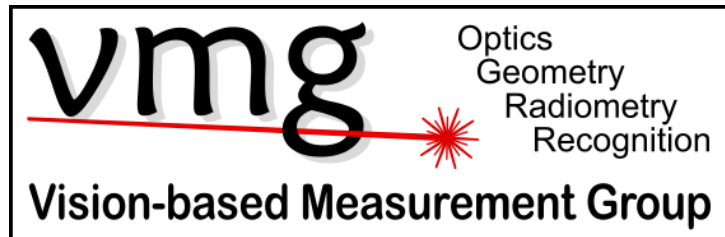


VISION:

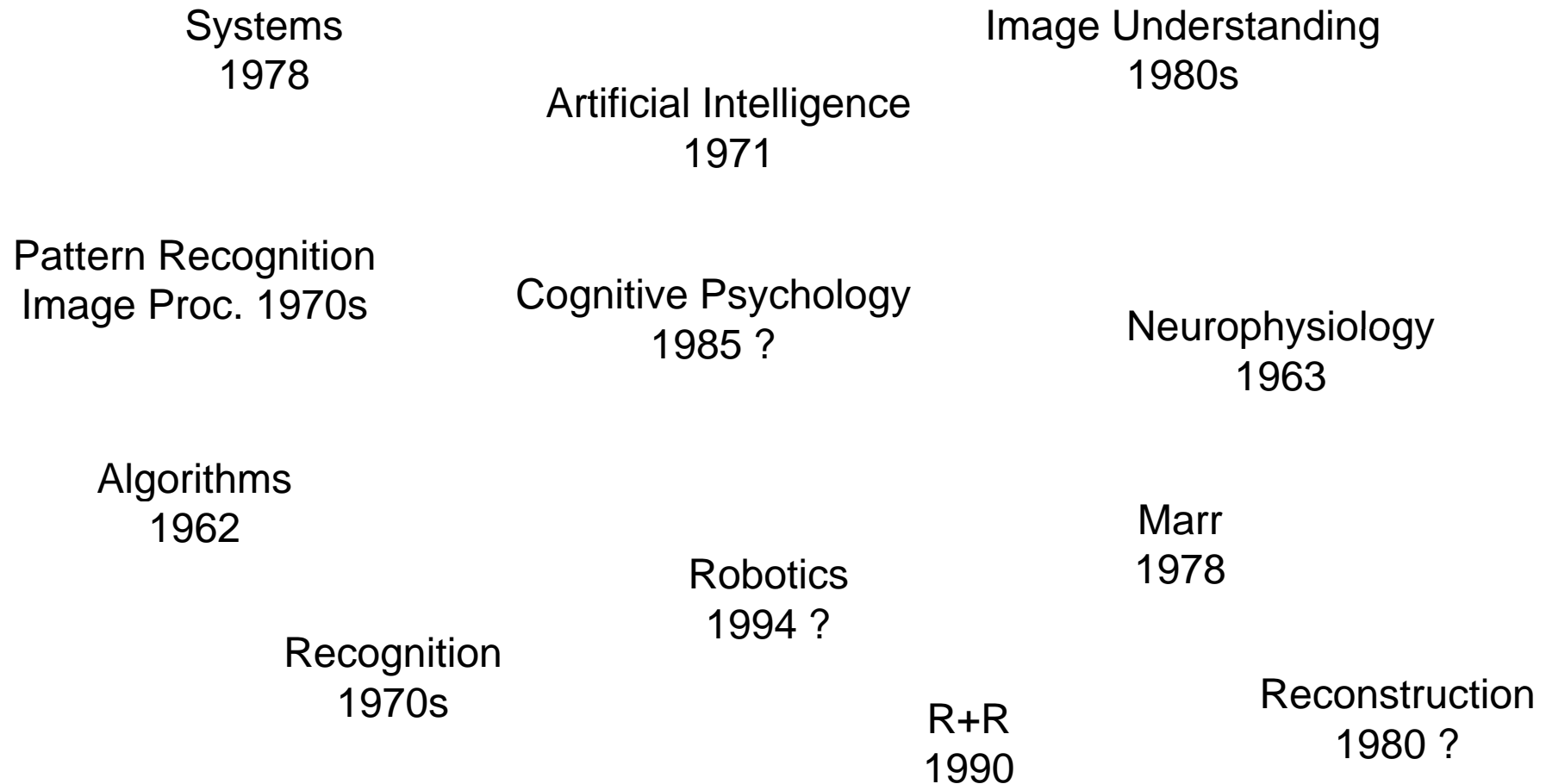
Axel Pinz



- Paradigms
- Systems
- Algorithms
- Applications
- Evaluation

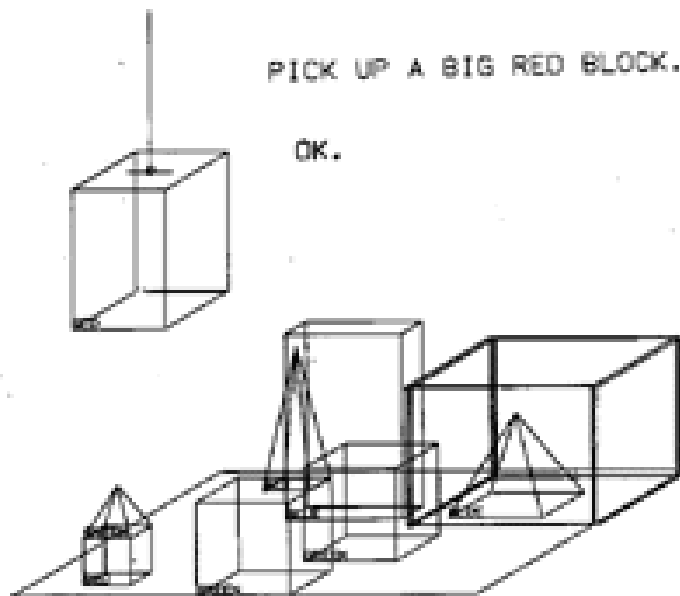
→ A Graph-based structure !

The Graph of Vision History

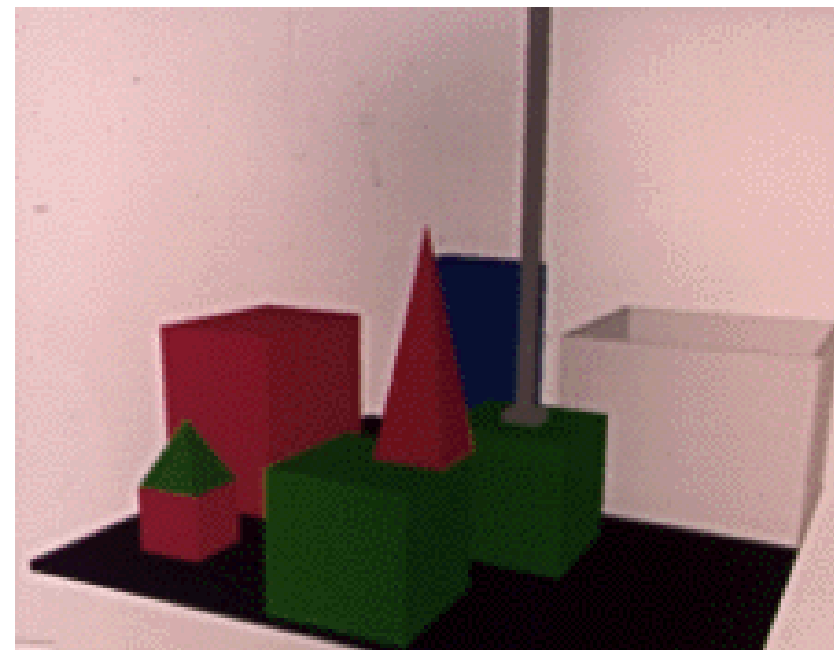


Artificial Intelligence

- SHRDLU [Winograd 1971]
- Blocks world



original screen display (MIT AI Lab)



later color rendering (Univ. Utah)

Blocks World – 1970s

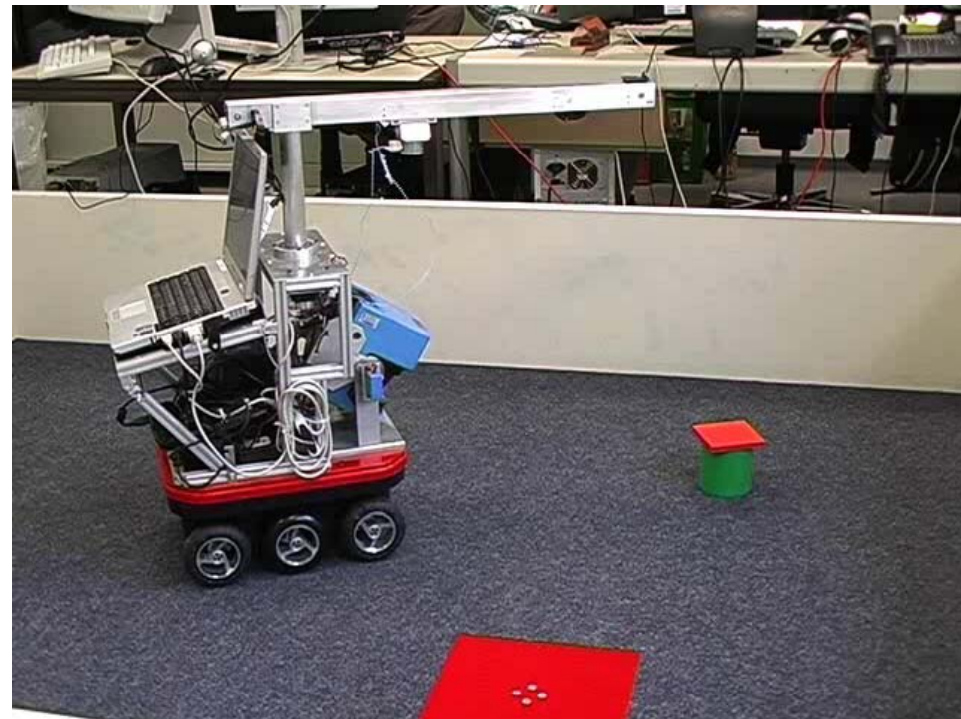
- Natural language understanding
 - Spatio-temporal reasoning
 - LISP
 - Visual input would be nice
-
- Vision as an “enabling sensory module” in AI
 - Patrick Henry Winston “The Psychology of Computer Vision” 1975
 - Marvin Minsky “A framework for representing knowledge” 1975
 - Frames
 - Frame Representation Language FRL

Blocks World – Today

- Still an issue of ongoing research...
- EU FP 6 project MACS
- 2004-2008



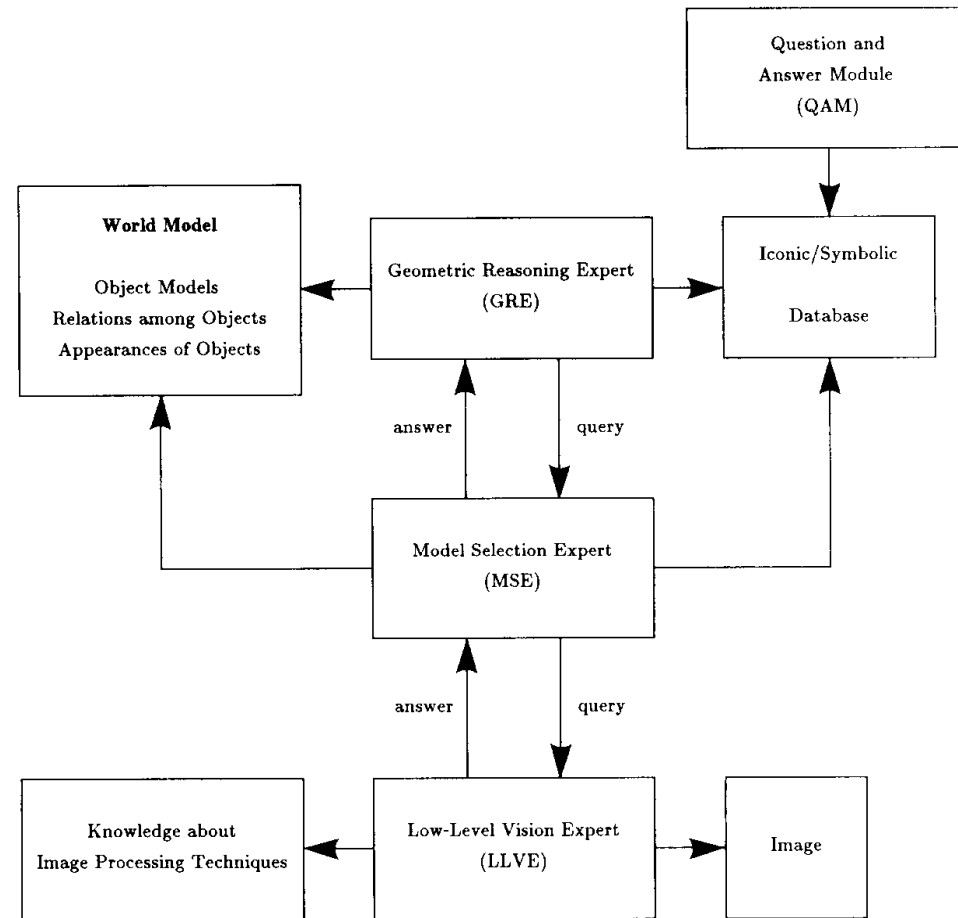
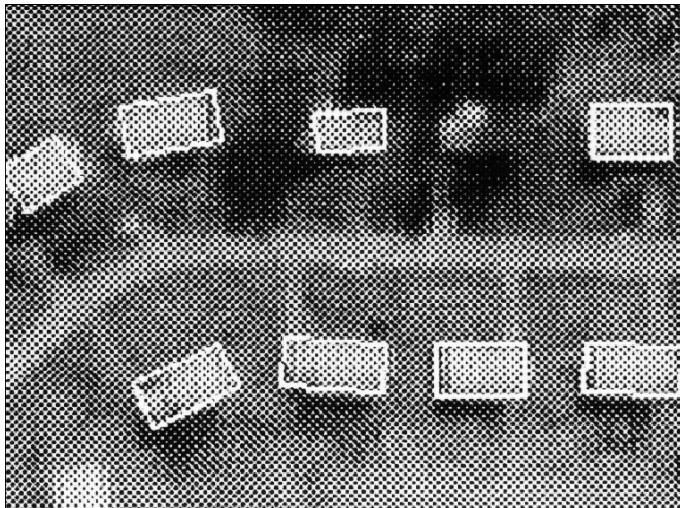
→ Learning !
→ “Affordances” ...



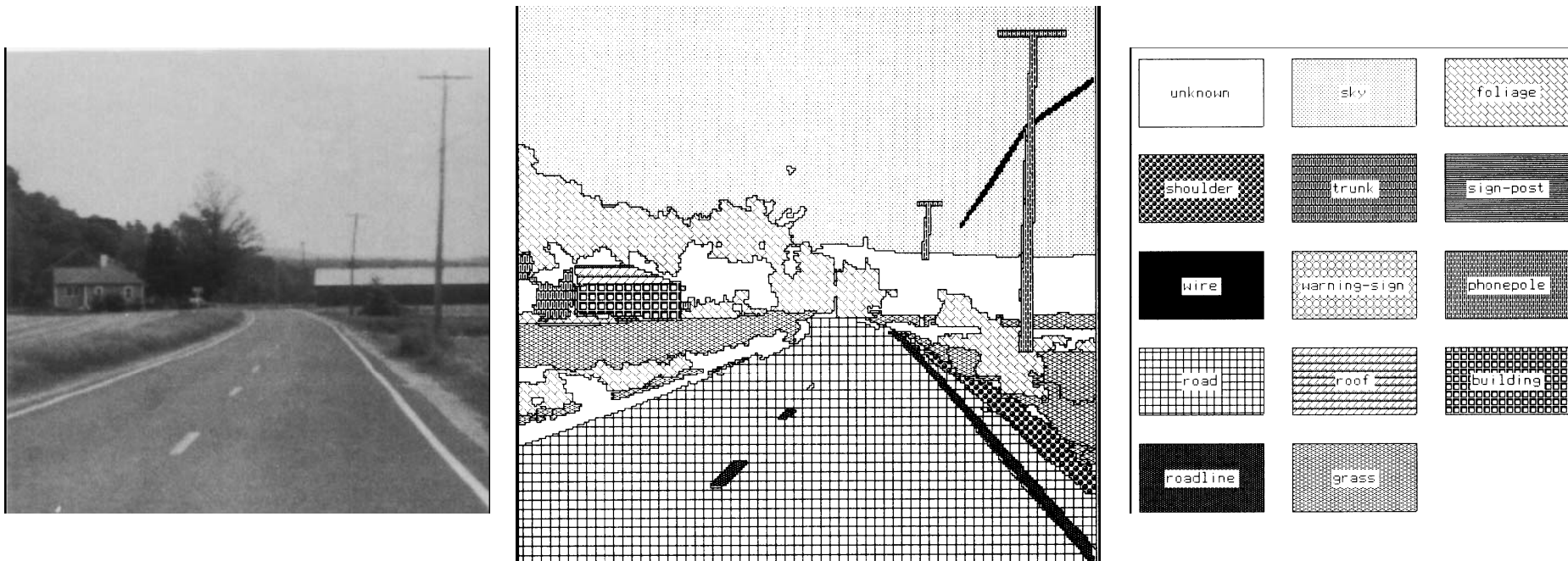
AI Methods for Vision – 1980s

- “Image Understanding”
- Expert systems
 - SIGMA Aerial Image Understanding
 - Matsuyama, Hwang, ...
- Reasoning, Blackboard system
 - VISIONS, Draper, OHM, ...
- Schema learning
 - SLS, Draper, ...

SIGMA [Matsuyama]



VISIONS [Hanson+Riseman], SLS [Draper]



A Massachusetts "road"-scene

“Image Understanding” – 1980s

- Integrating bottom-up and top-down processes → combinatorial explosion
- “knowledge engineers”, hand-crafted knowledge-base → learning is a must

Systems 1990 – 200+

- “Computer Vision Systems”
 - [Hanson+Riseman 1978]
- UMass VISIONS → KBVision
 - AAI Amerinex Artificial intelligence
 - AAI Amerinex Applied Imaging
- Grimson → Cognex
 - PatMax
- IUE
 - TargetJr Joe Mundy
- Matlab
- OpenCV
- toolboxes, libraries

Robotics

- Brooks: “building brains for bodies”
 - Autonomous Robots 1:7-25, 1994
- Learning!
- Interaction!
- Reward by survival....!

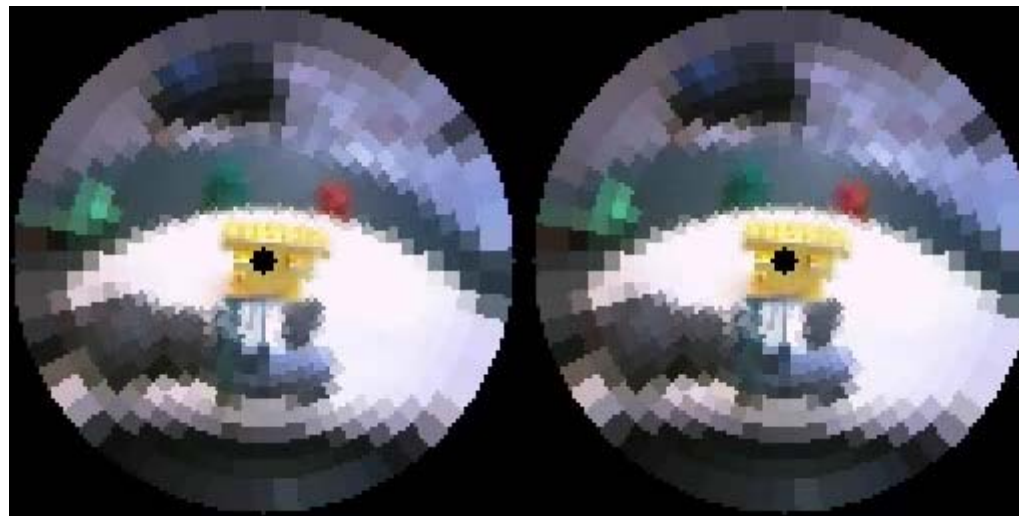
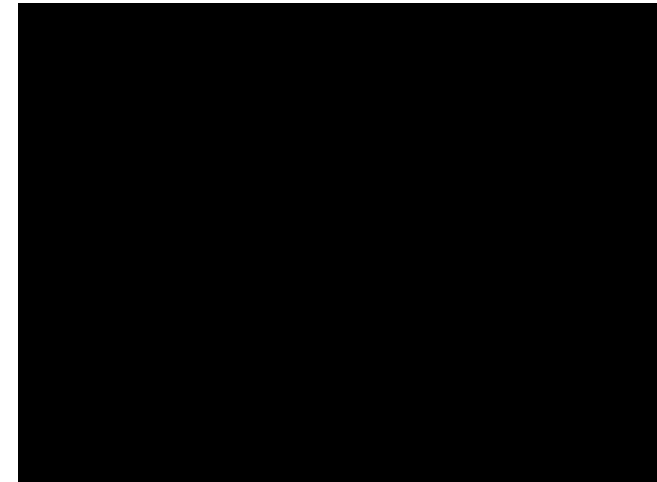
- Active, purposive, qualitative
 - Bajcsy 1988
 - Aloimonos 1989

Babybot – Lira Lab [Sandini]



learning to push

learnt

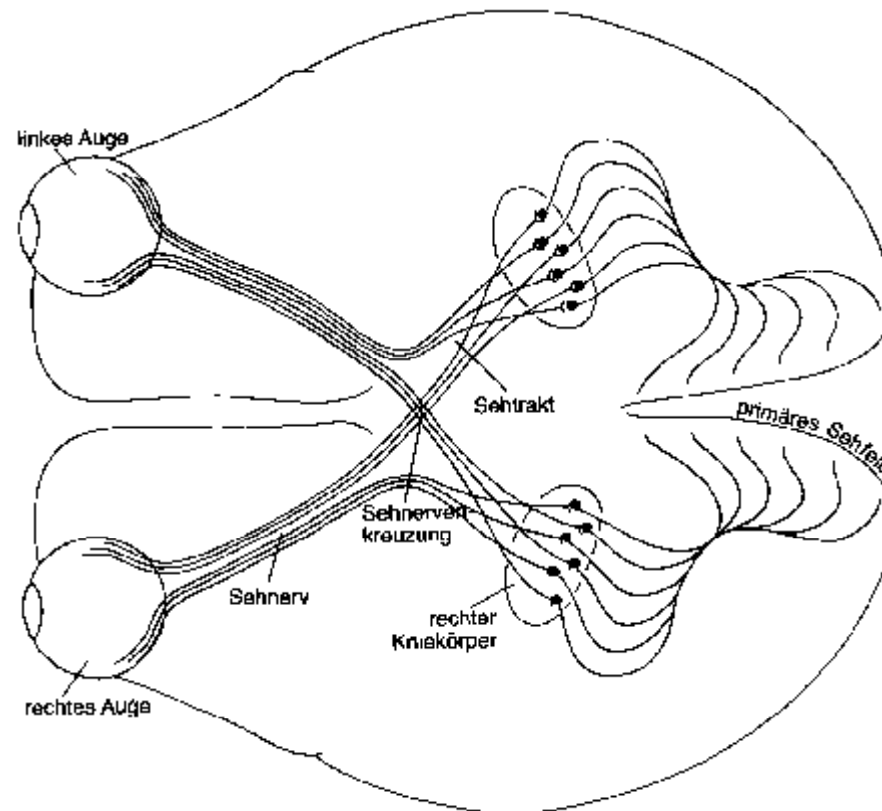


Babybot's view

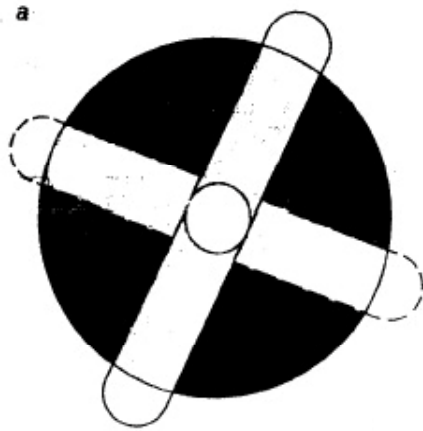
Neurophysiology

- Retina → visual cortex
- Layered representation in visual cortex
- Receptive fields
- Hubel+Wiesel
 - Gradients
 - Oriented “edgels”
 - At various scales
 - Bottom-up grouping → recognition
 - Visual pathway → stereo, reconstruction

Visual Pathway [Hubel+Wiesel]



Receptive Fields [Hubel+Wiesel]



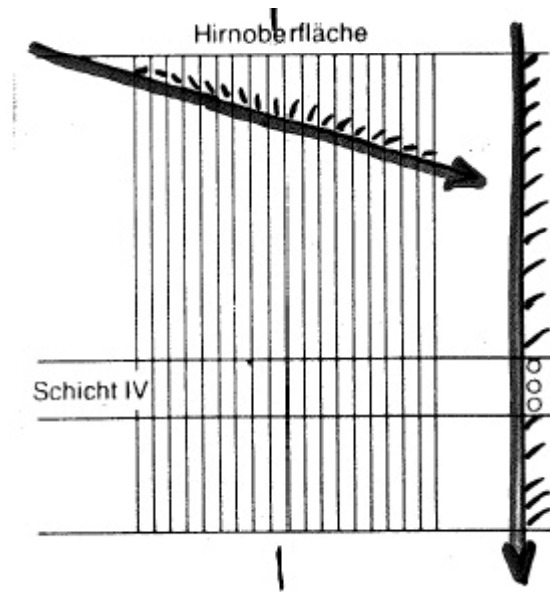
on-off



simple



complex

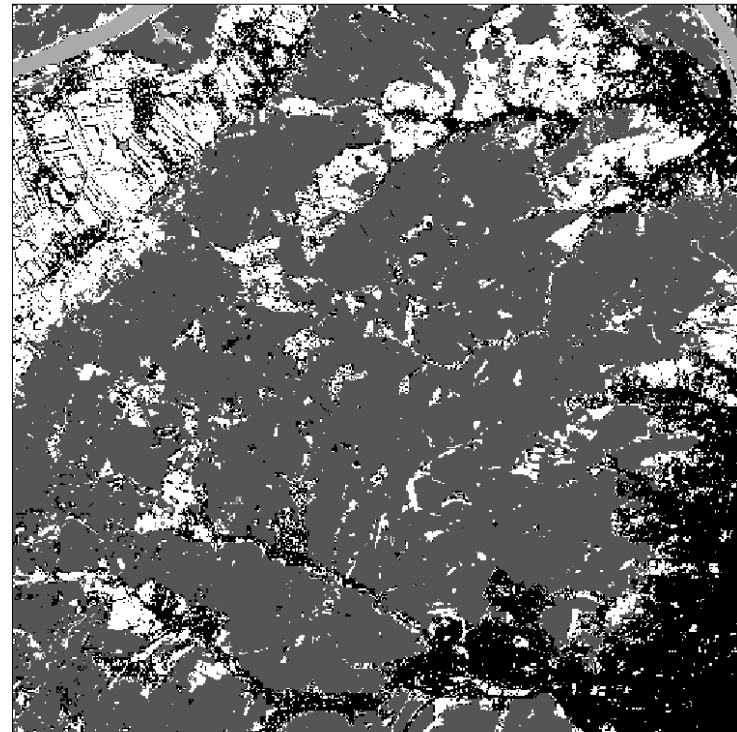
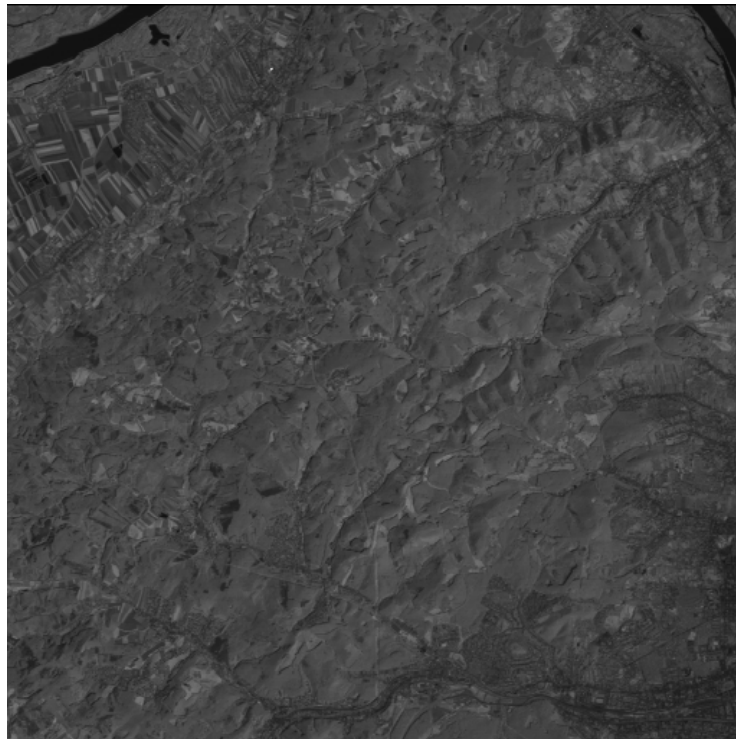


Pattern Recognition + Image Processing – PRIP

- Structural [Pavlidis 1972, Fu 1982]
- Statistical [Fukunaga 1990]

- Representation (regular, irregular, scale,...)
pyramids, scale space, graphs,...
- Feature sets, feature selection, classifiers,
learning,...
- 2D (discrete) signal processing

Maximum Likelihood Classification



Scale Space [Lindeberg]

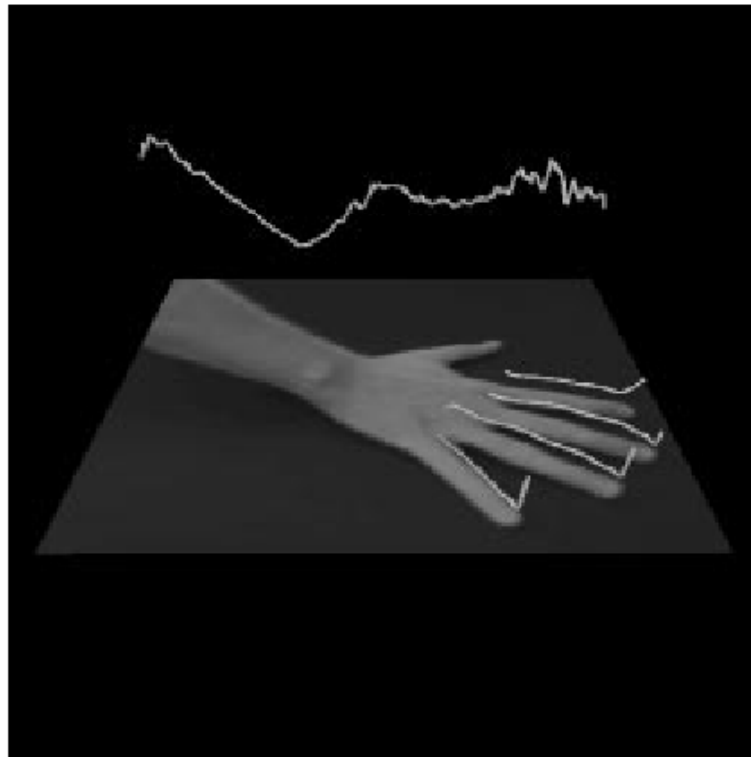
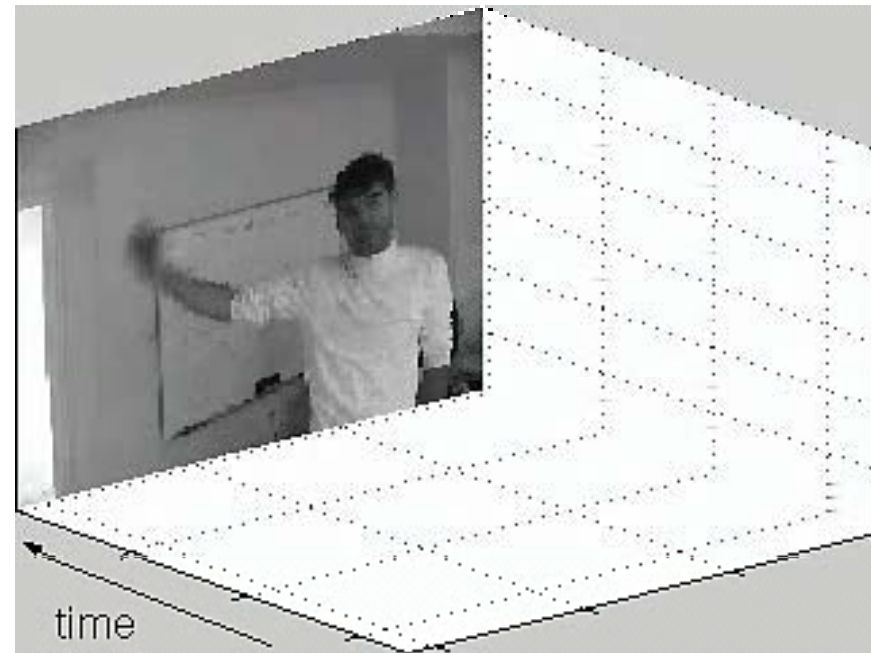


Figure 10: Three-dimensional view of the five strongest ridge curves in scale-space.



Space + time
[Laptev + Lindeberg]

Cognitive Psychology

- Perceptual grouping
 - Lowe 1985
 - Sarkar 1994
 - LLVE Matsuyama
- Qualitative volumetric models (Geons)
 - Biederman 1985, Bergevin+Levine 1988
 - Dickinson 1992: Aspect Hierarchy
- A renaissance in today's shape models for category detection...

My Favorite Example [Lowe]

Very hard !!!

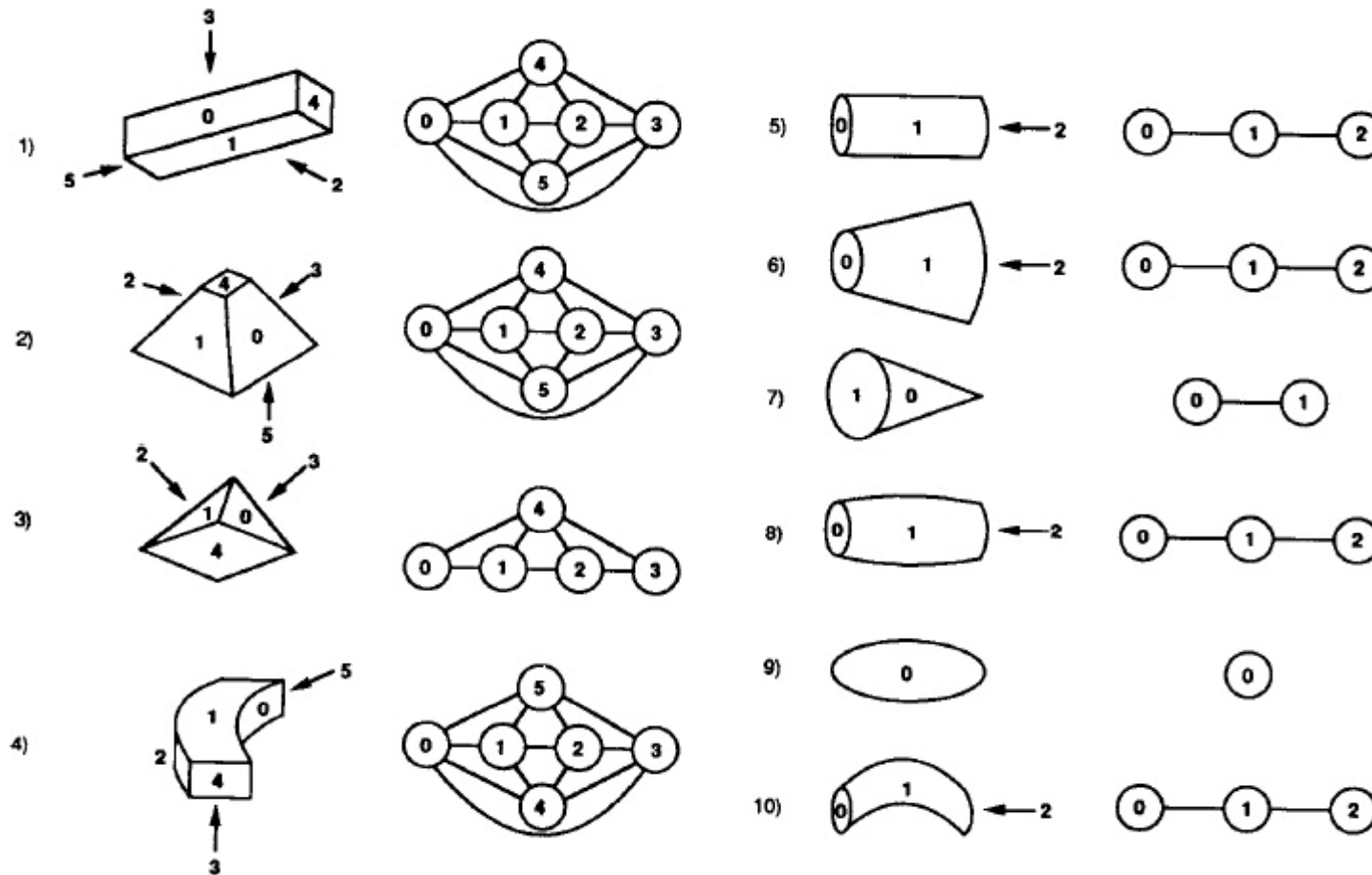


My Favorite Example [Lowe]

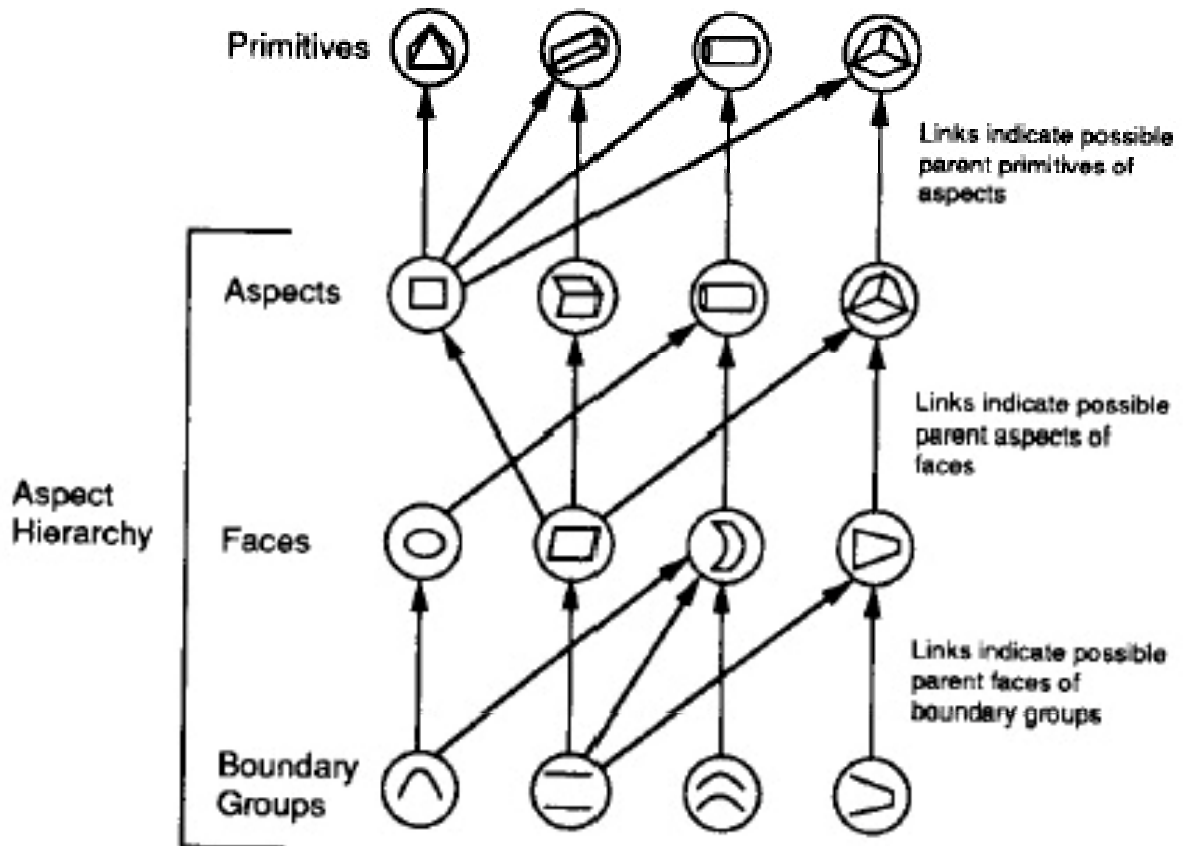
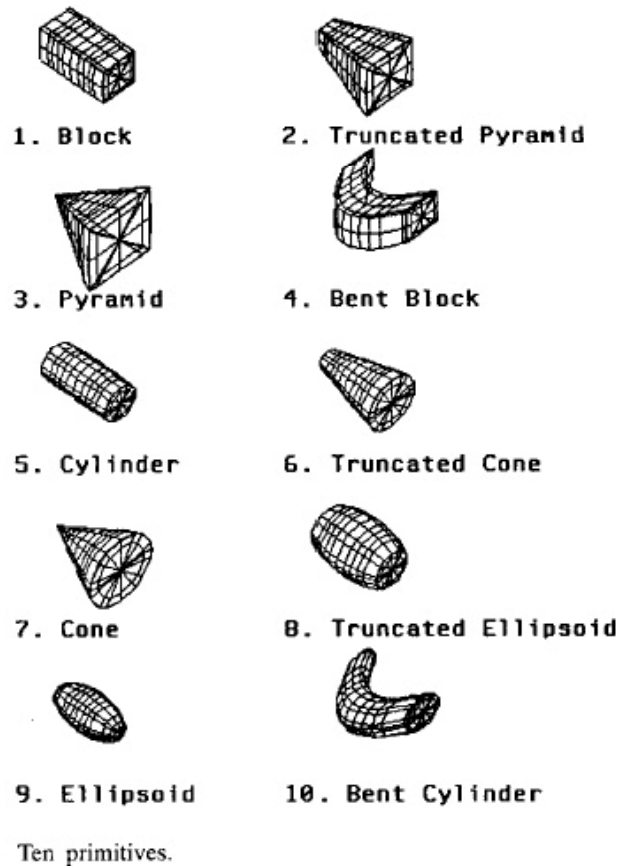
Very easy !!!



Geons [Biederman], Aspect Hierarchy [Dickinson]



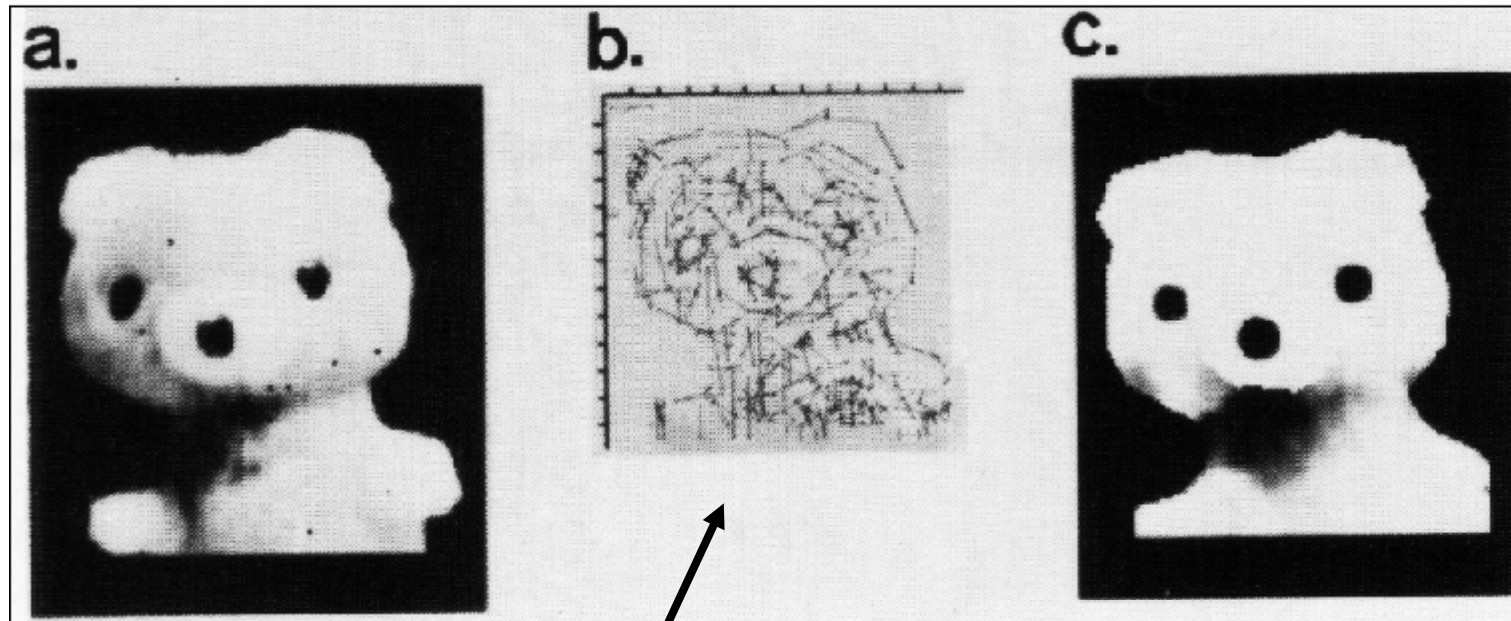
Geons [Biederman], Aspect Hierarchy [Dickinson]



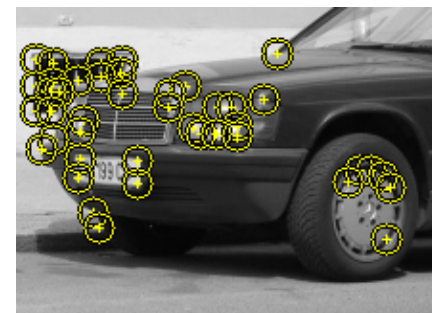
The Marr Paradigm – 1978-?

- Systems approach to vision → computational model
 - Algorithm, data, hardware implementation
 - Low level modules
- Representational levels
 - Primal sketch → saliency
 - 2-1/2 D → reconstruction !
 - 3D object model → is it really required?
- “reconstruction school”
- “visual modules”, “shape from X”
- Zerroug 1994
- Structure + Motion 200x

Marr (~1978) – “primal sketch”

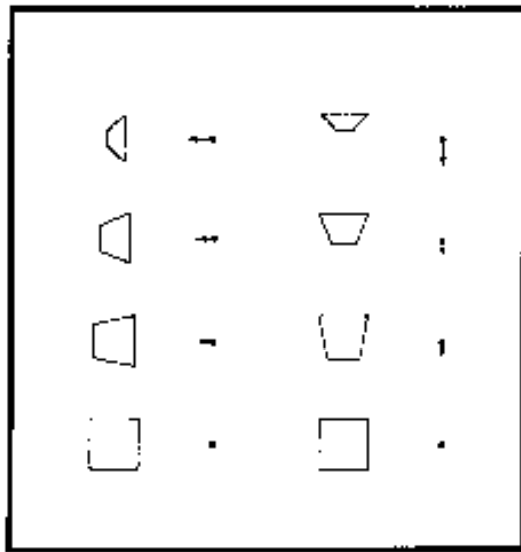


“saliency” !

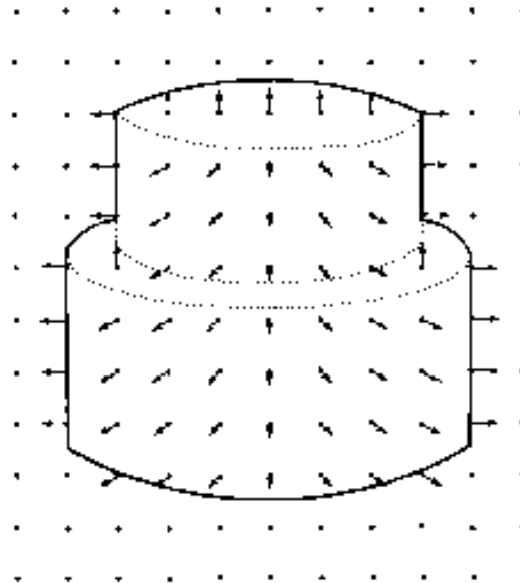


“Harris” corners

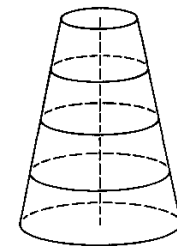
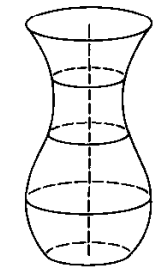
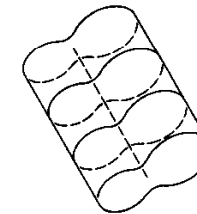
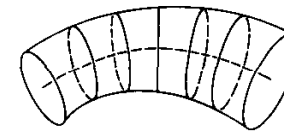
Marr (~1978) – “2-1/2-D sketch”, “generalized cone”



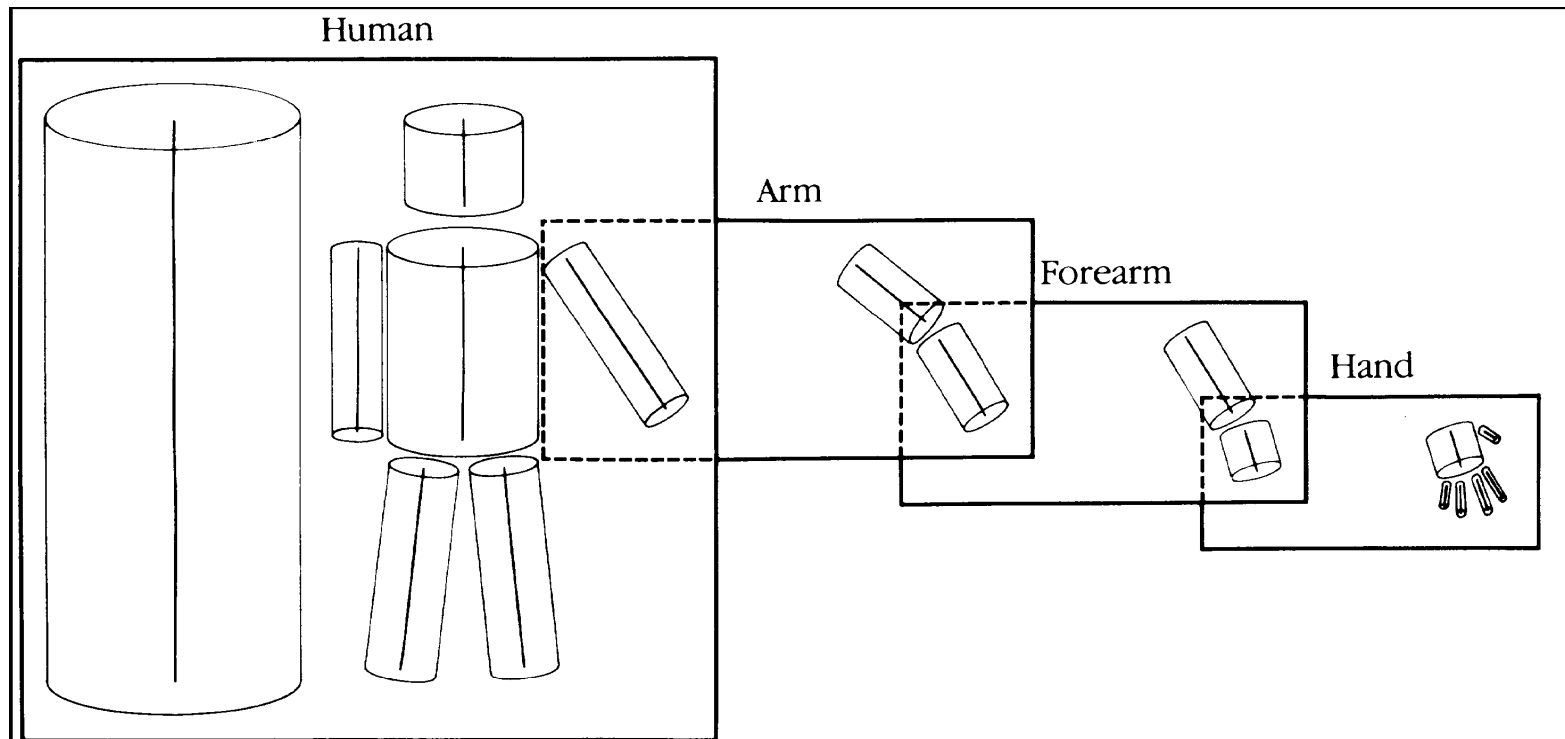
(a)



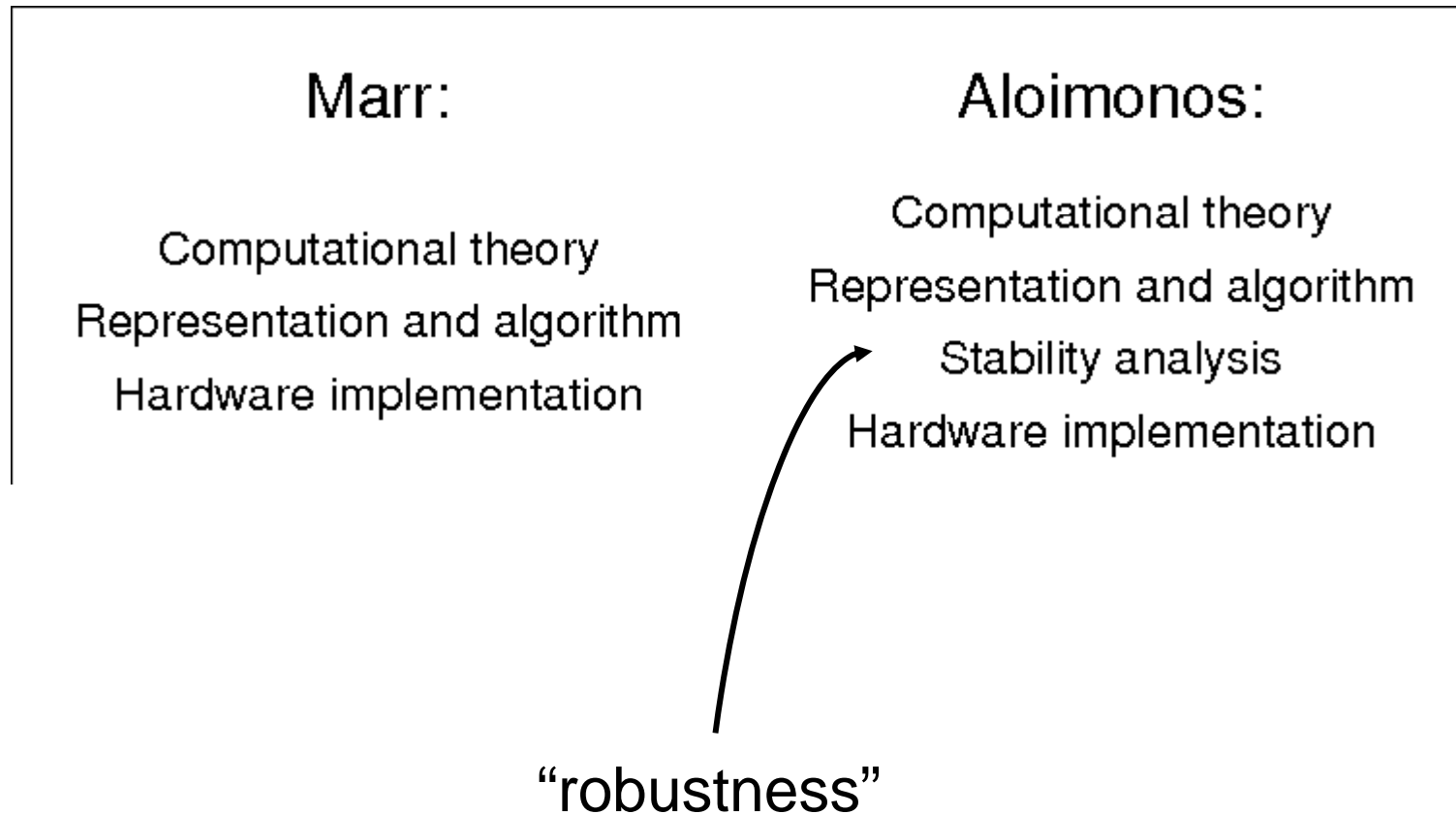
(b)



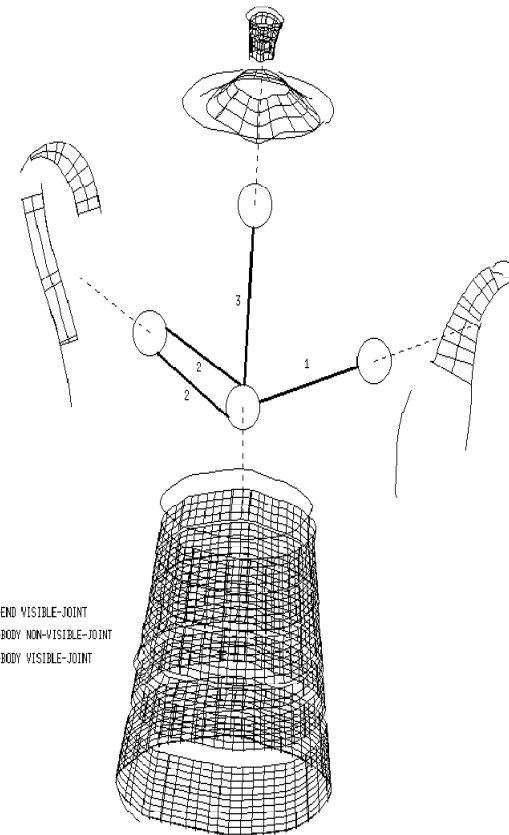
Marr (~1978) – 3D hierarchical models



Marr ~1978, Aloimonos ~1988



Zerroug 1994



3 : END-TO-END VISIBLE-JOINT
2 : END-TO-BODY NON-VISIBLE-JOINT
1 : END-TO-BODY VISIBLE-JOINT

Algorithms

- Convolution → kernels
- Mathematical morphology → structuring elements Serra
- Fourier transform → spectrum, phase, descriptors, Gabor filter bank
- Hough transform → voting space
- PCA
- Pyramids (regular, irregular) → graphs
- Scale space theory in CV Lindeberg
- Algebraic projective geometry Hartley, Zisserman
- Color (good chapters in Forsyth+Ponce, Burger+Burge)
- Texture → invariant moments (Hu ... Haralick ... VanGool)
- Interest point detection (Schmid, ...) → affine covariance
- Pattern Classification Duda+Hart
- Machine Learning Hastie et al.

Reconstruction

- Projective geometry
- Photogrammetry
- Essential, fundamental, multiview, ...
- Calibrated, uncalibrated, autocalibration, calibration-free, ...
- Structure and Motion

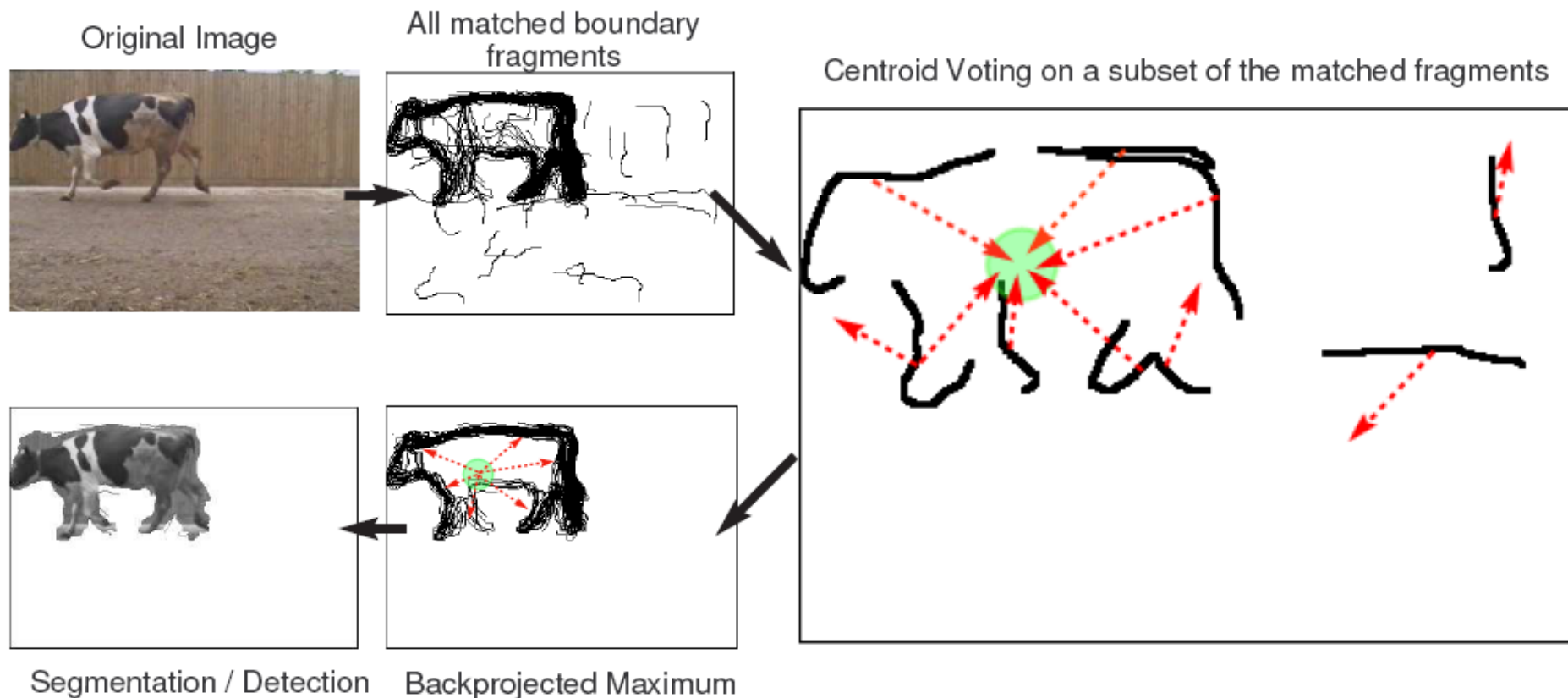
Structure and motion [Schweighofer 2008]



Recognition

- Specific objects
 - PCA [Murase+Nayar 95], Eigenfaces [Pentland 93]
- Categories
 - Generative [Fergus 2003], discriminative [Opelt 2004]
- Recognition vs. localization
- Features (descriptors), grouping
 - Shape, texture, color, proximity, similarity,...
- Saliency (detectors)
- Active
 - Motion, space+time ...

Shape-Based Category Detection [Opelt 2006]



“R + R”: Confluence of Recognition and Reconstruction

- “recognition and reconstruction schools will merge”
[Aloimonos+Shulman 1989]
- “Marr paradigm – use of 2-1/2D complicates the problem”
[Medioni et al. 2000] → tensor voting
- Spatio-temporal reasoning and representation



The Graph of Vision History

