CS410 Computer Vision

John Mc Donald
Department of Computer Science
National University of Ireland, Maynooth

Reading

[1] ***Machine Vision*** - Ramesh Jain, Rangachar Kasturi, and Brian G. Schunck. *McGraw-Hill, 1995*. 0-07-113407-7
[2] **Computer Vision: A Modern Approach** David A. Forsyth and Jean Ponce Prentice Hall, 2003. ISBN 0-12-085198-1.
[3] ***Machine Vision*** - David Vernon. Prentice Hall.
[4] ***A Guided Tour of Computer Vision*** - Vishvjit Nalwa. *Addison-Wesley.*
[5] ***Introductory Techniques for 3-D Computer Vision*** - Emanuele Trucco and Alessandro Verri. *Prentice Hall.*
[6] ***Robot Vision*** - Berthold K.P. Horn. *MIT Press.*

Why Study Computer Vision?

Esoteric Argument:-

- All naturally-occurring intelligent life-forms exhibit an ability to interact with and manipulate their environment in a coherent and stable manner.

- This interaction is facilitated by on-going intelligent interplay between perception and motion control (*i.e.* action).

Pragmatic Argument:-

Most manufacturers are concerned with the cosmetic integrity of their product; customers quite often equate quality of appearance with functional quality.

It is highly desirable therefore that :

- the product is checked visually before packaging and shipping,
- the Inspection Process be automated and effected without human intervention.

What is Computer Vision ?

The world we live in, and experience, is filled with an endless variety of objects and ***it is by looking and seeing that we come to know what is where*** in this world.

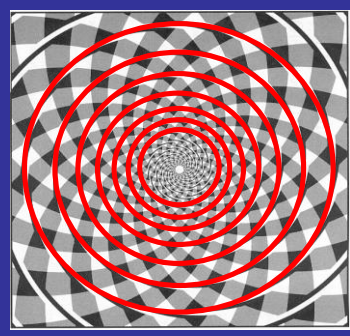
Vision is a means to an end: to know the world by looking

Computer Vision is exactly the same except that the medium by which the knowledge is gained is now a computational instrument rather than the brain of some living creature.

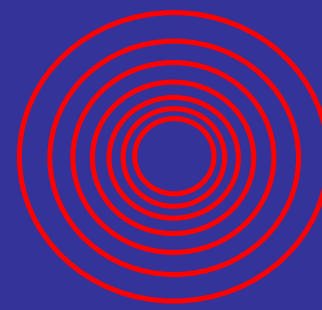
So should we concentrate our efforts on simulating the human visual system?

- Firstly, what we know of the human visual system is extremely limited, the only part of which is well understood being the eye itself (a subject we will revisit)
- Secondly, even though we have great faith in the human visual system, as we will see over the next few slides, it is a system which is far from infallible.

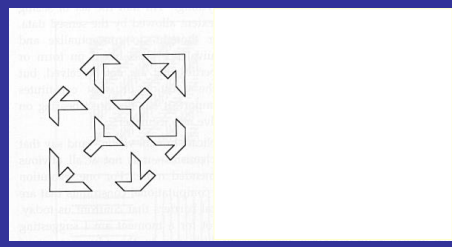
Visual Illusions: Fraser's Spiral



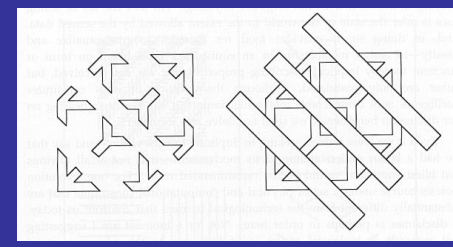
Fraser's Spiral

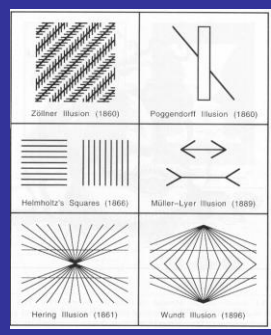


Kanizsa's Cube

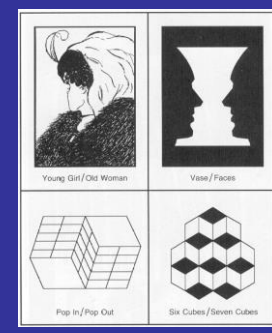


Kanizsa's Cube

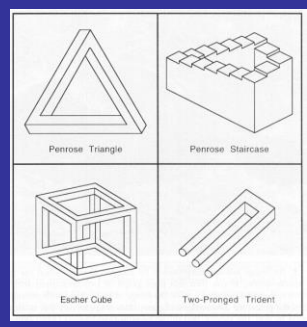


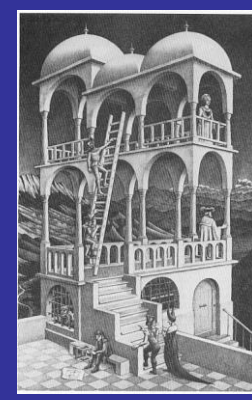


Visual Ambiguities



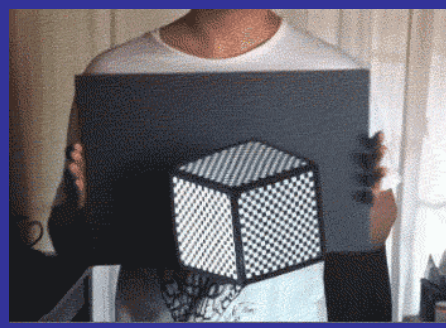
Visual Inconsistencies





- All the examples so far may seem contrived, and their success is dependent on their lack of realism
- Don't take too much consolation in this point, for illusions, ambiguities, and inconsistencies do arise in two dimensional visual projections of the *real world*

www.scifun.ed.ac.uk



Source: unknown

Conclusions??

- "Every image is the image of a thing merely for him/her who knows how to read it, and who is enabled by the aid of the image to form an idea of the thing" Herman von Helmholtz 1910
- Whereas we are quite forgiving when it comes to performance of humans, we are not quite as charitable when it comes to the performance of machines.

Computer Vision: Why is it so difficult?

Computer Vision is concerned with the physical structure of a three-dimensional world by the automatic analysis of images of that world.

Isn't this just the inverse to computer graphics?

The image is two-dimensional. We inevitably lose information in the projection process, *i.e.*, in passing from a 3D world to a 2D image.

The images are digital images :

- they are a discrete representation (*i.e.* they have distinct values at regularly sampled points)
- they are a quantised representation

(*i.e.* each value is an integer value)

Mathematicians would refer to vision as an ill-posed problem

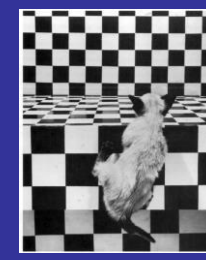
Computer Graphics is not!

Using Knowledge

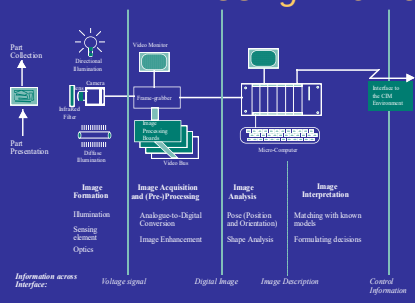
- So how do we get around this problem?
- Well even though the image does not contain enough information to reconstruct the scene..
- If we couple the information in the image with our knowledge (or assumptions) about the world we can then sufficiently constrain the situation to arrive at a single (or small number) of solutions.

Cues

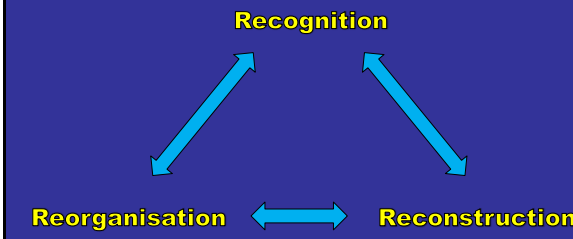
- Image formation
- Image features
- Shading
- Occlusion
- Texture
- Motion
- Multiple view geometry



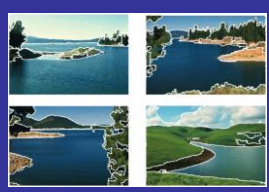
Using Knowledge



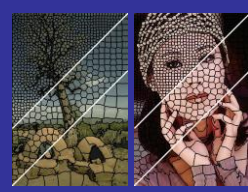
Another view: The 3 R's



Reorganisation



Comaniciu, D.; Meer, P., "Mean shift: a robust approach toward feature space analysis," PAMI, 24(5), 2002.



Achanta et al., SLIC Superpixels Compared to State-of-the-art Superpixel Methods, PAMI, 34(11), 2012.

Recognition

Everingham et al., The PASCAL Visual Object Classes (VOC) Challenge International Journal of Computer Vision, 88(2), 303-338, 2010

Recognition

Everingham et al., The PASCAL Visual Object Classes (VOC) Challenge International Journal of Computer Vision, 88(2), 303-338, 2010

Recognition

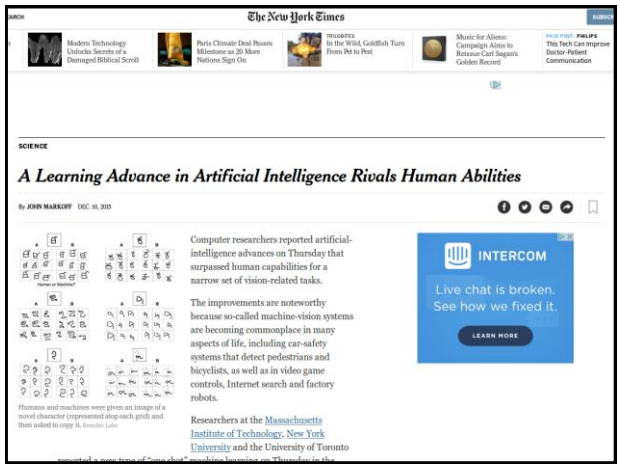
VOC Challenge:
~20K Images
20 Categories

Everingham et al., The PASCAL Visual Object Classes (VOC) Challenge International Journal of Computer Vision, 88(2), 303-338, 2010

ILSVR Challenge:
~1.4M Images
1000 Categories

Image Courtesy: A Karpathy

Russakovsky et al., ImageNet Large Scale Visual Recognition Challenge, 115(3), 211-252, 2015



The New York Times

Modern Technology Unlocks Secrets of a Damaged Biblical Scroll

Paris Climate Deal Passes Milestone as 20 More Nations Sign On

TRILOGIES In the Wild, Goldfish Turn From Pet to Pest

Music for Aliens: Campaign Aims to Reissue Carl Sagan's Golden Record

PAID POST: PHILIPS This Tech Can Improve Doctor-Patient Communication

SCIENCE

A Learning Advance in Artificial Intelligence Rivals Human Abilities

By JOHN MARKOFF DEC. 10, 2015

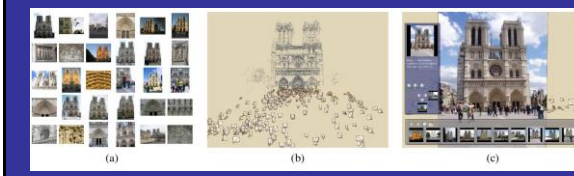
Computer researchers reported artificial-intelligence advances on Thursday that surpassed human capabilities for a narrow set of vision-related tasks.

The improvements are noteworthy because so-called machine-vision systems are becoming commonplace in many aspects of life, including car-safety systems that detect pedestrians and bicyclists, as well as in video game controls, Internet search and factory robots.

Researchers at the Massachusetts Institute of Technology, New York University and the University of Toronto

Humans and machines were given an image of a novel character (represented atop each grid) and then asked to copy it. Brenden Lake

Reconstruction



(a) (b) (c)

Noah Snavely, Steven M. Seitz, Richard Szeliski, "Photo tourism: Exploring photo collections in 3D," ACM Transactions on Graphics (SIGGRAPH Proceedings), 25(3), 2006, 835-846.

Photo Tourism

Exploring photo collections in 3D

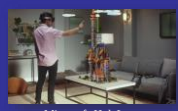
Noah Snavely Steven M. Seitz Richard Szeliski
University of Washington *Microsoft Research*

SIGGRAPH 2006

Real-time Dense Visual Mapping

[Whelan, Kaess, Fallon, Johannsson, Leonard, McDonald, RGBD '12.]

Application Areas


Microsoft HoloLens

DAQRI Smart Helmet


Dyson 360 Eye Vacuum

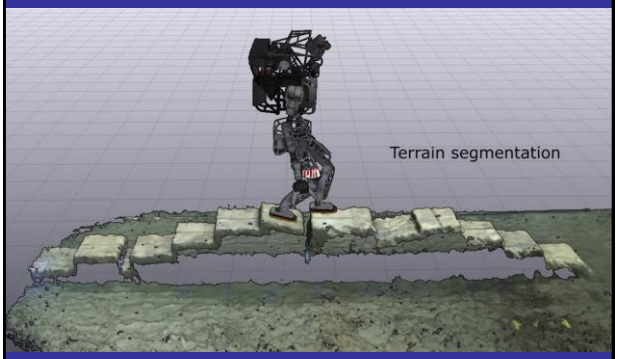

Google Project Tango


Google Self-Driving Car Project

Continuous Humanoid Locomotion Enabled by Online Footstep Planning and Stereo Fusion

MIT DARPA Robotics Challenge Team drc.mit.edu

 [Fallon et al.]



 [Fallon et al.]

What's ahead?

- Camera modelling and calibration
- Binary Image Processing
- Image Filtering
- Feature Detection
- Image Analysis