

Introduction to Computer Vision

Prof Bhavana N

SCHOOL OF C and IT

What is Computer Vision?

- **Computer vision** is the science and technology of machines that see.
- Concerned with the theory for building artificial systems that obtain information from images.
- The image data can take many forms, such as a video sequence, depth images, views from multiple cameras, or multi-dimensional data from a medical scanner



Computer Vision

Make computers understand images and videos.



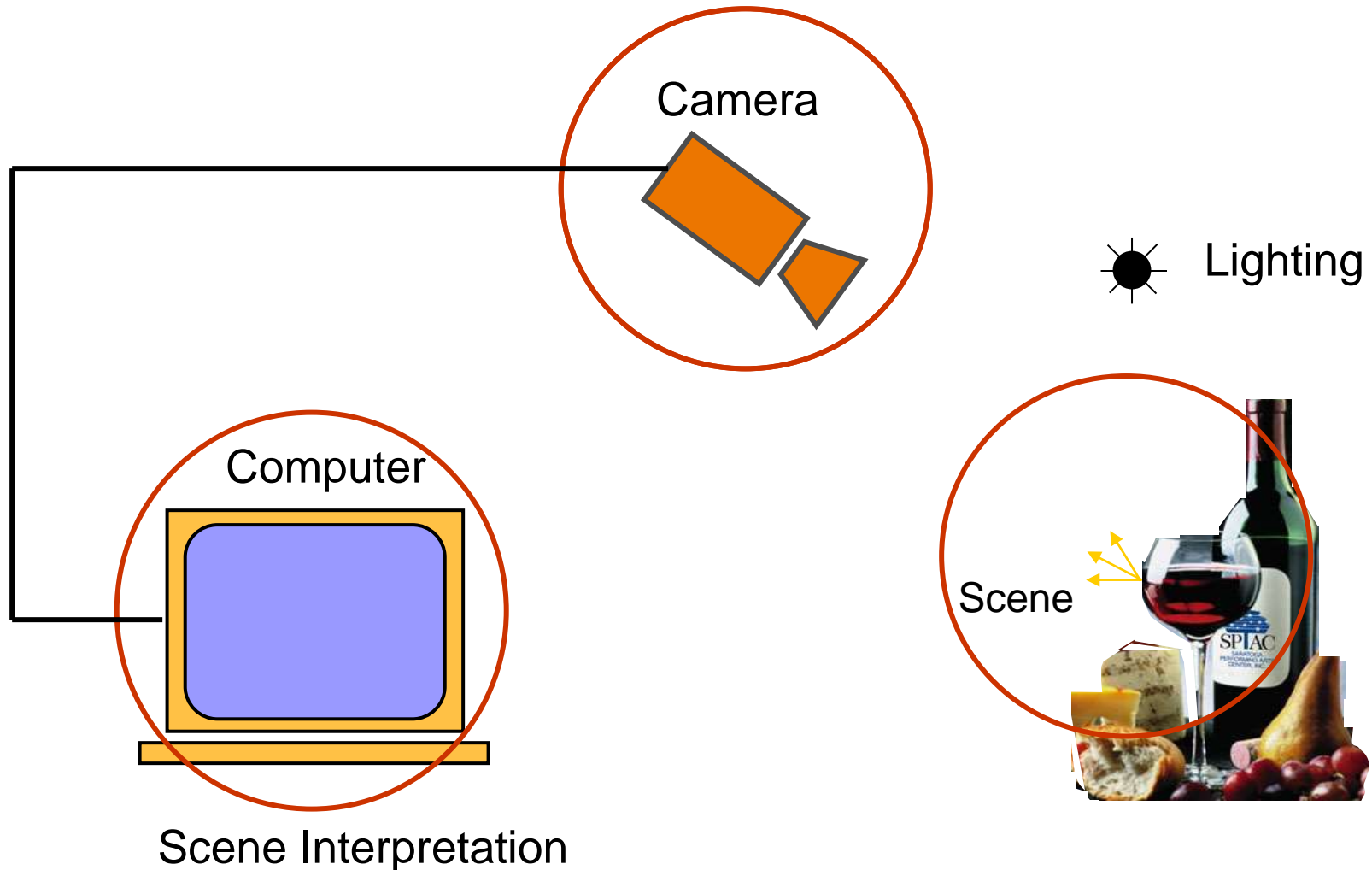
What kind of scene?

Where are the cars?

How far is the building?

...

Components of a computer vision system



Computer vision vs human vision

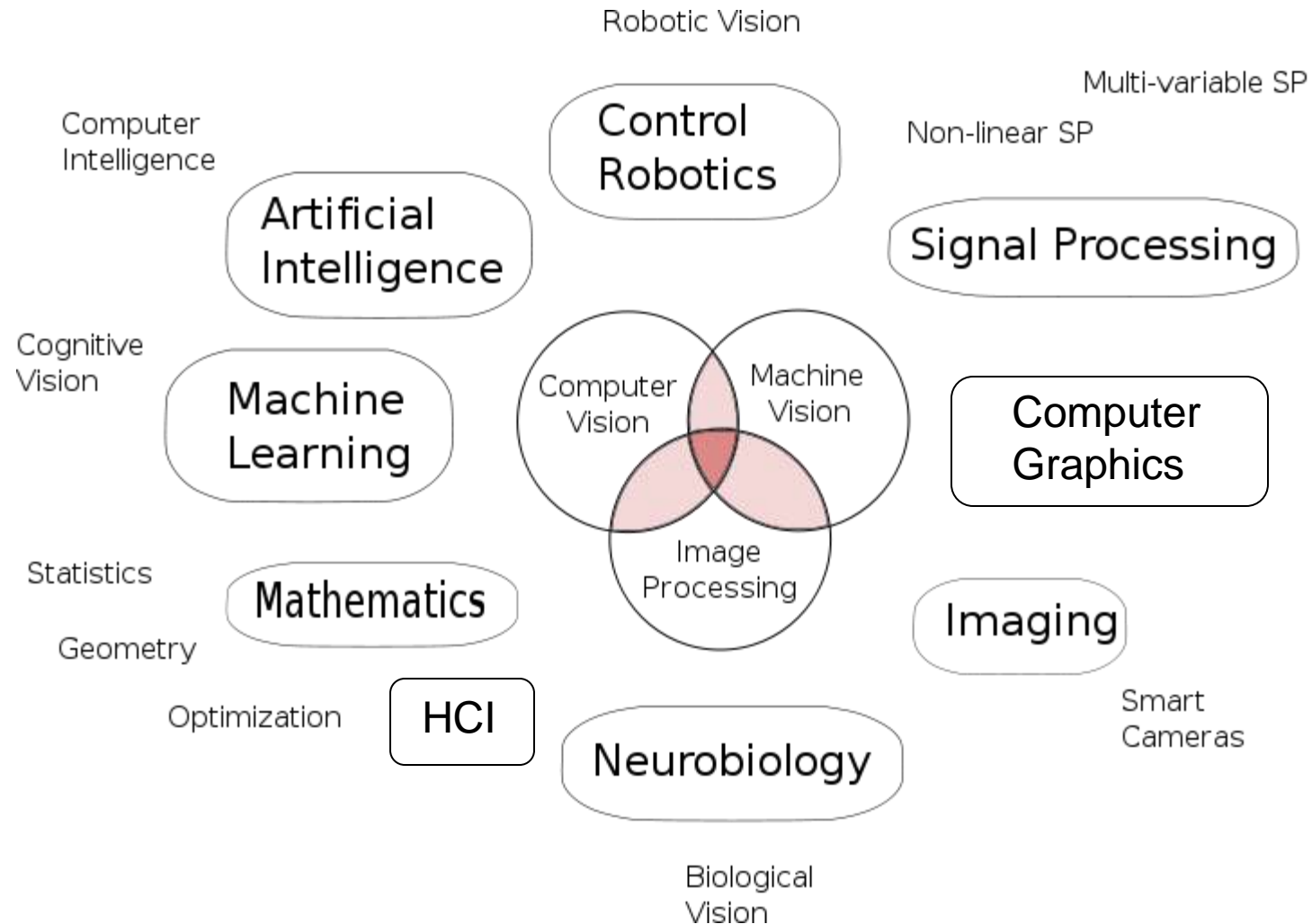


What we see

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

Vision is multidisciplinary



Why computer vision matters



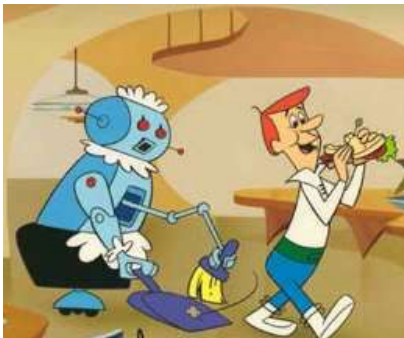
Safety



Health



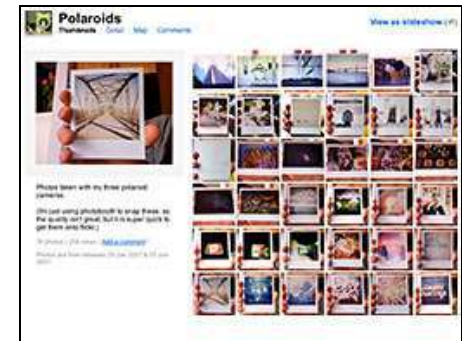
Security



Comfort



Fun



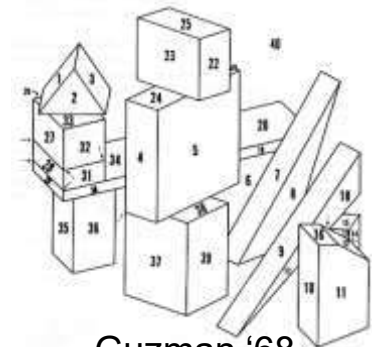
Access

A little story about Computer Vision

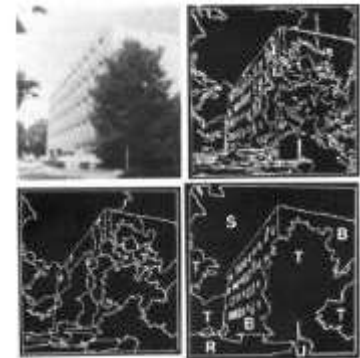
In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”. We now know that the problem is slightly more difficult than that. (Szeliski 2009, Computer Vision)

brief history of computer vision

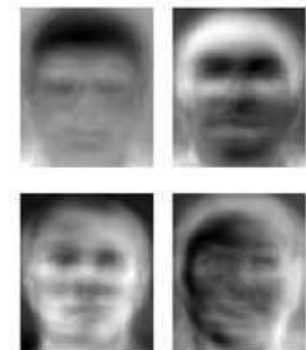
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis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



Guzman '68



Ohta Kanade '78



Turk and Pentland '91

How vision is used now

- Examples of state-of-the-art

Face detection

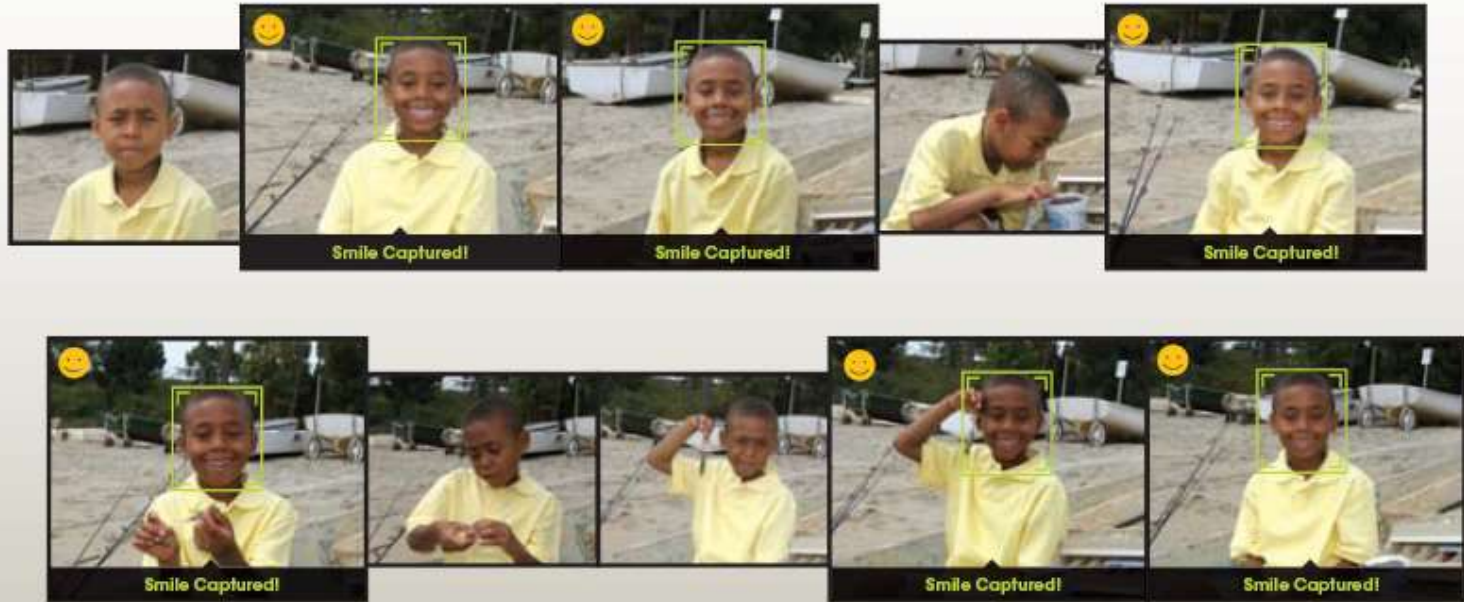


- Many new digital cameras now detect faces
 - Canon, Sony, Fuji, ...

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](#)

Object recognition (in supermarkets)



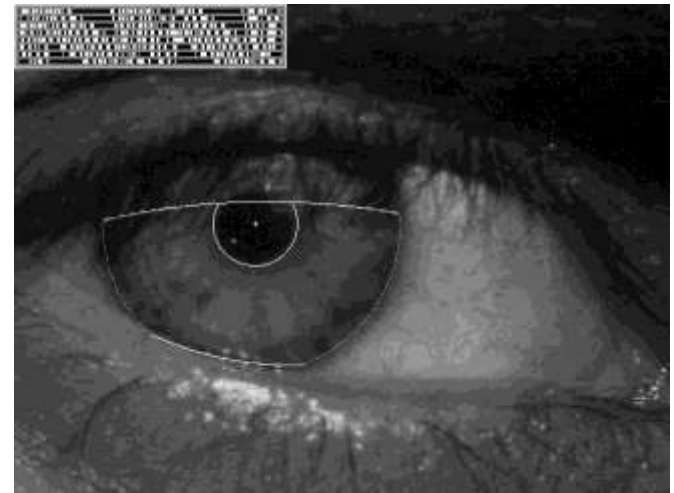
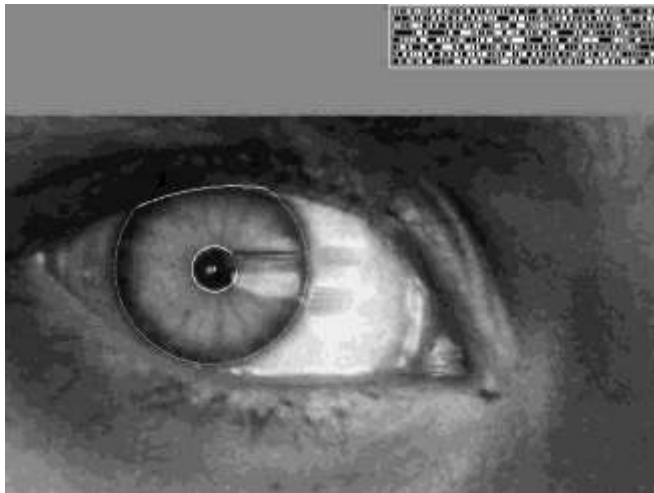
[LaneHawk by EvolutionRobotics](#)

“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



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



Login without a password...



Fingerprint scanners on many new laptops, other devices



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

Object recognition (in mobile phones)



[Point & Find](#), [Nokia](#)
[Google Goggles](#)

Sports



Sportvision first down line

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

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

Smart cars

Slide content courtesy of Amnon Shashua

The image is a screenshot of the Mobileye website. At the top, there are two navigation tabs: 'manufacturer products' and 'consumer products'. Below them is a large banner with the text 'Our Vision. Your Safety.' and a top-down view of a car with yellow beams representing camera fields of view. Labels point to 'rear looking camera', 'side looking camera', and 'forward looking camera'. To the right is a 'News' sidebar with two articles: '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'. Below the banner are three main product sections: 'EyeQ Vision on a Chip' with an image of a chip, 'Vision Applications' showing a pedestrian on a crosswalk, and 'AWS Advance Warning System' with a circular display showing a car icon and the number '0.8'. Each section has a 'read more' link. On the right, an 'Events' sidebar lists 'Mobileye at Equip Auto, Paris, France' and 'Mobileye at SEMA, Las Vegas, NV', also with a 'read more' link.

- [Mobileye](#) [[wiki article](#)]
 - Vision systems currently in many car models

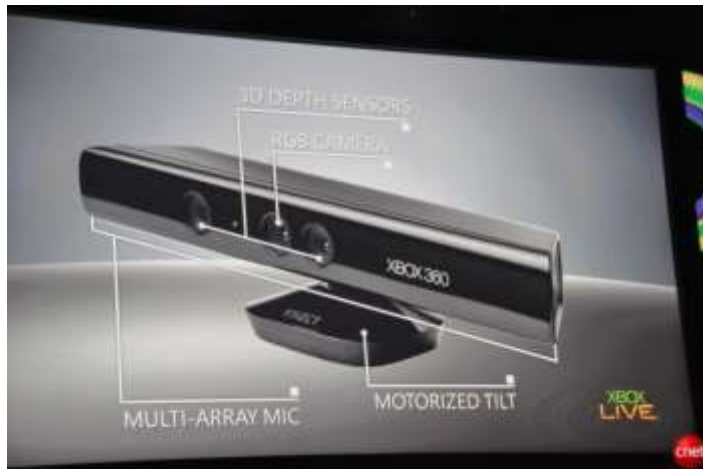
Google cars



<http://www.nytimes.com/2010/10/10/science/10google.html?ref=artificialintelligence>

Interactive Games: Kinect

- Object Recognition: <http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o>
- Mario: <http://www.youtube.com/watch?v=8CTJL5IUjHg>
- 3D: <http://www.youtube.com/watch?v=7QrnwoO1-8A>
- Robot: <http://www.youtube.com/watch?v=w8BmgtMKFbY>
- 3D tracking, reconstruction, and interaction: <http://research.microsoft.com/en-us/projects/surfacerecon/default.aspx>



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.

Mobile robots



NASA's Mars Spirit Rover

http://en.wikipedia.org/wiki/Spirit_rover



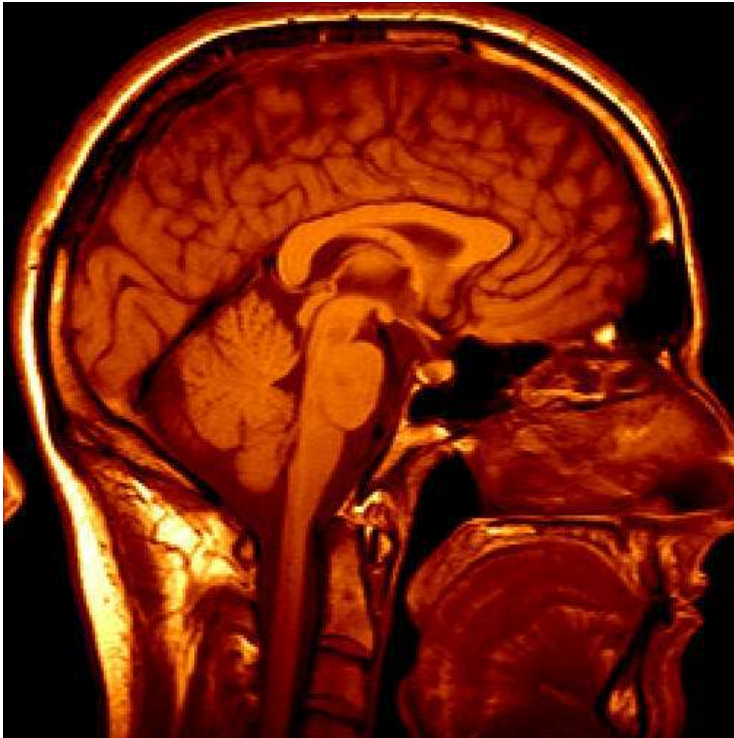
<http://www.robocup.org/>



Saxena et al. 2008

[STAIR](#) at Stanford

Medical imaging



3D imaging
MRI, CT



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

Object categorization



sky

building

flag

banner

face

wall

street lamp

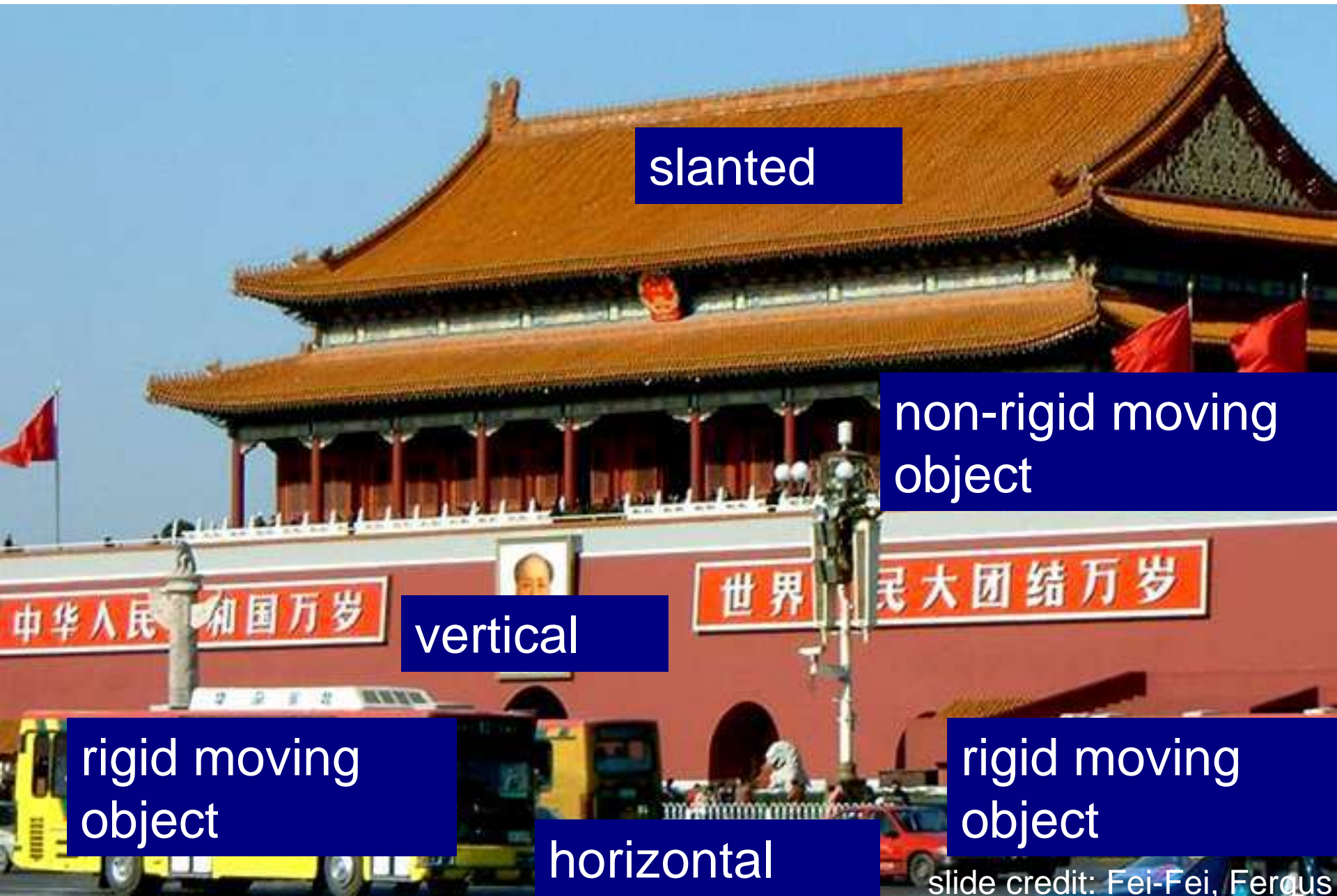
bus

bus

cars

slide credit: Fei-Fei, Fergus

Qualitative spatial information



slanted

non-rigid moving
object

vertical

rigid moving
object

rigid moving
object

horizontal

slide credit: Fei-Fei, Fergus

Challenges: viewpoint variation



Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus

Challenges: illumination

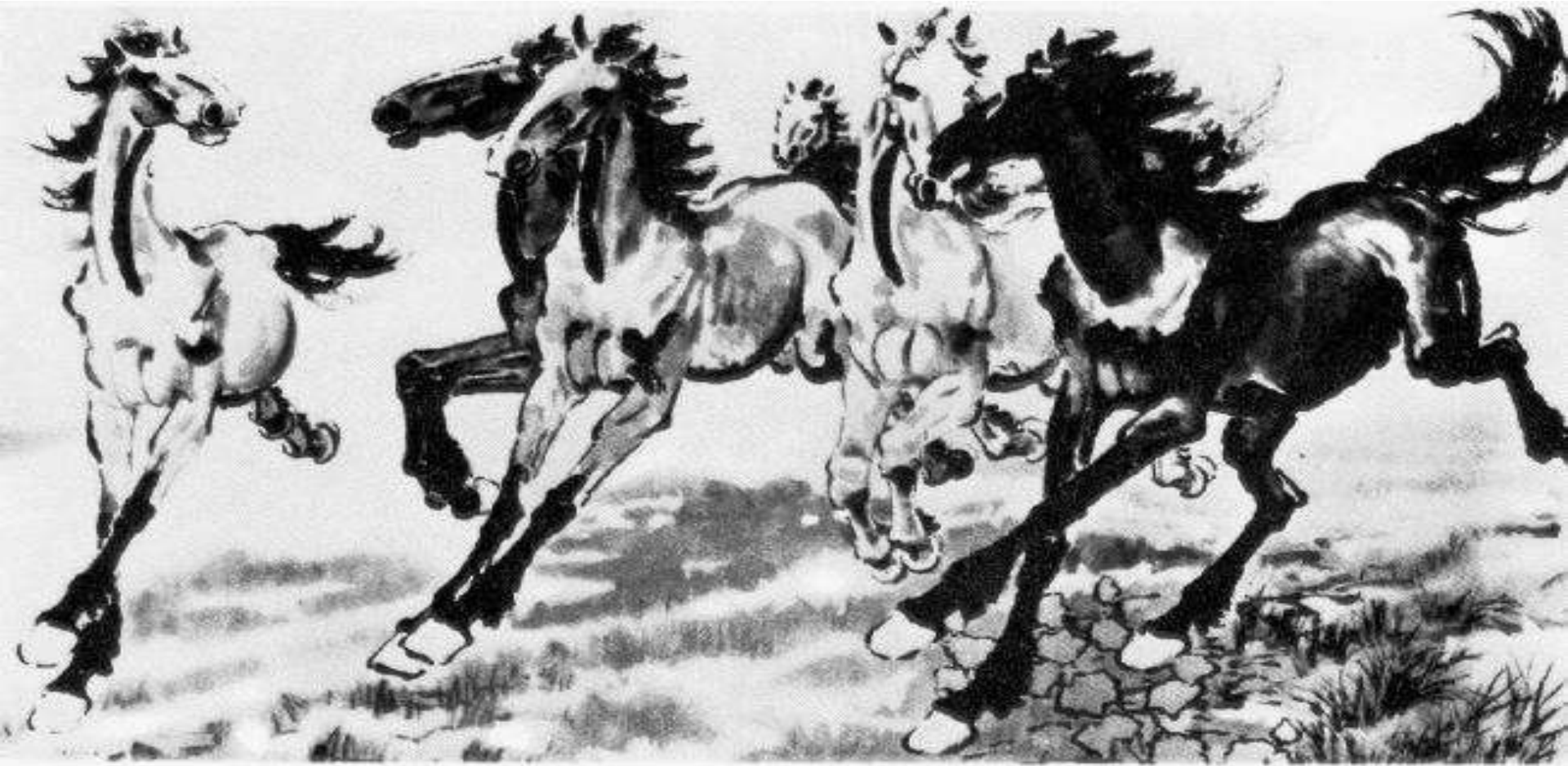


Challenges: scale



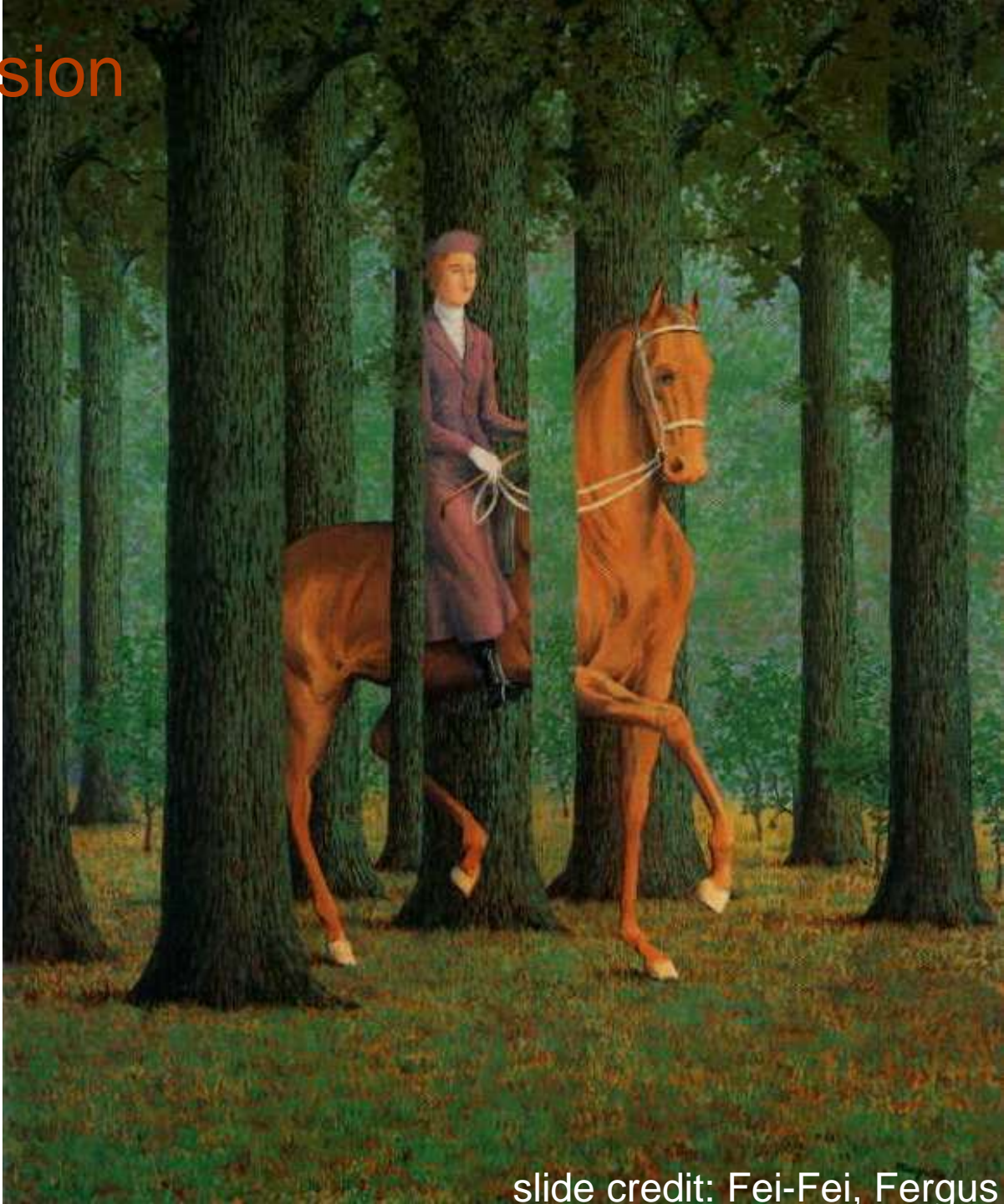
slide credit: Fei-Fei, Fergus

Challenges: deformation



Xu, Beihong 1943

Challenges: occlusion



Magritte, 1957

slide credit: Fei-Fei, Fergus

Challenges: background clutter



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

Challenges: object intra-class variation



slide credit: Fei-Fei, Fergus

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!



Computer Vision Publications

- Journals
 - IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)
 - #1 IEEE, Thompson-ISI impact factor: 5.96
 - #1 in both electrical engineering and artificial intelligence
 - #3 in all of computer science
 - Internal Journal of Computer Vision (IJCV)
 - ISI impact factor: 5.358, Rank 2 of 94 in “CS, artificial intelligence
 - IEEE Trans. on Image Processing
 - ...
- Conferences
 - Conf. of Computer Vision and Pattern Recognition (CVPR), once a year
 - International Conference on Computer Vision (ICCV), once every two years
 - Europe Conference on Computer Vision (ECCV), once every two years