



Computer Vision

Marc Pollefeys
Luc Van Gool



The team



Marc Pollefeys



Luc Van Gool



Bastien Jacquet
Federico
Camposeco-Paulsen

Danfeng Qin

Taha Koltukluoglu



What is computer vision?

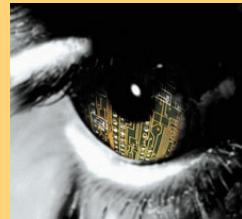


Done?



What is computer vision?

- Automatic understanding of images and video
 - Computing properties of the 3D world from visual data (*measurement*)
 - Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. (*perception and interpretation*)



Vision for measurement

Real-time stereo

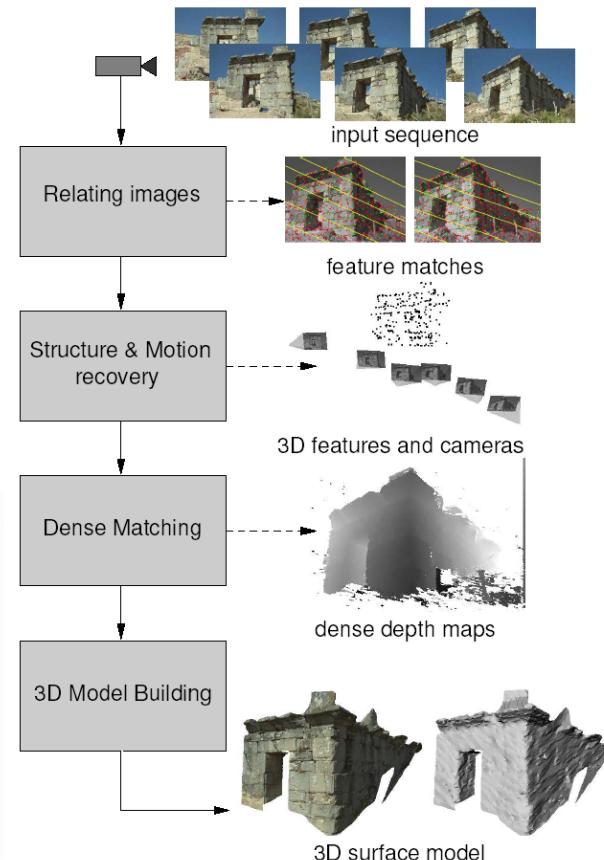


NASA Mars Rover

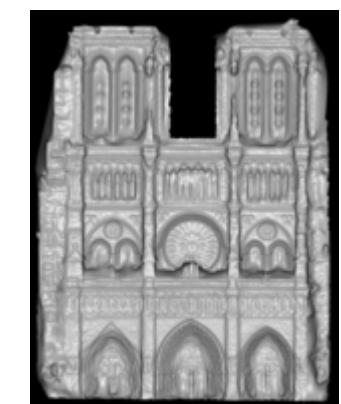
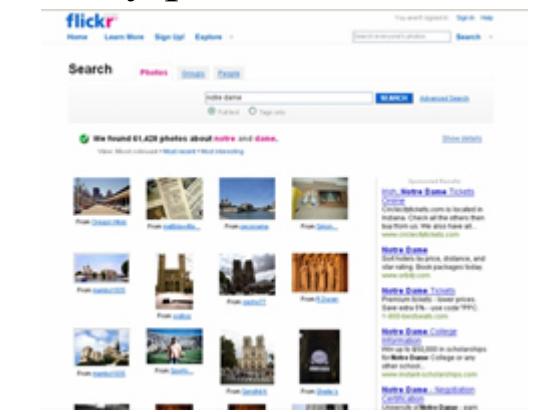


Pollefeys et al.

Structure from motion



Multi-view stereo for
community photo collections

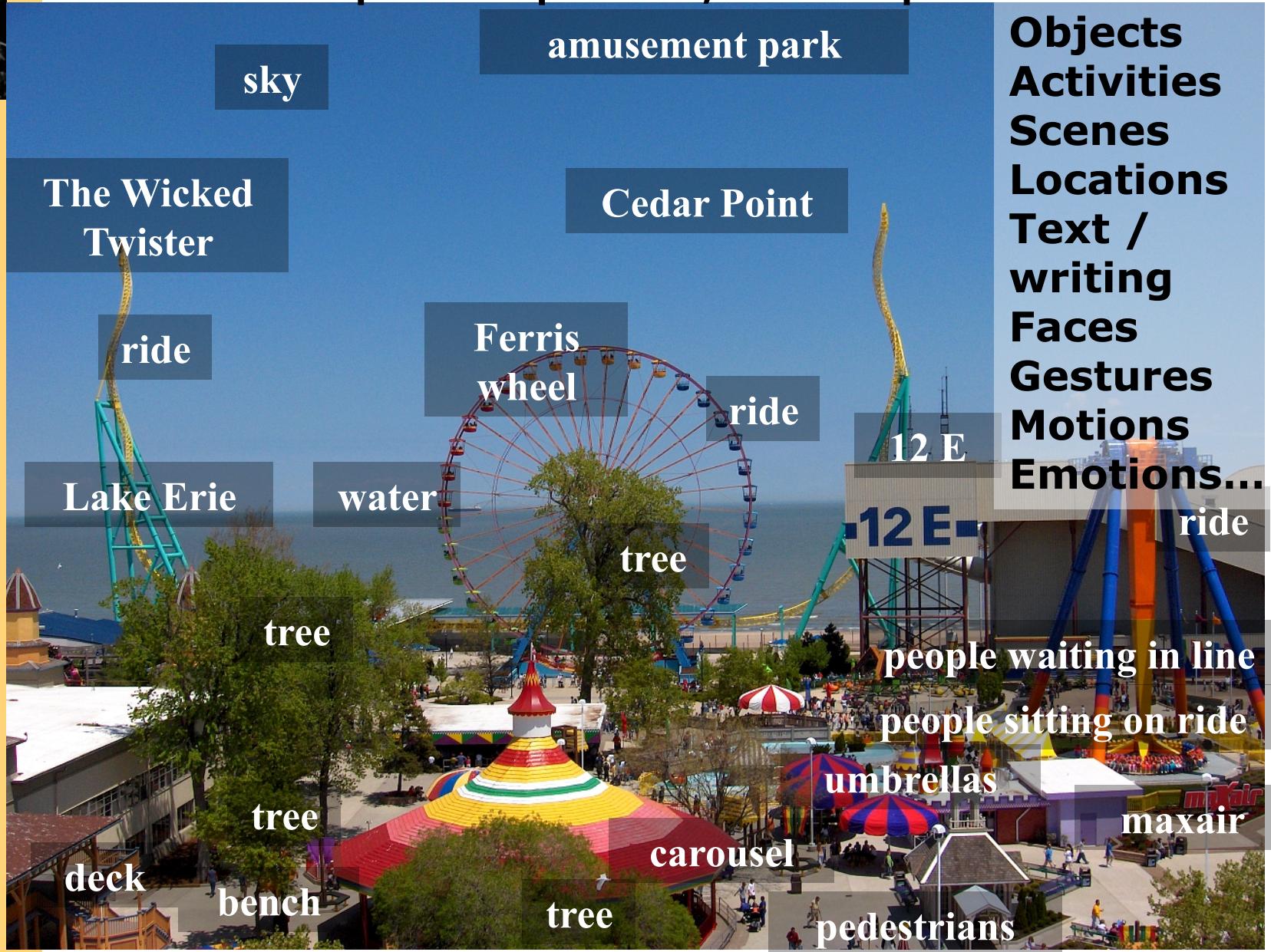


Goesele et al.

Slide credit: L. Lazebnik

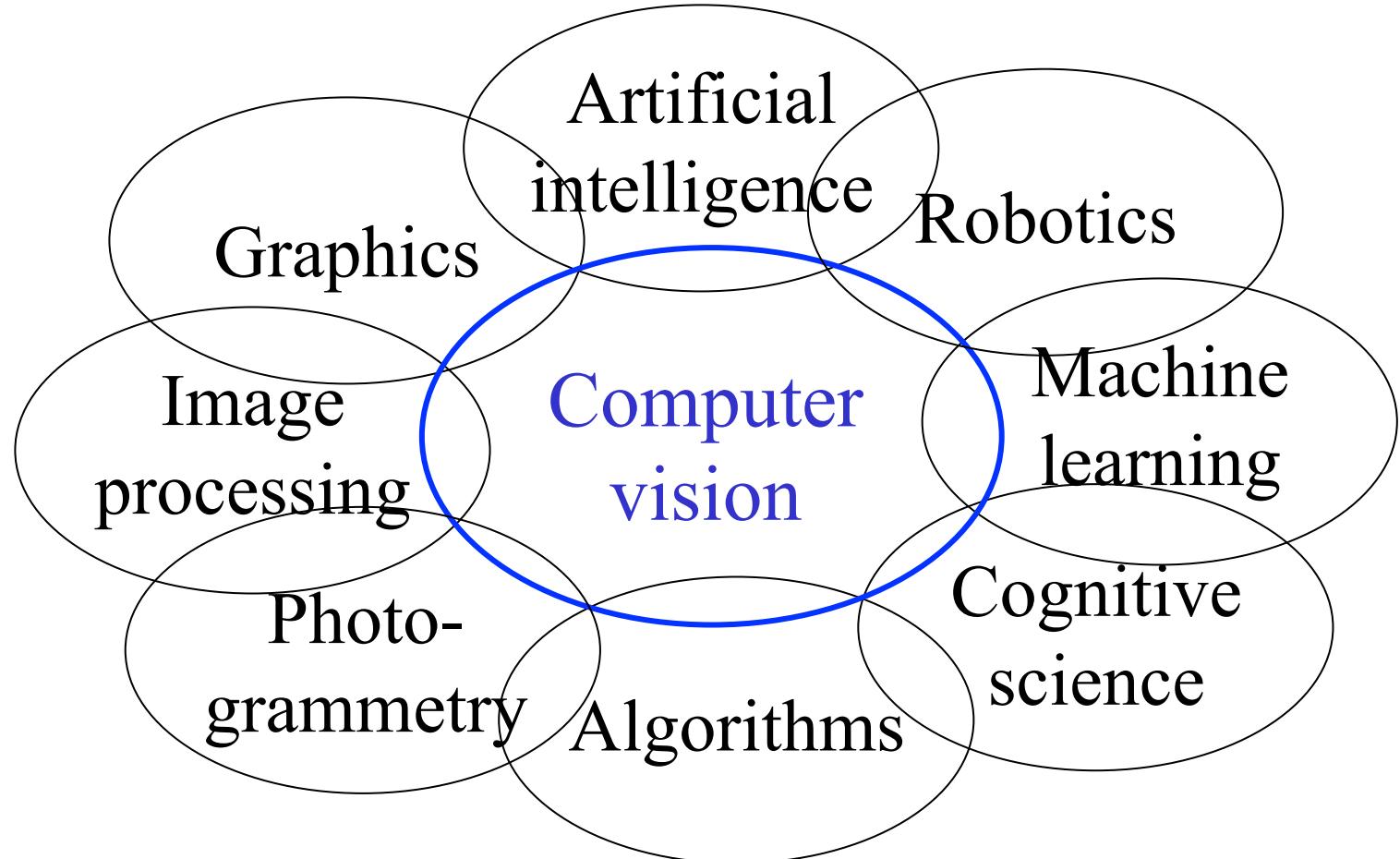


Vision for perception, interpretation



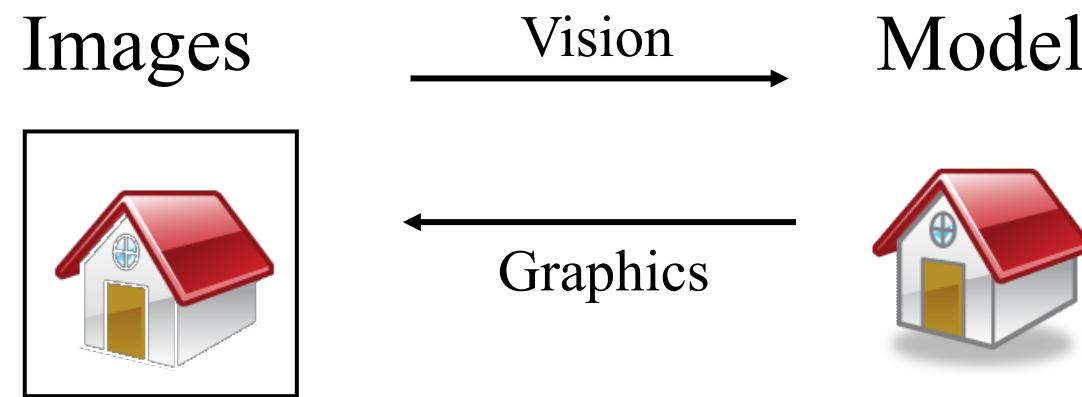


Related disciplines





Vision and graphics

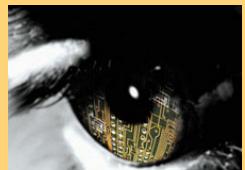


Inverse problems: analysis and synthesis.



Why vision?

- As image sources multiply, so do applications
 - Relieve humans of boring, easy tasks
 - Enhance human abilities: human-computer interaction, visualization
 - Perception for robotics / autonomous agents
 - Organize and give access to visual content



Why vision?

Images and video are everywhere!



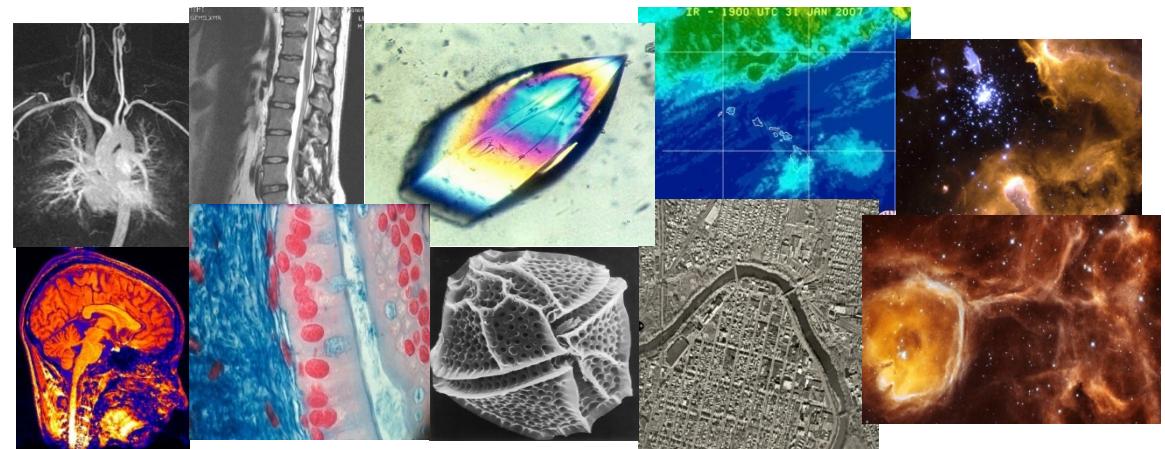
Personal photo albums



Movies, news, sports



Surveillance and security



Medical and scientific images



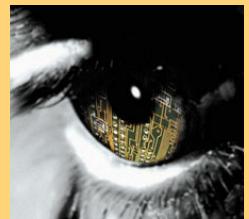
Why is vision hard?

Grayscale Image

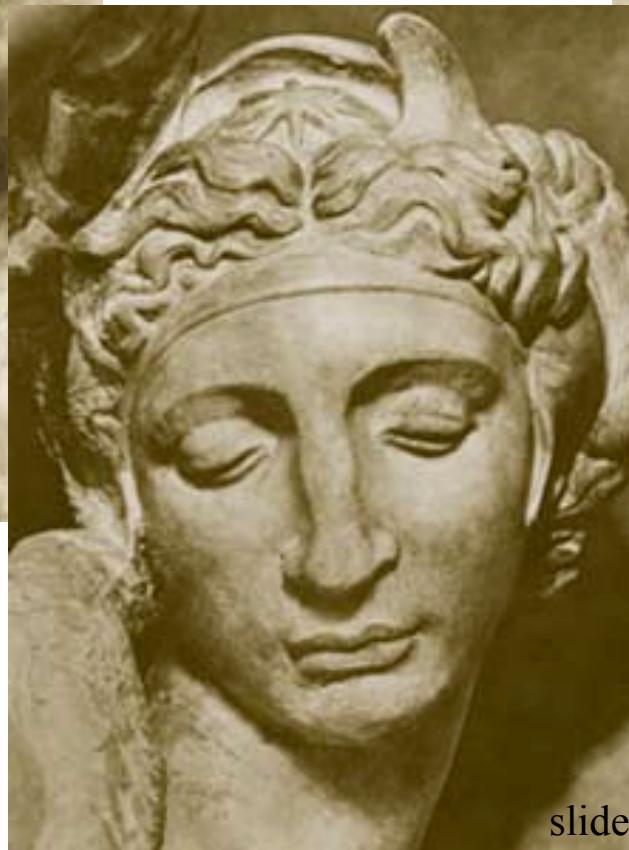


x =	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	
y =	41	210	209	204	202	197	247	143	71	64	80	84	54	54	57	58
	42	206	196	203	197	195	210	207	56	63	58	53	53	61	62	51
	43	201	207	192	201	198	213	156	69	65	57	55	52	53	60	50
	44	216	206	211	193	202	207	208	57	69	60	55	77	49	62	61
	45	221	206	211	194	196	197	220	56	63	60	55	46	97	58	106
	46	209	214	224	199	194	193	204	173	64	60	59	51	62	56	48
	47	204	212	213	208	191	190	191	214	60	62	66	76	51	49	55
	48	214	215	215	207	208	180	172	188	69	72	55	49	56	52	56
	49	209	205	214	205	204	196	187	196	86	62	66	87	57	60	48
	50	208	209	205	203	202	186	174	185	149	71	63	55	55	45	56
	51	207	210	211	199	217	194	183	177	209	90	62	64	52	93	52
	52	208	205	209	209	197	194	183	187	187	239	58	68	61	51	56
	53	204	206	203	209	195	203	188	185	183	221	75	61	58	60	60
	54	200	203	199	236	188	197	183	190	183	196	122	63	58	64	66
	55	205	210	202	203	199	197	196	181	173	186	105	62	57	64	63

How do we go from an array of numbers recognizing fruit?



Challenges: viewpoint variation



Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba



Challenges: illumination

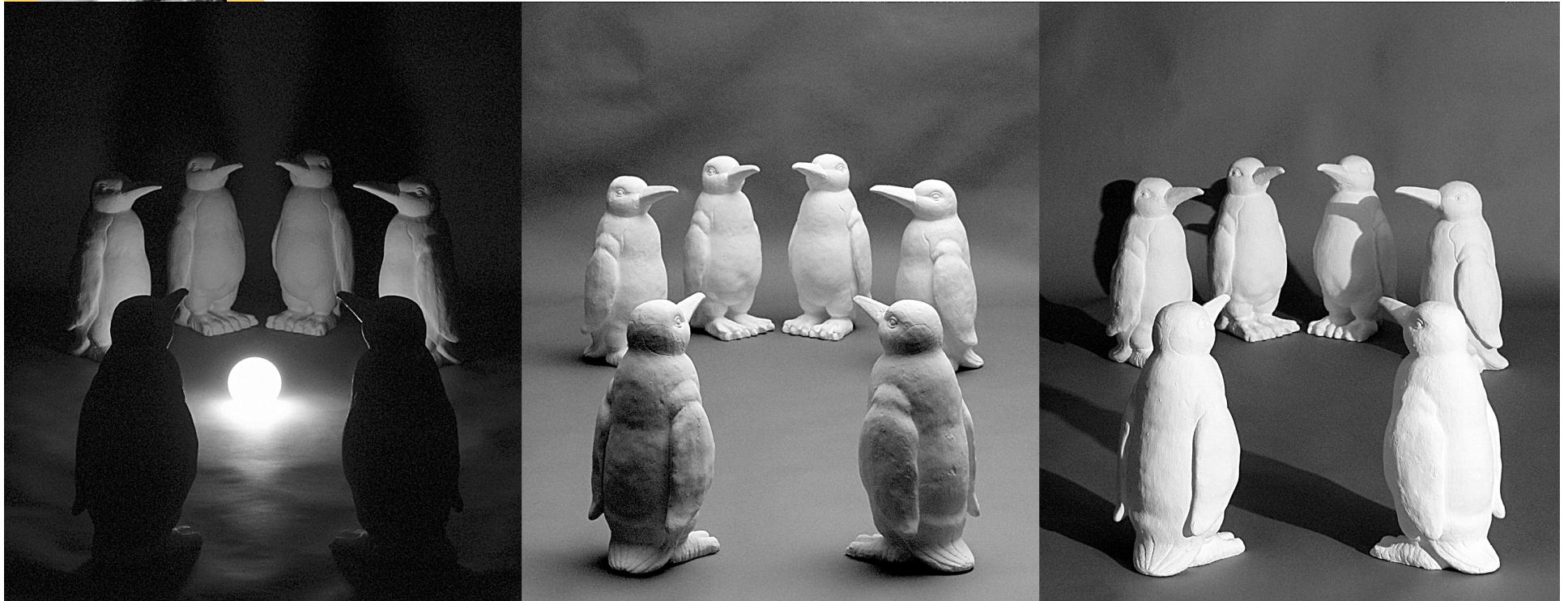


image credit: J. Koenderink



Challenges: scale



slide credit: Fei-Fei, Fergus & Torralba



Challenges: background clutter



Emperor shrimp and commensal crab on a sea cucumber in Fiji
Photograph by Tim Laman

slide credit: Svetlana Lazebnik
NATIONAL GEOGRAPHIC
reserved.



Challenges: Motion



slide credit: Svetlana Lazebnik



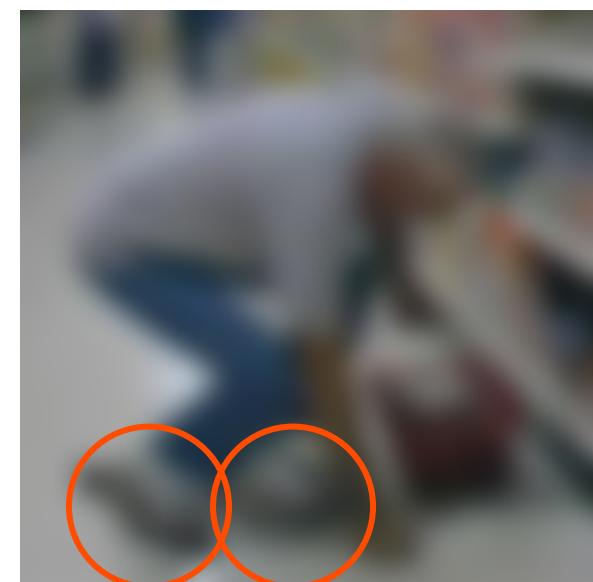
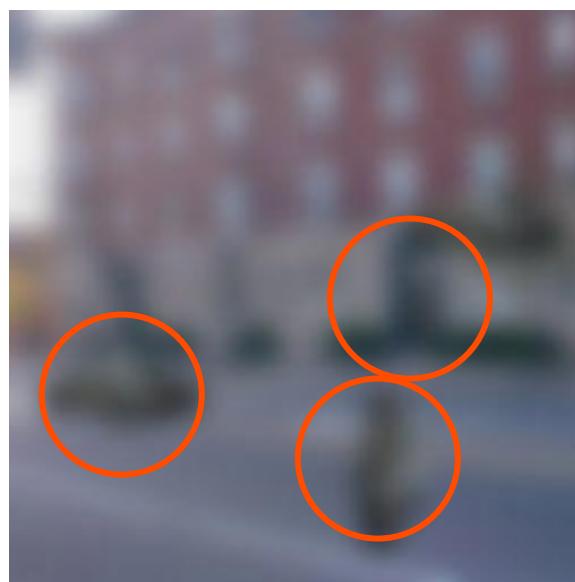
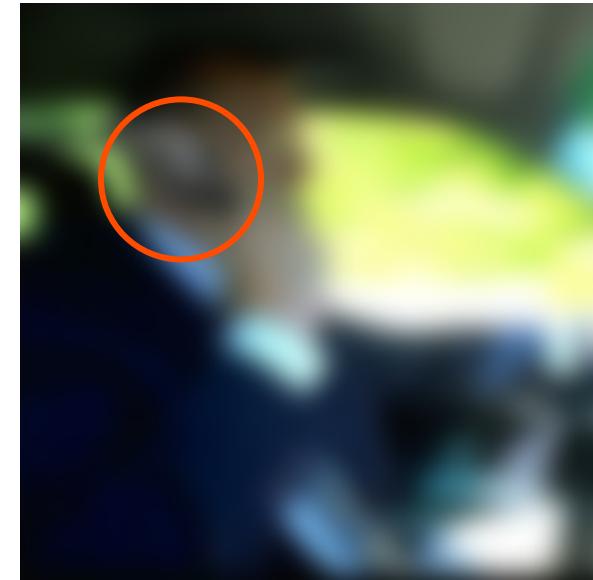
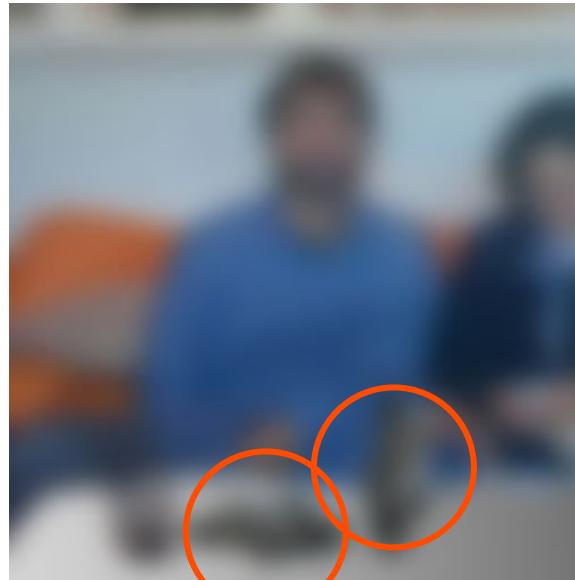
Challenges: object intra-class variation



slide credit: Fei-Fei, Fergus & Torralba



Challenges: local ambiguity



slide credit: Fei-Fei, Fergus & Torralba



Challenges or opportunities?

- Images are confusing, but they also reveal the structure of the world through numerous cues
- Our job is to interpret the cues!



Image source: J. Koenderink



Depth cues: Linear perspective





Shape cues: Texture gradient



slide credit: Svetlana Lazebnik
nationalgeographic.com



Position and lighting cues: Cast shadows



Source: J. Koenderink



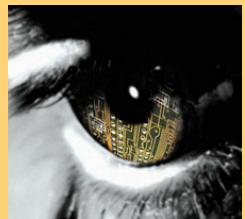
Grouping cues: Similarity (color, texture, proximity)



NATIONALGEORGIC.COM

© 2003 National Geographic Society. All rights reserved.

slide credit: Svetlana Lazebnik



Grouping cues: “Common fate”



Image credit: Arthus-Bertrand (via F. Durand)



Bottom line

- Perception is an inherently ambiguous problem
 - Many different 3D scenes could have given rise to a particular 2D picture



- Possible solutions
 - Bring in more constraints (more images)
 - Use prior knowledge about the structure of the world
- Need a combination of different methods



Vision Demo?



Terminator 2

we're not quite there yet....



Current state of the art

- The next slides show some examples of what current vision systems can do



Earth viewers (3D modeling)

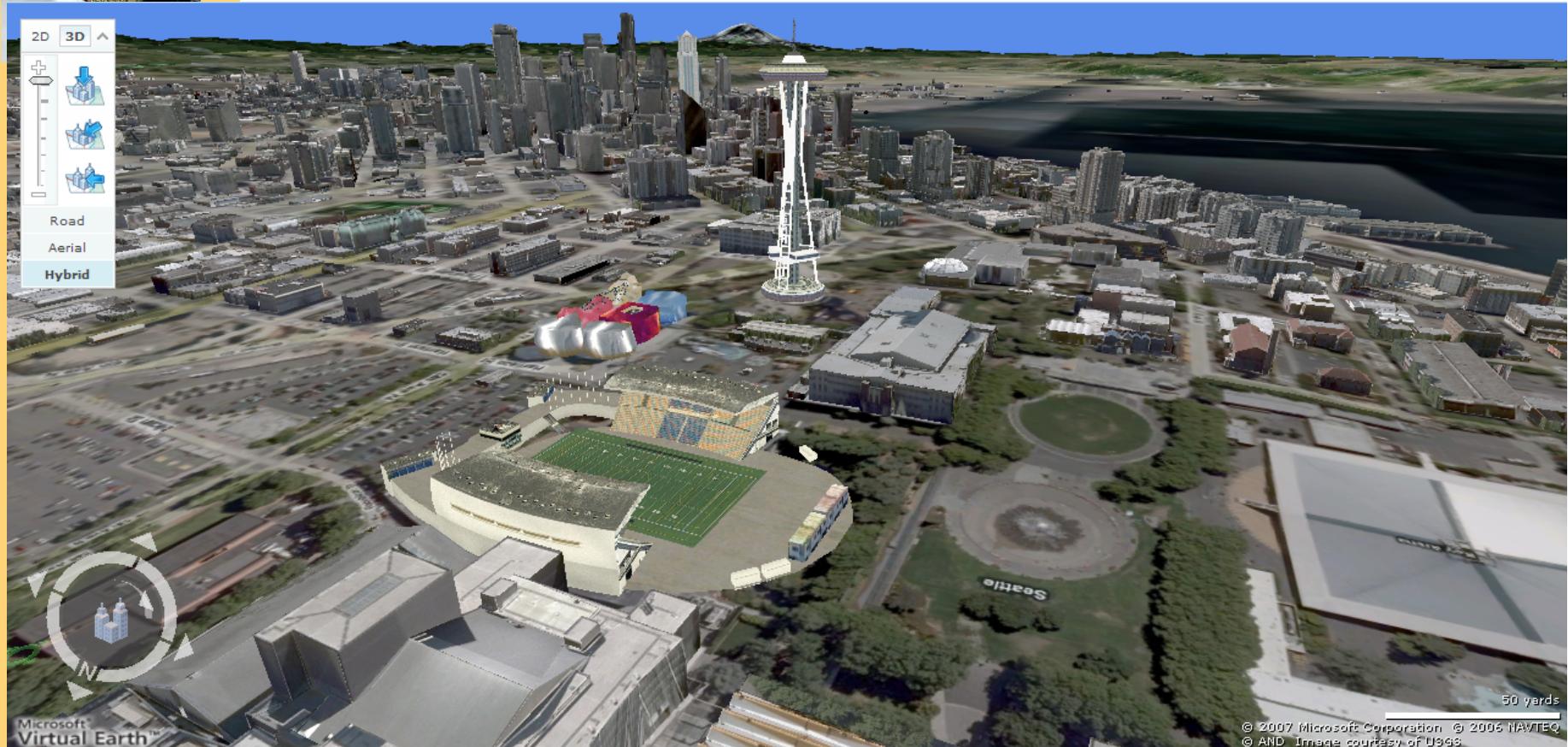


Image from Microsoft's [Virtual Earth](#)
(see also: [Google Earth](#))



Home

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

*"What if your photo collection was an entry point into the world,
like a wormhole that you could jump through and explore..."*

[Try it](#)



[Try the Tech 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/>

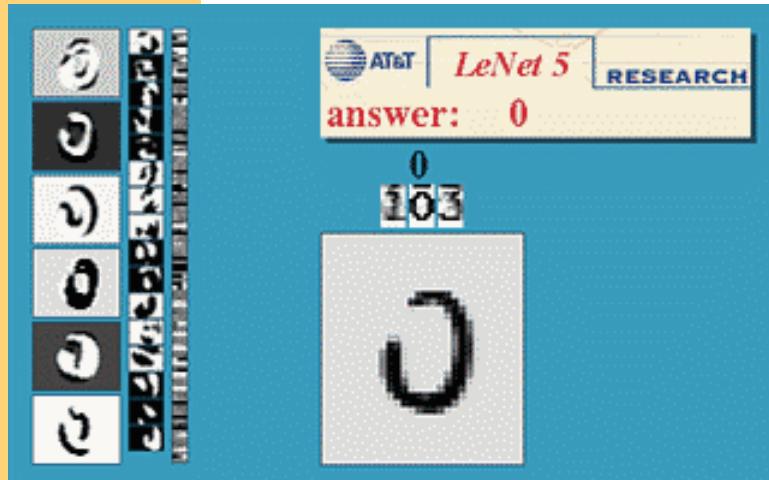
Based on [Photo Tourism technology](#) developed
by Noah Snavely, Steve Seitz, and Rick Szeliski



Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



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



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



Face detection



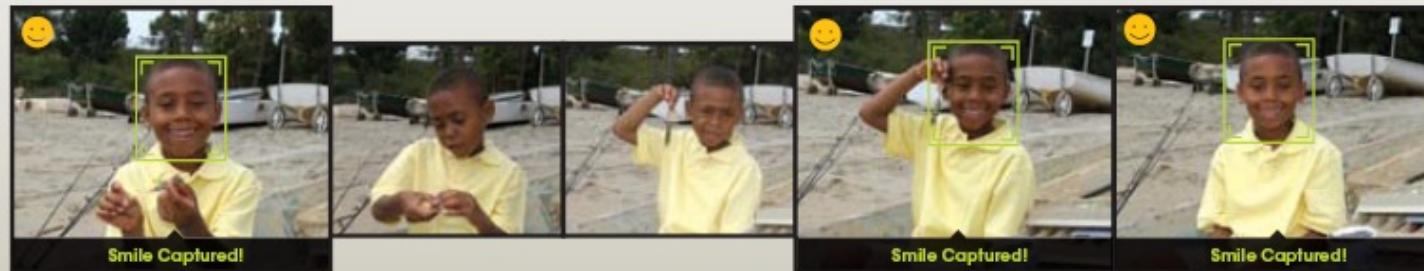
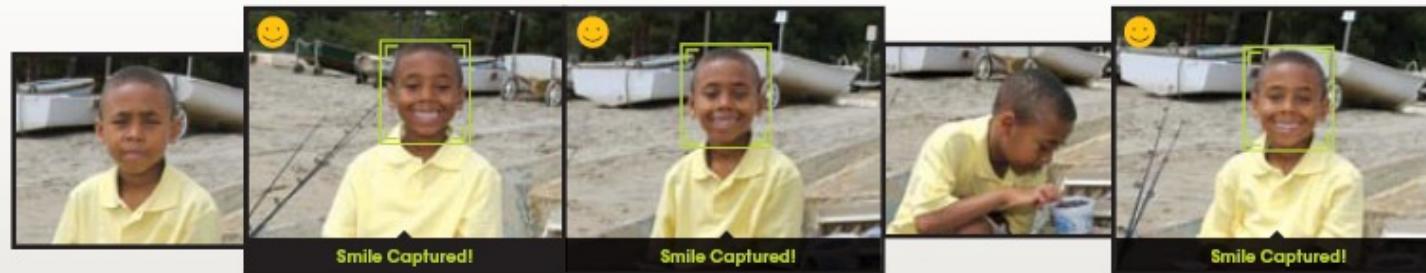
Many new digital cameras now detect faces
– Canon, Sony, Fuji, ...



Smile detection?

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.

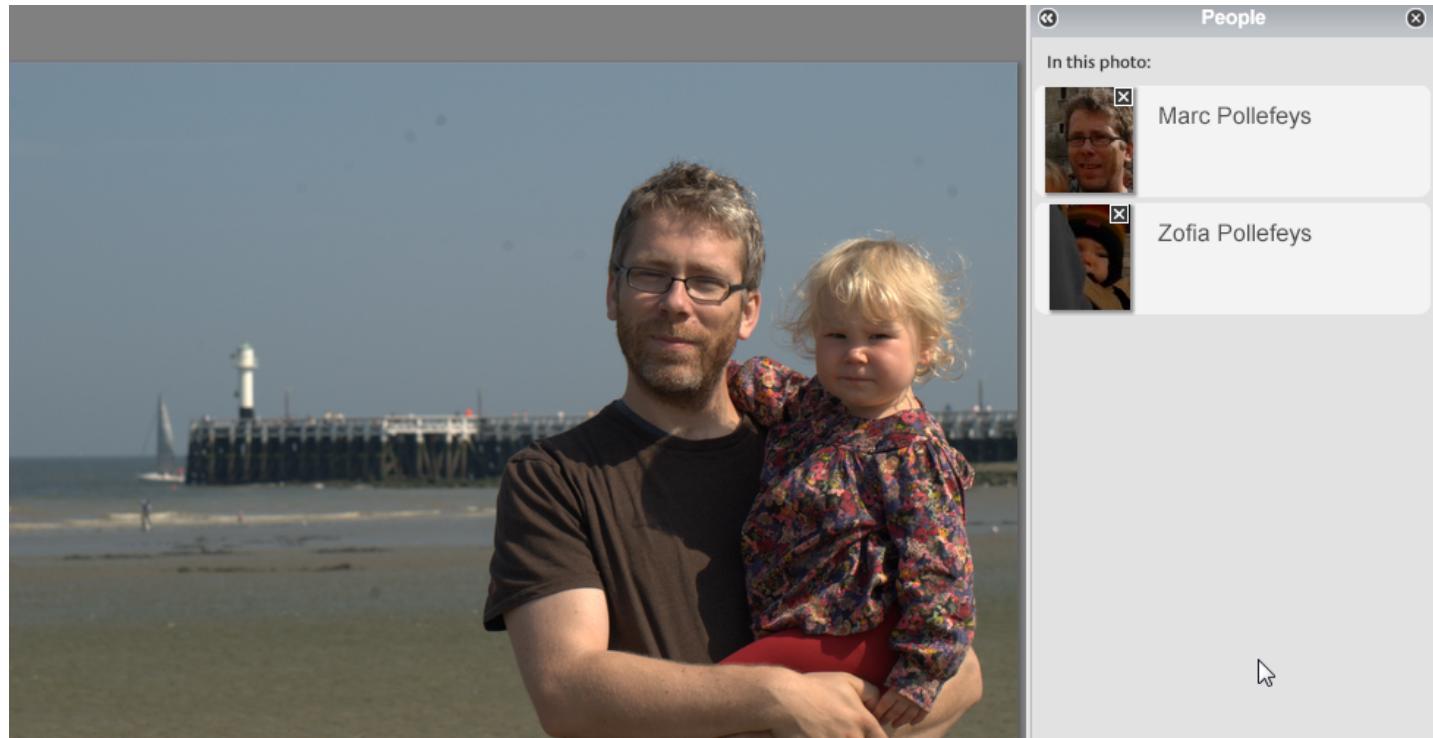


[Sony Cyber-shot® T70 Digital Still Camera](#)



Face recognition

- E.g. Picasa





Object recognition (in supermarkets)

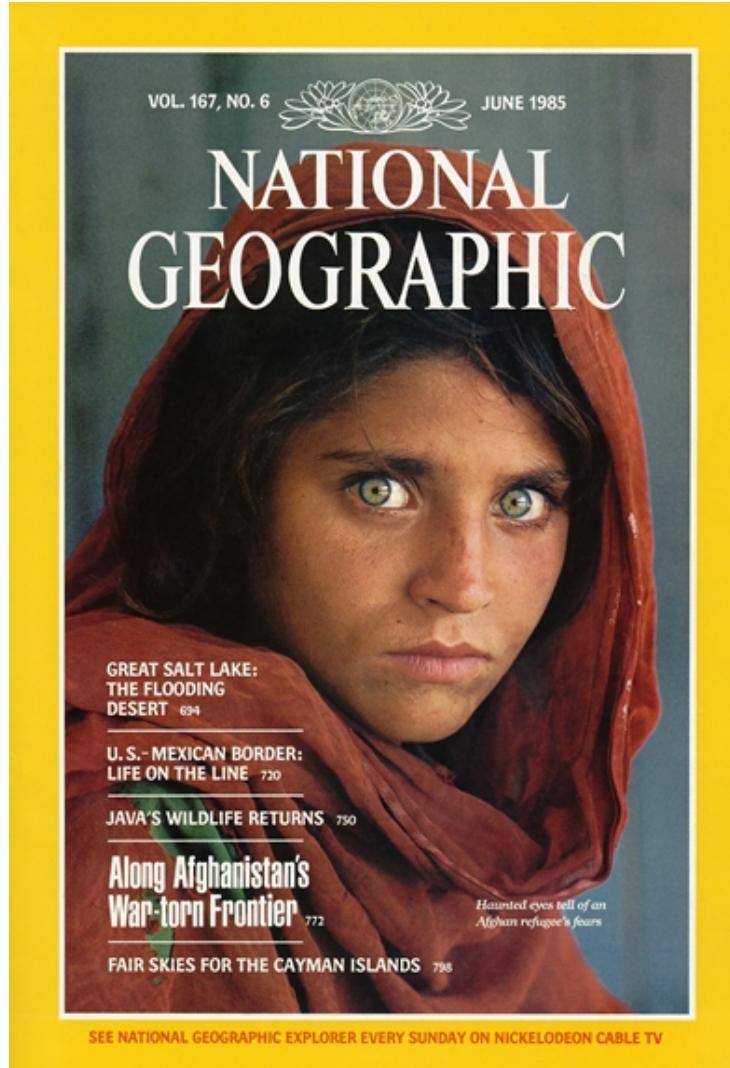


[LaneHawk by EvolutionRobotics](#)

“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... “



Face recognition



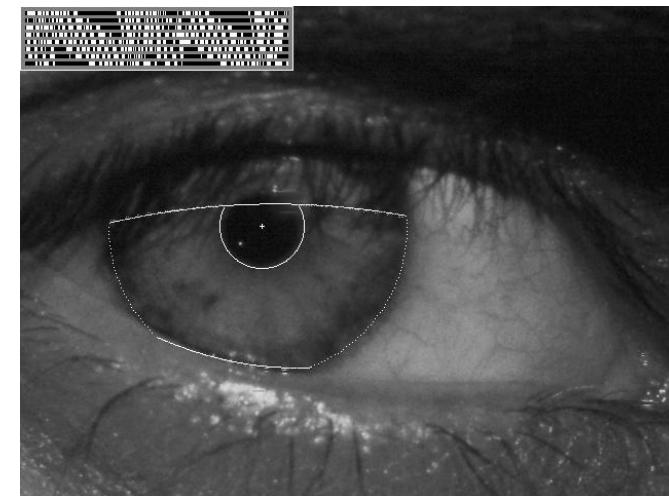
Who is she?

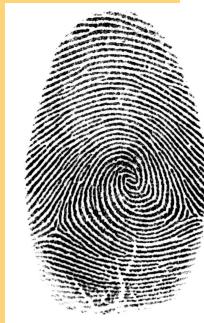
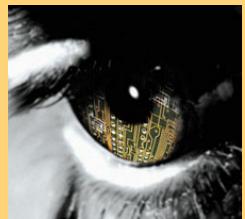


Vision-based biometrics



“How the Afghan Girl was Identified by Her Iris Patterns” Read the [story](#)

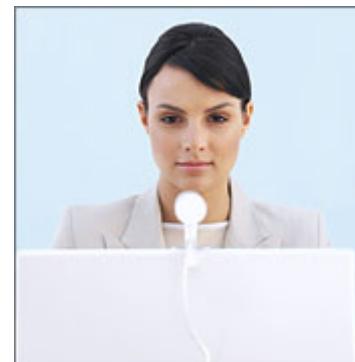




Fingerprint scanners on
many new laptops,
other devices

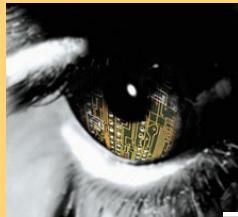


Login without a password...



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

Object recognition (in mobile phones)



- This is becoming real:
 -  Google Goggles
Use pictures to search the web.
 - [Point & Find](#), Nokia
 - [Kooaba](#) (ETH start-up)
 - [SnapTell](#)



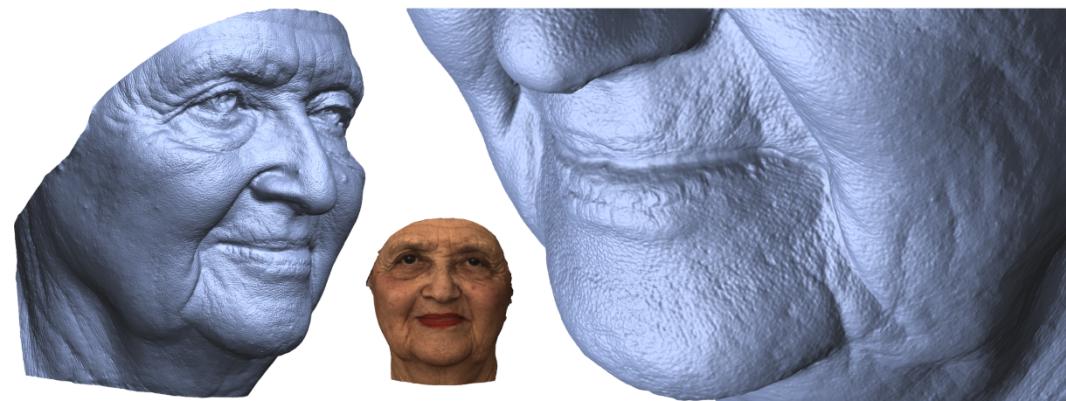
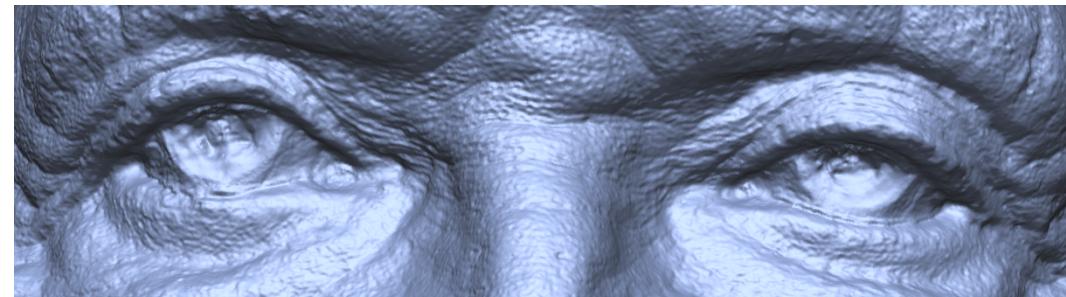
Special effects: shape capture



The Matrix movies, ESC Entertainment, XYZRGB, NRC



Special effects: shape capture





Special effects: motion capture



Pirates of the Caribbean, Industrial Light and Magic

Smart cars

The screenshot shows the Mobileye website's "Our Vision. Your Safety." section. It features a top-down view of a car with three cameras highlighted: a "rear looking camera" on the left, a "forward looking camera" on the right, and a "side looking camera" at the bottom. Below this, there are three main sections: "EyeQ Vision on a Chip" (with an image of a chip), "Vision Applications" (with an image of a pedestrian crossing a street), and "AWS Advance Warning System" (with an image of a dashboard display). To the right, there are "News" and "Events" columns.

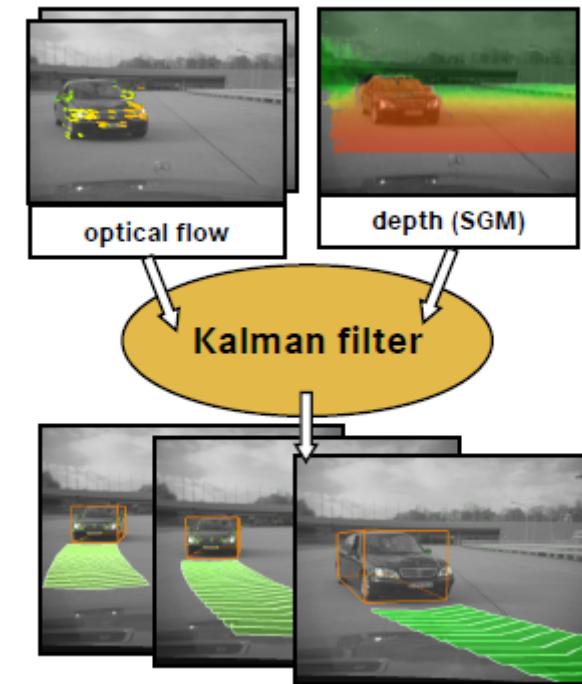
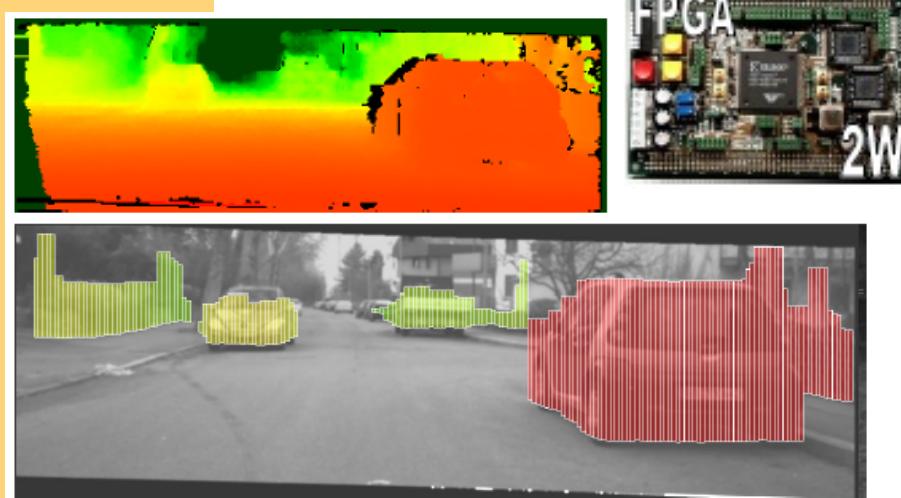
- News**
 - Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System
 - Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end
- Events**
 - Mobileye at Equip Auto, Paris, France
 - Mobileye at SEMA, Las Vegas, NV

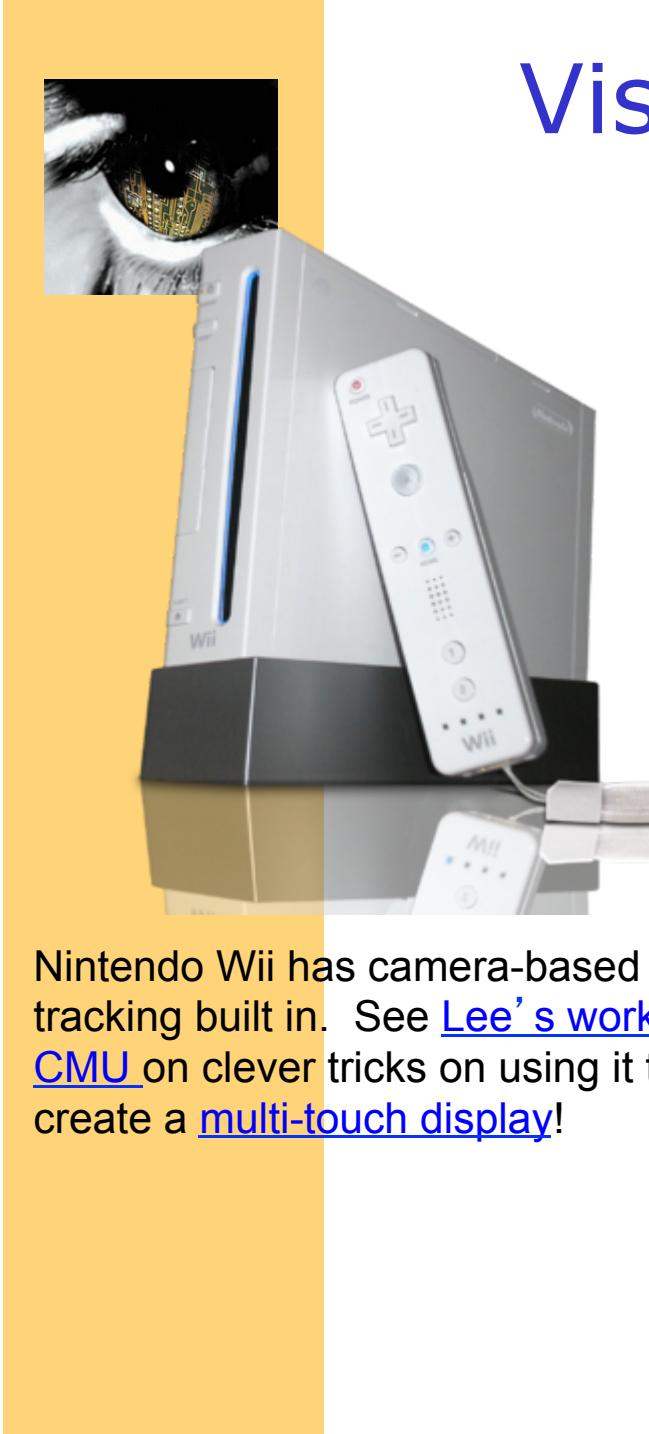
- Mobileye
 - Vision systems currently in high-end BMW, GM, Volvo models
 - By 2010: 70% of car manufacturers.
 - [Video demo](#)

Slide content courtesy of Amnon Shashua

Driver Assistance

Daimler stereo system





Nintendo Wii has camera-based IR tracking built in. See [Lee's work at CMU](#) on clever tricks on using it to create a [multi-touch display](#)!

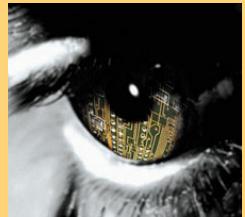
Vision-based interaction (and games)



[Digimask](#): put your face on a 3D avatar.



["Game turns moviegoers into Human Joysticks"](#), CNET
Camera tracking a crowd, based on [this work](#).



Vision-based interaction (and games)

MS project natal/Kinect



<http://www.xbox.com/en-US/kinect>



Vision in space

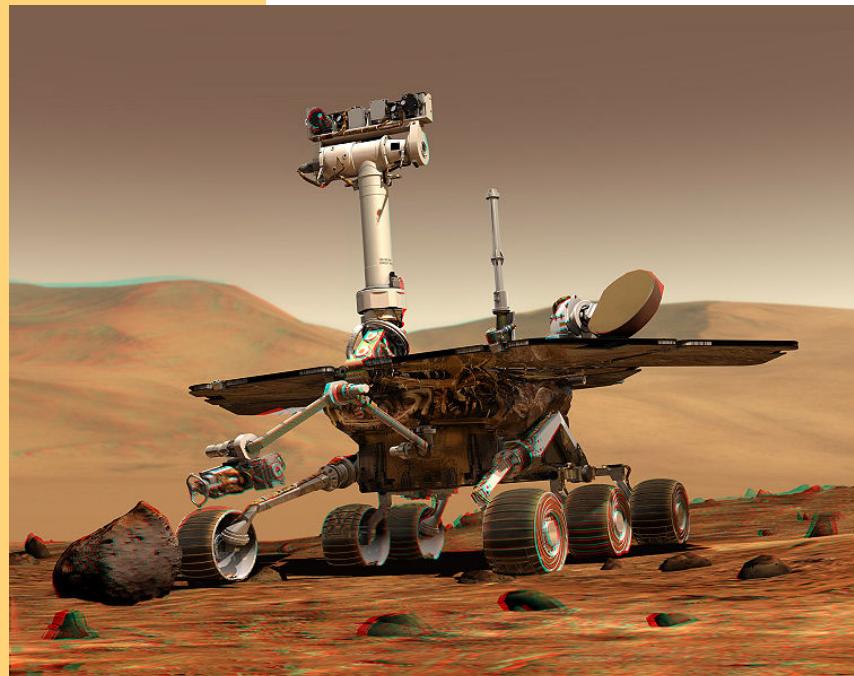
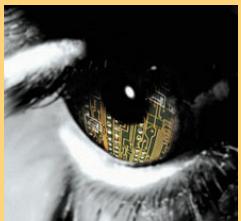


[NASA's Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

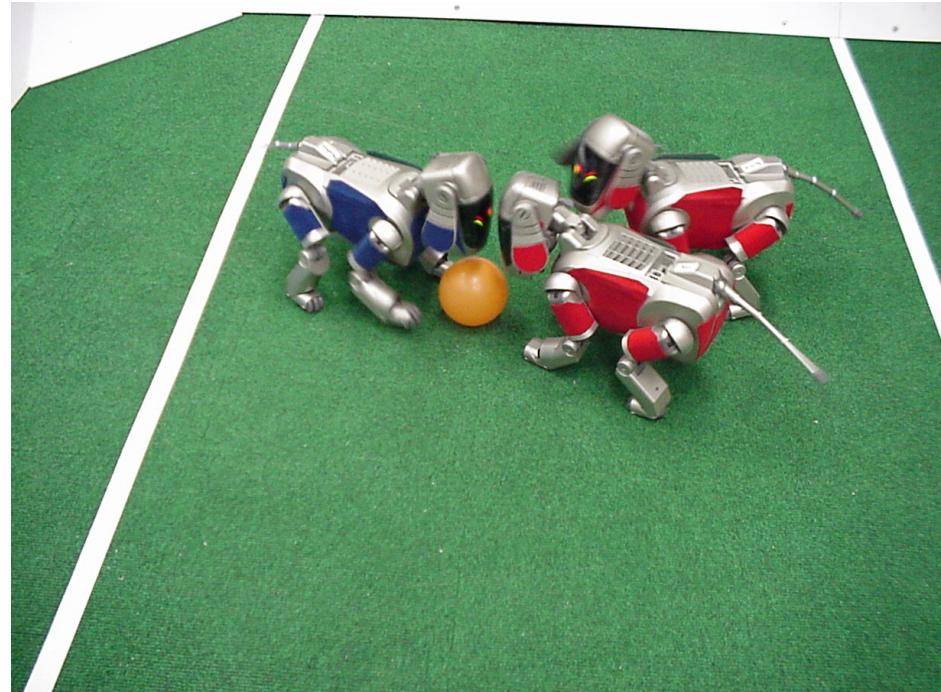
Vision systems (JPL) used for several tasks

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

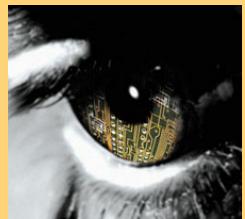
Robotics



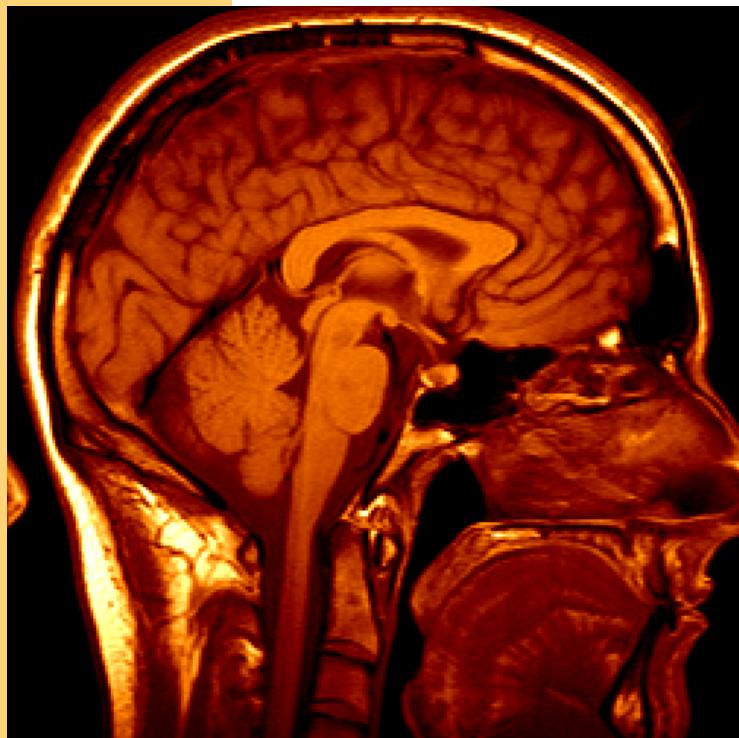
NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



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



Medical imaging



3D imaging
MRI, CT



Image guided surgery
[Grimson et al., MIT](#)



Current state of the art

- You just saw examples of current systems.
 - Many of these are less than 5 years old
- This is a very active research area, and rapidly changing
 - Many new apps in the next 5 years
- To learn more about vision applications and companies
 - [David Lowe](#) maintains an excellent overview of vision companies
 - <http://www.cs.ubc.ca/spider/lowe/vision.html>