Vision in Man and Machine

Part 1
Challenges of Biological and Machine Vision

Heiko Wersing

Honda Research Institute Europe GmbH

Content of this lecture

1. Challenges of biological and machine vision
2. The human visual pathways
3. *Tutorial 1: Feature representations in early visual cortex*
4. Principles of sensory coding
5. Principles of biological feature learning and adaptation
6. *Tutorial 2: Experiments on sparse coding for feature learning*
7. Object recognition in biological brains
8. Object recognition models
9. *Tutorial 3: Building visual feature hierarchies*
10. Gestalt perception
11. Models of perceptual grouping
12. *Tutorial 4: Perceptual grouping with the CLM*
13. Cognitive vision, learning, and humanoid robots
14. Summary and Conclusions
15. *Tutorial 5: Recognition experiments with feature hierarchies*

Vision is useful

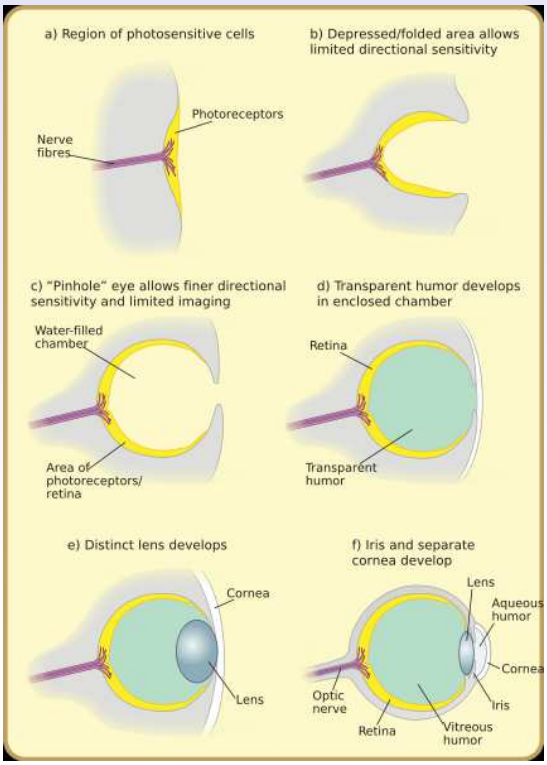
Vision is useful

What is visual perception ?

David Marr ,1982:
"The plain man's answer (and Aristotle's too) would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is".

- Visual perception concerns the acquisition of knowledge
  - Cognitive activity rather than physical measurement
- Acquiring knowledge about objects and events in the environment
  - Reference to an external reality
- Visual knowledge is obtained by extracting information
  - Formal description of visual processes
- Information comes from the light emitted or reflected by objects
  - Optical information is the basis of visual sensory input

The biological evolution of the eye

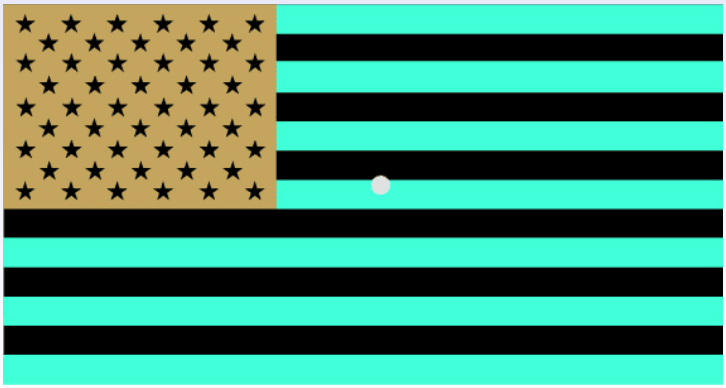


- Directed light perception turned out to be of great advantage for living organisms
- The eye is a part of the brain
- We see with the brain
- Estimate: More than half of the human brain is concerned with vision
- We can currently build machine eyes, but not machine brains

Perception as active construction

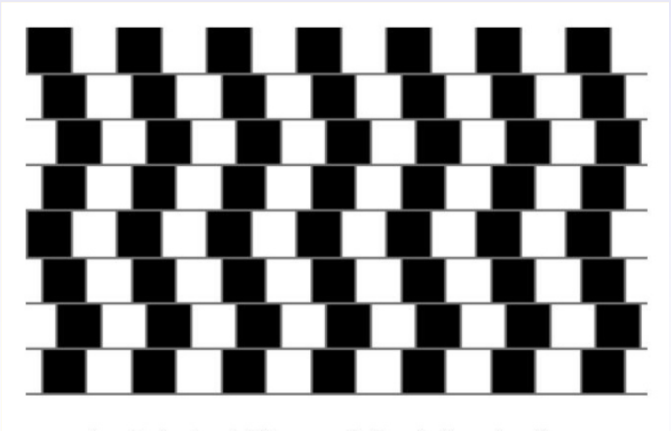
Perception is rather an active process of interpretation than a passive recording of visual information.

Aftereffects – An american illusion

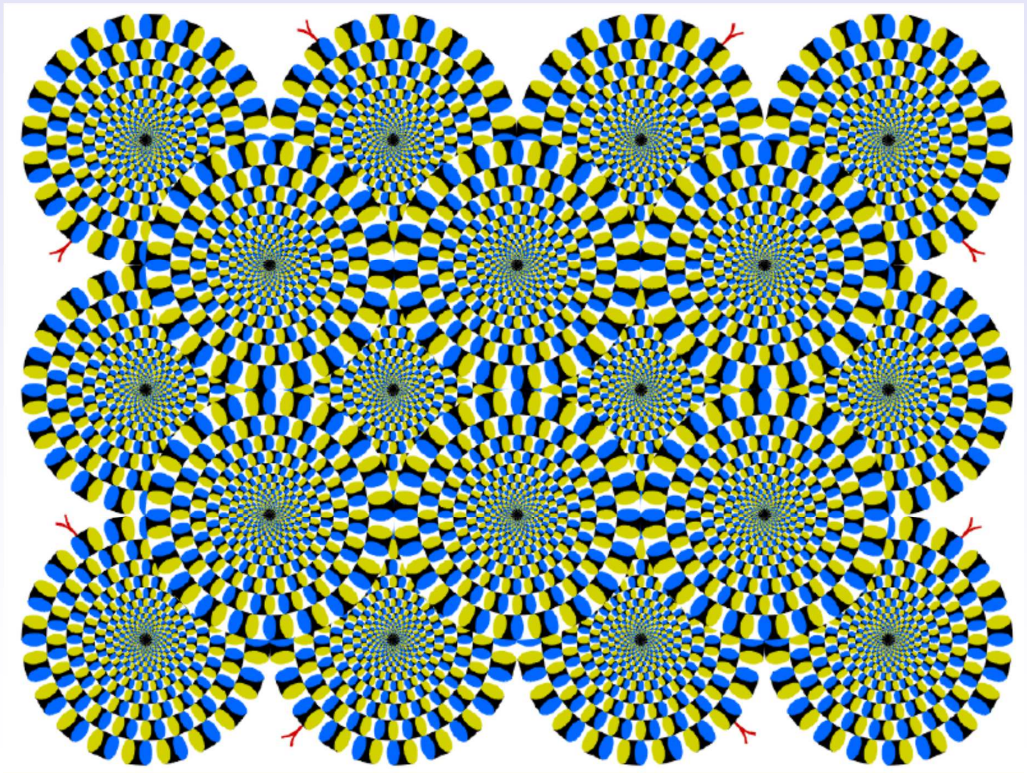


Aftereffects – An american illusion

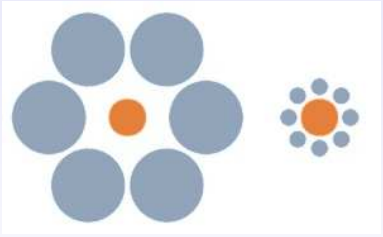
Structural illusions: The cafe wall illusion



Virtual Motion



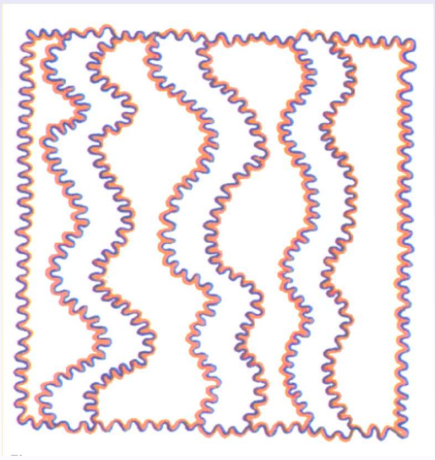
Size illusions

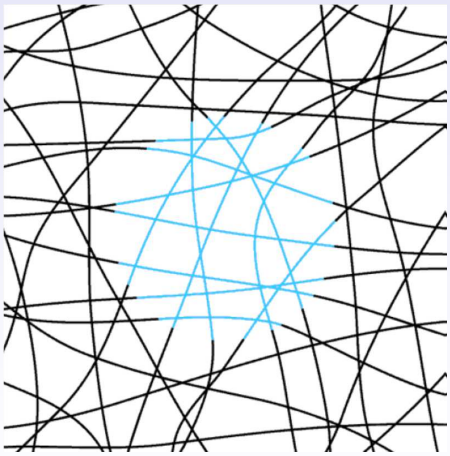


Ebbinghaus illusion

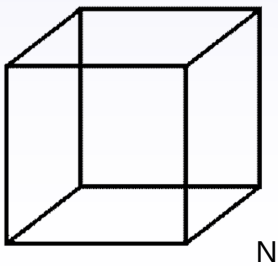
Moon illusion

Color filling in



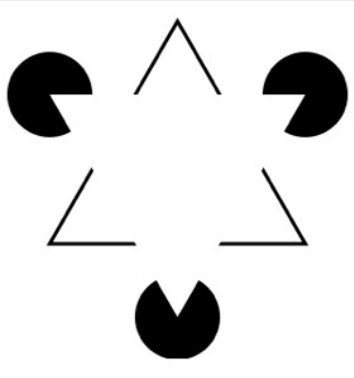


Ambiguous figures

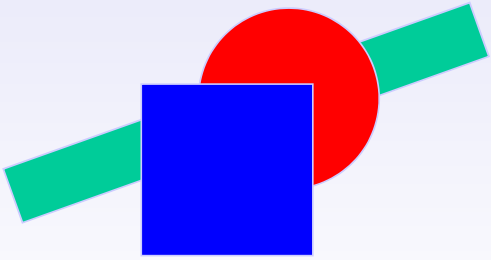


Necker cube

Visual completion



Kanizsa triangle



Perception as active construction

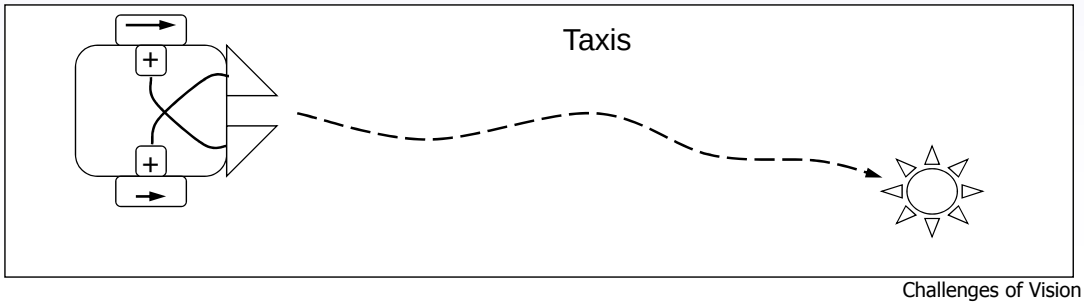
Perception is rather an active process of interpretation than a passive recording of visual information.

- Adaptation
  - Dark adaptation
  - Afterimages
  - Orientation aftereffects
- Optical Illusions
  - Structural illusions
  - Object sizes (e.g. moon illusion)
  - Apparent motion
- Ambiguous figures
  - Necker cube
  - Figure-ground separation
  - Multiple interpretations of the same image
- Visual completion
  - Occlusion completion
  - Transparency perception

Creating visually-driven behavior

- Braitenberg vehicles
  - Coupling of directional vision sensors to motors
  - Light-searching behavior: Crossing of links
  - Light avoiding behavior: Parallel links
  - Reactive system

- Is this visual perception ?



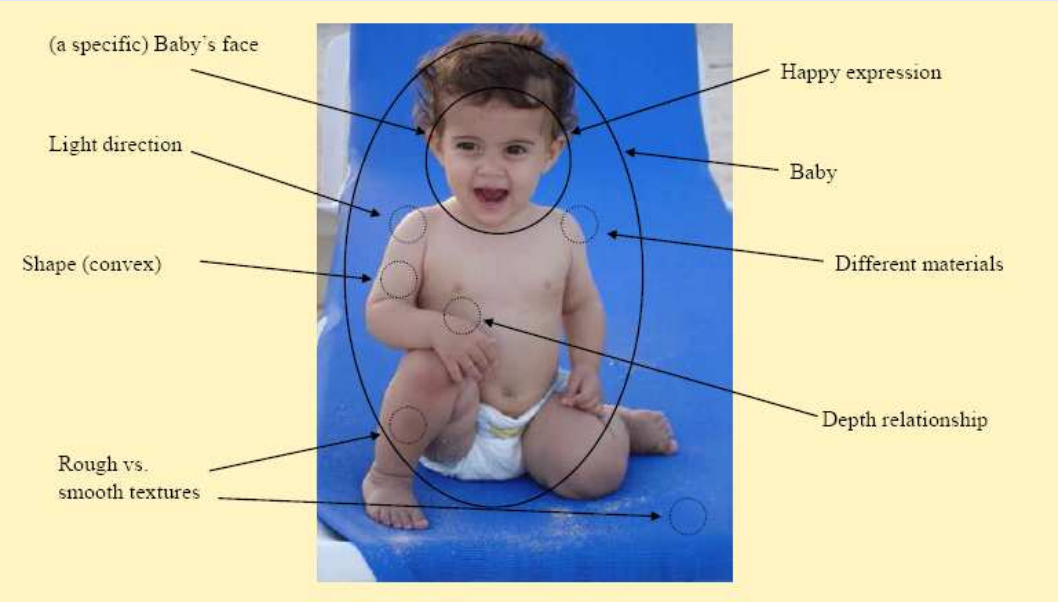
Vision seems to be easy

- All people can "see" equally well (but only few can solve hard mathematical problems, play good soccer, or play good chess)
- Babies can "see"
- Really primitive animals can "see"
- We "see" effortlessly (at least it feels this way)
- Vision is immediate
- Vision appears to be flawless

→Vision for computers has turned out be a tough problem

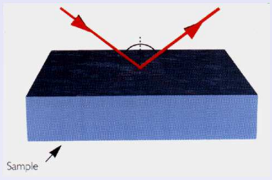
→ Why ?

Interpreting an image

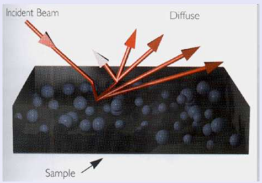


Characteristics of optical information

- Illumination
  - Point sources
  - Diffuse illumination
  - Interaction with surfaces
    - Absorption
    - Reflection
      - Specular
      - Matte
    - Transmission

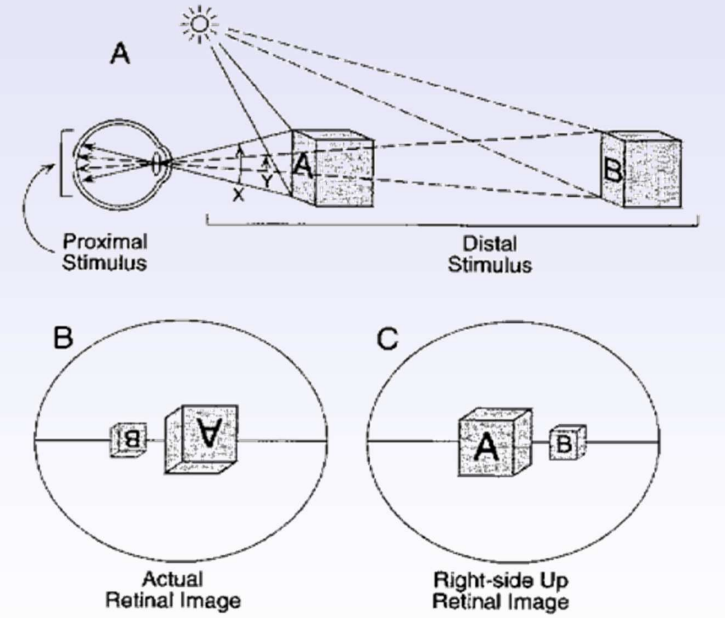


specular

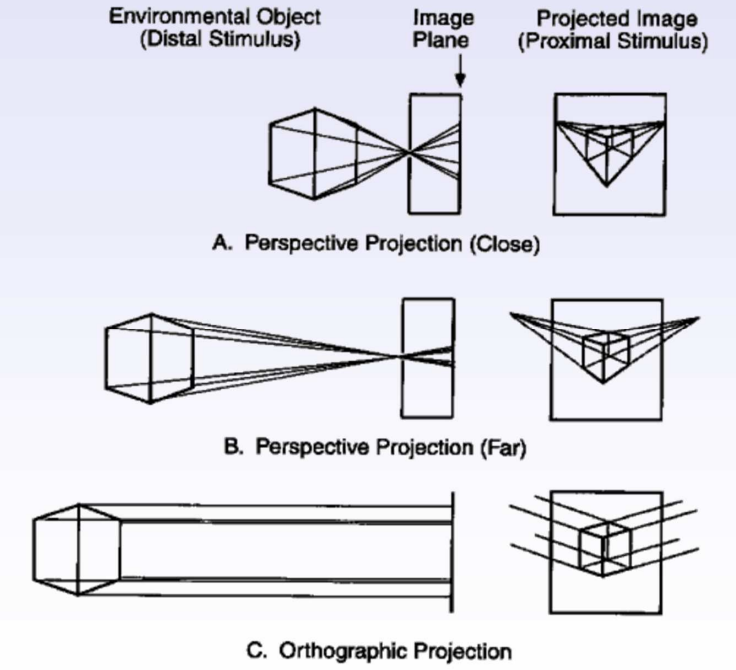


matte

Proximal and distal stimulus

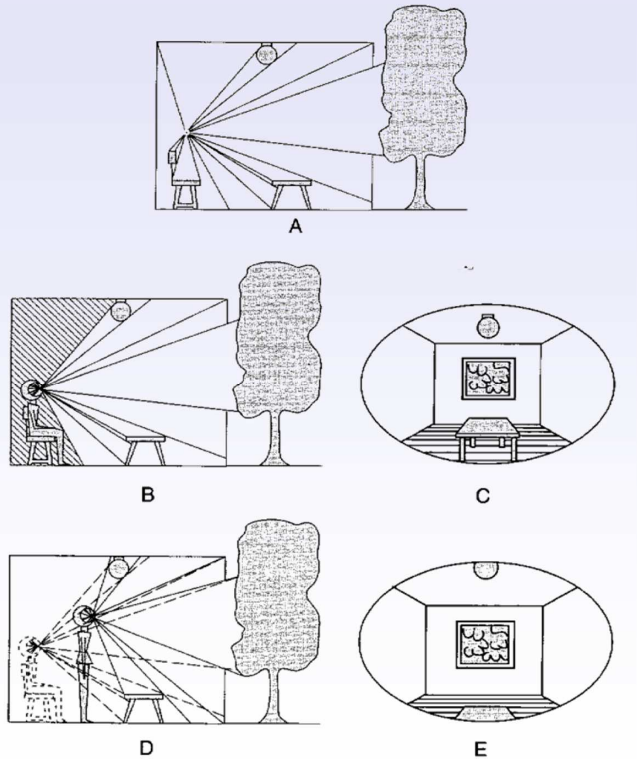


Optical projections 3D→2D



Characteristics of optical information

- Ambient optic array (Gibson)



Computer vision

David Marr ,1982:
"The plain man's answer (and Aristotle's too) would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is".

- Artificial systems that obtain information from images
  - Object recognition and tracking
  - Scene reconstruction
  - Motion processing and optical flow
  - Pattern matching
  - Localization and mapping

Current application fields of computer vision

Applications:

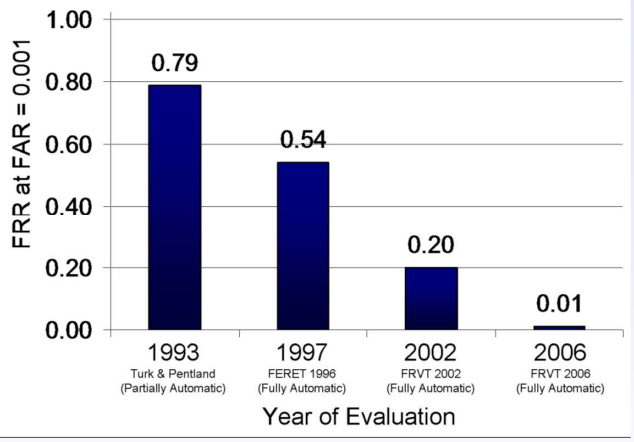
- Automobile driver assistance
  - Lane keeping
  - Adaptive cruise control
  - Driver monitoring for sleepiness
- Digital photography
  - Face detection
- Film and Video: Sports Analysis
  - Tracking of balls and players
- Movie production
  - Markerless Tracking of gestures and mimics
- Games
  - Gesture tracking , e.g. Sony EyeToy
- Industrial automation and inspection
  - Checking parts, surfaces, materials
  - Counting/checking goods, food, fruit, vegetables
  - Textile inspection

Current application fields computer vision

Applications:

- Biometrics
  - Fingerprint, Iris, Face
  - Identification / Security
- Surveillance
  - People tracking
- Image search
  - Image Database search
- Biomedical image processing
  - Quantitative analysis, cell counting, tissue analysis
- Traffic management
  - Car license plate reading

Face Recognition



- Face recognition has achieved dramatic improvement over the last 15 years
- FRVT 2006: Computer vision algorithms exceed human performance for an 80 person matching task
- But works only in cooperative situations

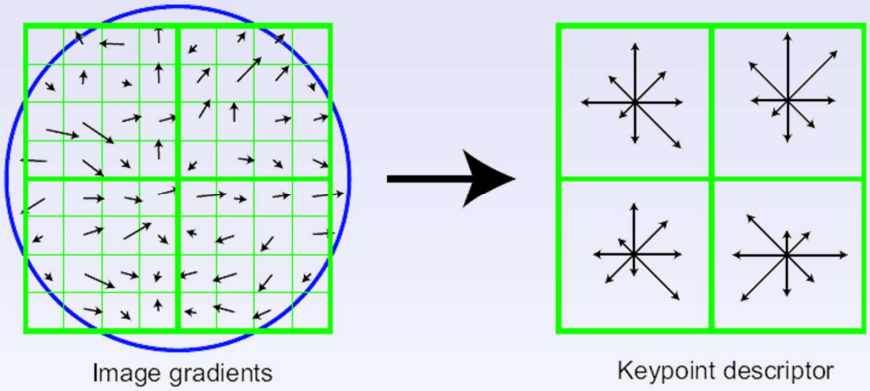
Object recognition

- Established approach: SIFT-based recognition
- Compute focus points
- Compute Scale-invariant feature transform at focus points
- Match features

- Newer variant: Histogram of Oriented Gradients (HoG)

- Smartphone Apps: e.g. LabelMe (HoG), KlickTel (SIFT)

Orientation and descriptor computation



- Orientation
  - Maximum of gradient orientation histogram (36 bins)
- Descriptor
  - Set of gradient orientation histograms (8 bins each)
  - E.g.: 2x2 histograms on 8x8 samples
  - Real: 4x4 histograms on 16x16 samples

Example images with marked keypoints