Designing Artificial Minds

Harri Valpola

Computational neuroscience group

Laboratory of computational engineering

Helsinki University of Technology

http://www.lce.hut.fi/~harri/

“The great end of life is not knowledge but action”

—Thomas Henry Huxley (1825-1895)

Sea Squirts — Our Distant Cousins

- Sea-squirts are common marine animals
- Two stages of development: larva and adult
- Larvae look very much like tadpoles

Picture:
http://www.jgi.doe.gov/News/ciona_4panel.jpg

No Movement → No Brain

Picture:
http://www.gulfspecimen.org/images/LeatherySeaSquirt.jpg

Smooth, Elegant, Skillful

Picture:
http://www.africanbushsafaris.com/fotos%20touren/Oryx.jpg

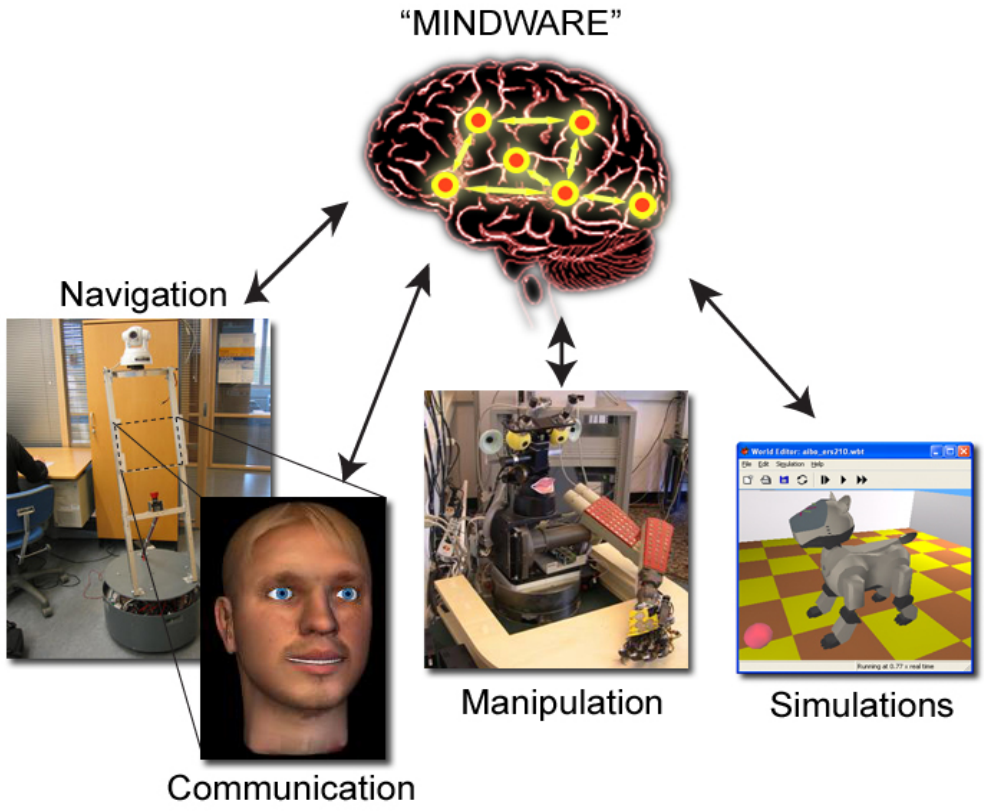
Picture:
http://www.ebroadcast.com.au/blahdocs/uploads/tiger_running_sml_2784.jpg

Evolution of the Motor System

- Spinal cord: “simple” reflexes and rhythmic movement
- Brain stem: more complex reflexes
- Cerebellum / midbrain structures: complex motor coordination
- Basal ganglia: action selection
- Hippocampal formation: navigation
- Neocortex: integration and planning

Methods

- Synthetic approach: learning by building
- Neural network simulations
- Real and simulated robots

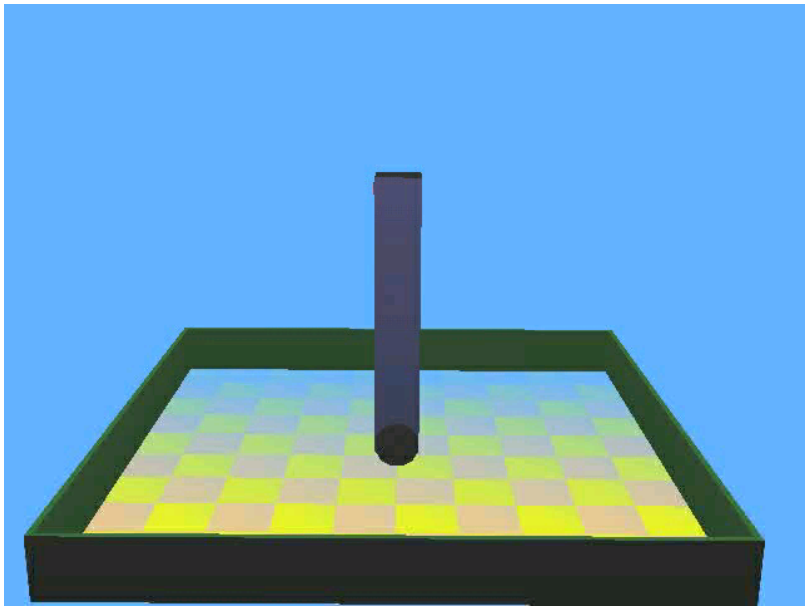


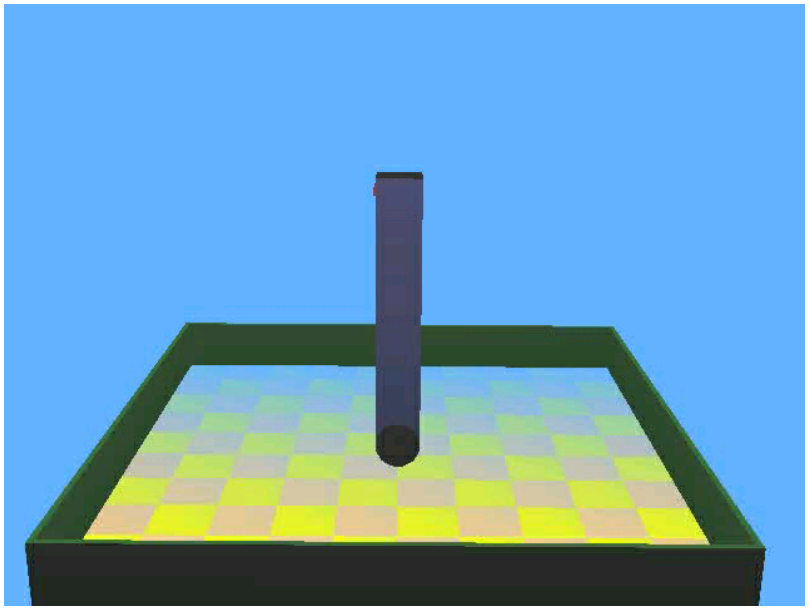
Early Motor System

- Spinal cord: “simple” reflexes and rhythmic movement
- Brain stem: more complex reflexes
- Cerebellum / midbrain structures: complex motor coordination
- Basal ganglia: action selection
- Hippocampal formation: navigation
- Neocortex: integration and planning

Prediction and Anticipation

Adaptive Motor Control Based on a Cerebellar Model





Prediction and Anticipation

Self-Supervised Learning in Control

- Corrections are made by a large number of “reflexes” (spinal cord, brain stem, cortex / basal ganglia).
- Cerebellar system learns to control using the reflexes as teaching signals.

Picture of cerebellar system

Reflexes

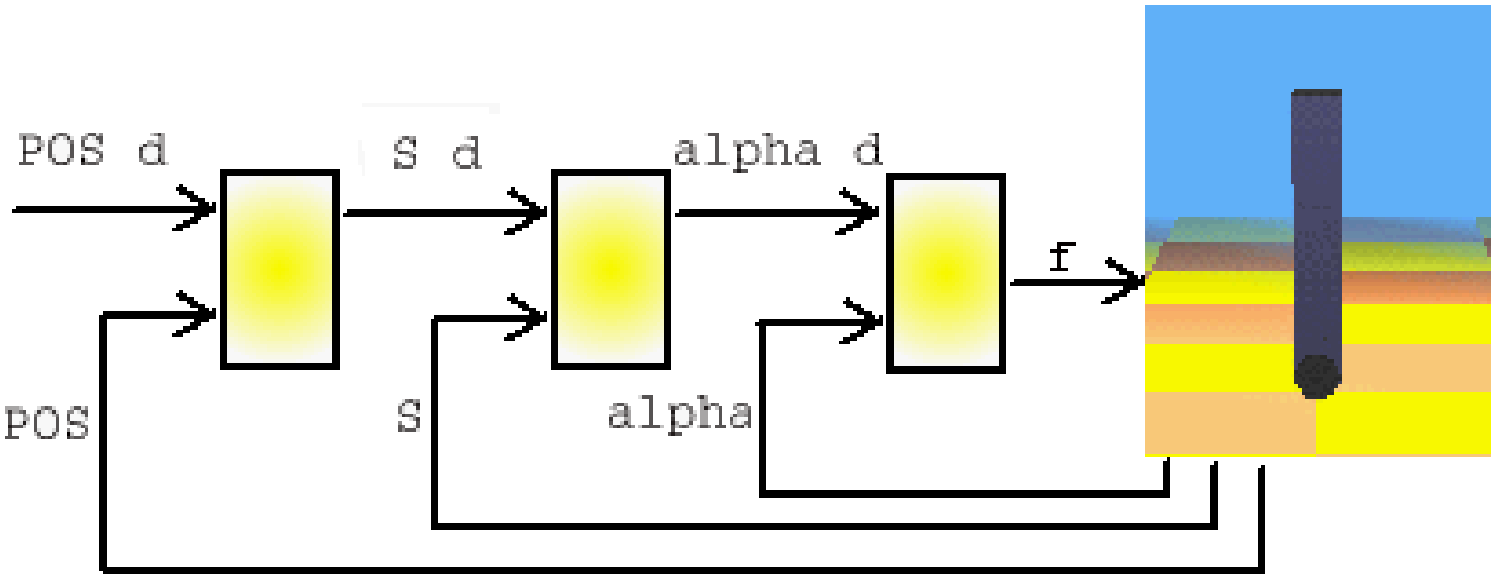
Stretch reflex

Opto-kinetic reflex

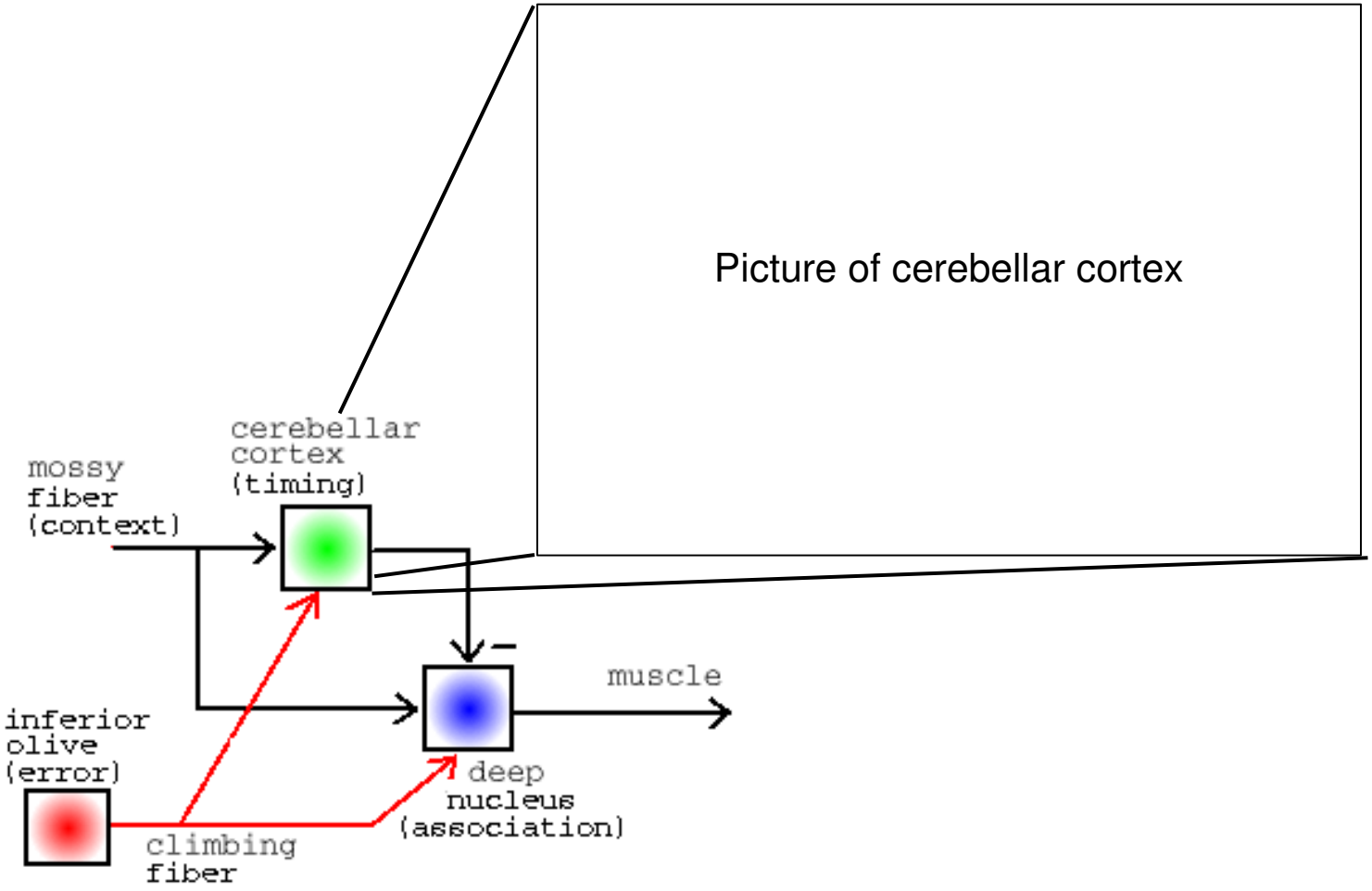
Picture:
http://www.inma.ucl.ac.be/EYELAB/neurophysio/perception_action/vestibular_optokinetic_reflex_fichiers/image004.jpg

Picture:
http://www.cs.stir.ac.uk/courses/31YF/Notes/musstr.jpg

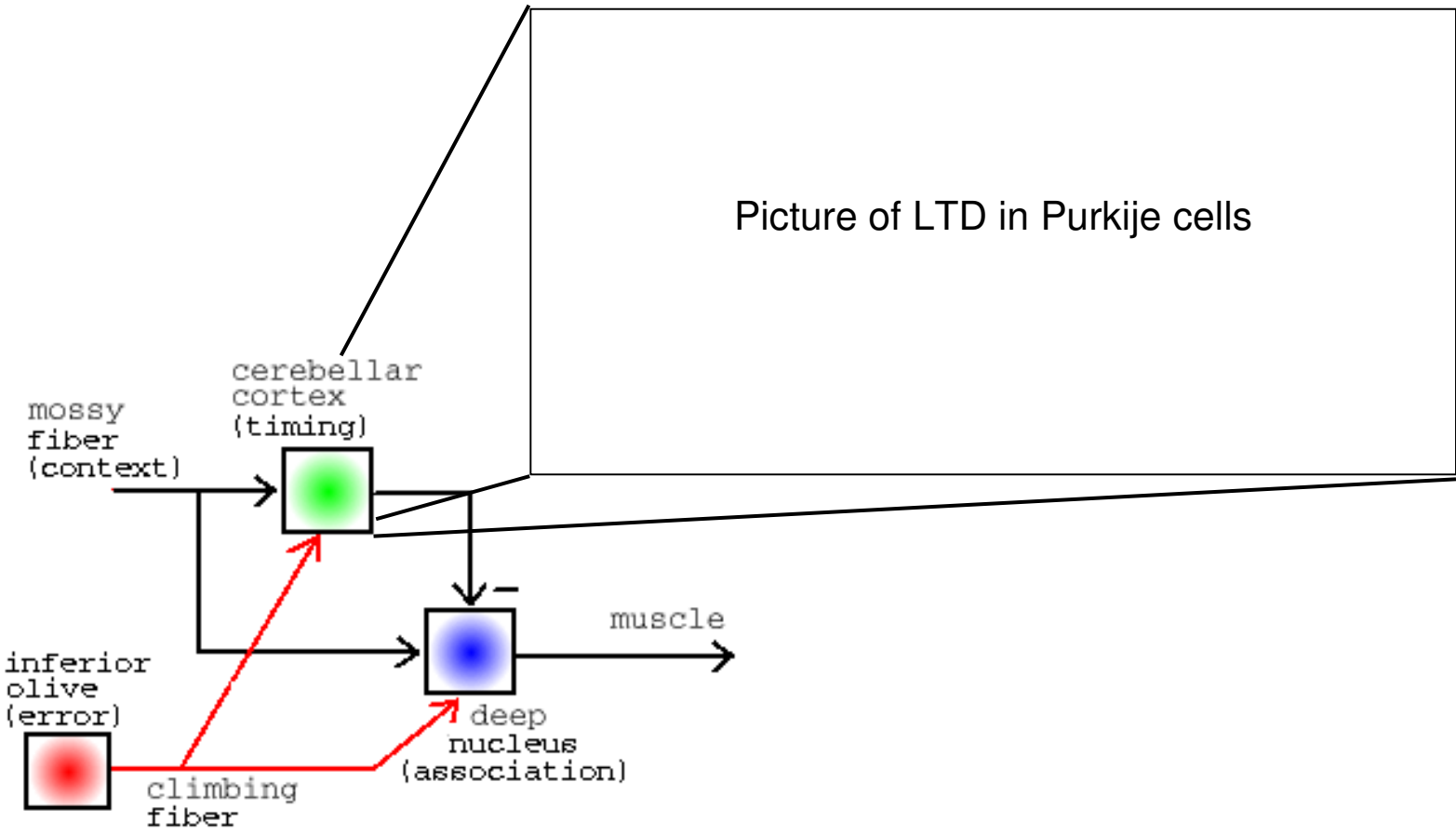
Robot “Reflex”



The Cerebellar System



Long-Term Depression (LTD) Guided by Climbing Fibres



Vestibulo-Ocular Reflex

Picture:
http://www.uq.edu.au/nuq/jack/VOR.jpg

System-Level Computational Neuroscience

Questions to be answered:

- What kind of components are needed for a cognitive architecture?
- What are different algorithms good for and how they can be combined?

The brain is a good solution to these questions ➔ Try to understand its algorithms on system level (level of behaviour)

“Without knowledge action is useless and action without knowledge is futile”

—Abu Bakr (c. 573-634)

Components for a Cognitive System

Basal ganglia: selection, reinforcement learning (trial-and-error learning)

Hippocampal formation: one-shot learning, navigation, episodic memory

Neocortex:

- Represents the state of the world — including oneself
- Invariant representations, concepts
- Attention / selection (both sensory and motor)
- Simulation of potential worlds = planning and thinking
- Relations and other structured representations (akin to symbolic AI)

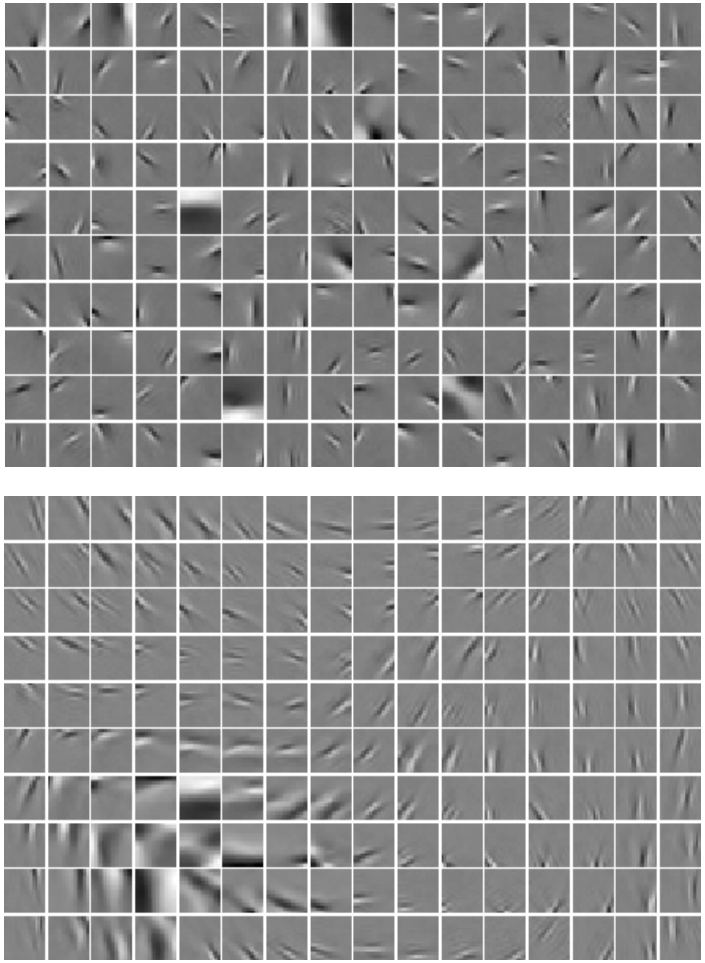
Neocortex

- A hierarchy of feature maps: increasing levels of abstraction
- Bottom-up and top-down/lateral inputs treated differently
- Local competition
- Long-range reciprocal excitatory connections

Picture:
http://www.pigeon.psy.tufts.edu/avc/husband/images/lsocrtx.gif

Representations for Natural Images by Independent Component Analysis (ICA)

http://www.cis.hut.fi/projects/ica/imageica/



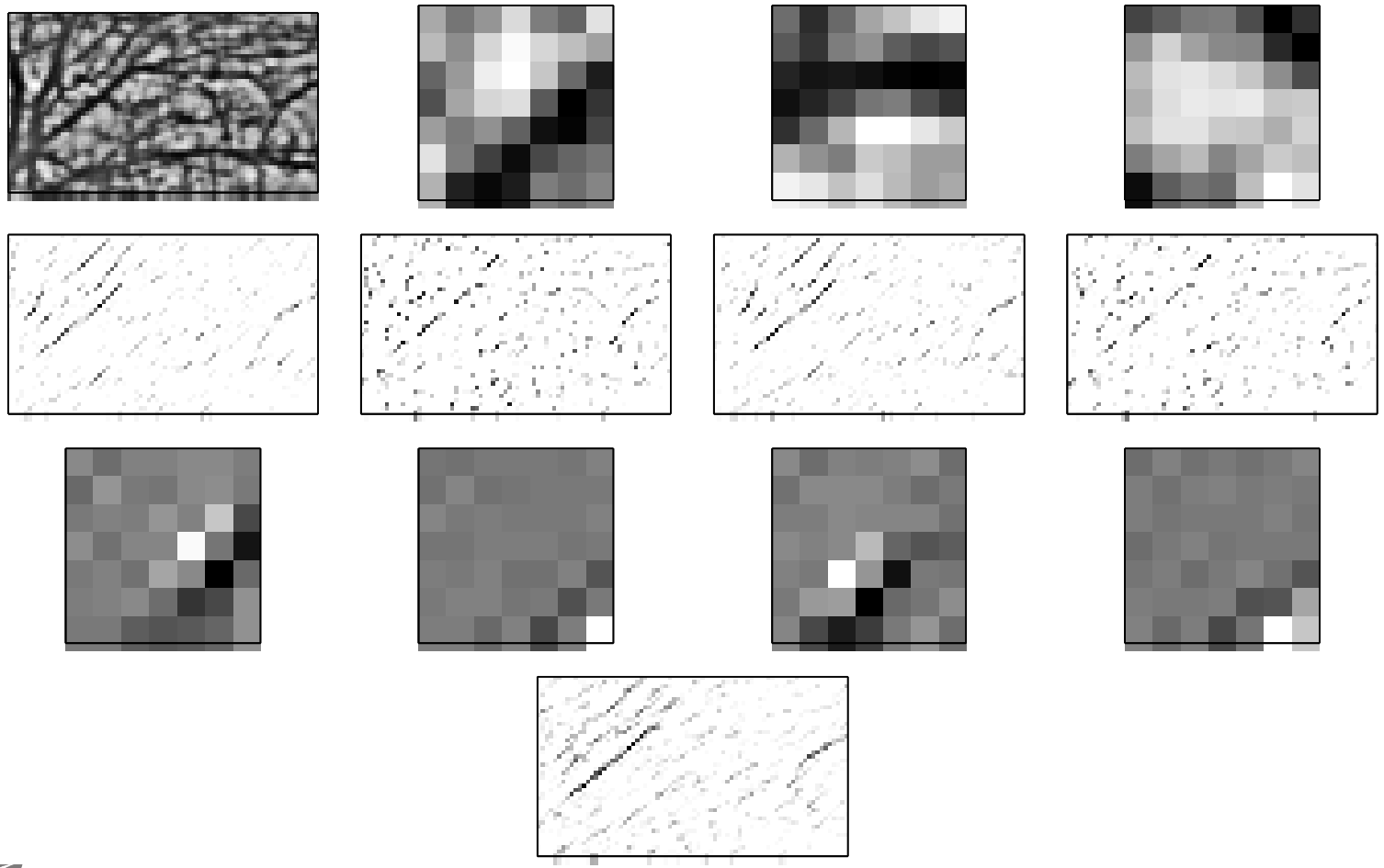
- ICA is an example of unsupervised learning.
- Can learn something like V1 simple cells.

Invariant features

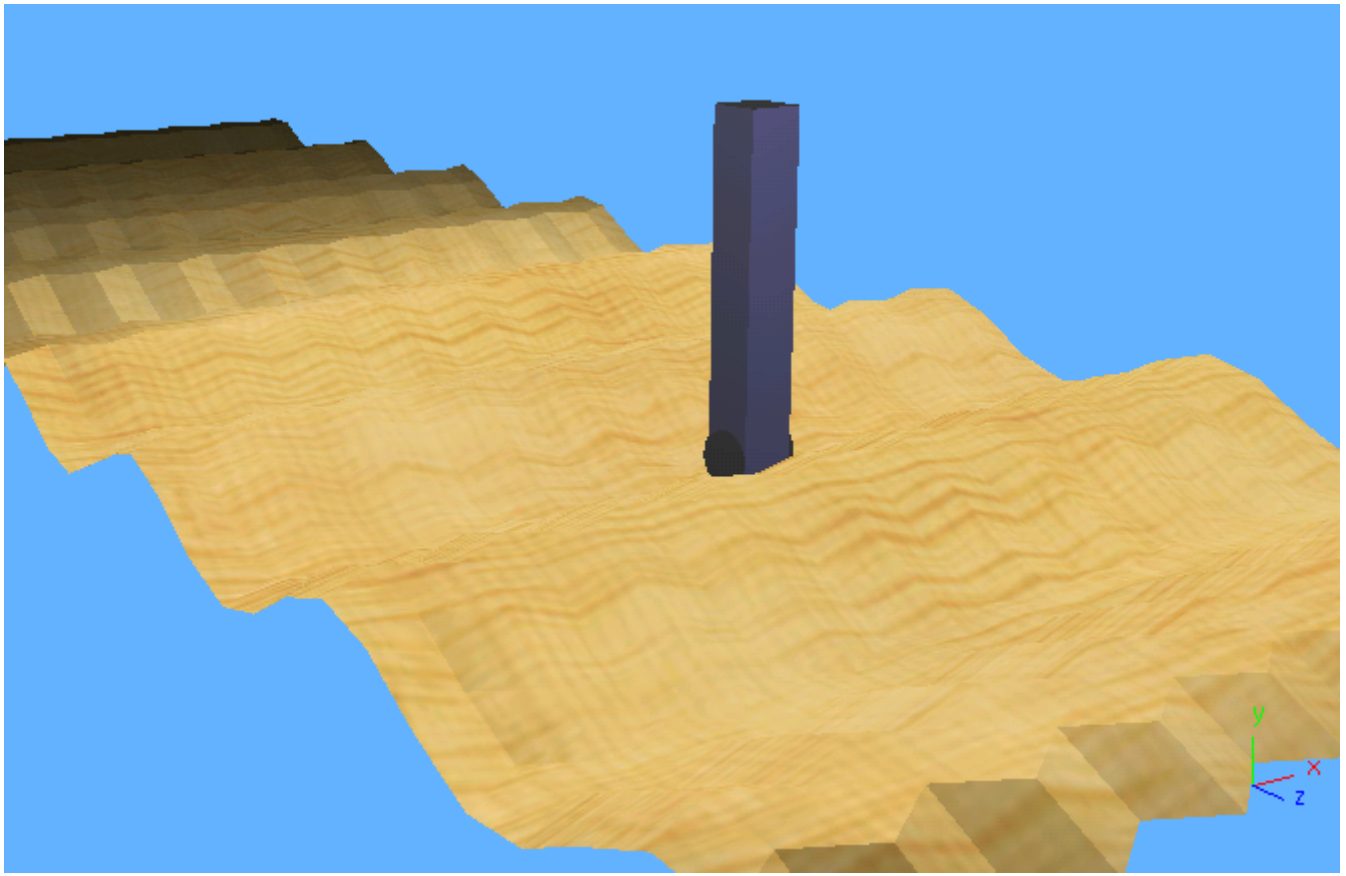
- Group simple features into complex in a hierarchical model.

Picture:
http://cs.felk.cvut.cz/~neurony/neocog/en/images/figure3-1.gif

”Complex Cells” from Images



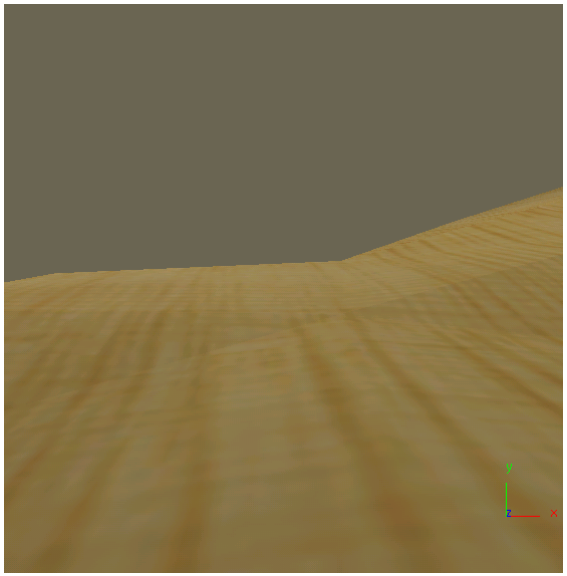
Abstractions and Meaning



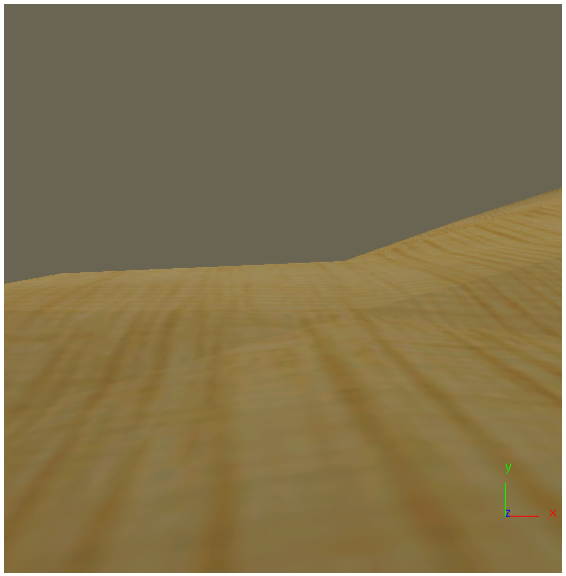
Abstractions and Meaning

- Once we have motor output, we can learn which information is important and meaningful

Left camera



Right camera



Relevance to Human Enhancement

How about mind prostheses? New senses (like web-sense)?

Knowing how the brain works would certainly be useful for prostheses, but for healthy persons…

- Input to the brain is easiest to deliver through existing senses — content matters, not the channel
- Output from the brain through motor system is more limited → implanted electrodes might surpass this capacity
- I expect intelligent tools to be far more common than prosthetic devices for a long time, but this doesn't mean their societal impact would be any smaller

Brain is a good solution for an **engineering** problem

Neuroscience

Picture:
http://britton.disted.camosun.bc.ca/escher/drawing_hands.jpg

Technology

KIITOS! THANK YOU!