Computer Vision

Introduction

Andrea Torsello
torsello@dsi.unive.it

Getting in contact

- Student -> Teacher

  torsello@dais.unive.it

- Teacher -> Strudents
  - course web site

    http://www.dsi.unive.it/~atorsell/Visione

Texts

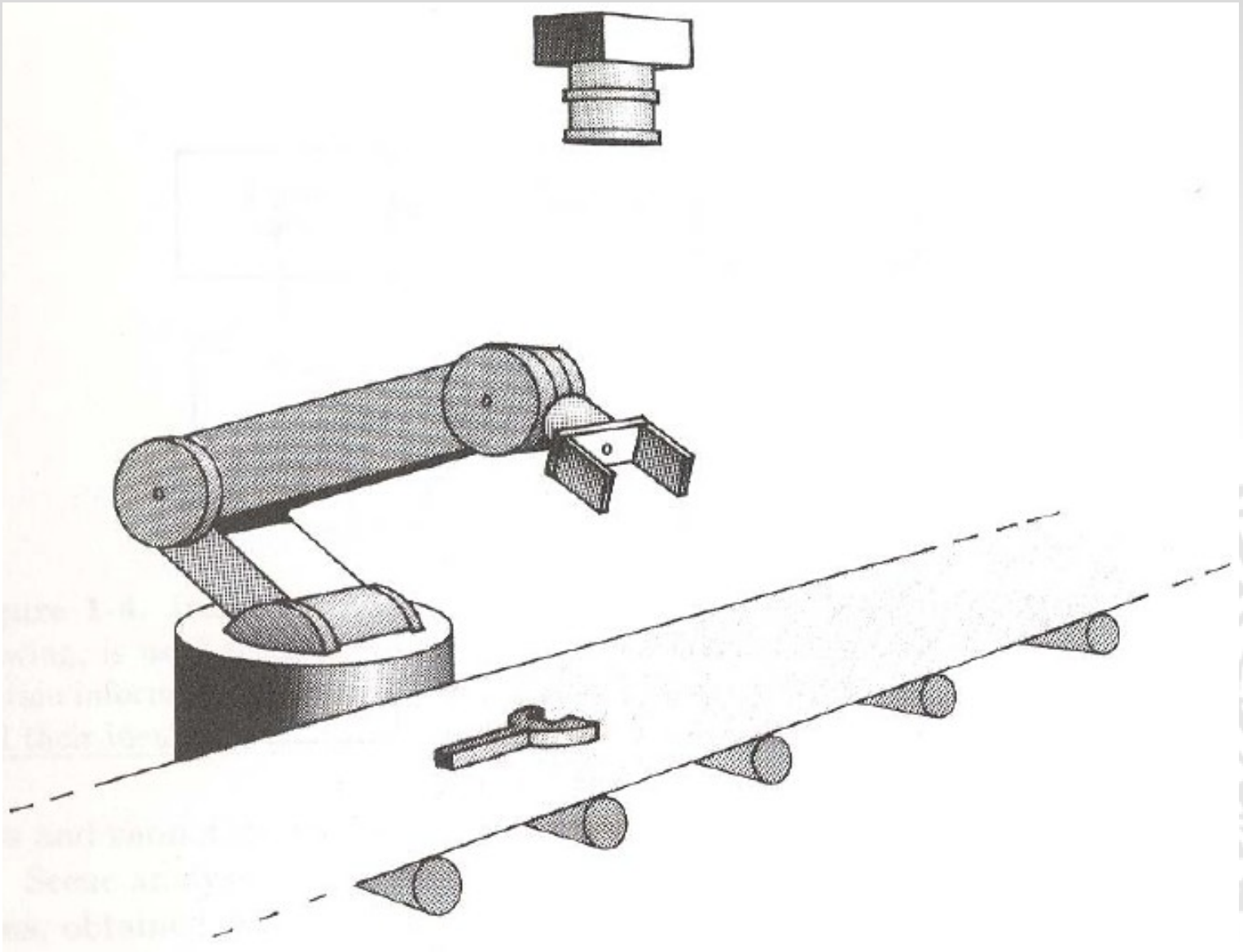
- R. C. Gonzalesz and R. E. Woods. Digital Image Processing. Prentice-Hall, 2002
- D. Forsyth and J. Ponce. Computer Vision. A Modern Approach. Prentice-Hall, 2002
- E. Trucco and A. Verri. Introductory Techniques for 3D Computer Vision. Prentice-Hall, 1998
- R. Szeliski. Computer Vision. Springer 2011.

Goals

What does it mean to see?

Computer vision:

Build a system that “sees”



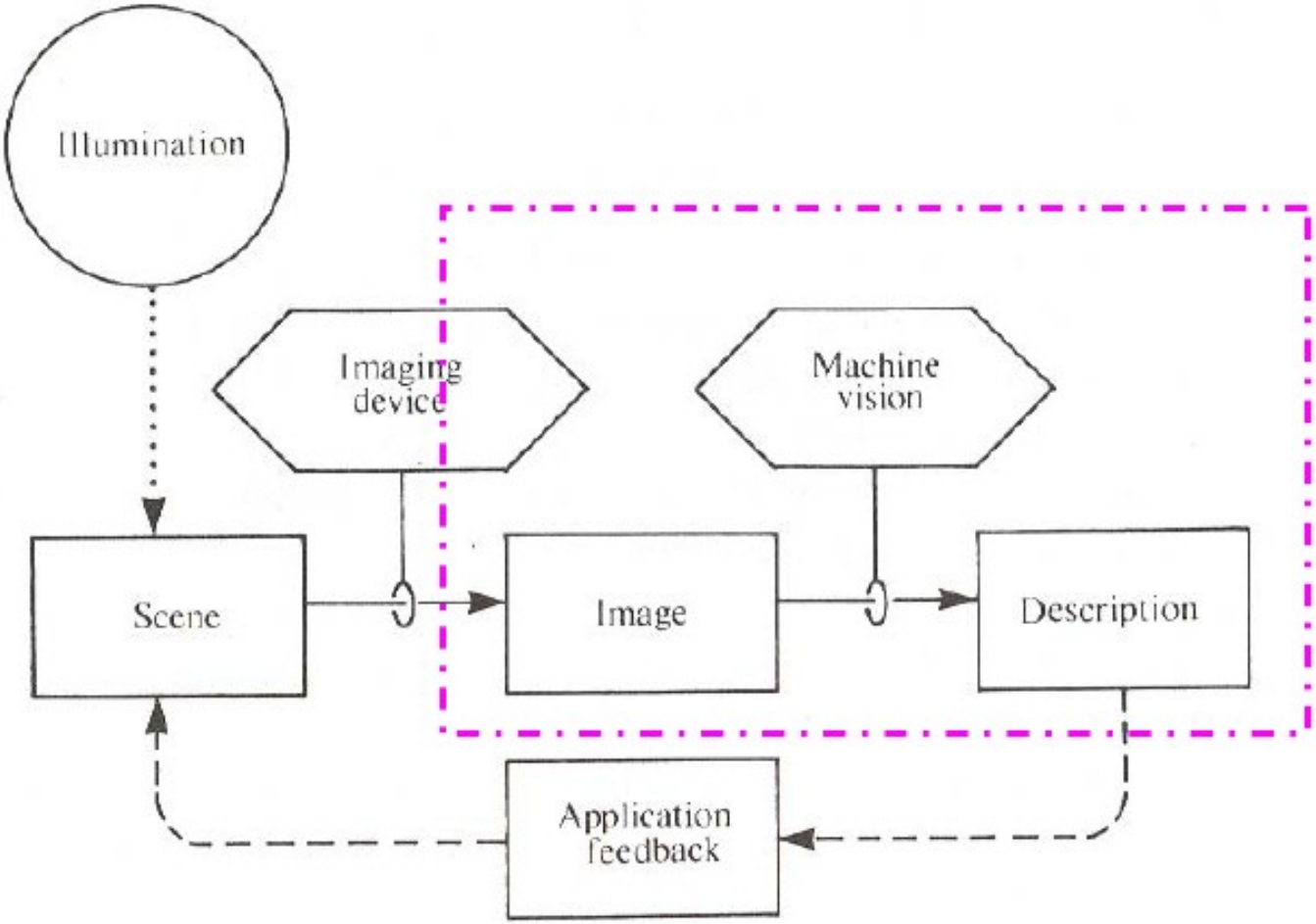
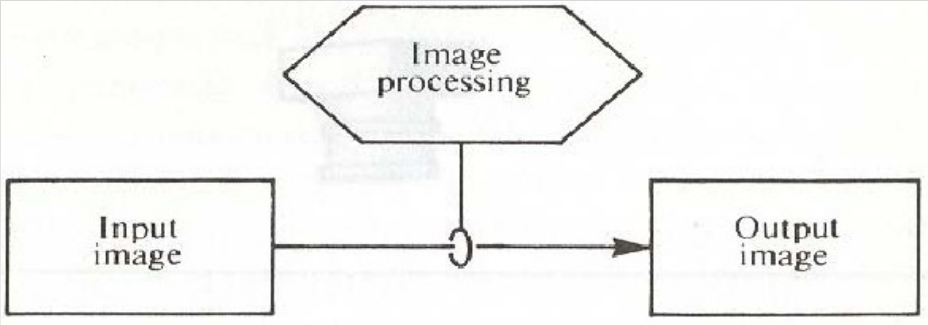


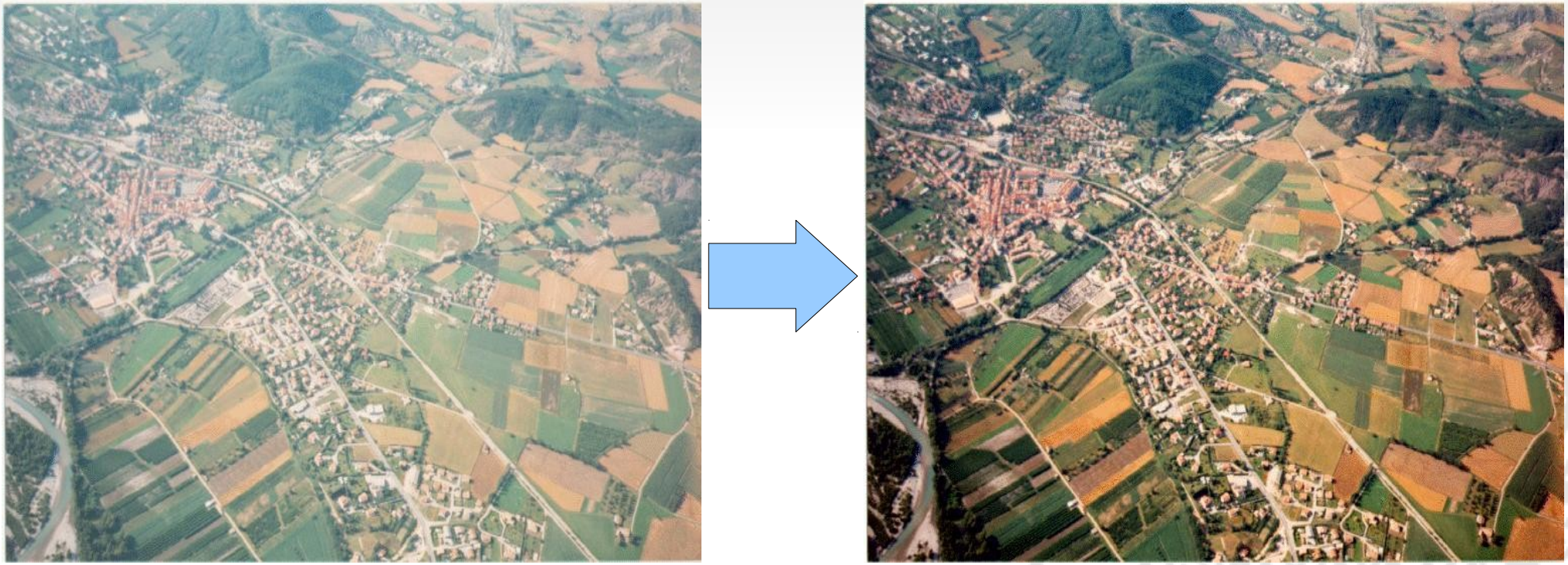
Figure 1-2. The purpose of a machine vision system is to produce a symbolic description of what is being imaged. This description may then be used to direct the interaction of a robotic system with its environment. In some sense, the vision system's task can be viewed as an inversion of the imaging process.

Related Fields

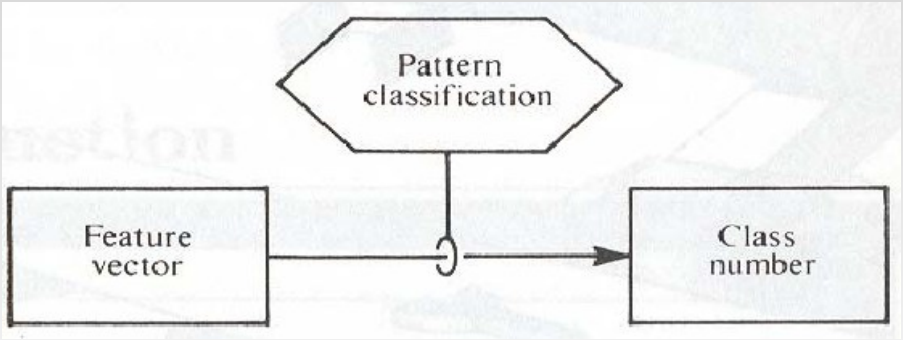
- Image processing
- Pattern recognition (Machine learning/Data mining)
- Scene analysis

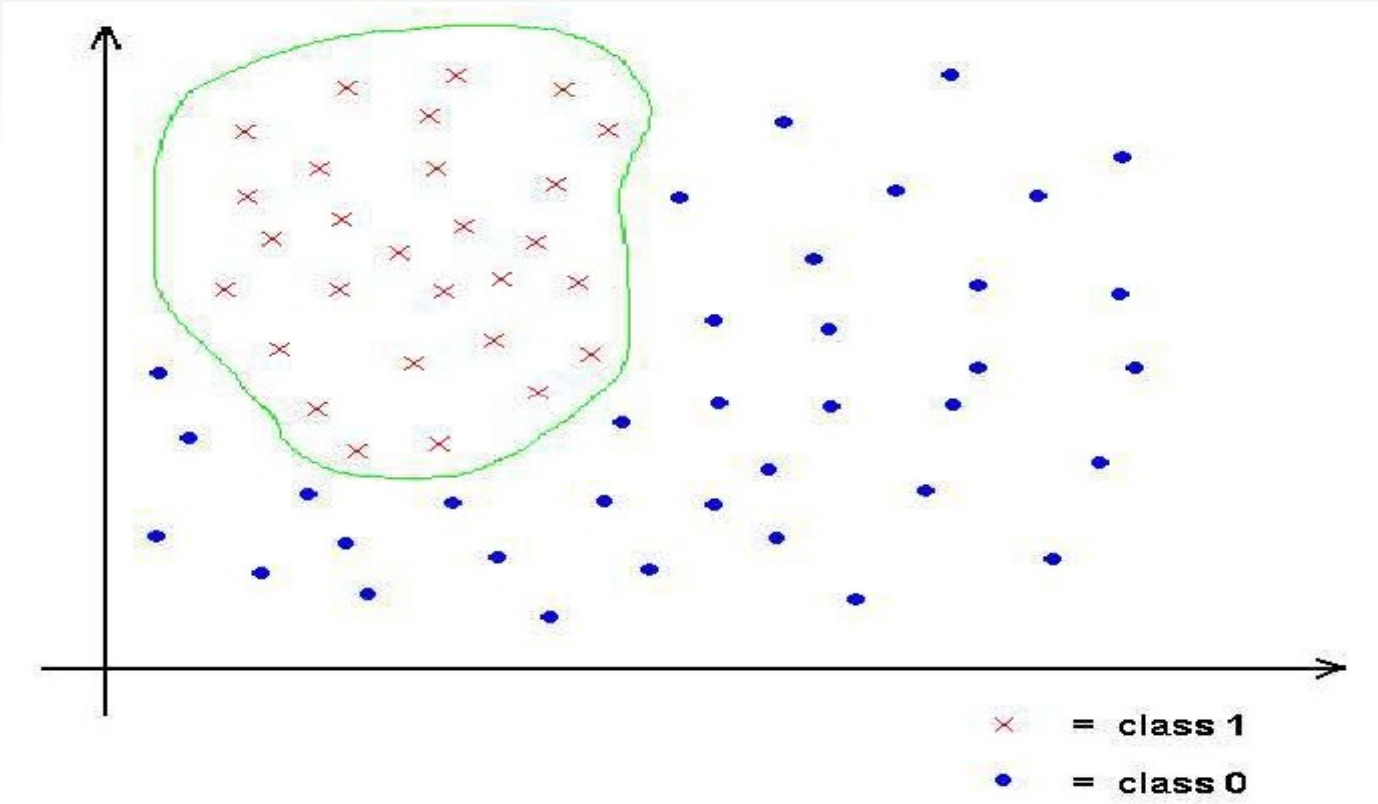
Image Processing



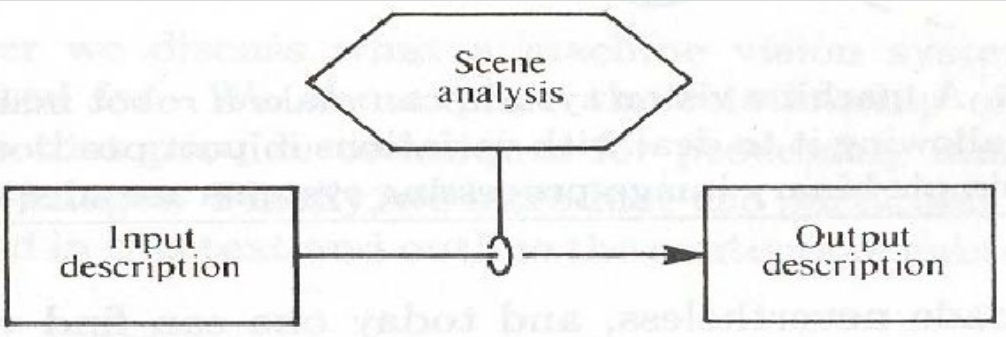


Pattern recognition





Scene Analysis



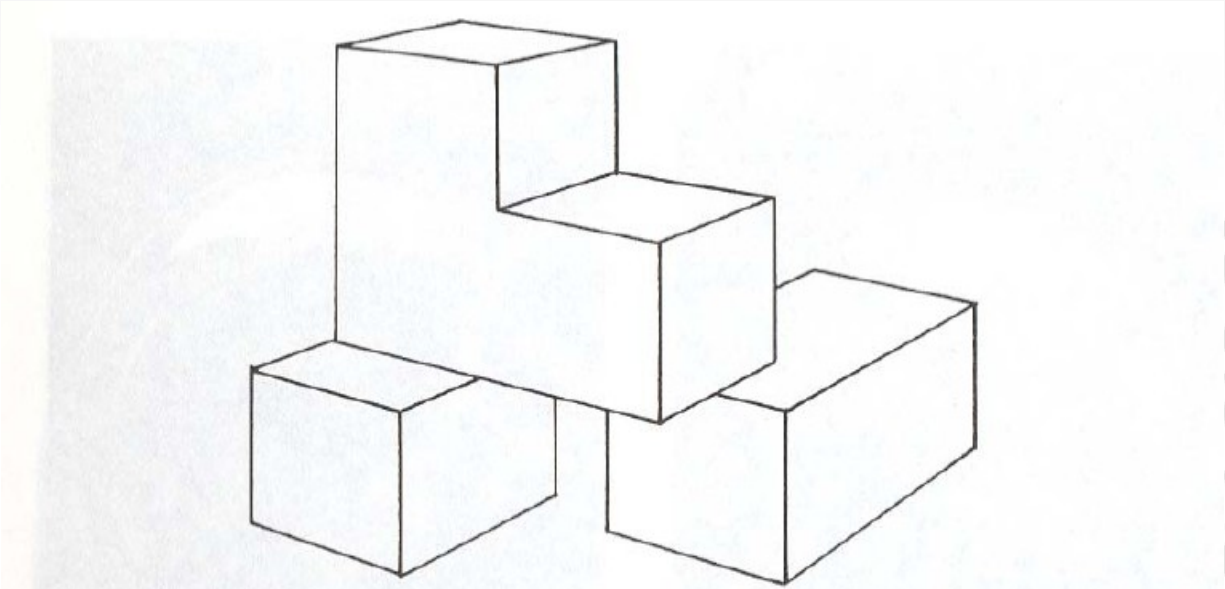


Figure 1-4. In scene analysis, a low-level symbolic description, such as a line drawing, is used to develop a high-level symbolic description. The result may contain information about the spatial relationships between objects, their shapes, and their identities.

Taxonomy

High/Medium/Low level vision

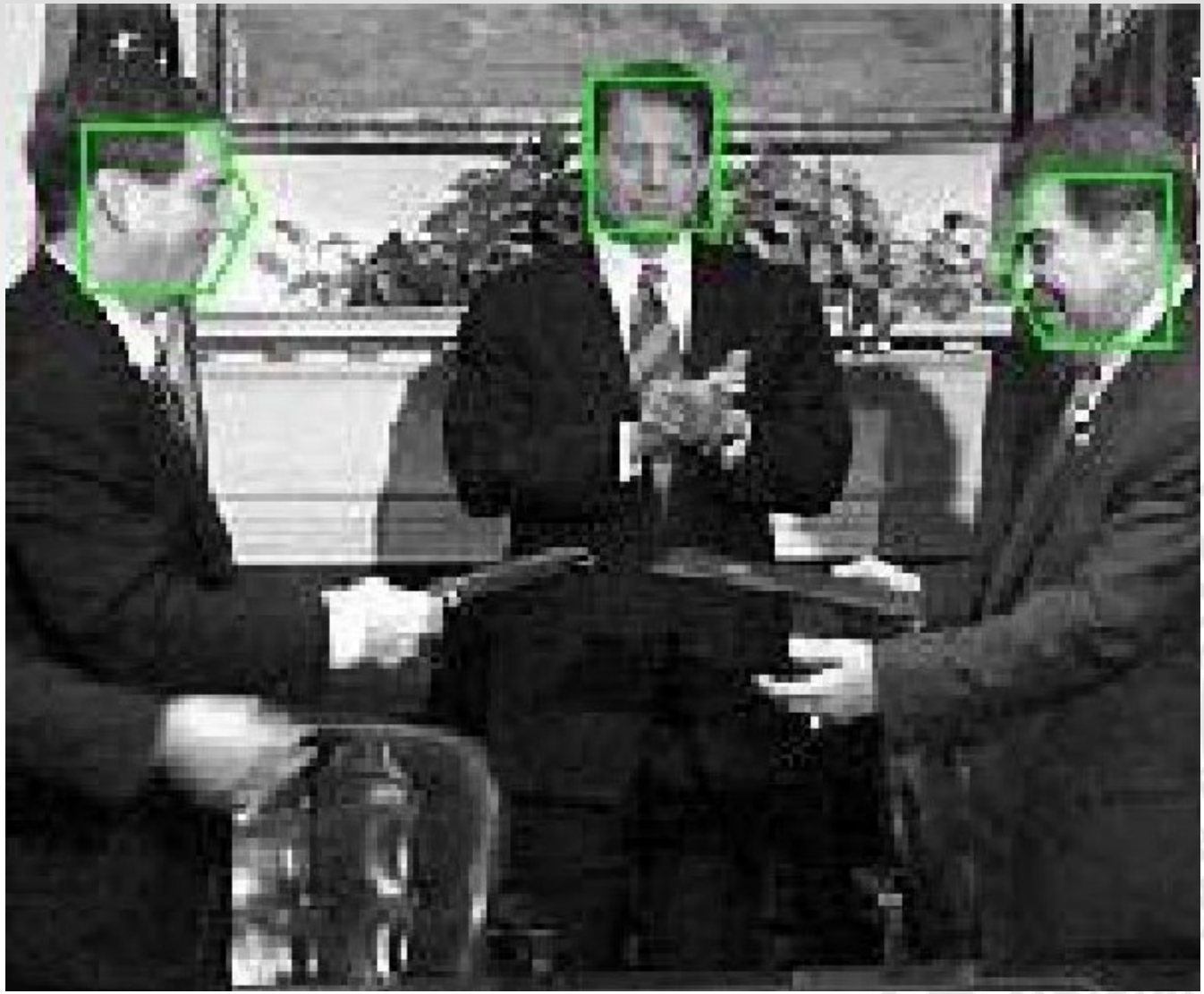
- **Low** (image restopration, contrast enhancement, …):
  - input=image, output=image
- **Medium** (segmentaion, shape recogntition, …):
  - input=images, output=attributes/features
- **High** (Image analisys, scene understanding, …):
  - input=images, output=concepts

Some applications of Computer Vision

Image search

From a search for horse pix in 100 horse images and 1086 non-horse images

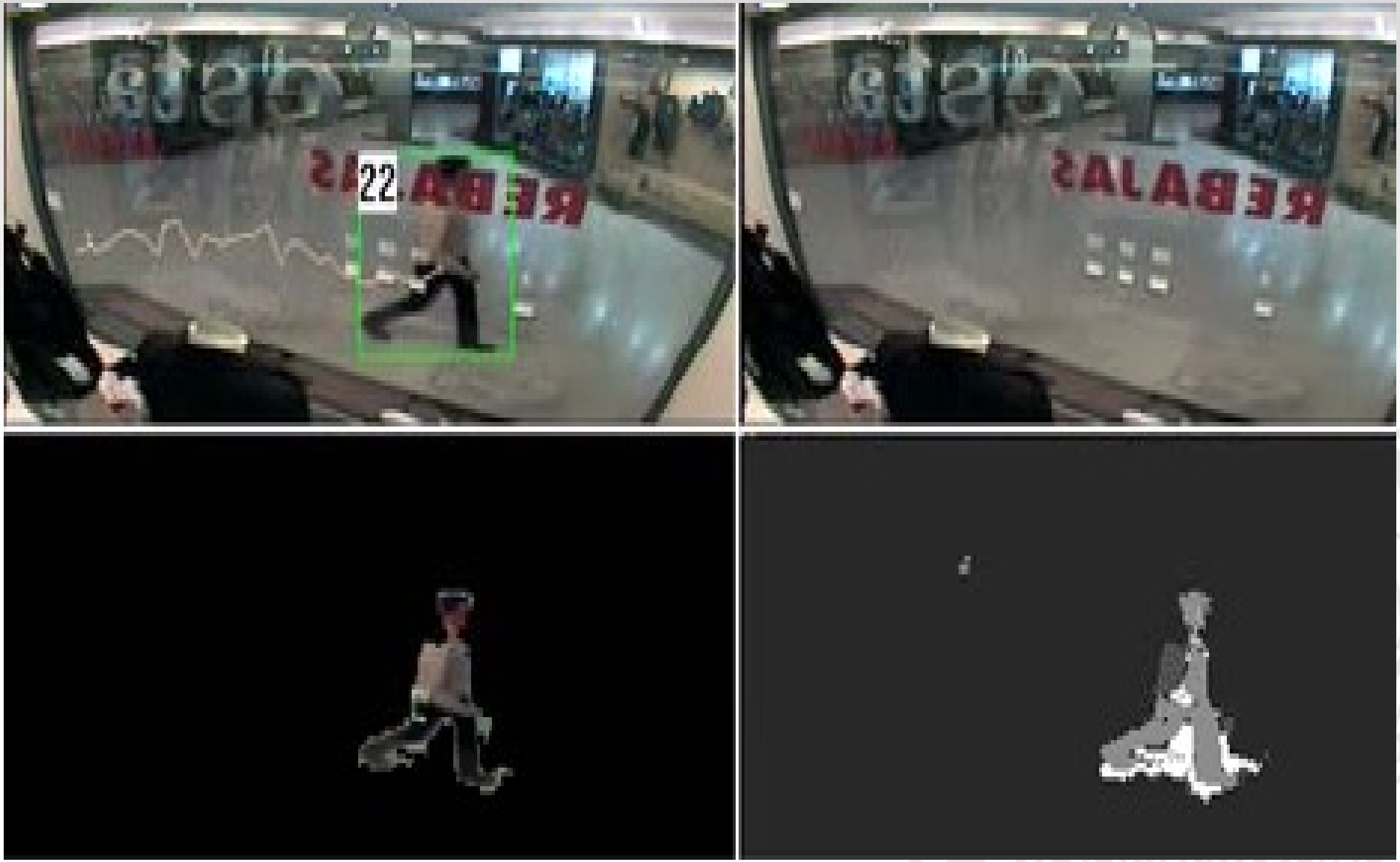
Face Detection



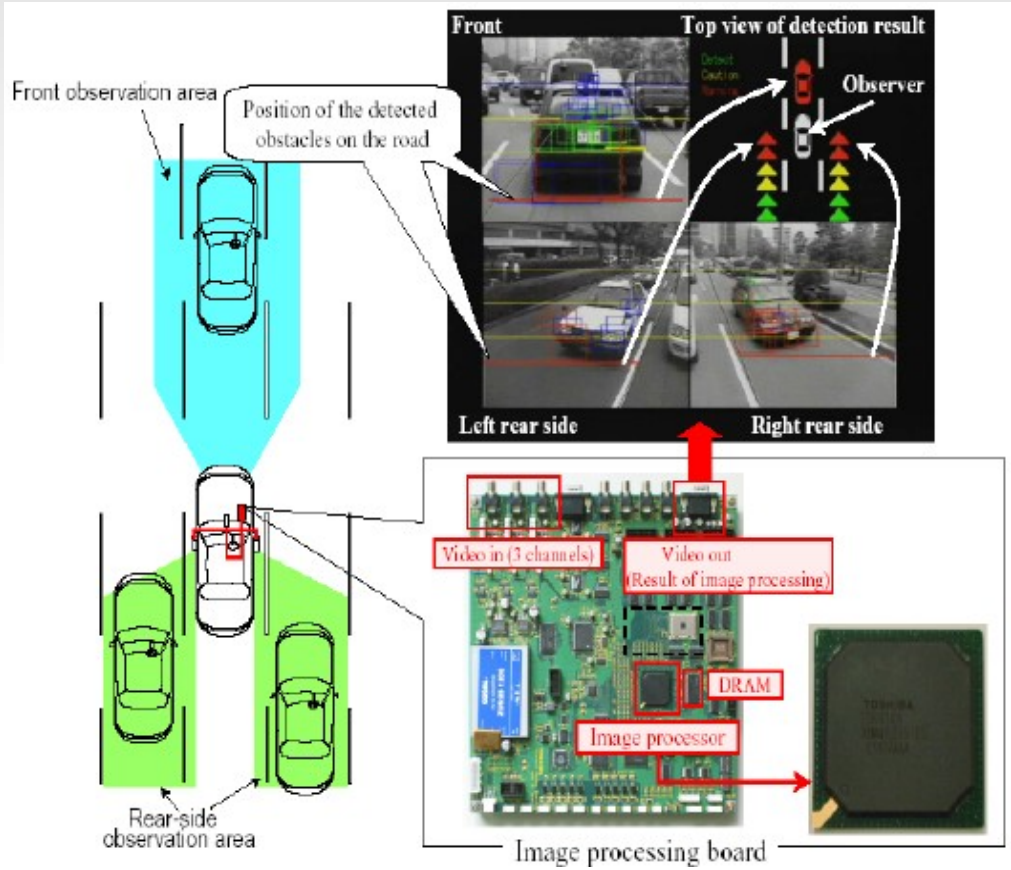
Face Recognition

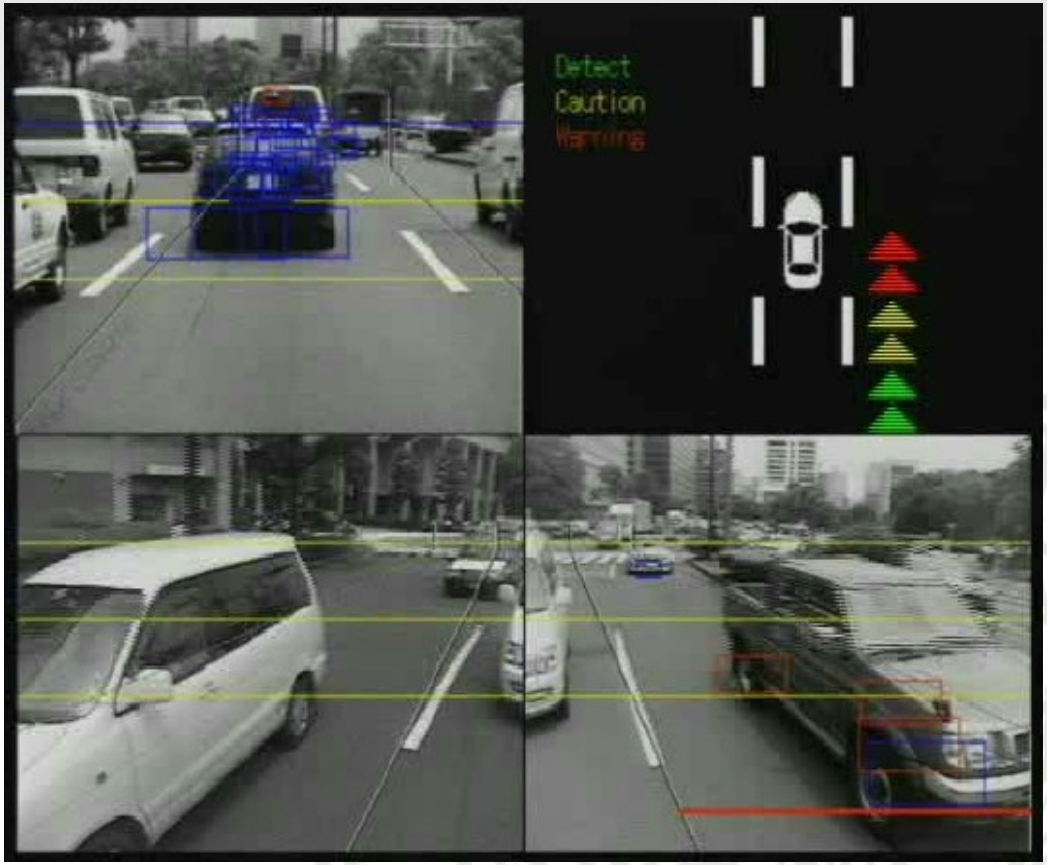
Robotics

Surveilance



Driving Assistance





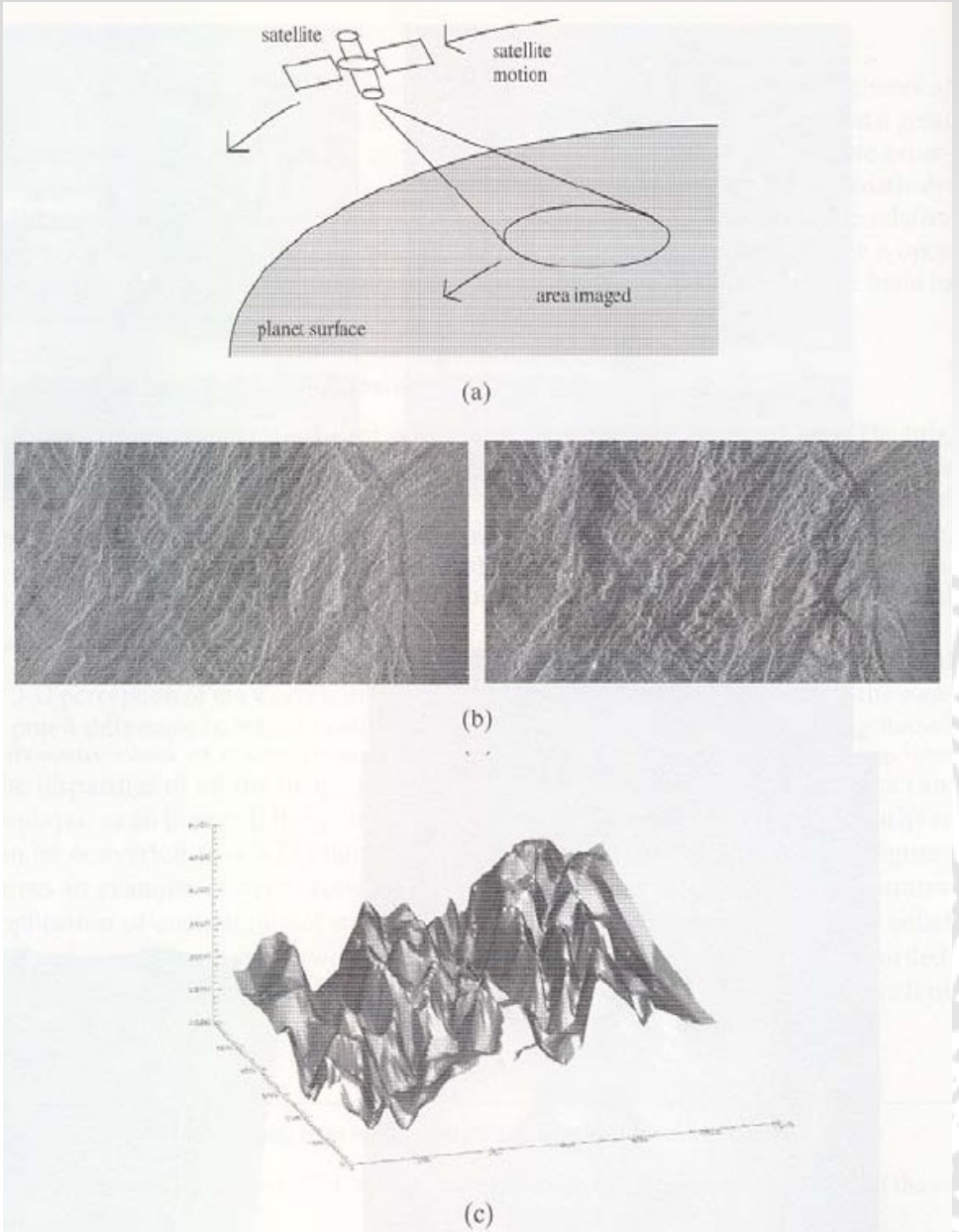
Autonomous Vehicles

Tartan racing (Carnegie Mellon University, General Motors Corporation, Caterpillar, Continental e altri)

vinctori DARPA Urban Challenge 2007

Satelite surveys

- Stereo reconstruction of the surace of Venus from a pair of satelite images
- (Institute for Computer Graphics and Vision, Technical University of Graz, Austria)



Medical Image Analysis

Digitalization

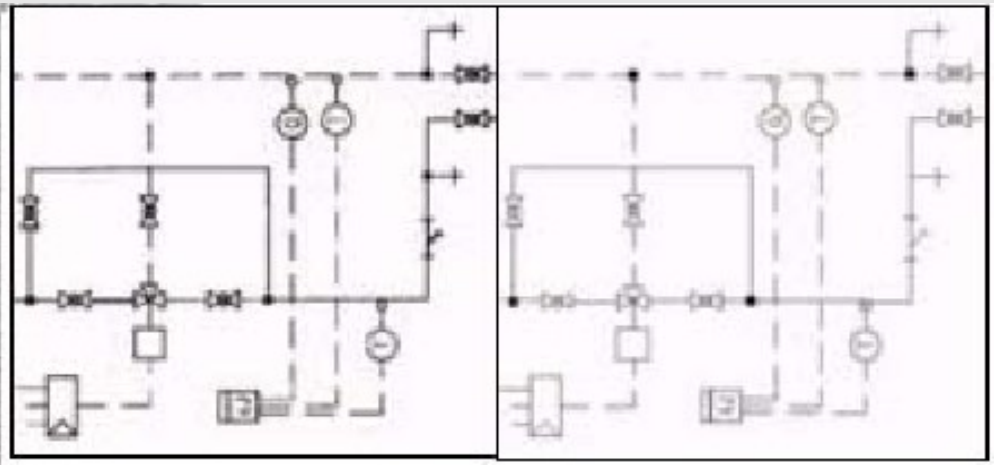
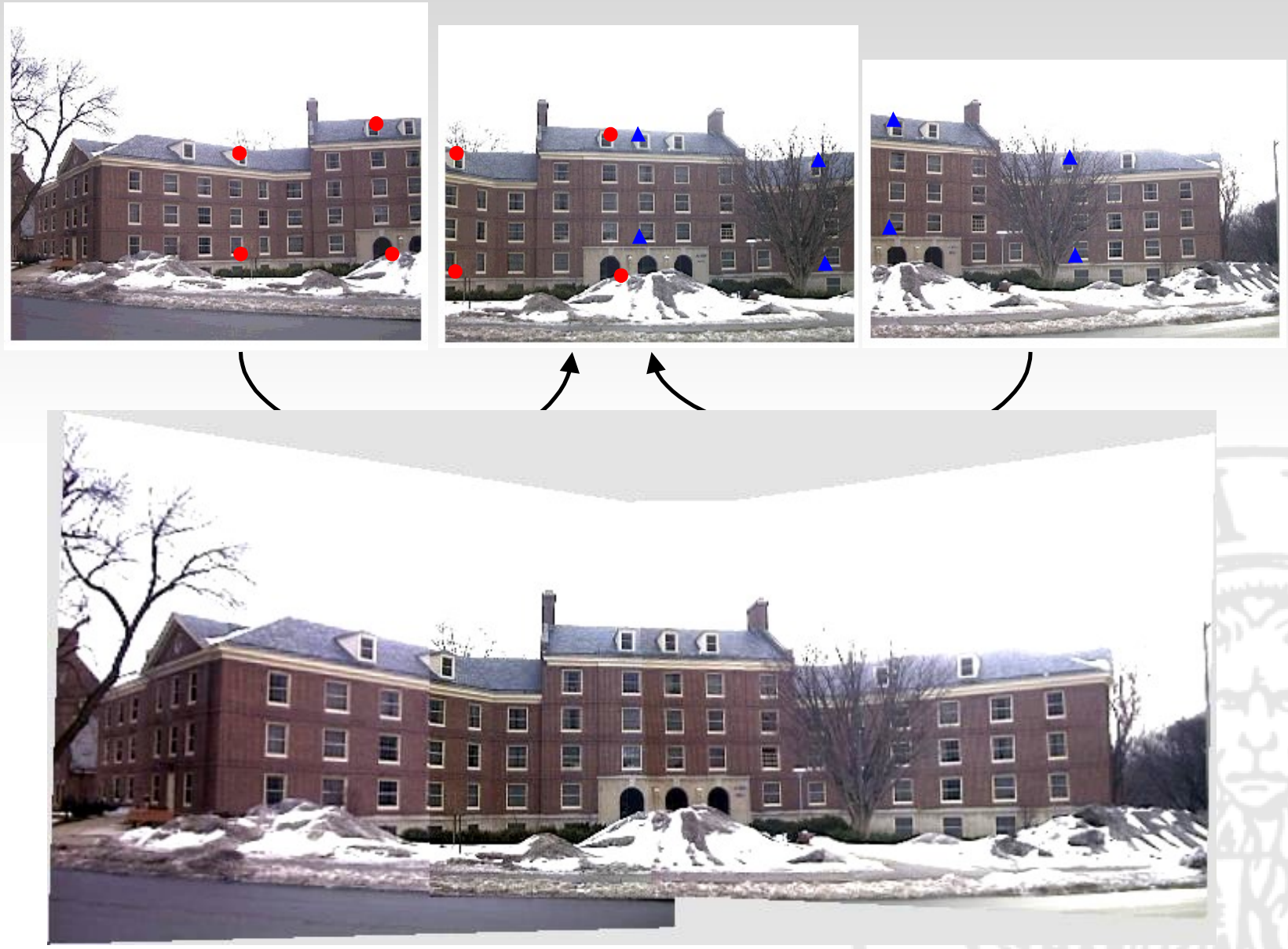


Immagine raster

Immagine vettoriale

Digital Photography

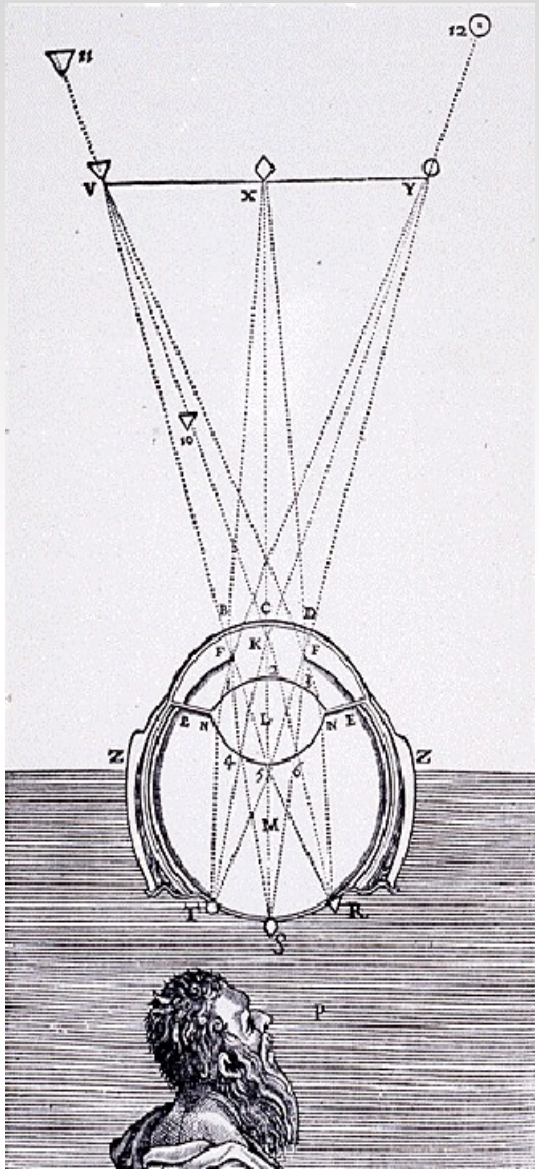


Cinema & TV

Some Hystory

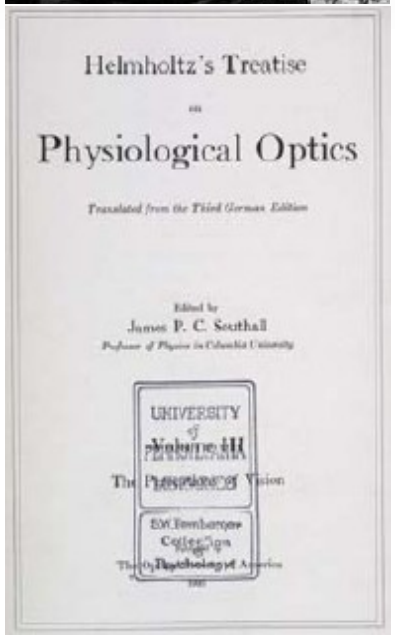
- Emissionist theory (tactile vision)
  - The eye emits rays that travel through space and hit objects. The contact of the ray and the objects causes the visual sensation.
    - Pythagoras, Empedocles, Euclid
- Intromissionist theory
  - Object continuously emit to the surounding space imaeges of them selves. This images (eidola) enter the eye throught the pupil, thus revealign themselves.
    - Democritus, Epicurus, Lucretius
- Persistence of retinal images (“after-image”)
  - Ibn al-Haytham (Alhazen) (965-1039 AD)

Keplero and the theory of retinal images



- 1604 Kepler develops the modern theory of retinal images.

Perception as unconscious conclusions Helmholtz



The psychic activities that lead us to infer that there in front of us at a certain place there is a certain object of a certain character, are generally not conscious activities, but unconscious ones. In their result they are equivalent to a *conclusion*, to the extent that the observed action on our senses enables us to form an idea as to the possible cause of this action; although, as a matter of fact, it is invariably simply the nervous stimulations that are perceived directly, that is, the actions, but never the external objects themselves. But what seems to differentiate them from a conclusion, in the ordinary sense of that word, is that a conclusion is an act of conscious thought. An astronomer, for example, comes to real conscious conclusions of this sort, when he computes the positions of the stars in space, their distances, etc., from the perspective images he has had of them at various times and as they are seen from different parts of the orbit of the earth. His conclusions are based on a conscious knowledge of the laws of optics. In the ordinary acts of vision this knowledge of optics is lacking. Still it may be permissible to speak of the psychic acts of ordinary perception as *unconscious conclusions*, thereby making a distinction of some sort between them and the common so-called conscious conclusions. And while it is true that there has been, and probably always will be, a measure of doubt as to the similarity of the psychic activity in the two cases, there can be no doubt as to the similarity between the results of such unconscious conclusions and those of conscious conclusions.

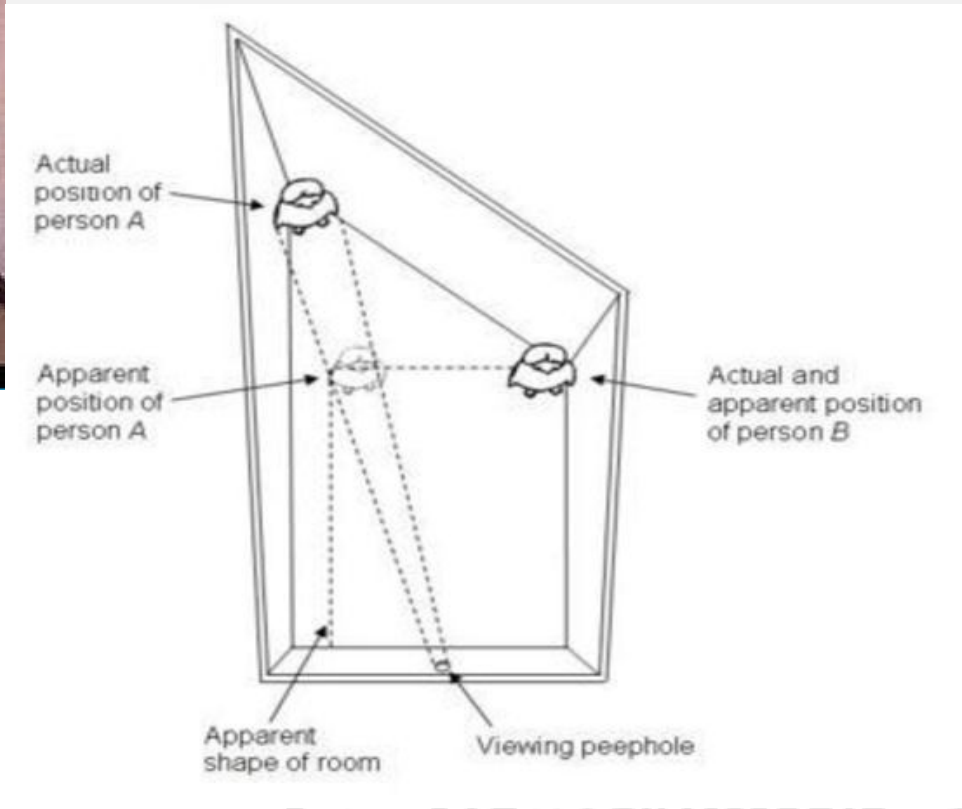
Recent models...

- James J. Gibson (1904-1979) and the ecologic approach
- "The belief of the empiricists that the perceived meanings and values of things are supplied from past the experience of the observer will not do. But even worse is the belief of nativists that meanings and values are supplied from the past experience of the race by innate ideas."
(Gibson, 1979)

- David Marr and the computational approach
- Three levels:
  - Computational
  - Algorithmic
  - Implementative (hardware)

Illusions

Ames' room


Actual position of person A
Apparent position of person A
Actual and apparent position of person B
Apparent shape of room
Viewing peephole

Necker Cube

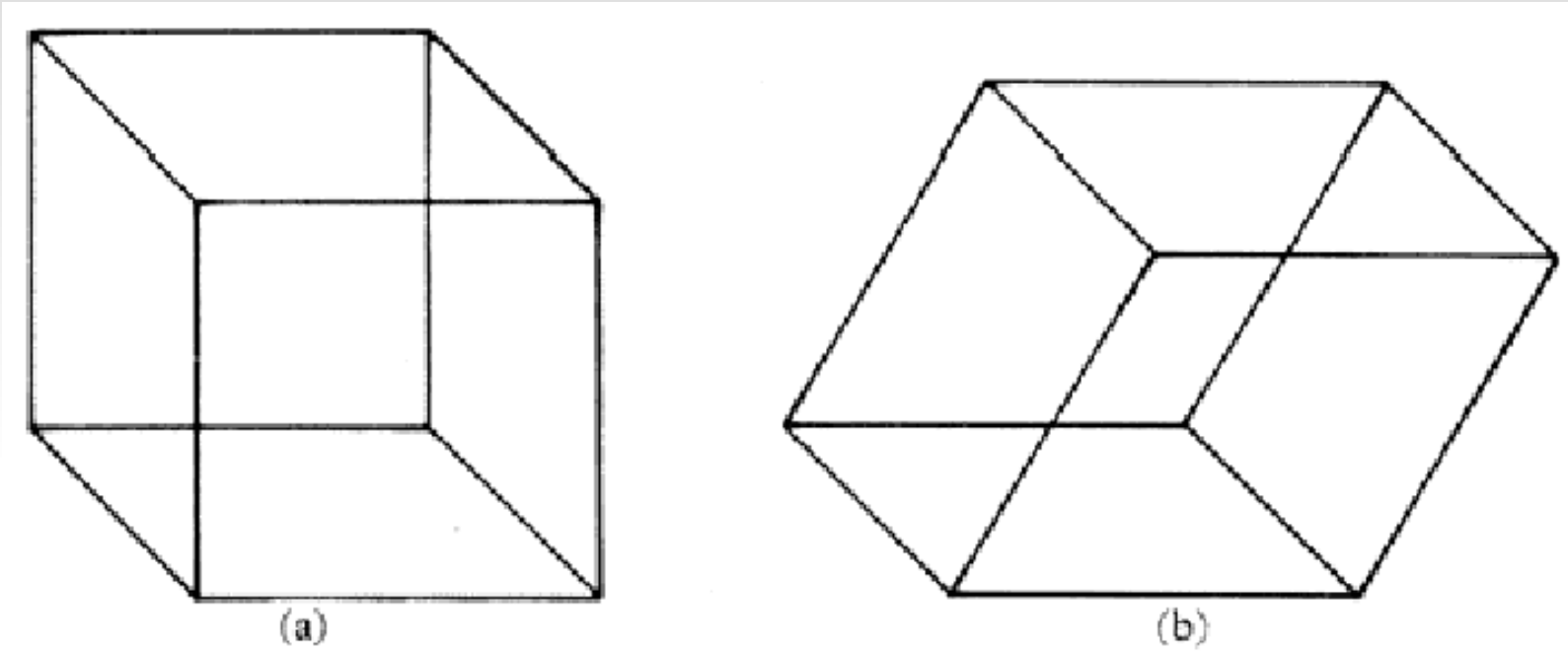
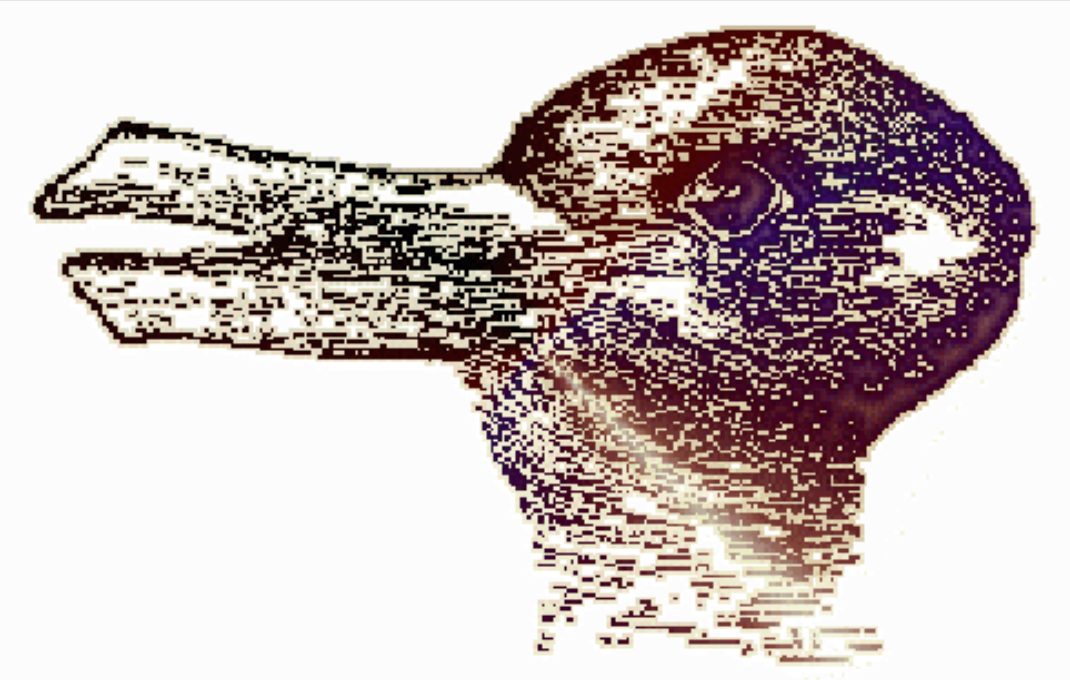


FIG. 2 (a) Necker cube. This is the most famous of many depth-ambiguous, figures. (When presented with no background it changes in shape with each reversal, the apparent back being larger than the apparent front face.) (b) Necker rhomboid. This is the original form, presented by L. A. Necker in 1832.

Bistable images



- Jastrow duck/rabbit

Contextual perception of length

- Mueller-Lyer illusion

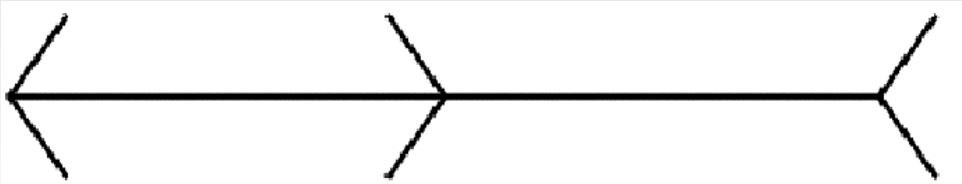


FIG. 5 Müller-Lyer arrows figure 1889. The most famous illusion: the outward-going 'arrow heads' produce expansion of the 'shaft' and the inward-going heads contraction.

- Ponzo illusion

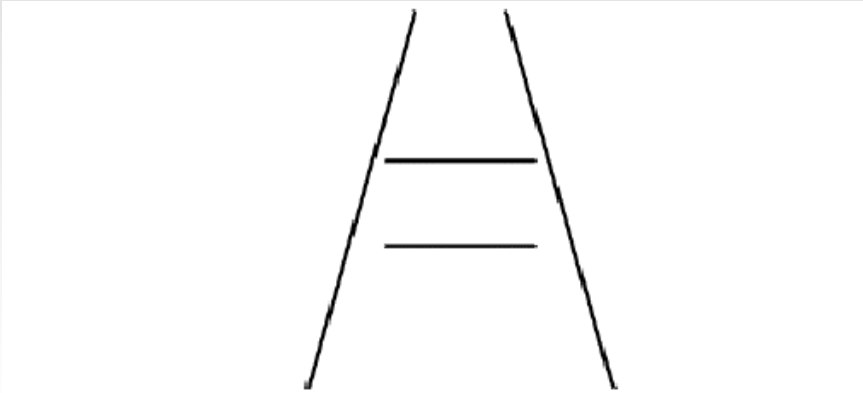
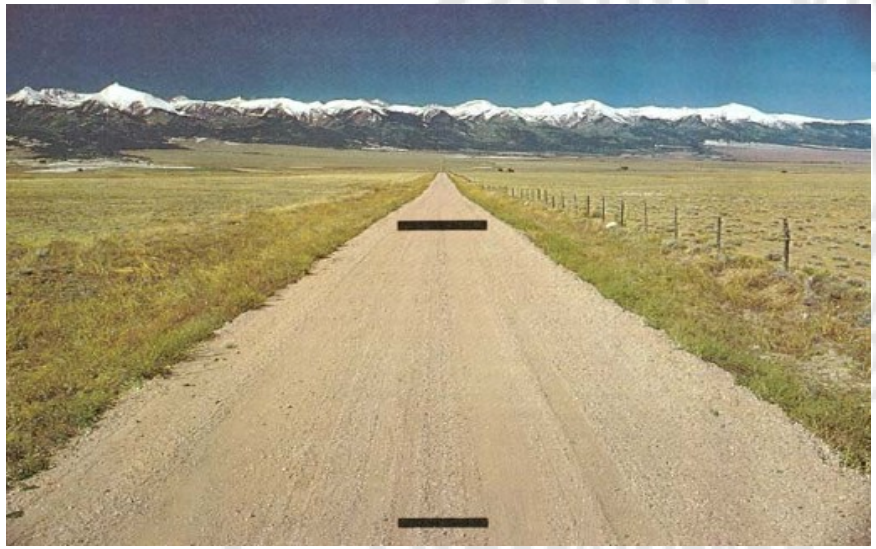
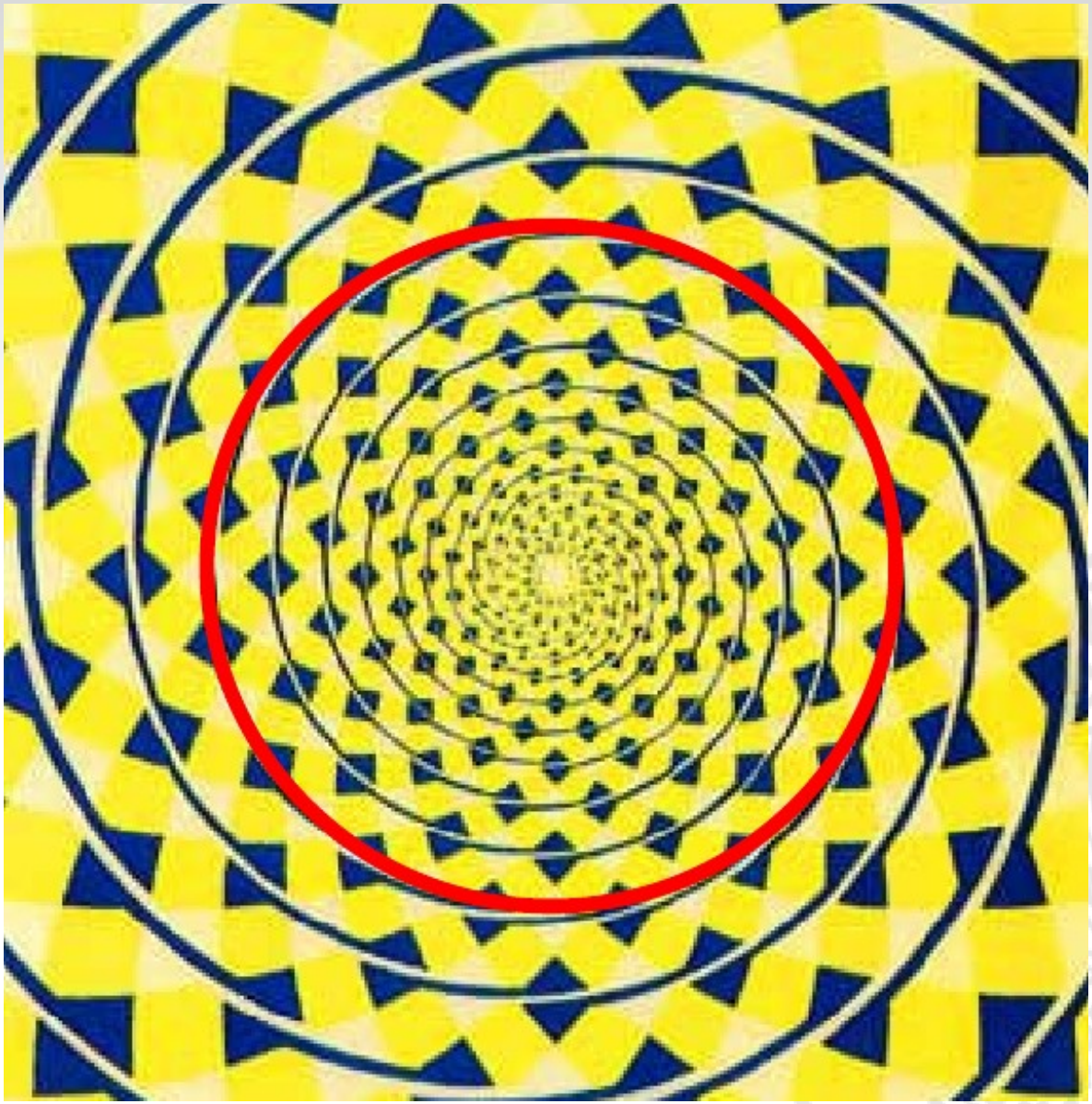


FIG. 6 Ponzo figure. The upper of the parallel lines is expanded with respect to the lower.



Fraser Spiral



Escher's impossible images

Bottom-up or Top-down?

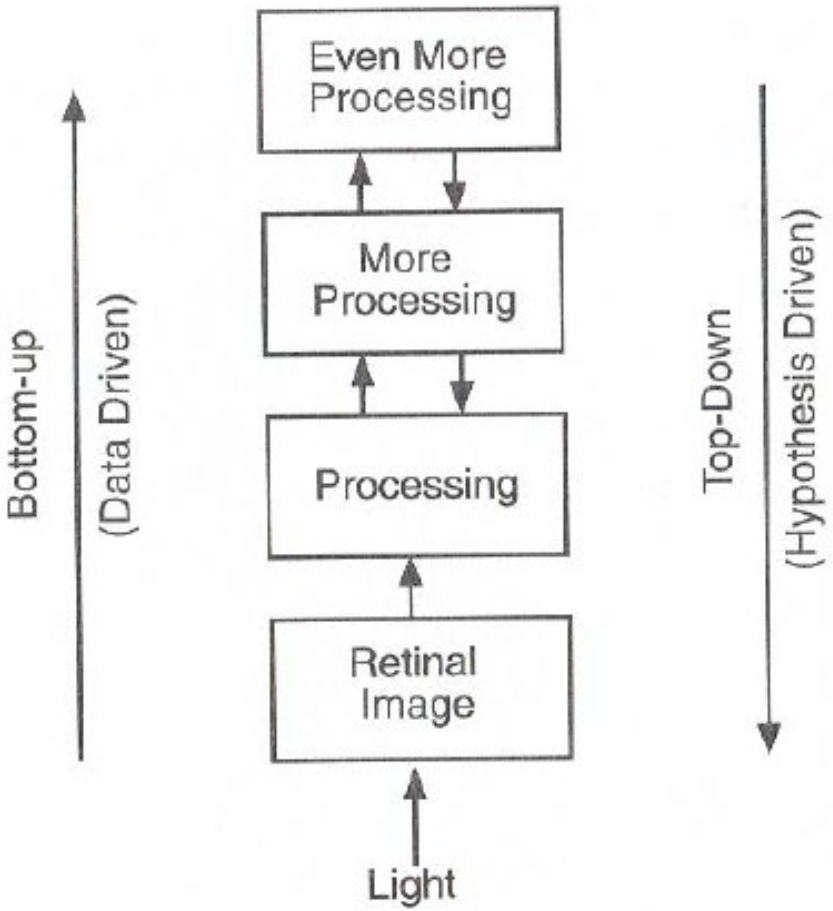


Figure 2.3.11 Bottom-up versus top-down processing. The two directions of processing are referred to as *bottom-up* (or *data driven*) from lower to higher levels of processing and *top-down* (or *hypothesis driven*) from higher to lower levels of processing.

Bottom-up or Top-down?