

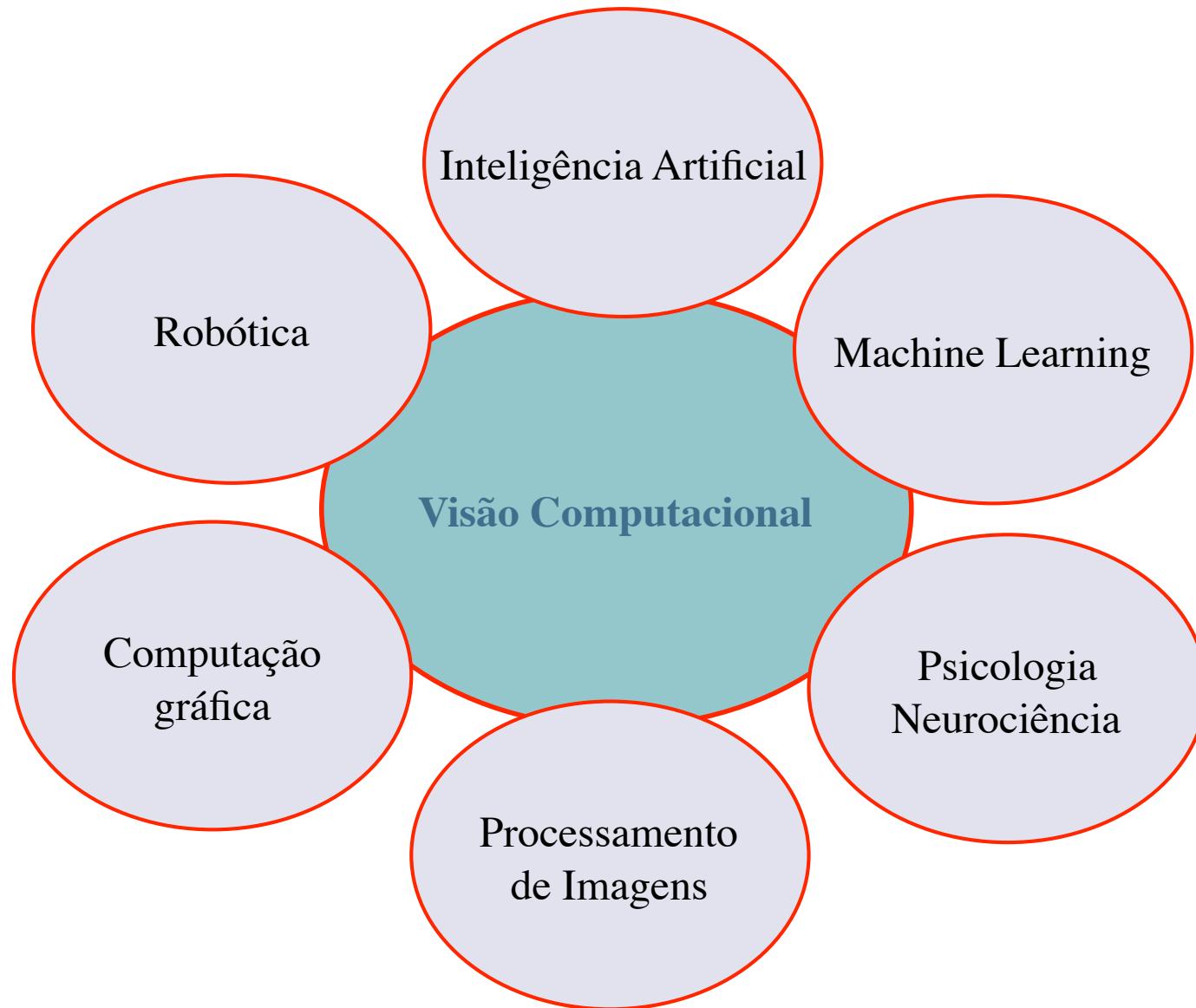
# Visão Computacional

## Aula 01

Introdução

Computer Vision - A Modern Approach  
Set: Introduction to Vision  
Slides by D.A. Forsyth  
Adapted by Flávio Vidal

# O que é Visão Computacional?



# Por quê estudar Visão Computacional?

- Imagens e filmes estão por toda parte;
- Crescimento exponencial das aplicações envolvendo o tema:
  - Representação de edifícios em 3D a partir de imagens;
  - Vigilância automatizada;
  - Pós-processamento de filmes (Ex.: Efeitos especiais);
  - Detecção de faces;
- Diversos e atrativos “mistérios científicos”
  - Como o processo de reconhecimento de objetos funciona?
- Entendimento das complexidades da visão humana;

# Propriedades da Visão Computacional

- Vislumbrar aplicações “no futuro”:
  - Evitar acidentes em rodovias federais:
    - Monitorar a direção do veículo, e em caso de situações de perigo (ex.: choque frontal) tomar a direção do carro para evitar o acidente;
  - Monitorar fluxo veicular em grandes centros
    - Gerenciar os semáforos de grandes cruzamentos e vias urbanas expressas de acordo com o fluxo, para minimizar o “engarrafamento”;

# Propriedades da Visão Computacional

- Representações 3D são facilmente obtidas:
  - Existem diversas metodologias;
  - Viável
    - Para humanos (desvio de obstáculos; planejamento de captura de objetos, etc...)
    - Construção de modelos para filmes;
  - Essas metodologias incluem:
    - Multiple views (movimento, estereoscopia);
    - Textura;
    - Projeção de sombras (shading);

# Propriedades da Visão Computacional

- Seres humanos facilmente conseguem “ver” distinção:
  - “Reconhecimento de diversos objetos”;
  - Diferenças entre um “peixe” e uma “bicicleta”;
  - Uma pessoa famosa;
  - Situações de risco;
  - Grande mistério:
    - Como construir apps capazes de distinguir estas situações baseadas somente na informação visual (imagem)?

# Divisão da Visão Computacional

- Dividido em sete grandes partes:
  - Características da formação (“físicas”) das Imagens
    - Câmeras;
    - Iluminação;
    - Cores
  - Visão Computacional para “uma” imagem
    - Representação de pequenas partes da imagem;
    - Influência da variação de sombreamento;
    - Representação de Textura;
  - Visão Computacional para “várias” imagens
    - Múltipla vistas;
    - Estereoscopia;
    - Estrutura para Movimento;

# Divisão da Visão Computacional

- Dividido em sete grandes partes:
  - Visão Computacional em Nível Intermediário
    - Segmentação
    - Rastreamento
  - Visão Computacional em Alto Nível
    - Geométrico
    - Probabilístico
  - Reconstrução 3D de várias “vistas”
  - Reconhecimento Padrões



# Alguns Exemplos...

(by University of California @ Merced)

# Earth viewers (3D modeling)

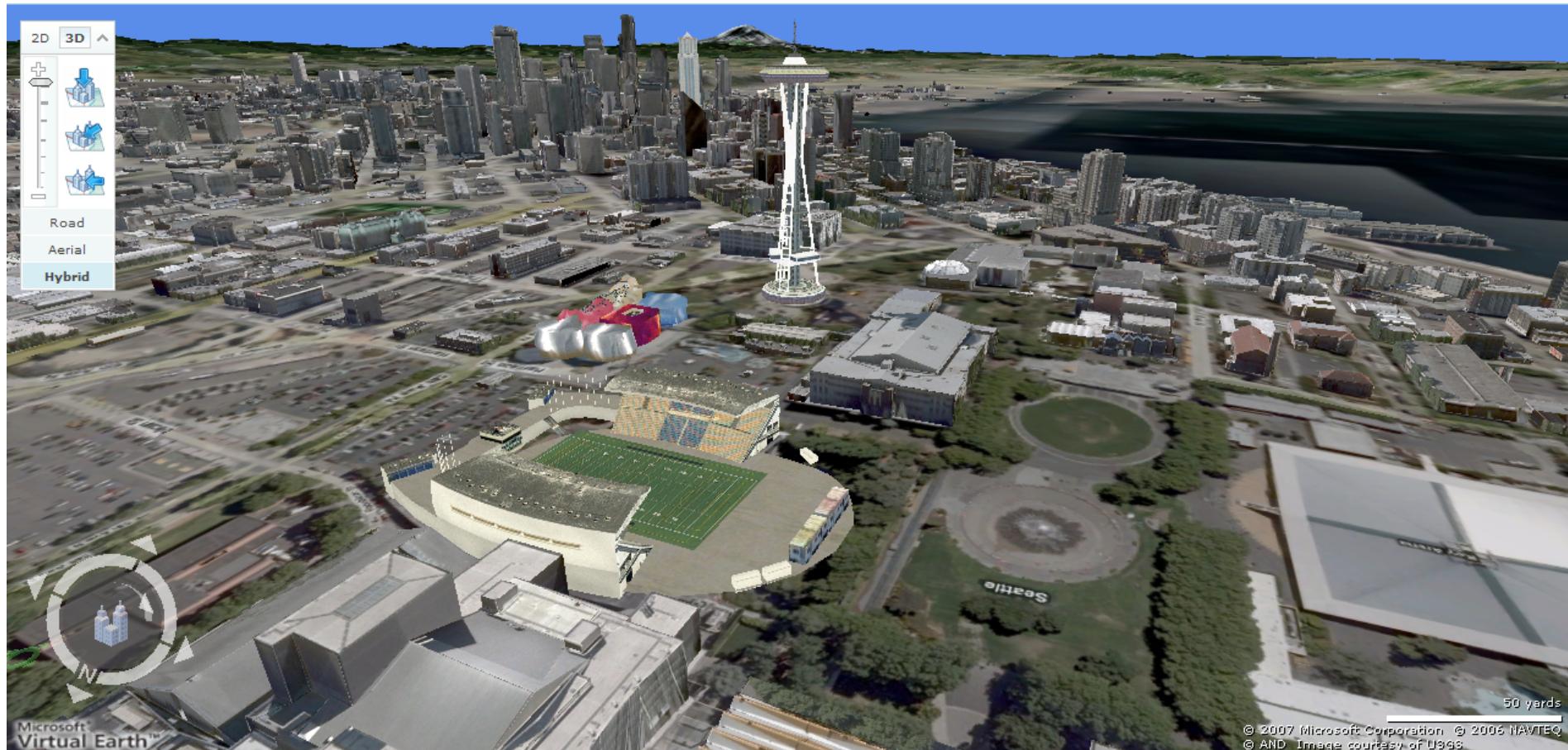
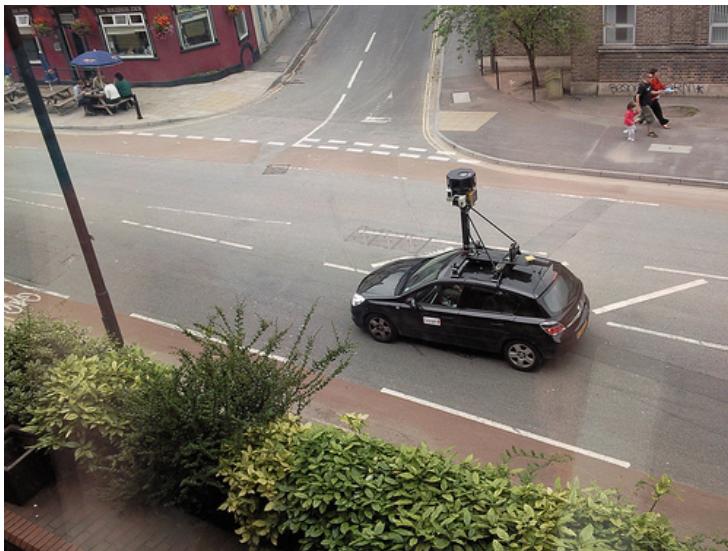
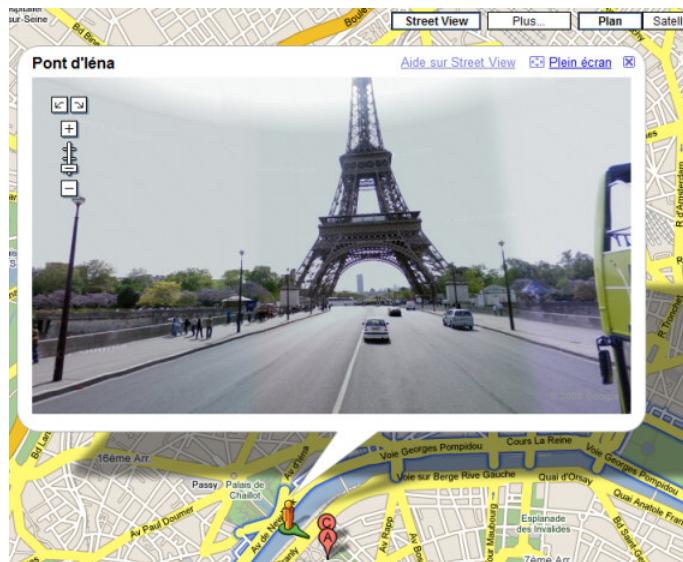


Image from Microsoft's [Virtual Earth](#)  
(see also: [Google Earth](#))

# Google streetview



# Photosynth

The screenshot shows the Microsoft Live Labs Photosynth website. At the top left is the Photosynth logo with a green leaf icon. The main menu on the left includes links for Home, Try it, What is Photosynth?, Collections, Team blog, Videos, System requirements, About us, and FAQ. The central content area features a quote: "What if your photo collection was an entry point into the world, like a wormhole that you could jump through and explore..." with a "Try it" button. Below this is a 3D reconstruction of St. Mark's Basilica in Venice, Italy, surrounded by trees. A large orange "Try the Tech Preview" button is at the bottom. A descriptive paragraph at the bottom explains the technology: "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://labs.live.com/photosynth/>

<http://www.youtube.com/watch?v=p16frKJLVi0>

# Optical Character Recognition (OCR)



4 YCH428



4 YCH428



4 Y C H 4 2 8

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

# Face detection

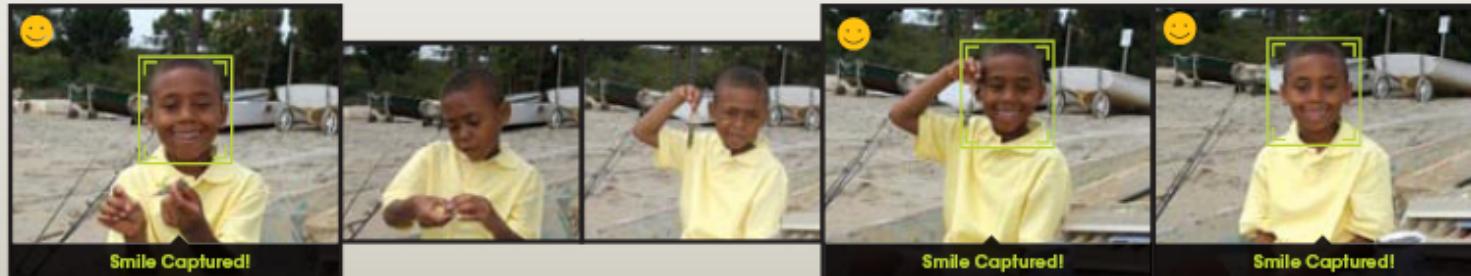
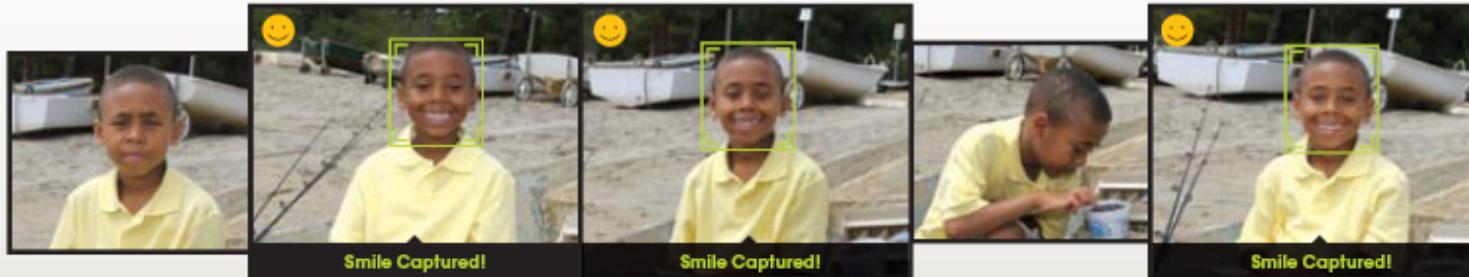


- Várias marcas de câmeras digitais
  - Canon, Sony, Fuji, ...

# Smile detection (by Sony)

## 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](#)

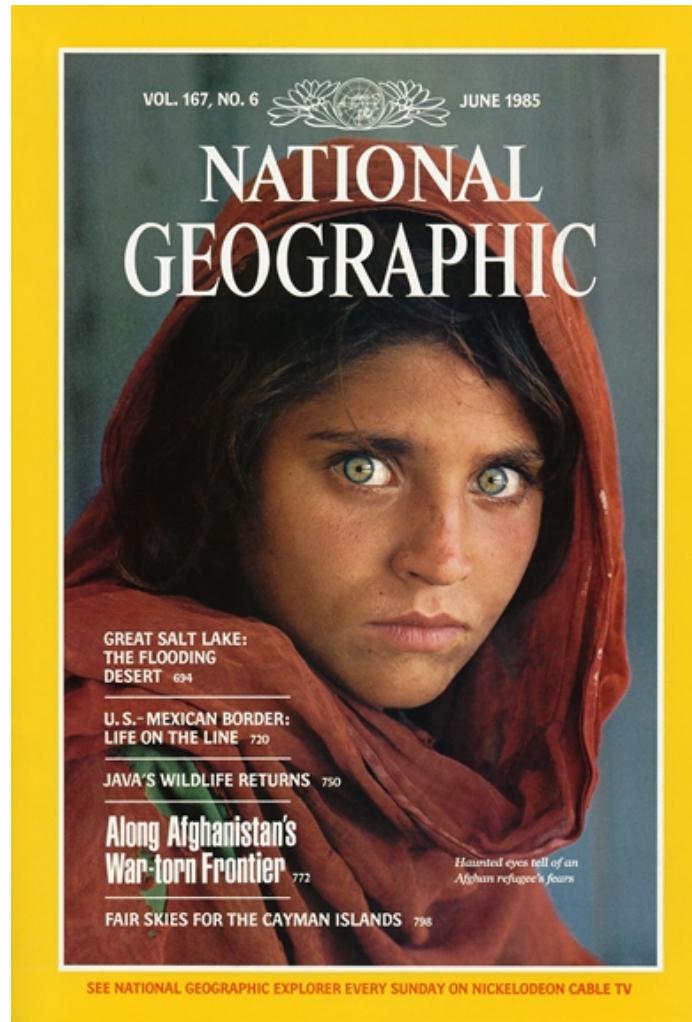
# Object Recognition (Ex.: Supermercados)



## 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...”

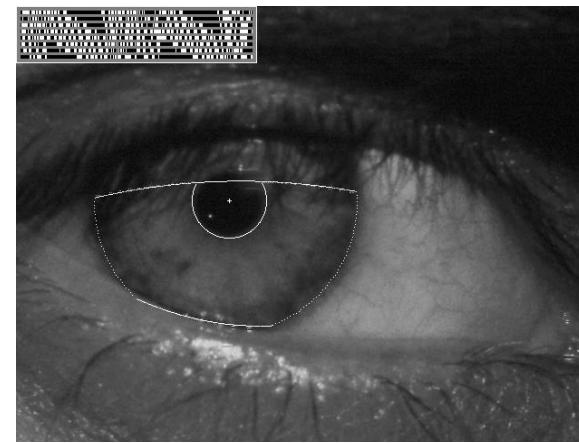
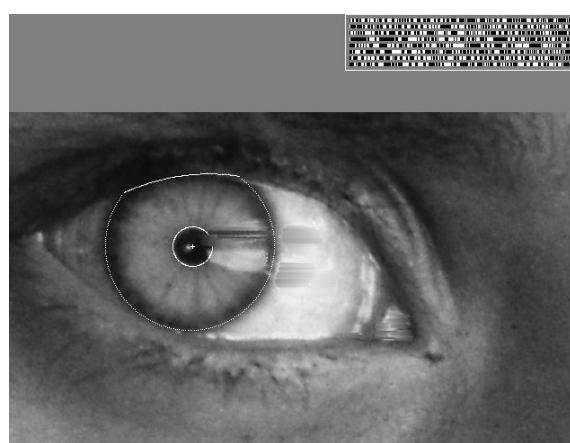
# Face recognition



# Vision-based biometrics



Identificada pelo padrão da Iris - <http://www.cl.cam.ac.uk/~jgd1000/afghan.html>



# Login sem um password...



<http://www.sensiblevision.com/>

# Object recognition (Mobile)



- Já é uma realidade...
  - **Lincoln** Microsoft Research
  - Point & Find, Nokia, NTT Docomo

# Efeitos Especiais



<http://www.youtube.com/watch?v=J5ryLMZTO5M>

# Efeitos especiais



*Pirates of the Caribbean*, Industrial Light and Magic <http://www.cl.cam.ac.uk/~jgd1000/afghan.html>

# Análise esportiva



<http://www.youtube.com/watch?v=UyPU2l9rdvo>

# Smart cars

The screenshot shows the Mobileye website homepage. At the top, there are two tabs: "manufacturer products" and "consumer products". Below them is a main heading "Our Vision. Your Safety." with a background image of a car from above, showing three cameras: "rear looking camera" (top left), "forward looking camera" (top right), and "side looking camera" (bottom). Below this are three sections: "EyeQ Vision on a Chip" (with an image of a chip), "Vision Applications" (with an image of a person walking), and "AWS Advance Warning System" (with an image of a dashboard display). To the right, there are two columns: "News" (listing articles like "Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System" and "Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end") and "Events" (listing "Mobileye at Equip Auto, Paris, France" and "Mobileye at SEMA, Las Vegas, NV").

- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers.
  - <http://www.mobileye.com/>

# Vision-based interaction (games)



# Vision-based HCI

- Reatrix: <http://www.youtube.com/watch?v=QzsQKULMbiU>



# Gaming

- Sony EyeToy



[http://www.youtube.com/watch?  
v=AOXohr4XE-4&feature=related](http://www.youtube.com/watch?v=AOXohr4XE-4&feature=related)

- Microsoft Natal (Kinect)



[http://www.youtube.com/watch?  
v=1BRSfCuLYHc](http://www.youtube.com/watch?v=1BRSfCuLYHc)

# Motion capture

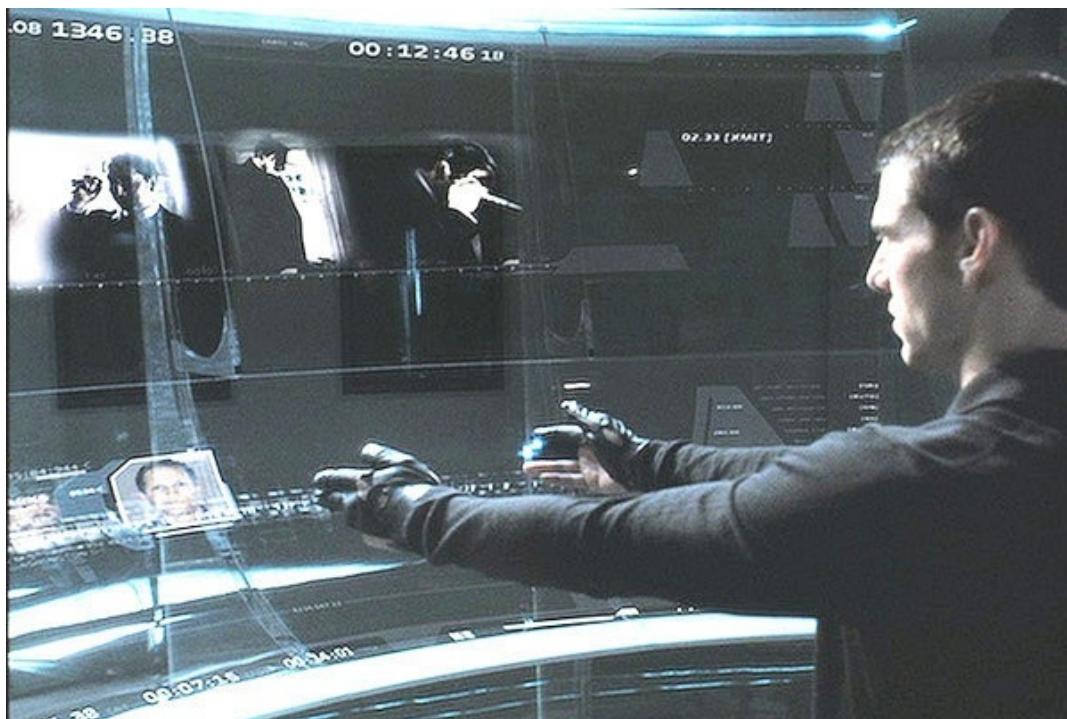
- Marker-based motion capture
  - <http://www.youtube.com/watch?v=VoyT8mwg9nc>
- Organic motion
- <http://www.organicmotion.com/>



# Looking at people

- Hand gesture
- Head pose
- Expression
- Identity

<http://www.youtube.com/watch?v=NwVBzx0LMNQ>



# Visão no espaço sideral...



[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

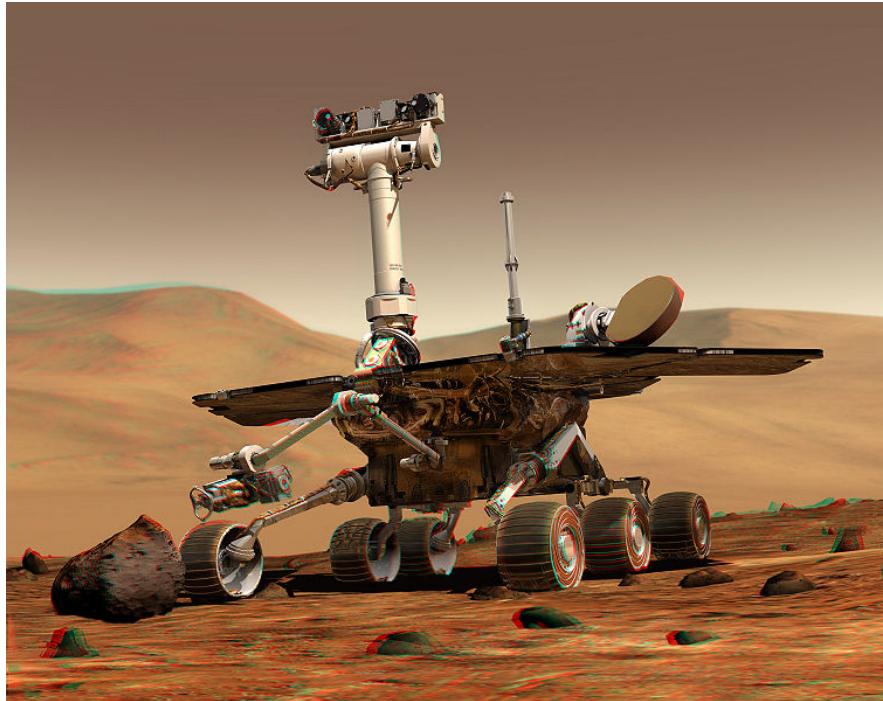
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking

# Gigapan

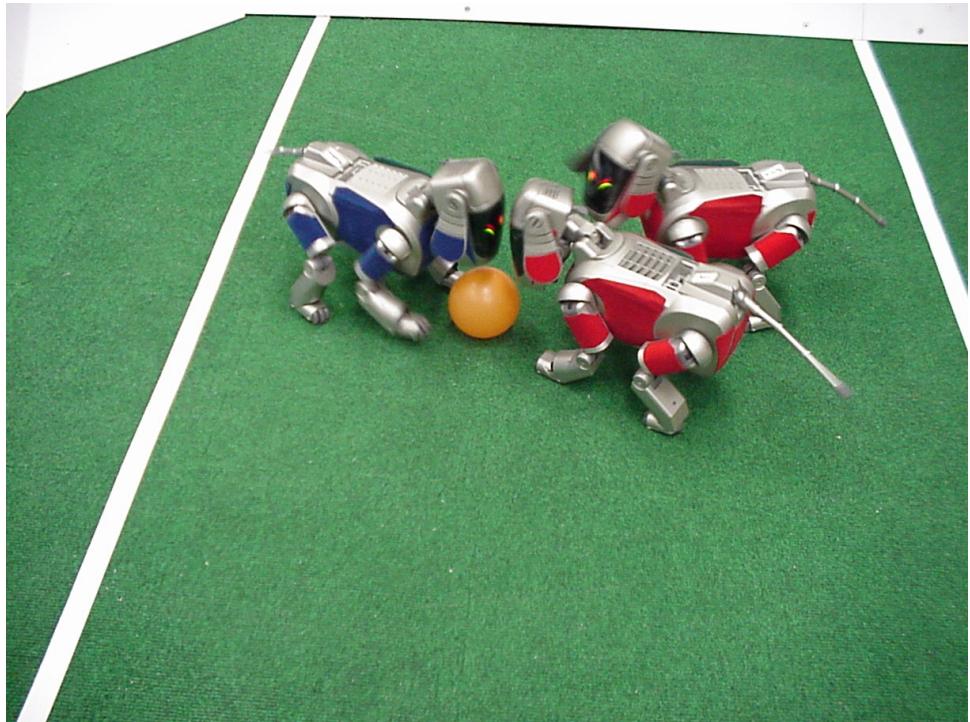
- <http://www.gigapan.org/index.php>



# Robótica

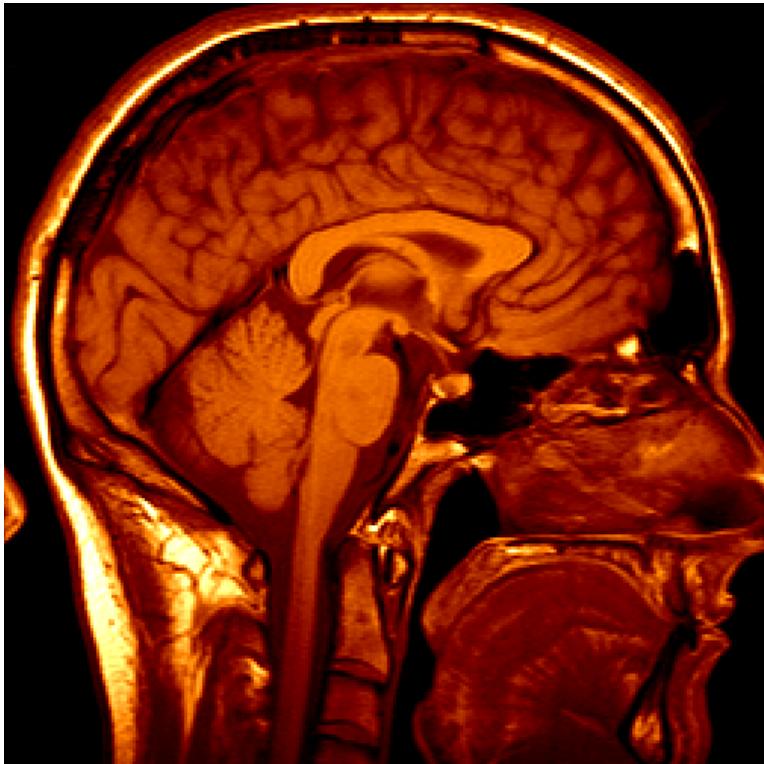


[http://en.wikipedia.org/wiki/Spirit\\_rover](http://en.wikipedia.org/wiki/Spirit_rover)



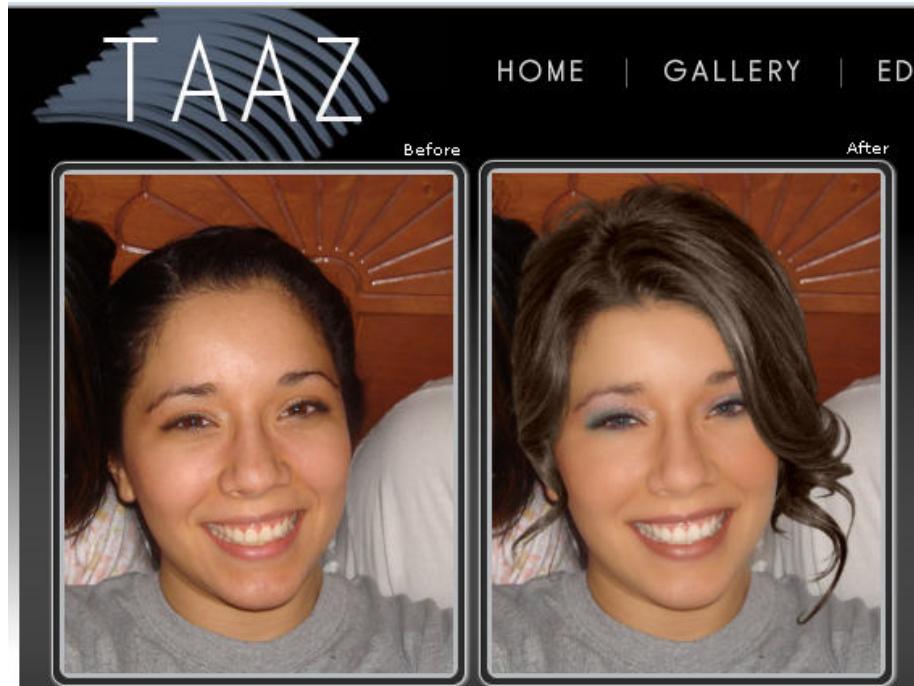
<http://www.robocup.org/>

# Medical imaging

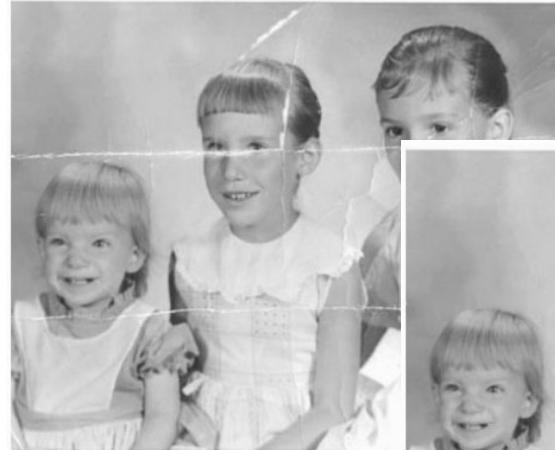


[Grimson et al., MIT](#)

# Digital cosmetics



# Inpainting



Since 1699, when French explorers landed at the great bend of the Mississippi River and celebrated the first Mardi Gras in North America, New Orleans has brewed a fascinating mélange of cultures. It was French, then Spanish, then French again; then sold to the United States. Through all these years, and even into the 1900s, others arrived from everywhere: Acadians (Cajuns), Africans, indige-



# Deblurring



=

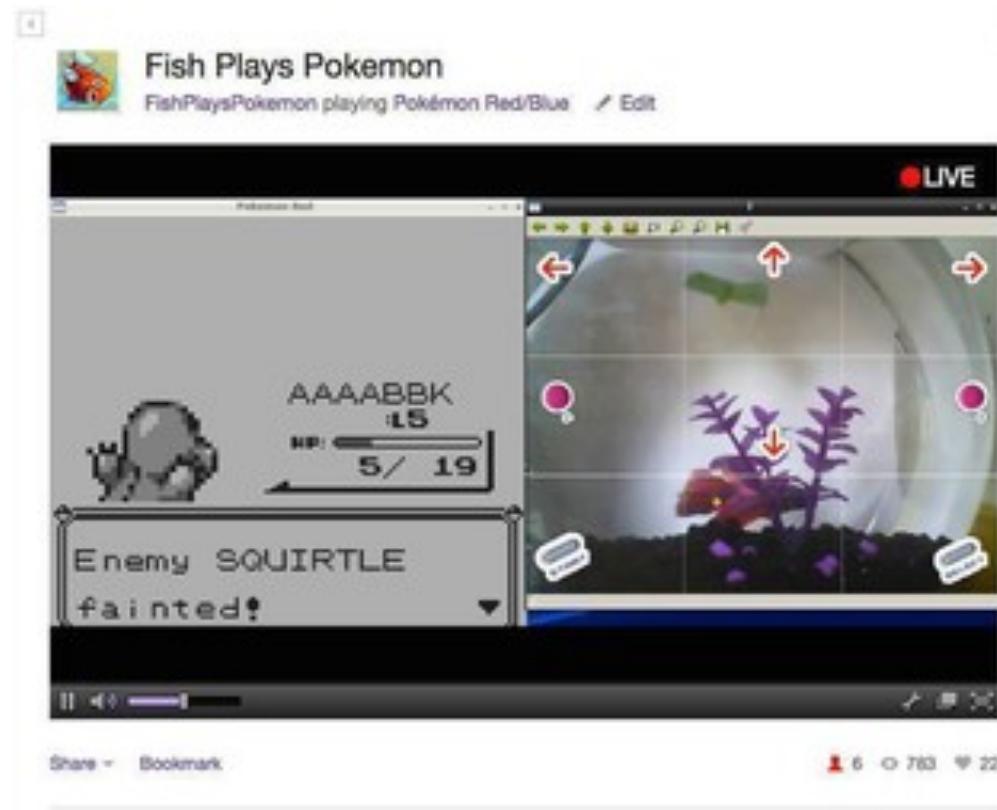


⊗



# Ou um jogo “bobo”...

- Peixe jogando Pokemon
  - <http://fi.twitch.tv/fishplayspokemon>



# Próxima aula...

- Ferramental Computacional - OpenCV