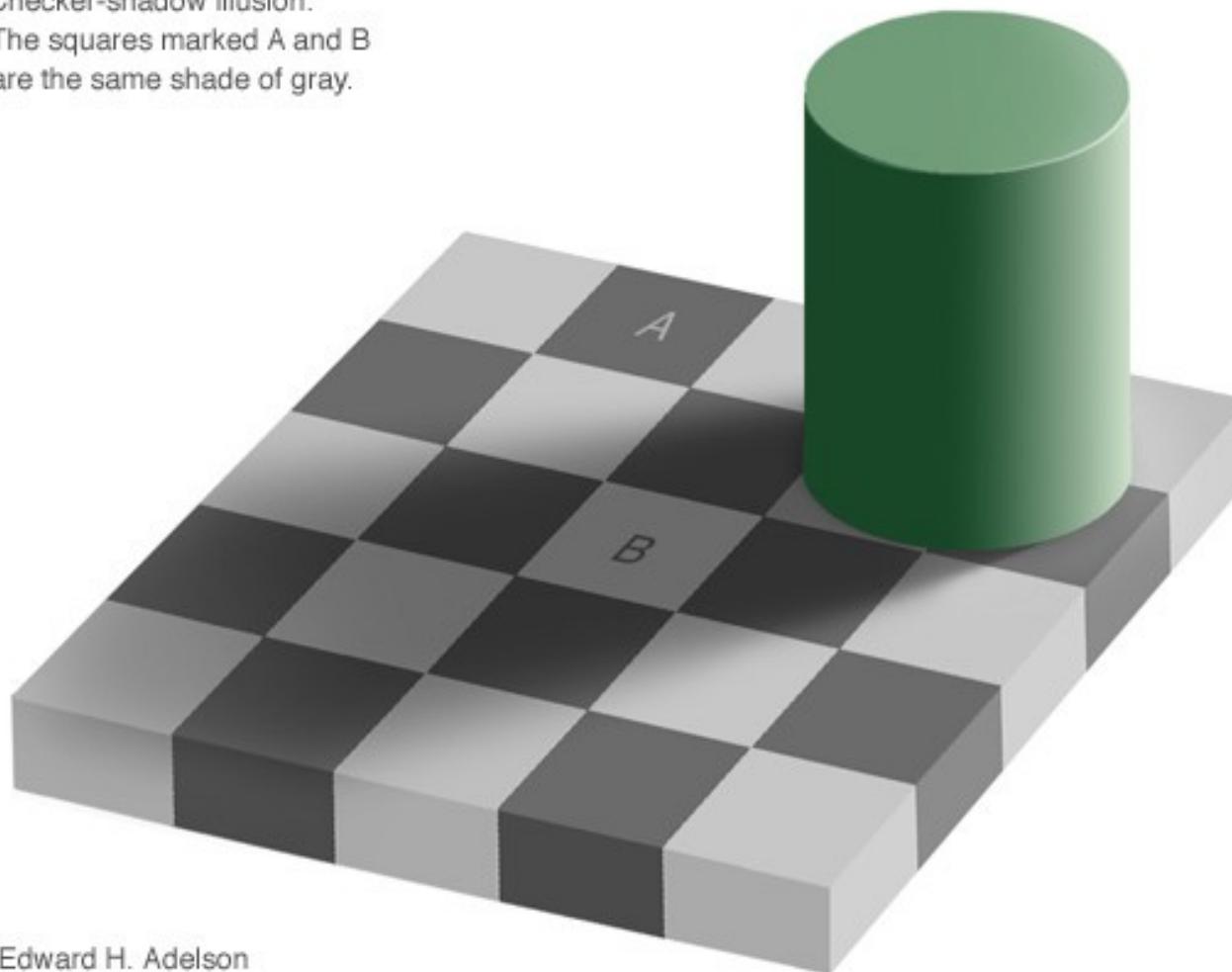
50



Copyright [A.Kitaoka](#) 2003

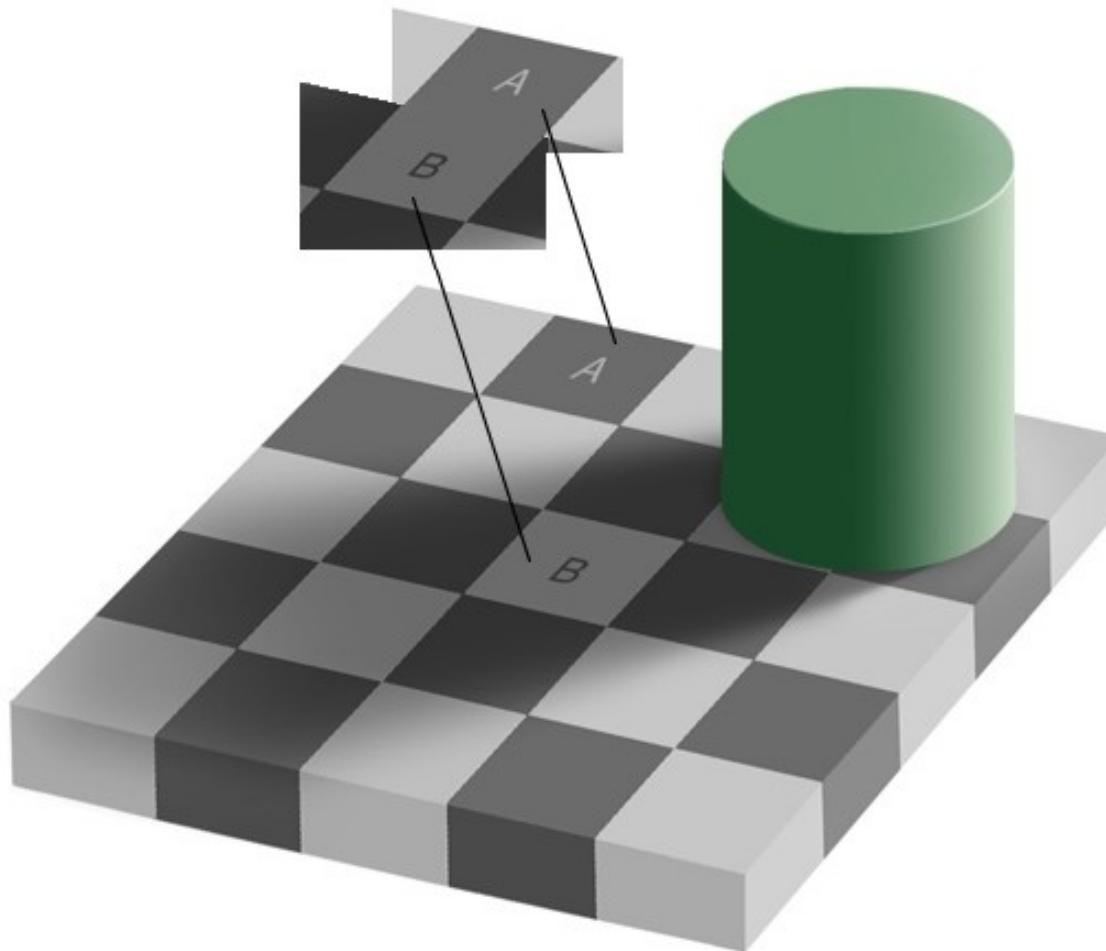
# Visual illusions

Checkers shadow illusion:  
The squares marked A and B  
are the same shade of gray.



Edward H. Adelson

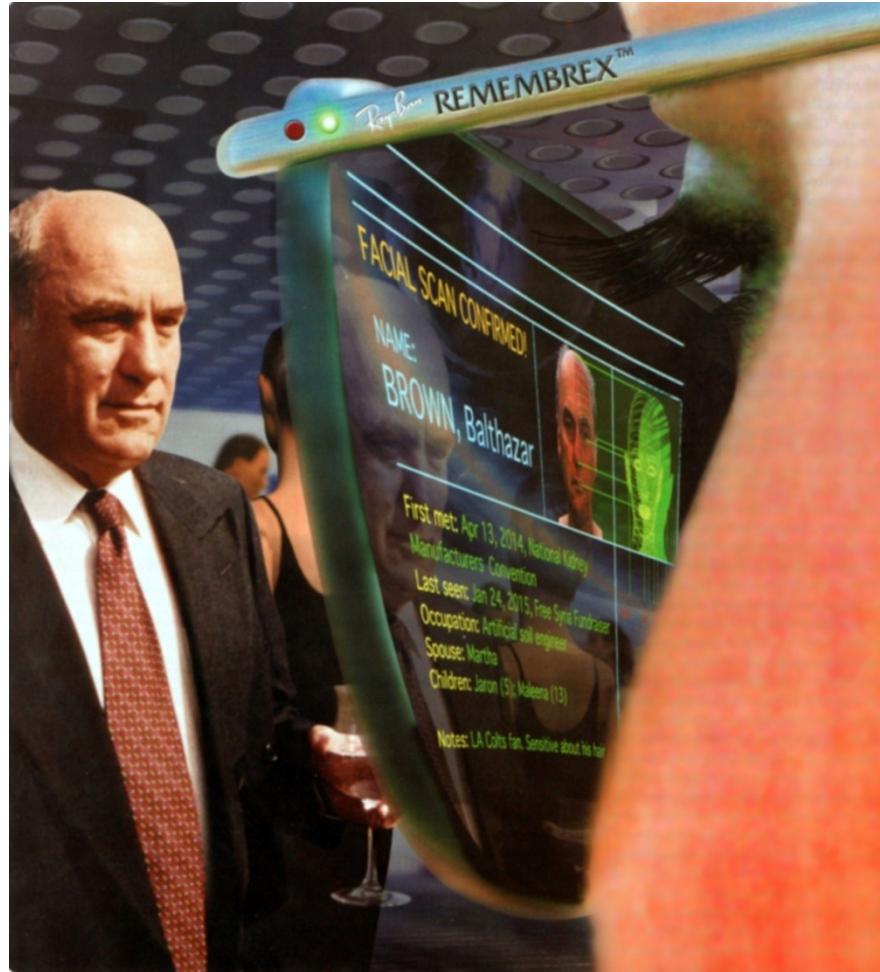
# Visual illusions



# Today

1. What is Artificial Vision?
  2. A little bit of history
  3. Main problems of Artificial Vision
  4. Difficulties of the Artificial Vision
  5. Applications
  6. Subject organization
- 
- Bibliography:
    - Szeliski, CV: A&A, Ch 1.0 (Introduction)

# Is Artificial vision useful?



# Computer vision today

Only a few examples:

- Most have less than 15 years!
- A very active and changing area
- Expected numerous applications the next 5-10 years!
  - Computer Vision companies

## Contents

[The most popular Computer Vision Companies in 2022](#)

1. SenseTime
2. MegVii
3. viso.ai
4. NAUTO
5. Verkada
6. Tractable
7. Airobotics
8. Hawk-Eye Innovations
9. Trigo
10. Movidius
11. Standard Cognition
12. Orbital Insight
13. Regna
14. AnyClip
15. Bossa Nova Robotics
16. Descartes Labs
17. Zebra Medical Imaging
18. Neuromation
19. Shield AI
20. EyeSight
21. Onfido

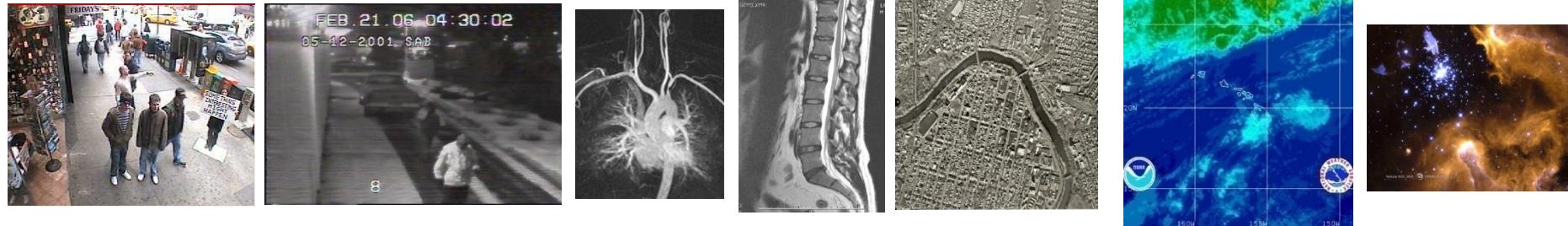
In the following, this article will feature 21 AI leaders of today.



Real-world computer vision in [smart cities](#). – Built on [Viso Suite](#)

# Why to study Computer Vision?

- Each second million of images are captured and stored



- Numerous real applications

# Deep learning

[ABOUT](#)[PRICING](#)[DEVELOPER](#)[CAREERS](#)[CONTACT US](#)[SIGN UP NOW](#)

\*By using the demo you agree to our [terms of service](#)



## Predicted Tags

no person    table    elegant  
indoors    luxury    furniture  
fashion    decoration    tableware  
party



# Deep learning

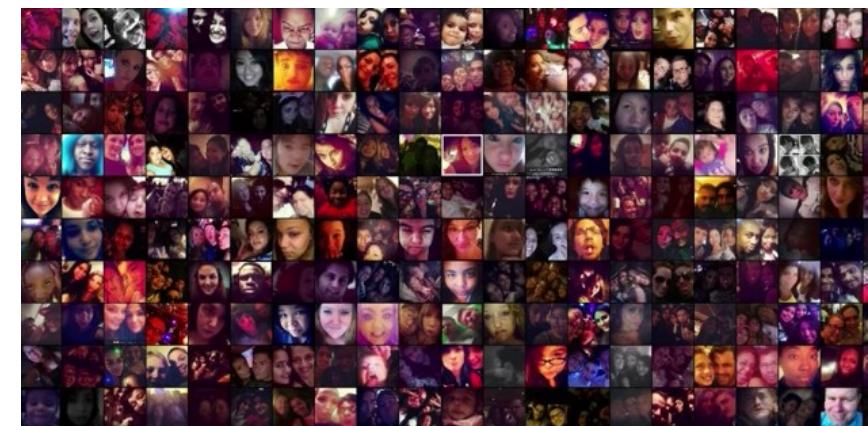


Classified as HIGH aesthetic images



Classified as LOW aesthetic images

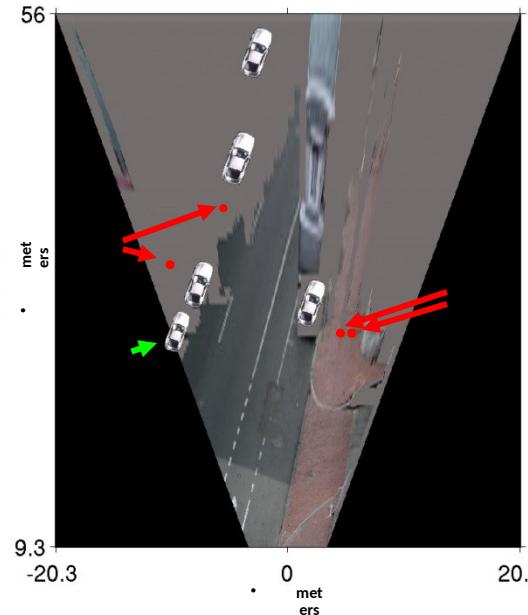
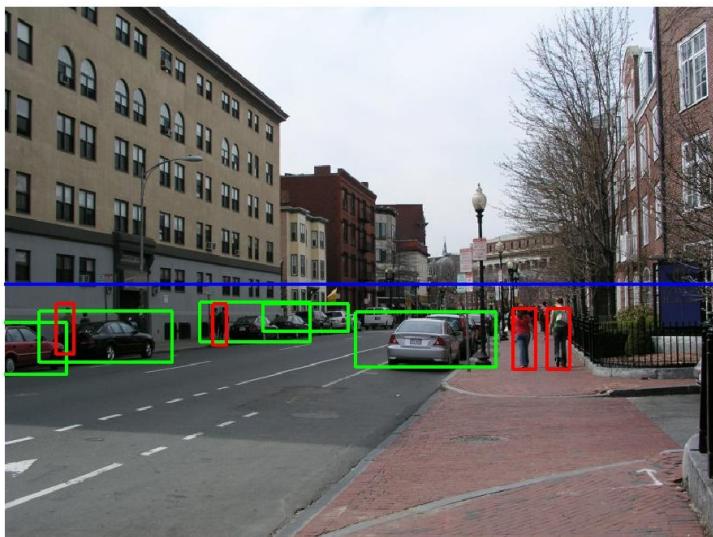
**SHARE** **TWEET** **DOWNLOAD**



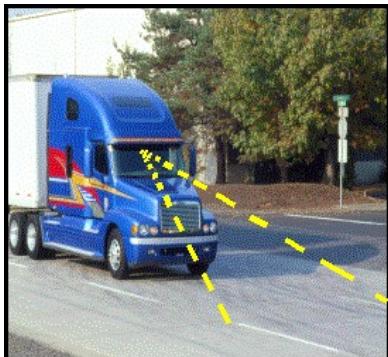
Worst 300 out of 50,000 selfies, as judged by the Convolutional Neural Network.

# Assisted driving

- Pedestrian and car detection

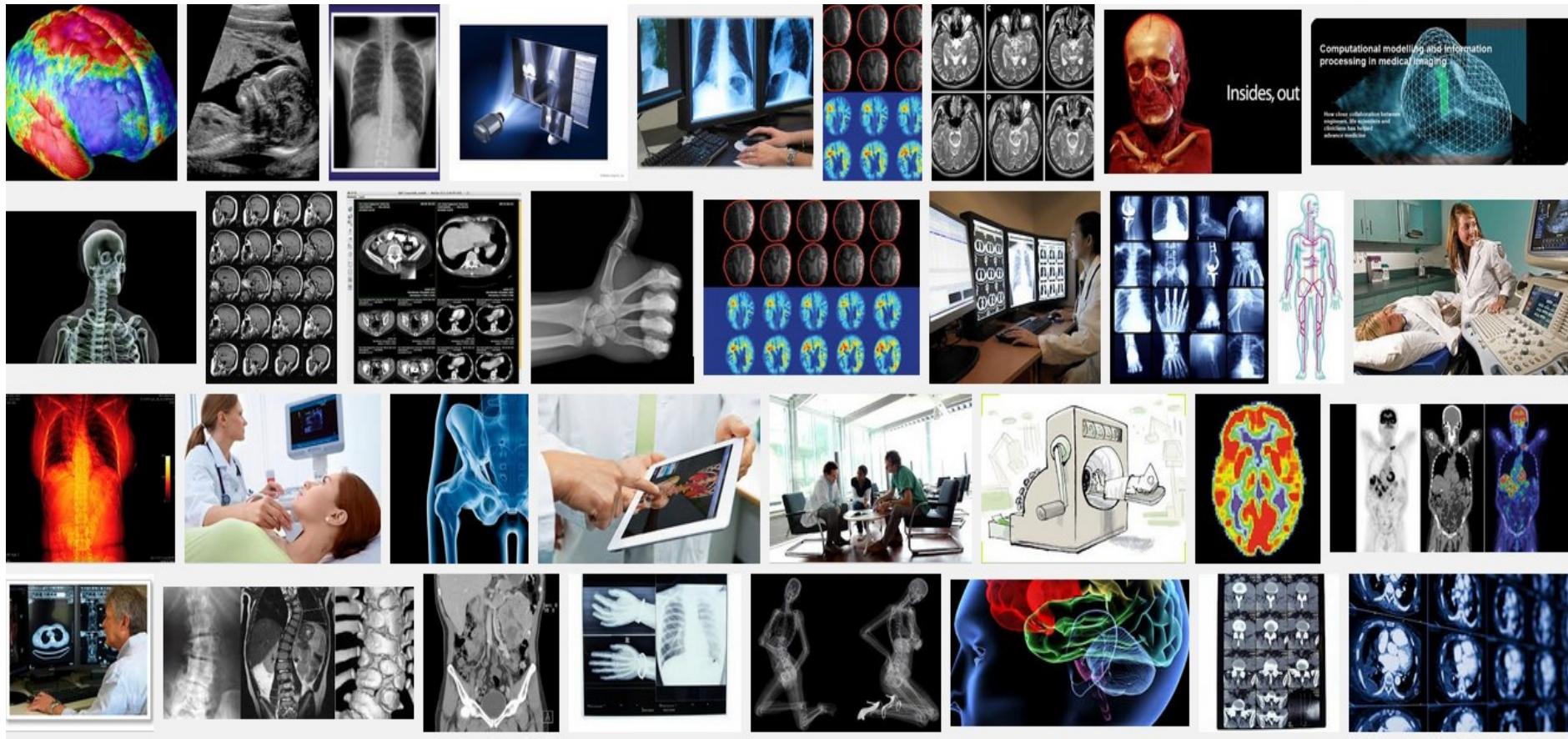


- Lane detection



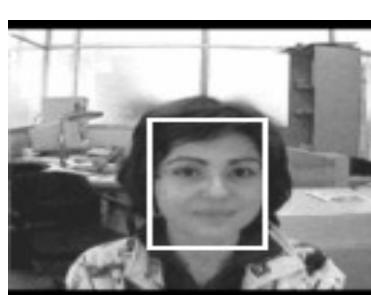
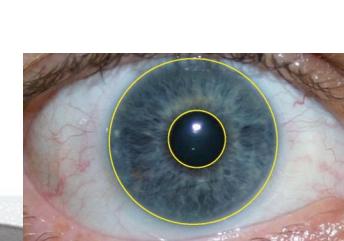
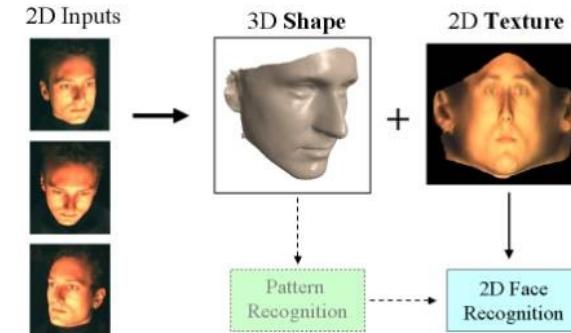
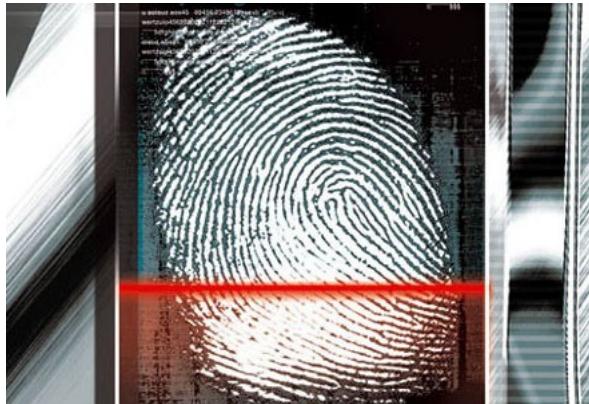
- Collision warning systems with adaptive cruise control,
- Lane departure warning systems,
- Rear object detection systems,

# Medical images and health

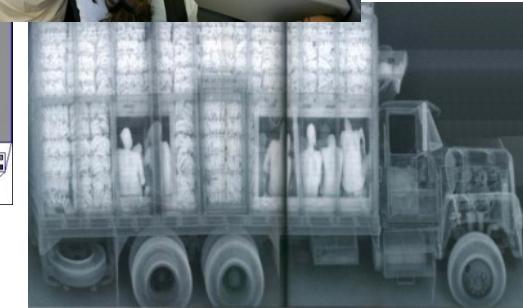
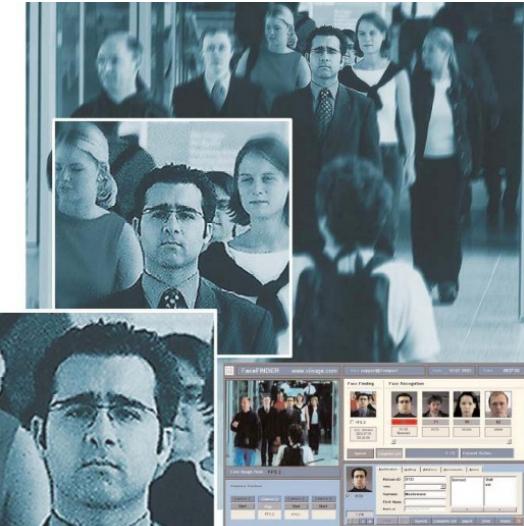


# Surveillance and security: biometric systems

- Fingerprint identification, facial recognition, iris and retinal scan, hand geometry, geometry of the ear, signature recognition, voice identification, identification of the DNS, the smell of human characteristics or recognition of typing motion (not a privilege of science fiction anymore).



# Surveillance and security: biometric systems



# Media and Entertainment

- Gaming interfaces



- Sony EyeToy (for PlayStation) allows players to interact with games using motion, color detection and sound.



- GestureFX.  
Gesture Interfaces  
ground, wall or  
table.

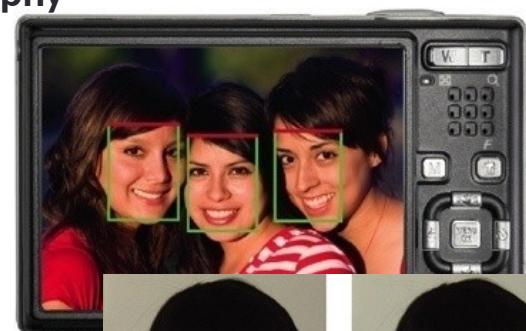


- Augmented Reality



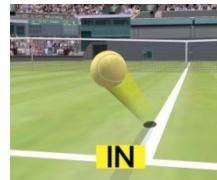
- Digital Photography

- Photosynth  
(Microsoft  
LiveLabs)

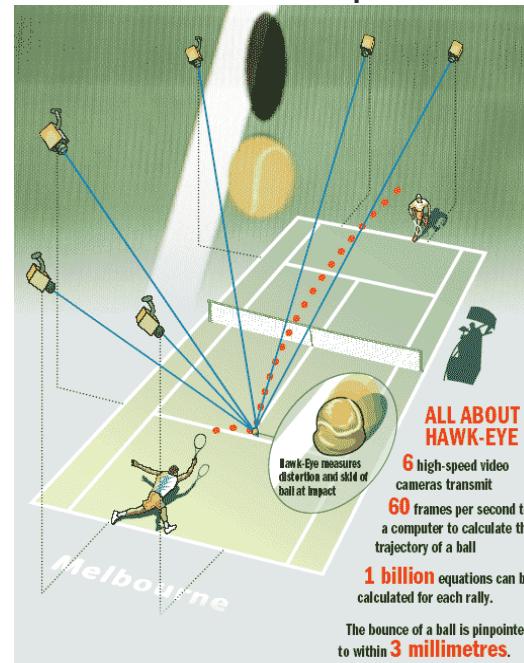


# Media and Entertainment

- Augmented TV



- Hawk Eye: tracking the ball in game sports.



- PVI: Virtual publicity in real television pictures.



# Mobile Computer Vision

- Client services
- For visually impaired
- Health services



- Android Developer Challenge (Google): Barcode Reader for price list and product information



- Tourism

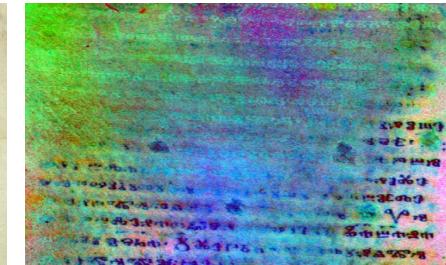
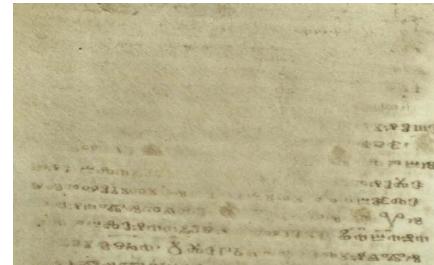


# Cultural inheritance

- Arqueology



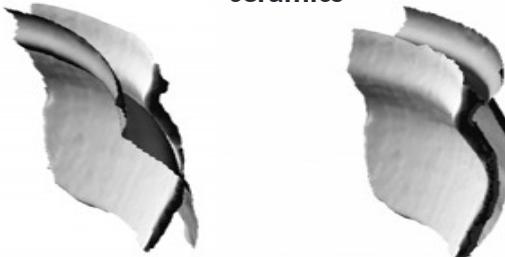
- Old documents



- Manuscript recognition



- 3D reconstruction of ceramics

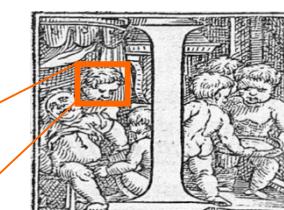
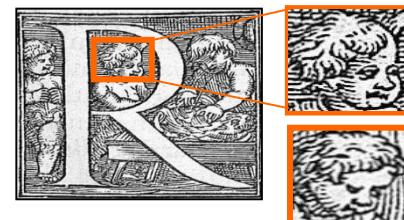


siguiendo lo que en los primeros tiempos del  
siglo se llamaba el libro de la  
vida y de la fisionomía de aquello

antiguos ciudadanos, que en Castilla se llamaban

caballeros y en otros países puebl baj renombrados  
sobresalían en carácter religioso, virtud y caballería; pero  
de su grandeza y importancia social, de su perseverancia

- Restauration

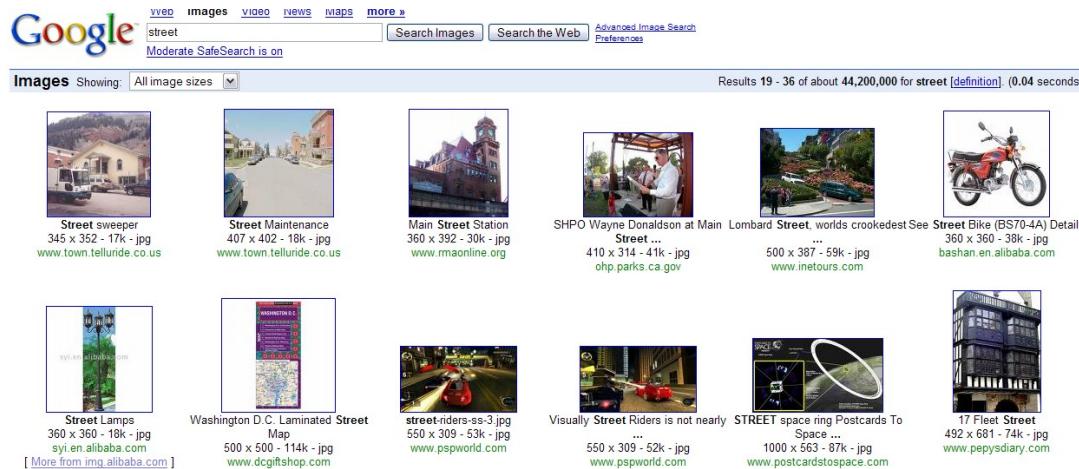


- Engravings analysis

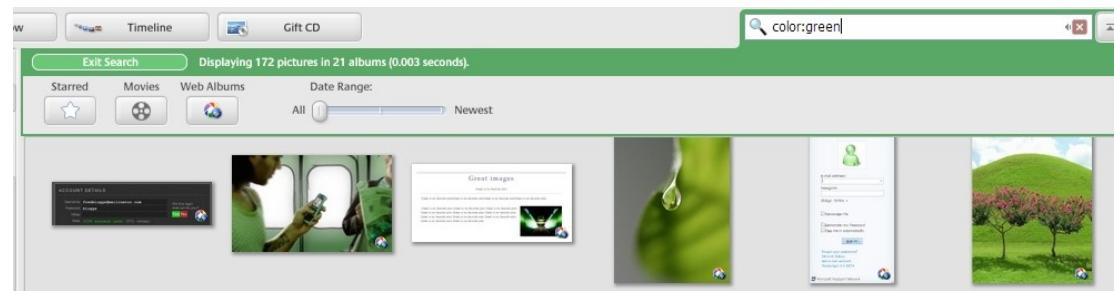
# Improving online search



- Query:
  - street



- Organizing photo collections



• Slide credit Fei-Fei, Fergus, Torralba CVPR07 Short Course

# Amazon farm



# Today:

- 1. What is Artificial Vision?**
- 2. A little bit of history**
- 3. Main problems of Artificial Vision**
- 4. Difficulties of the Artificial Vision**
- 5. Applications**
- 6. Subject organization**

# In this subject we will:



# Subject content

## Presentation and introduction

- The subject. What is Computer Vision?

## Features detection

- Edge detection. Differential operators. Discrete approximations. Discontinuities.
- Detection and localization of contours. Differential operators. Structure of the image.
- Visual characteristics (HOG, Harris, SIFT), etc.

## Shape

- Contours and Segmentation
- Texture analysis and image retrieval.

## Recognition

- Object detection.
- Recognition based on appearance. Face recognition.
- Convolutional Neural Networks

# Laboratory environment

- Python & Anaconda
- Numpy
- Scikit-image
- Matplotlib



The screenshot shows the official website for Anaconda, Inc. The header features the Anaconda logo and navigation links for "What is Anaconda?", "Products", "Support", "Community", "About", "Resources", and a prominent green "Download" button. Below the header, a large banner image shows hands typing on a keyboard with various Python data science library logos (SciPy, NumPy, Pandas, Jupyter, H2O.ai, TensorFlow, etc.) overlaid. The central text reads "The Most Popular Python Data Science Platform". Below the banner, three key statistics are displayed: "4.5M+ Users", "1,000+ Data Science Packages", and "150+ Enterprise Customers". A descriptive paragraph at the bottom explains Anaconda's role as the world's most popular Python data science platform, mentioning its open source projects like Anaconda, NumPy, and SciPy, and its flagship product, Anaconda Enterprise.

With over 4.5 million users, Anaconda is the world's most popular Python data science platform. Anaconda, Inc. continues to lead open source projects like Anaconda, NumPy and SciPy that form the foundation of modern data science. Anaconda's flagship product, Anaconda Enterprise, allows organizations to secure, govern, scale and extend Anaconda to deliver actionable insights that drive businesses and industries forward.

# Laboratories

Will be based on the material seen in theory and problems classes.

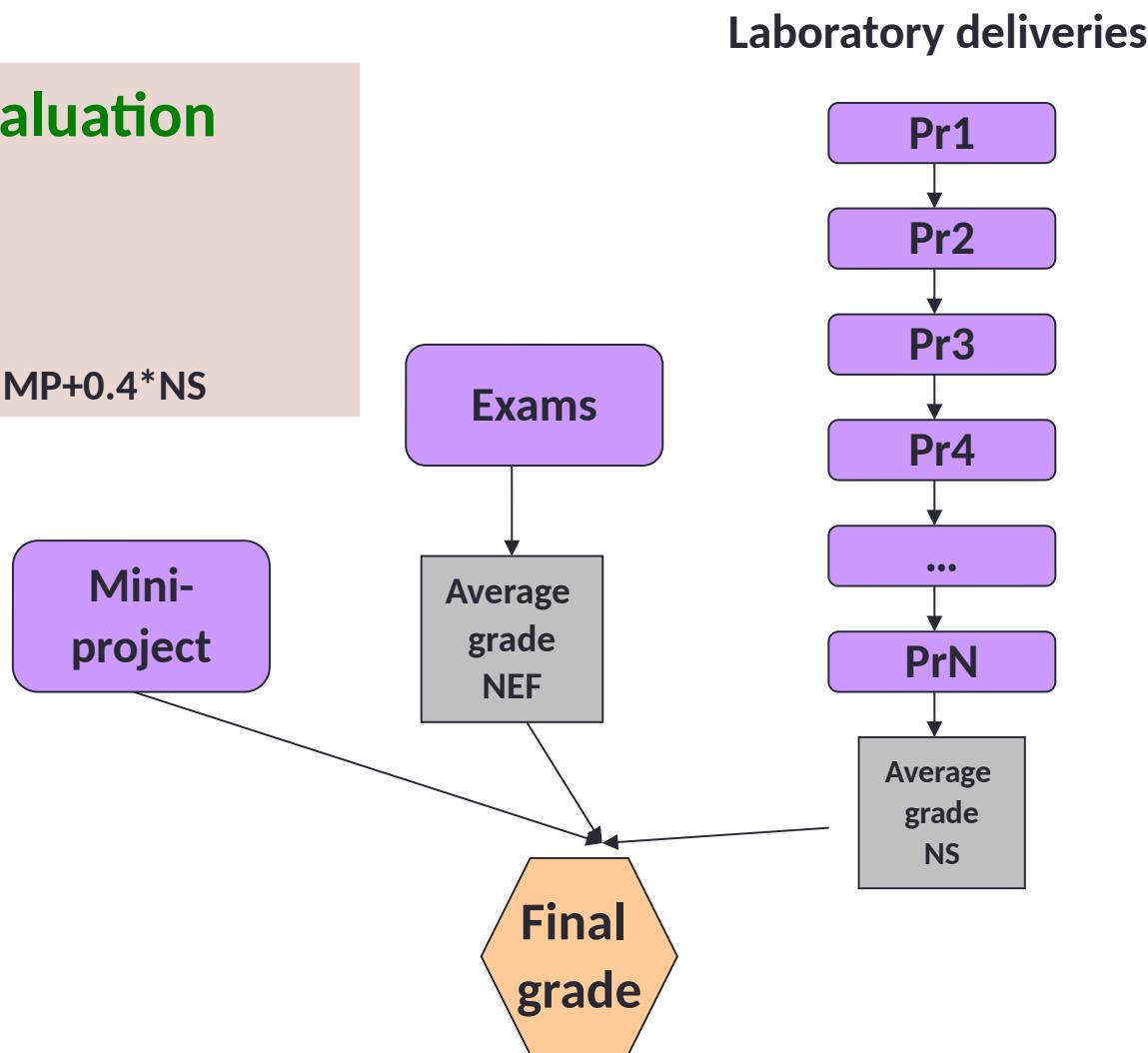
- Some deliveries may include small abstracts, techniques, etc.

English:

- Theory slides
- Additional material:
  - Videos and videolectures
  - Articles and other literature

## Continuous Evaluation

$$\text{NEC} = 0.5 * \text{NEF} + 0.1 * \text{MP} + 0.4 * \text{NS}$$



## Avaluació continuada

La nota Final de l'assignatura, Nota\_Final, es calcularà segons la següent fórmula:

$$\text{Nota_Avaluació_Continuada} = 0.5 * \text{Nota_Examens} + 0.1 * \text{Nota_Adicional} + 0.4 * \text{Nota_Sessions_Pràctiques}$$

on:

*Nota\_Examens* és la nota de les proves parciales.

*Nota\_Sessions\_Pràctiques* és la nota promitja de diferents lliuraments de pràctiques. Aquests lliuraments són presencials.

*Nota\_Adicional* és la nota que contempla diferents activitats que es poden incloure per a valorar l'assignatura (nota del miniprojecte).

$$\text{Nota_Examens} = 0.6 * \text{Nota_Examens_Teoria} + 0.4 * \text{Nota_Examens_Pràctiques}$$

on:

*Nota\_Examens\_Teoria* és la nota de l'examen de teoria dels parciais.

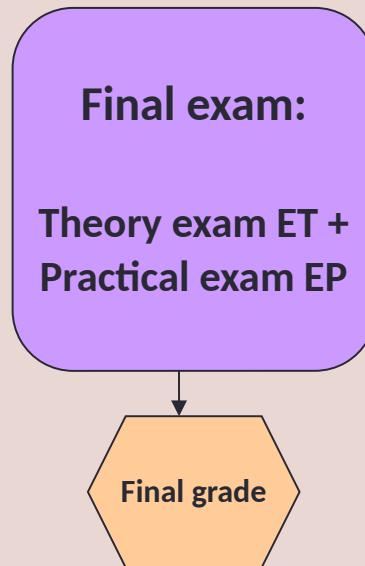
*Nota\_Examens\_Pràctiques* és la nota de l'examen pràctic dels parciais.

Per a poder fer el càlcul de la Nota\_Final, és condició imprescindible que:

- l'estudiant obtingui: *Nota\_Examens\_Teoria*  $\geq 4.0$  i *Nota\_Examens\_Pràctiques*  $\geq 4.0$
- s'aprovin per separat els dos parciais!

Qualsevol intent de frau realitzat durant el curs comportarà l'aplicació de la normativa acadèmica general de la UB i l'inici d'un procés disciplinari.

## Single Evaluation



$$FG=0.6*ET+0.4*EP$$

Approved if  $ET \geq 4$ ,  $EP \geq 4$  and Final grade  $\geq 5$ .

## Avaluació ònica (Finals d'Octubre - Termini)

L'estudiant que es vulgui acollir a l'avaluació ònica haurà de sol·licitar-ho a la secretaria de la Facultat dins dels terminis establerts a cada curs acadèmic.

Per avaluació ònica, ha de passar una prova final on s'avaluaran conceptes teòrics i pràctics de l'assignatura de la següent manera

$$\text{Nota\_Final} = 0.6 * \text{Nota\_Teoria} + 0.4 * \text{Nota\_Pràctiques}$$

On:

**Nota\_Teoria:** nota de l'examen final on s'inclouen aspectes a avaluar propis de l'avaluació ònica de la part de teoria,

**Nota\_Pràctiques:** nota de la part pràctica (preguntes i problemes sobre els continguts de les pràctiques de l'avaluació continuada, problemes específics pràctics de coneixements, etc.).

Per a poder fer el càlcul de la **Nota\_Final**, és condició imprescindible que l'estudiant obtingui:

**Nota\_Teoria >= 4.0 i Nota\_Pràctiques >= 4.0**

## Revaluació:

Un alumne té dret d'anar a l'examen de revaluació si  $NF \geq 3.5$ . L'examen de revaluació tindrà part teòrica (RT) i part pràctica (RP) i la nota final de l'assignatura serà:

- $NF = 0.6 * RT + 0.4 * RP$ .

## Observacions:

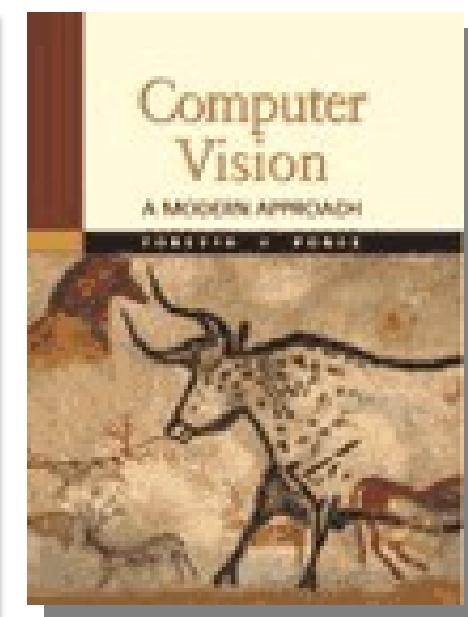
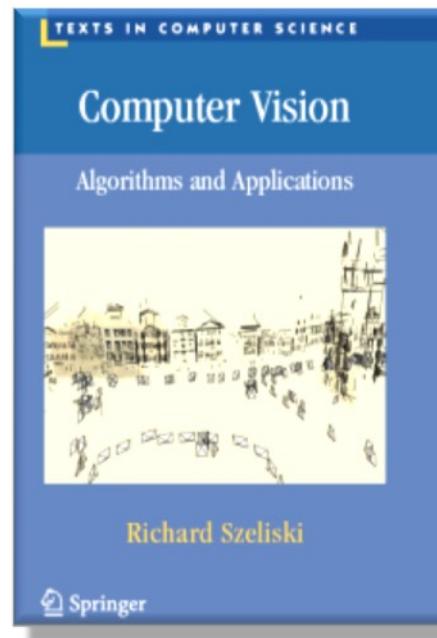
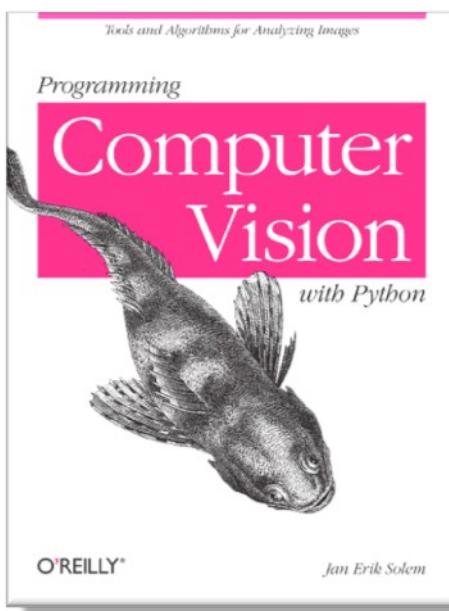
- a) Les pràctiques no són recuperables. Una pràctica no presentada té una nota 0.
- b) Un "No presentat" en l'avaluació continuada és quan no s'ha presentat cap pràctica ni s'ha fet cap parcial.
- c) Les notes no es guarden d'un any per a un altre.
- d) Si es suspèn alguna de les parts d'activitats (teoria o pràctiques) en l'avaluació continuada durant el semestre, s'ha de fer l'examen sencer; no es guarda la part aprovada.
- i) Algunes de les pràctiques poden ser pràctiques avaluables, on el treball pràctic es defineix i es lliura al final de la sessió.
- f) Els estudiants que fan la revaluació renuncien la nota anterior obtinguda.

# Bibliography

Programming Computer Vision with Python: Tools and algorithms for analyzing images 1st Edition, [Jan Erik Solem](#), O'Reilly, 2012.

Computer Vision: Algorithms and Applications (c) 2010 [Richard Szeliski](#), Microsoft Research <http://szeliski.org/Book/>

David A. Forsyth, Jean Ponce, *Computer Vision: A Modern Approach*, Prentice Hall, Hardcover, August 2012, II edition.



# Mini-projects

Groups of 4 students should be formed to work on a mini-project. The title and the group should be submitted before 15/11/2022.

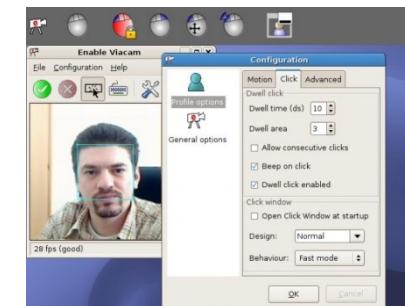
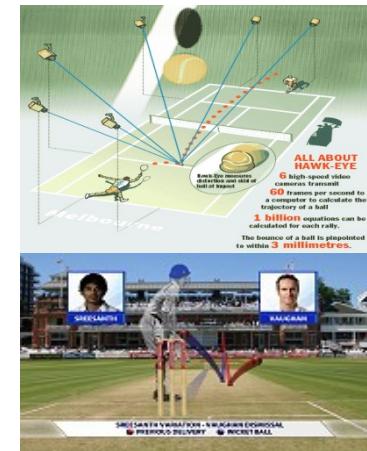
Memory of 5 pages should explain:

- target
- environment / context
- information about the implementation,
- alternatives and extensions to the application as seen in the course material.
- Introduction to Virtual Campus.

These mini-projects will have a weight of 10% of the final grade of continuous evaluation.

- The projects will be presented in the last two lectures.

# Miniprojects



# Additional material

- Get inspired for the miniprojects!

# Computer Vision APIs

Nyckel

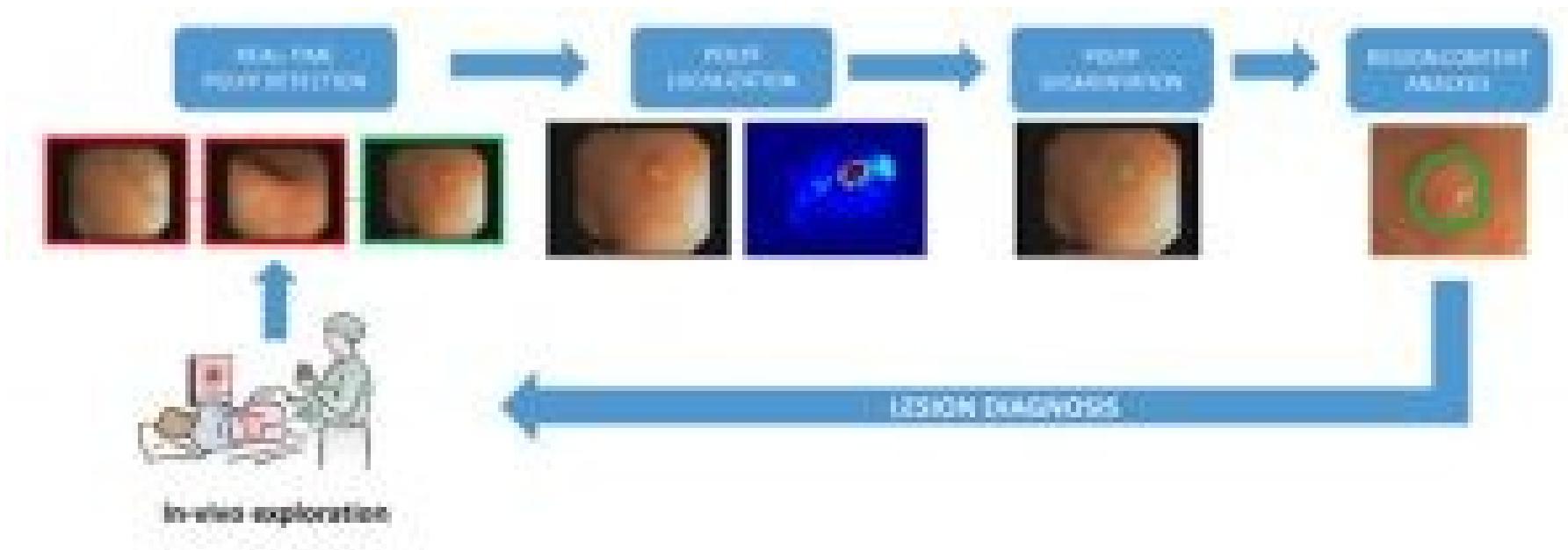
[Book a Demo](#)

[Home](#) [API](#) [Blog](#) [Pricing](#) [Sign in](#)

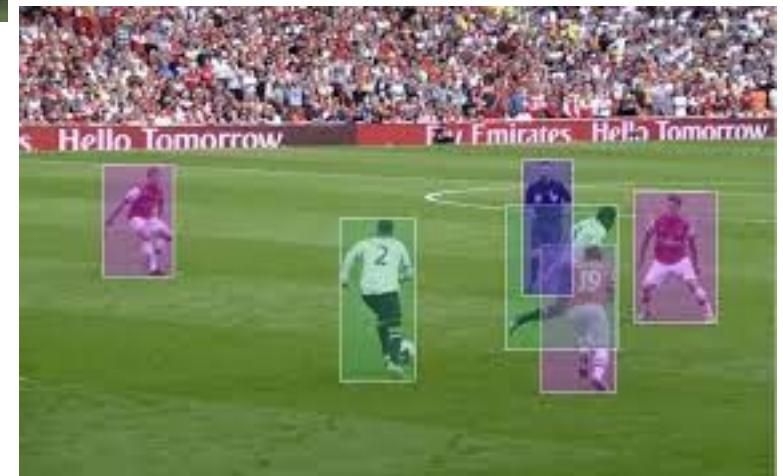
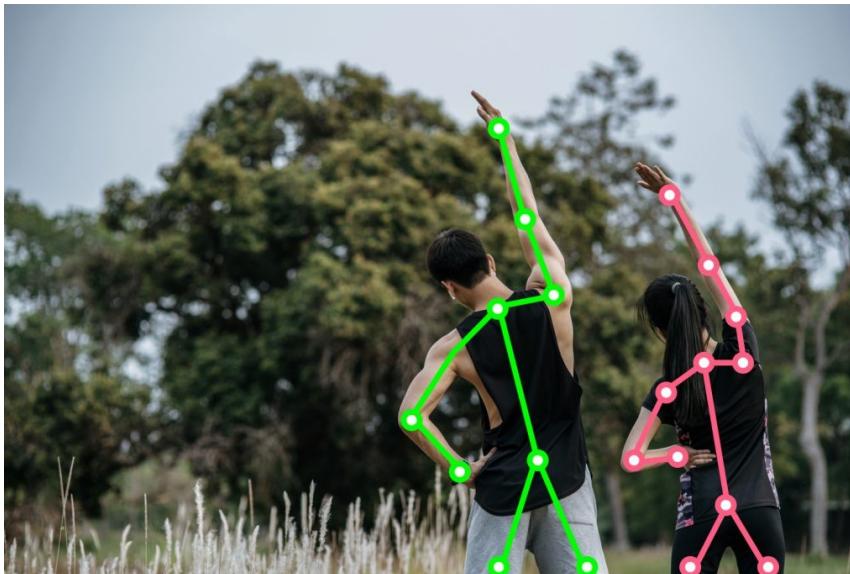
The image shows a screenshot of the Nyckel API dashboard. On the left, there's a large blue header area with the text "Lightning Fast Computer vision API for developers". Below this, a button says "Try it for Free →". In the center, there are three examples of food images with their classification results: "Dessert" (99.8%), "Dinner" (98.3%), and "Appetizer" (97.1%). To the right of these examples is a bar chart showing the accuracy of the API across different categories. The categories and their accuracy percentages are:

Category	Count	Accuracy (%)
All	50 of 111	98.0%
Dessert	36	100.0%
Appetizer	18	100.0%
Dinner	42	100.0%
Drinks	34	97.6%

# Computer Vision in Healthcare



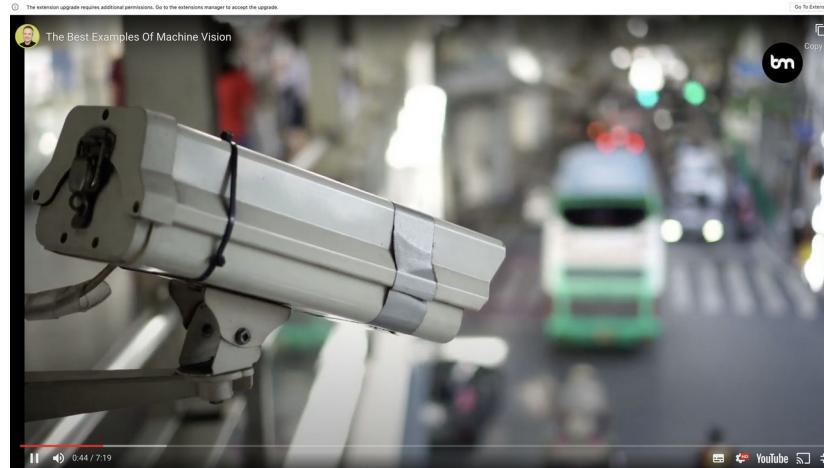
# Human pose estimation



# How Computer Vision Can Help Fight Climate Change

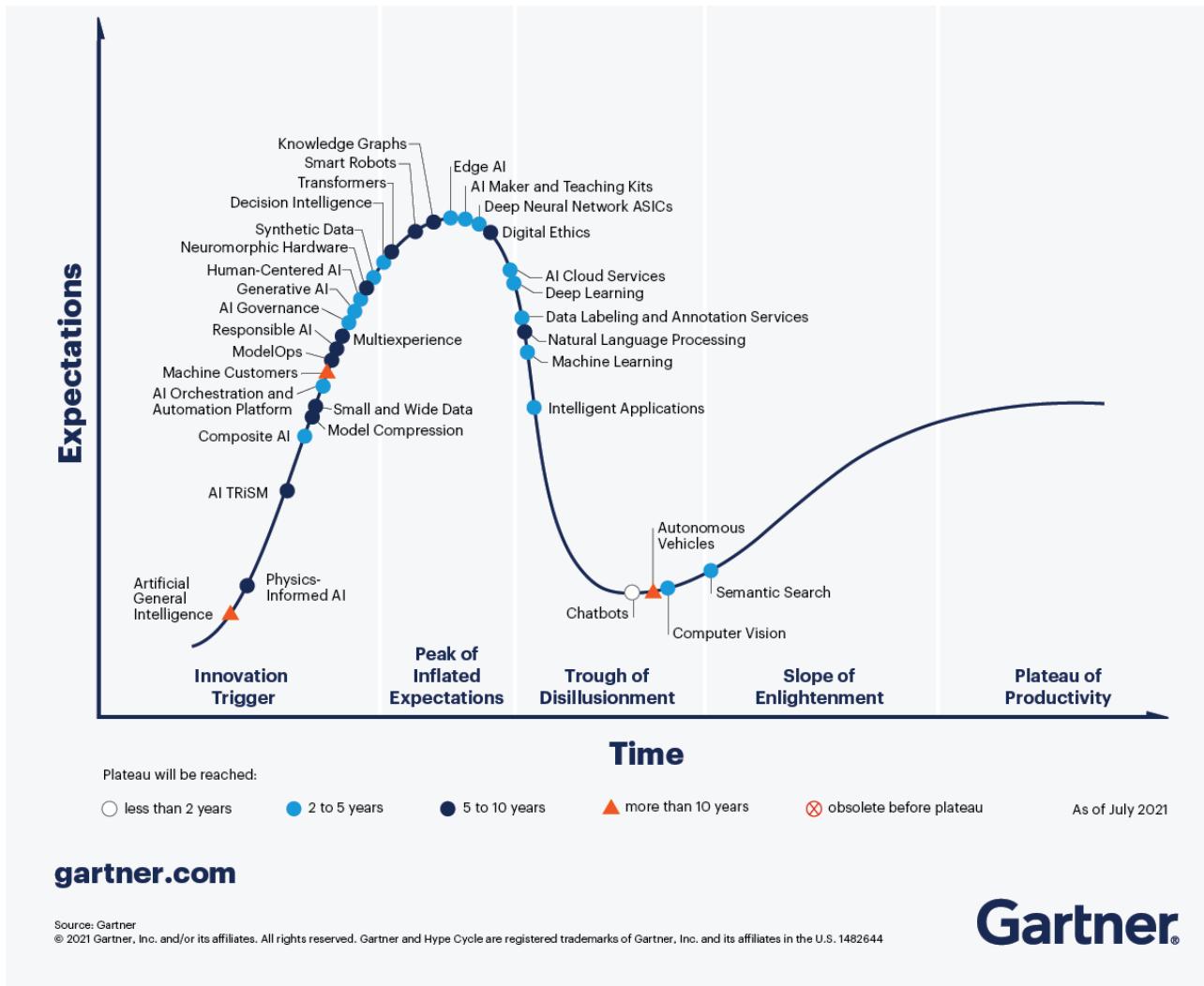


# The biggest Computer Vision applications

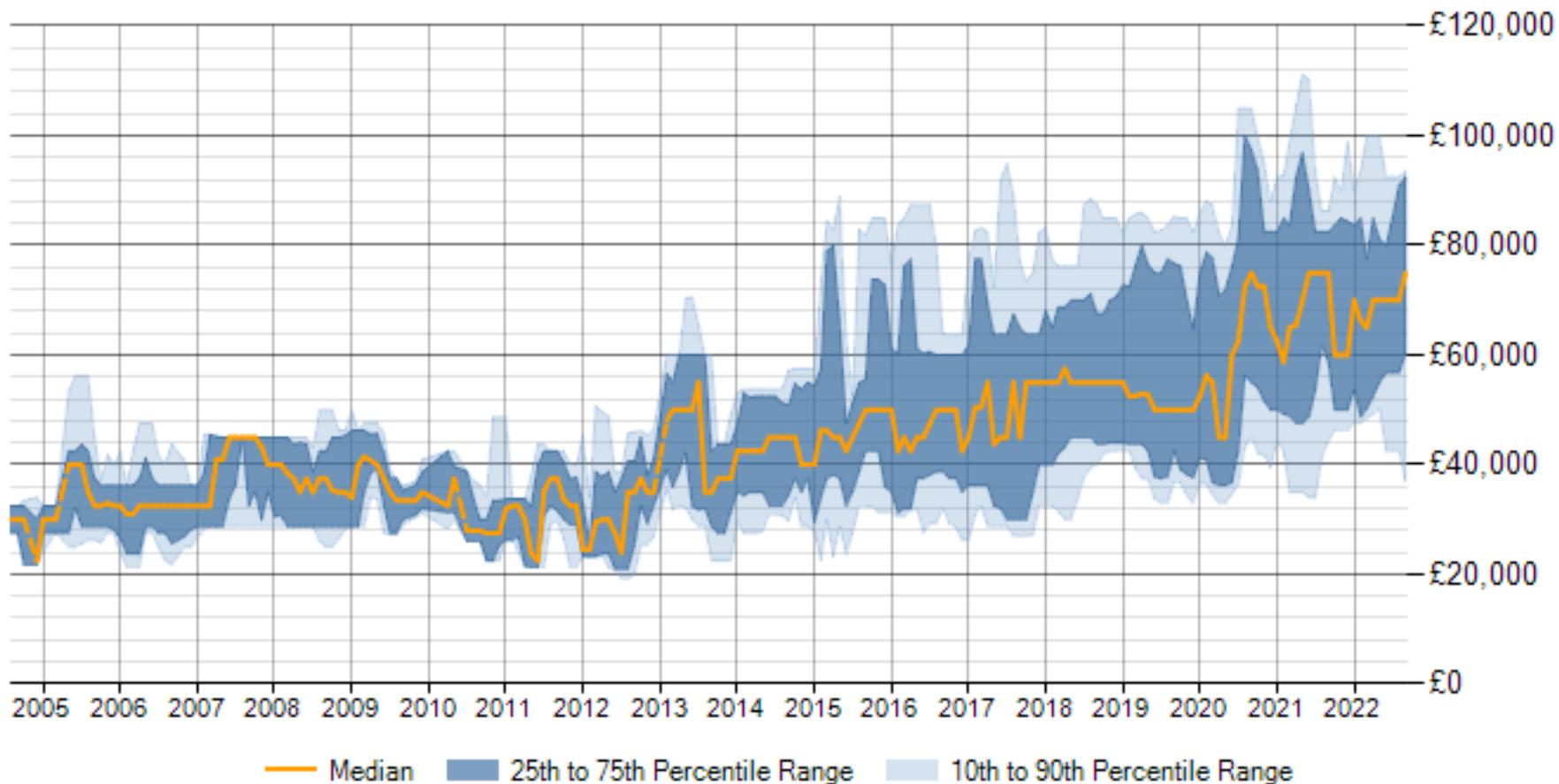


- [Everything happening in Computer Vision that you should know](#)
- [4 Steps to Start Machine Learning with Computer Vision](#)
- [Artificial Intelligence is expected to be a \\$60 billion industry by 2025. Join the industry by learning specialized skills in the most transformative AI field: Computer Vision.](#)
- [Top 10 Computer Vision Frameworks You Need To Know In 2022](#)
- [Augmented Intelligence: AR, VR, MR](#)
- [27+ Most Popular Computer Vision Applications and Use Cases in 2022](#)
- [Computer Vision Applications: A Detailed Look](#)
- [TOP 100 MACHINE LEARNING PROJECT IDEAS FOR TECH ENTHUSIASTS](#)
- [Computer Vision: Applications in Manufacturing, Surgery, Traffic, Satellites, and Unlabelled Data Recognition](#)
- [Computer Vision for Recognition of American Sign Language](#)
- [Real-Time Computer Vision Applications](#)
- [The 87 Most Popular Computer Vision Applications for 2023](#)
- [Everything happening in Computer Vision that you should know](#)
- [ETHICAL ISSUES IN TOPICAL COMPUTER VISION APPLICATION](#)
- [Top 7 Computer Vision Applications in Marketing](#)
- [What is Computer Vision? The Ultimate Guide](#)
- [\(Computer\) Vision without Sight](#)
- [Top Emerging Computer Vision Trends 2022](#)

# Hype cycle for Artificial Intelligence, 2021

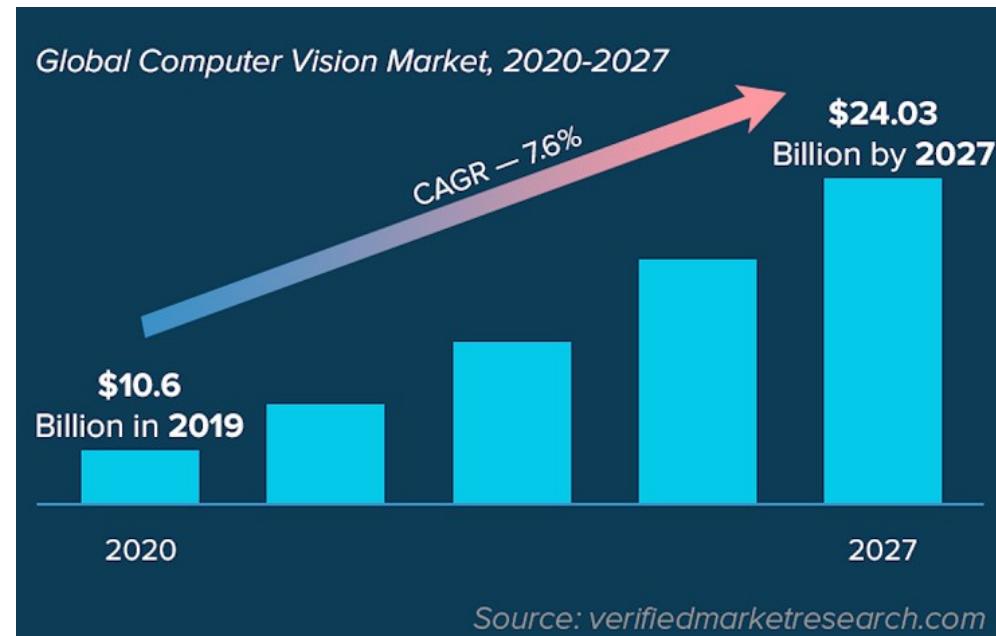
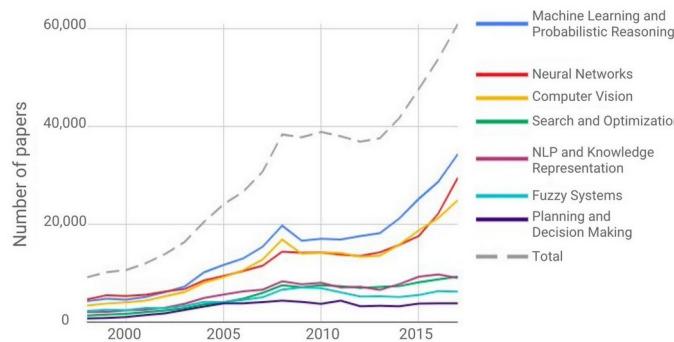


# SALARY OF CV EXPERTS



# Number of AI papers in Scopus

Number of AI papers on Scopus by subcategory (1998–2017)  
Source: Elsevier



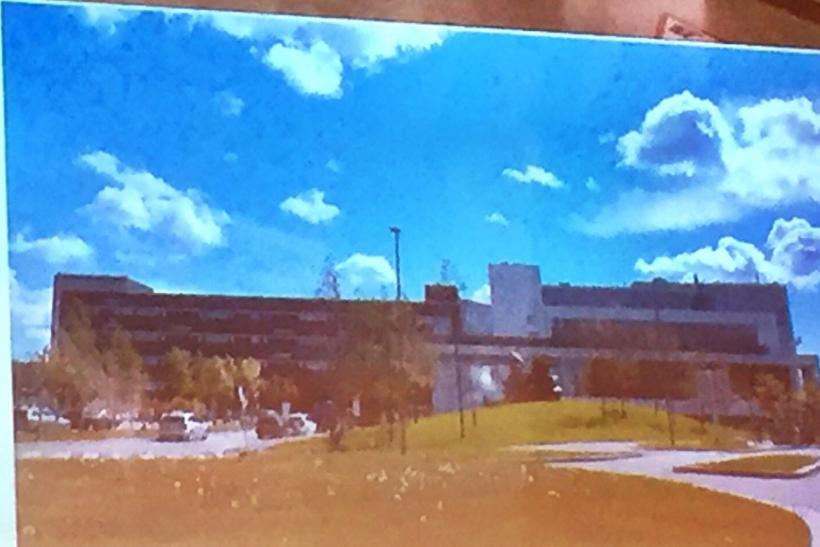
# Some more references of interest



- <https://www.youtube.com/watch?v=Tzv9T8wTykU>
- <https://www.youtube.com/watch?v=PT7Q9u9m5es>
- <https://www.youtube.com/watch?v=ty5QtslaAHM>
- <https://www.youtube.com/watch?v=XZAwtbkELRA>
- <https://www.youtube.com/user/CVOwebsite>
- [https://www.youtube.com/watch?v=\\_uwZrzLFizE](https://www.youtube.com/watch?v=_uwZrzLFizE)
- <https://www.youtube.com/watch?v=H4arwvdC7Z4>
- <https://www.youtube.com/watch?v=NHThkxUP-S8>
- <https://www.youtube.com/watch?v=1Ugo2KEV2XQ>
- <https://www.youtube.com/watch?v=5rQBH1TH9pA>



# Robots in hospitals



# Internet of Things

**Robotic surgery is already in the mainstream of modern surgery**



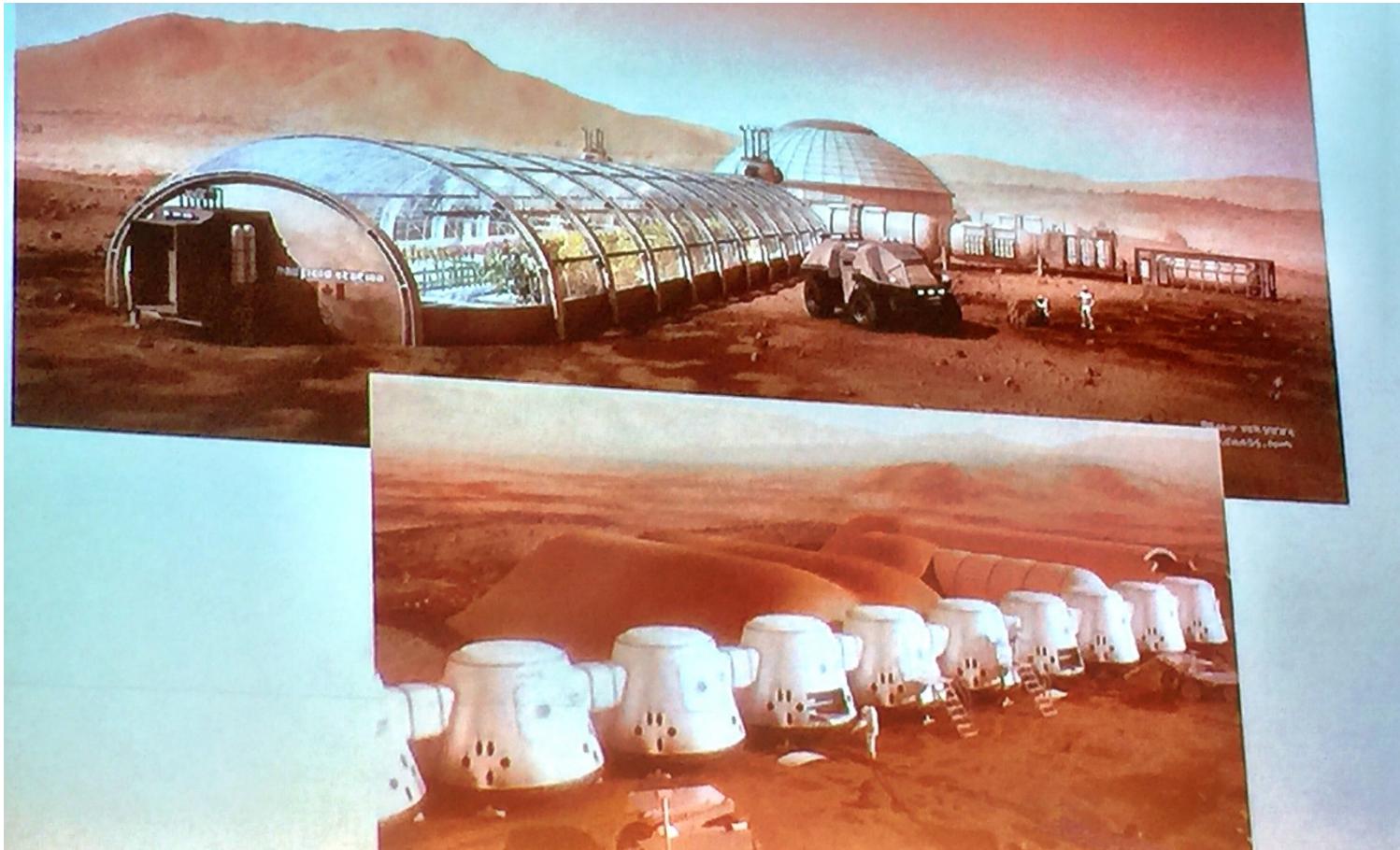
**Interventions likely to be deployed via IOT robots early**

- Physiotherapy for paralyzed patients (including exoskeletons)
- Drone based emergency response (e.g., defibrillation)
- Home assistants for senior citizens - including medication dispensation

Strictly Private & Confidential

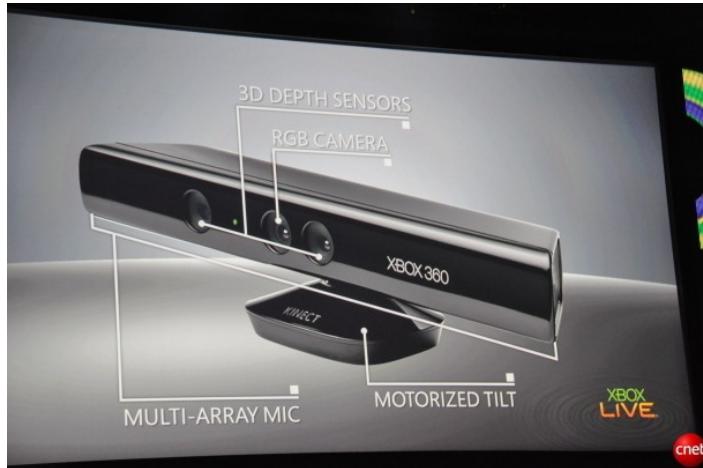
Manipal Hospital

# Visiting Mars

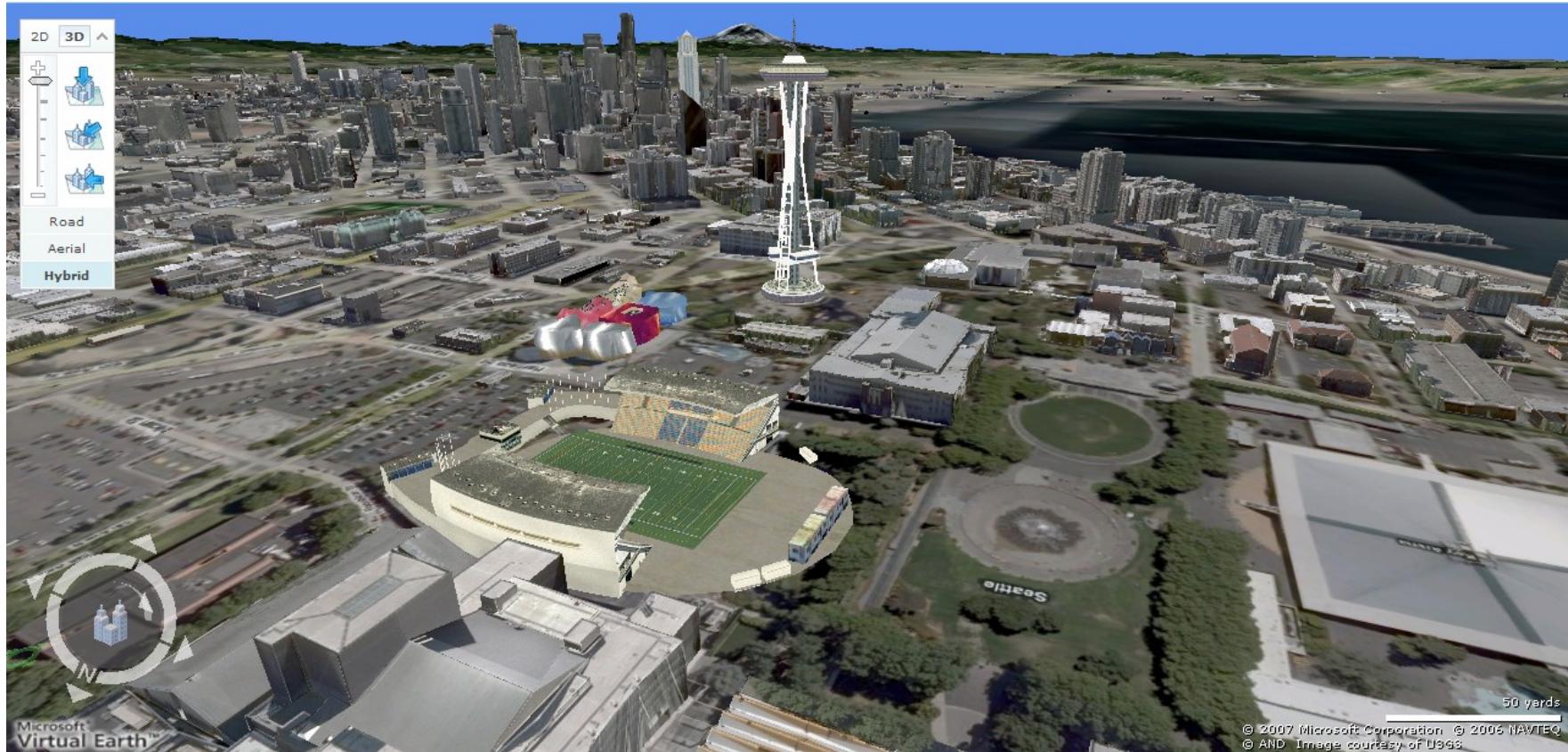


# Interactive Games

- Object Recognition
- Mario
- 3D
- Robot, etc.



# 3D viewers



Earth viewers (modelatge 3D), Imatge des de Microsoft's [Virtual Earth](#)  
(veure també: [Google Earth](#))

# Photosynth

The screenshot shows the Microsoft Live Labs Photosynth website. At the top left is the Photosynth logo with a stylized green leaf icon. To its right, the text "Microsoft Live Labs" and "Photosynth" is displayed. On the left side of the main content area, there is a sidebar with a blue background containing a navigation menu:

- Home
- Try it
- What is Photosynth?
- Collections
- Team blog
- Videos
- System requirements
- About us
- FAQ

The main content area features a large image of a cathedral with multiple domes and intricate architecture, set against a dark background with glowing particles. Above the image is a quote:  
*"What if your photo collection was an entry point into the world,  
like a wormhole that you could jump through and explore..."*

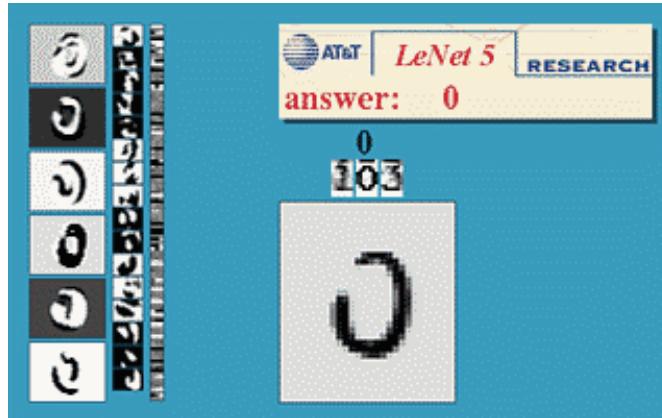
Below the image is a small "Try it" button. At the bottom of the main content area is a yellow "Try the Tech Preview" button.

The text below the image describes the Photosynth Technology Preview: "The Photosynth Technology Preview is a taste of the newest - and, we hope, most exciting - way to **view photos** on a computer. Our software takes a large collection of photos of a place or an object, analyzes them for similarities, and then displays the photos in a reconstructed **three-dimensional space**, showing you how each one relates to the next."

<http://photosynth.net/>  
Photo Tourism technology

# Optical character recognition (OCR)

- Technology to convert scanned docs to text
  - If you have a scanner probably came with OCR software



Digits recognition, AT&T labs

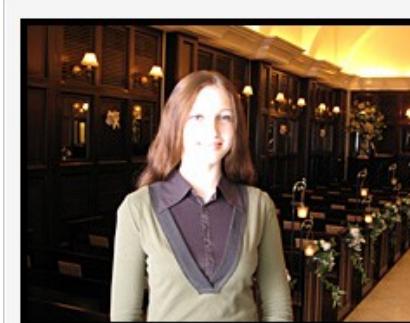
4YCH428  
4YCH428  
4YCH428



Plates recognition

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

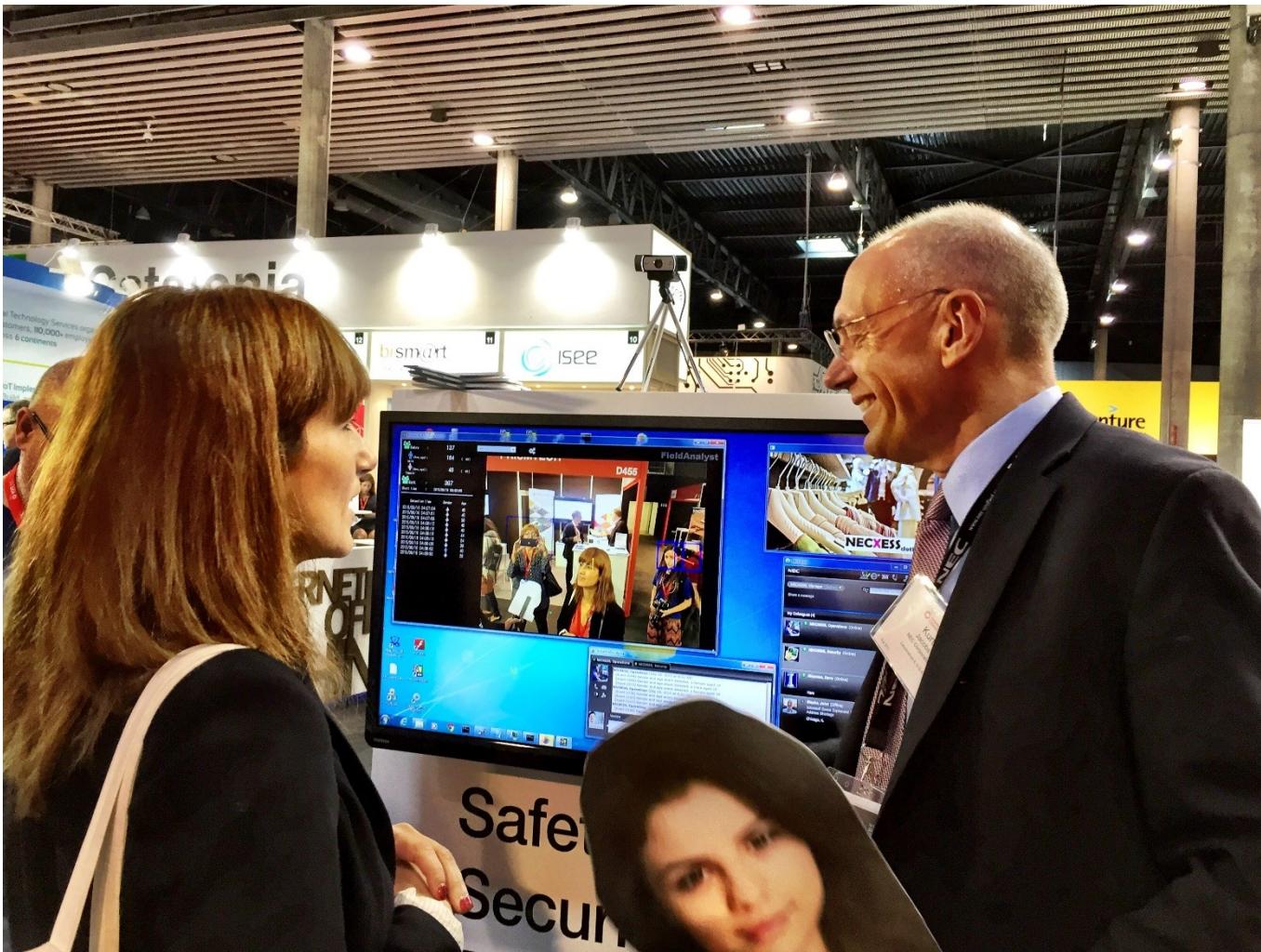
# Face detection



[Face priority AE] When a bright part of the face is too bright

- Almost all digital cameras: Canon, Sony, Fuji, ...

# Face analysis



NEC Corporation

# Google maps



barcelona

Buscar en Maps

Mostrar opciones de búsqueda

[Imprimir](#) [Enviar](#) [Enlazar](#)636 Gran Via de les Corts Catalanes, Barcelona, CT, España  
La dirección es aproximada.

Haz doble clic para ir

VITALICIO SEGUROS



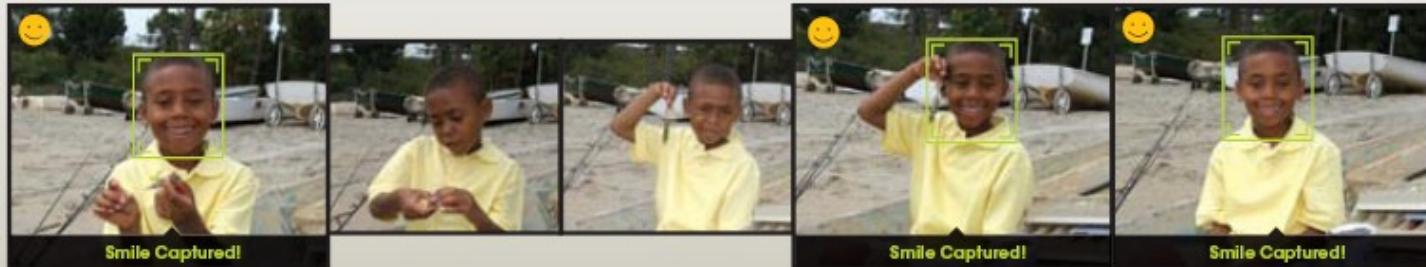
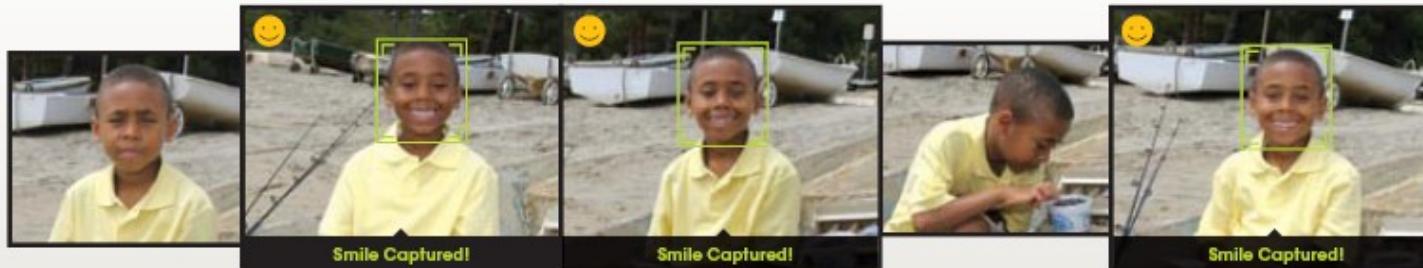
# Picasa



# 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](#)

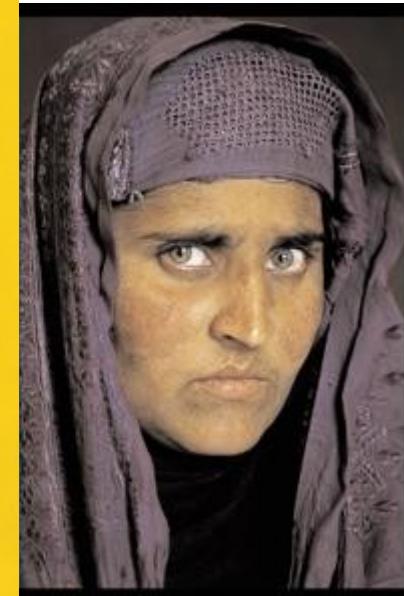
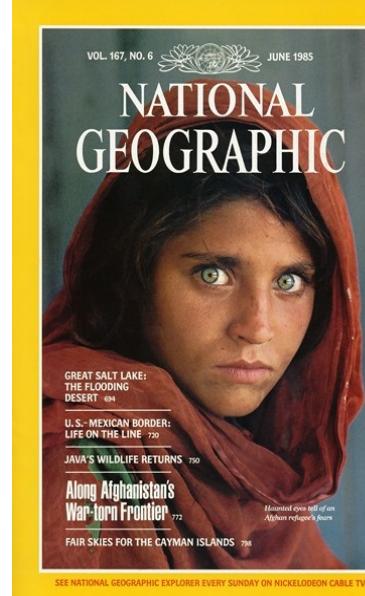
# Object recognition in supermarkets



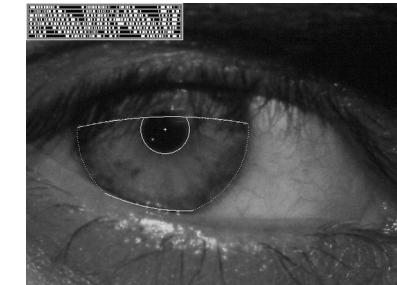
## LaneHawk by EvolutionRobotics

"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



"As an Afghan girl was identified by his employer iris" ,... Read the story

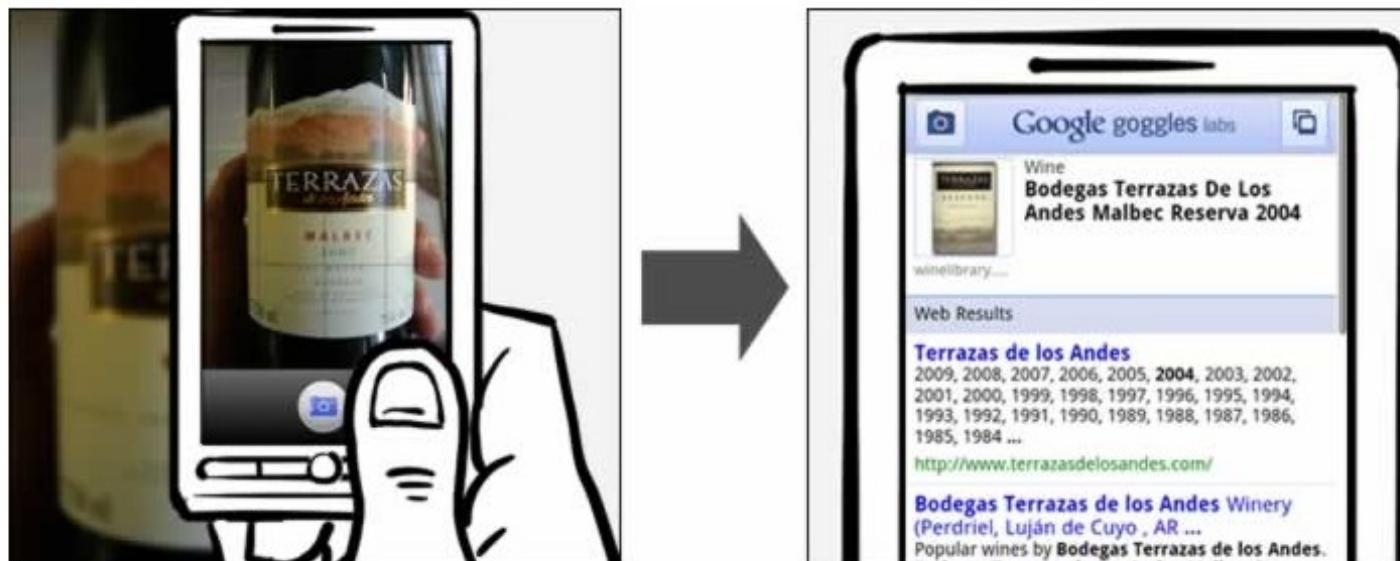


[Discover the story about the Afghanistan's girl and iris recognition](#)

# 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

# Mobile visual search: iPhone Apps

*Query Images*



Perspective



Zoom



Rotation



Coverage



Lighting



Occlusion



Blur



Zoom

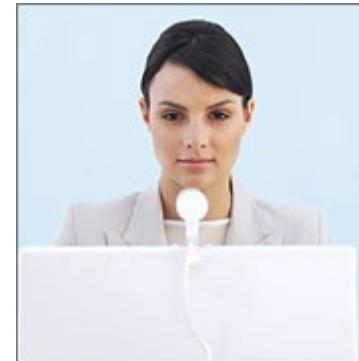
*Matched Image*



# Login without password...



Fingerprints for many  
portable scanners or other  
devices



[Face recognition](#)

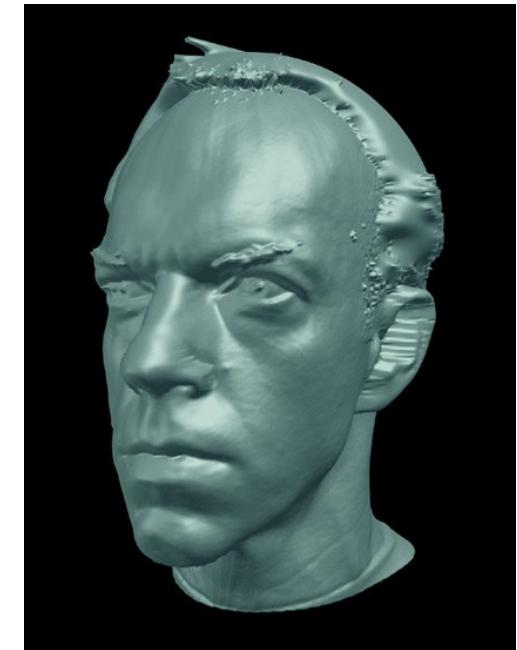
# Mobile recognition



- Microsoft Research
  - Point & Find, Nokia

Lincoln

# Visual effects: shape capture



Matrix, ESC Entertainment, XYZRGB, NRC

# Visual effects: shape capture



*Pirates of the Caribbean*, Industrial Light and Magic

# Sport



Sportvision first down line

# Intelligent cars

►► manufacturer products      consumer products ◀◀

## Our Vision. Your Safety.

rear looking camera      forward looking camera      side looking camera

**EyeQ** Vision on a Chip

> read more

**Vision Applications**  
Road, Vehicle, Pedestrian Protection and more

> read more

**AWS** Advance Warning System

> read more

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](#)

Events

> [Mobileye at Equip Auto, Paris, France](#)

> [Mobileye at SEMA, Las Vegas, NV](#)

> [read more](#)

- Mobileye

- Vision systems in BMW, GM, Volvo, Volkswagen
- In 2010: 70% of the manufactured cars.
- video

# Google cars



- Oct 9, 2010. "[Google Cars Drive Themselves, in Traffic](#)". [The New York Times](#). John Markoff
- June 24, 2011. "[Nevada state law paves the way for driverless cars](#)". [Financial Post](#). Christine Dobby
- Aug 9, 2011,  
["Human error blamed after Google's driverless car sparks five-vehicle crash"](#). [The Star \(Toronto\)](#)

# Vision based interaction



Nintendo Wii. Work of [Lee's](#) from CMU how to use:  
multi-touch display!



Digimask: put the face of a person to the avatar.



["Game turns moviegoers into Human Joysticks"](#), CNET  
A camera following the audience:[aquest treball](#).

# Vision in the 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) for:

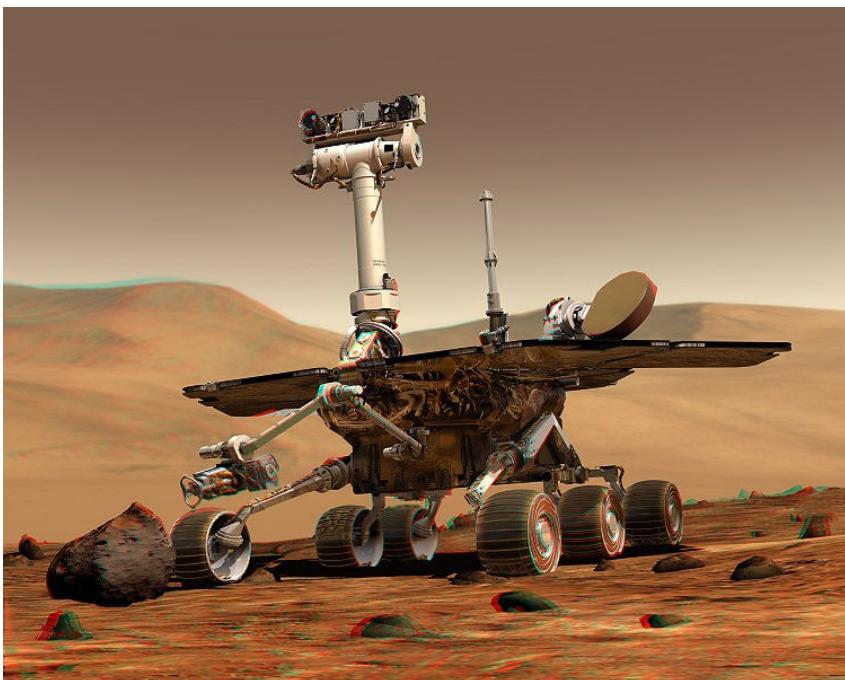
- Panorama stitching
- 3D Terrain Modeling
- Obstacle detection, tracking the position
  - “[Computer Vision on Mars](#)” by Matthies et al.

# Industrial robots

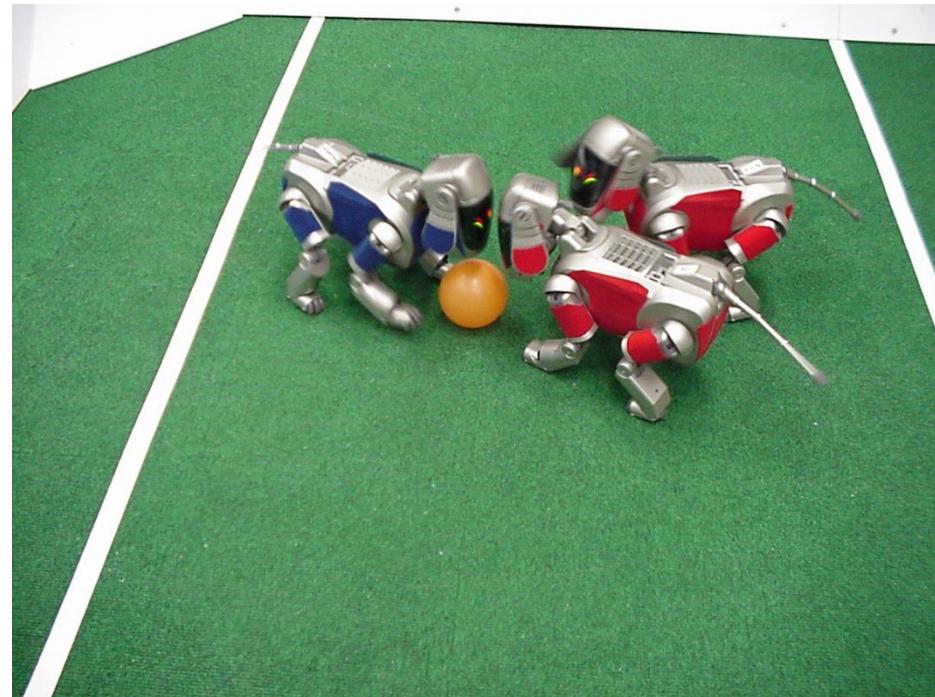


- Vision-guided robots position nut runners on wheels

# Robotics



NASA's Mars Spirit Rover  
[http://en.wikipedia.org/wiki/Spirit\\_rover](http://en.wikipedia.org/wiki/Spirit_rover)

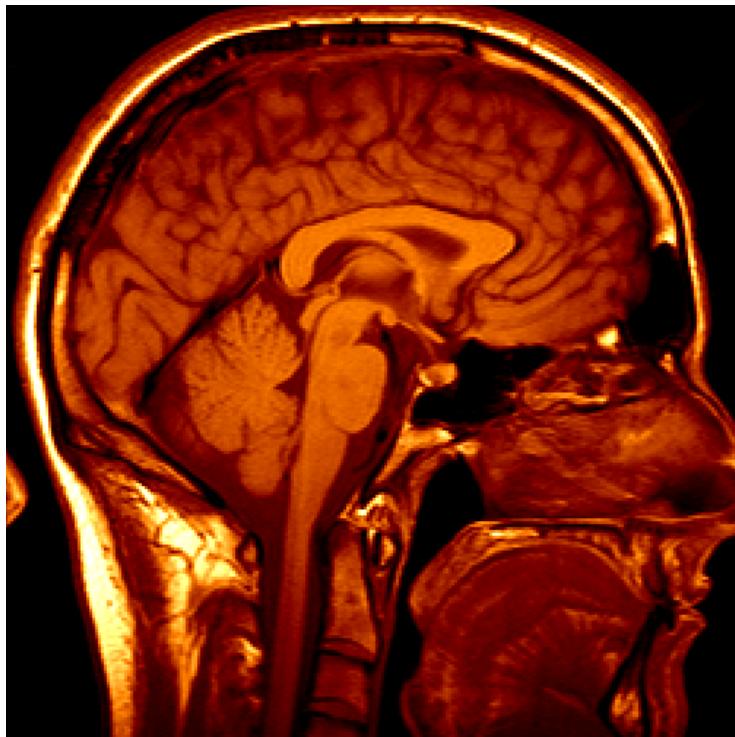


<http://www.robocup.org/>

# Elder people assistance



# Medical diagnosis and intervention



3D Images  
MRI, CT



[Grimson et al., MIT](#)

# Photo Tourism

The screenshot shows a Microsoft Internet Explorer browser window with the URL <http://phototour.cs.washington.edu/>. The page title is "Photo Tourism" with the subtitle "Exploring photo collections in 3D". On the left, there is a logo for the University of Washington Computer Science & Engineering department. In the center, there are three main visual components: (a) a grid of thumbnail images of various buildings, (b) a sparse 3D point cloud model of a cathedral, and (c) a 3D viewer showing a detailed 3D reconstruction of the Notre Dame Cathedral. A Microsoft logo is visible in the top right corner of the browser window.

Photo Tourism  
Exploring photo collections in 3D

University of Washington  
Computer Science & Engineering

(a)

(b)

(c)

Microsoft

Live Demo

\*New!\* See our work on **Finding Paths through the World's Photos**.

Our structure from motion code is also now available at the **Bundler** homepage.

# Snaptell

The screenshot shows a web browser window displaying the <http://snaptell.com/demos/DemoLarge.htm> page. The page has a light blue header bar with the SnapTell logo and navigation links: blog, partners, contact us, newsletter, and customer login. Below the header is a navigation menu with links: Solution, Technology, Customers, Campaigns, News & Events, and Company. A breadcrumb trail indicates the user is at Home > Demo. The main content area features a large orange header with the word "SnapTell". Underneath, there are two sections: "iPhone Application" and "With any Camera Phone". The "iPhone Application" section describes the SnapTell Explorer app available on the AppStore. The "With any Camera Phone" section explains how users can snap a picture of a product cover and send it to fun@snaptell.com for instant information. At the bottom of the page is a promotional image for Michael Franti & Spearhead's album "All Rebel Rockers".

**SnapTell**

blog | partners | contact us | newsletter | customer login

Solution | Technology | Customers | Campaigns | News & Events | Company

Home > Demo

## SnapTell

**iPhone Application**

Our new iPhone application SnapTell Explorer is on the AppStore. Read more on our blog. Great descriptions and reviews have been posted at AppVee and TMCnet.

**With any Camera Phone**

**Snap** a picture of the cover of any DVD, CD, Book or Video game, **Send** it to fun@snaptell.com, and within seconds **Get** information on the product. Our image matching database includes millions of cover pictures.

Michael Franti & Spearhead - All Rebel Rockers

Michael Franti &  
**SPEARHEAD**  
ALL REBEL ROCKERS