Oscillatory Neural Network for Image Segmentation with Biased Competition for Attention

Tapani Raiko and Harri Valpola

School of Science and Technology

Aalto University

(formerly Helsinki University of Technology)

Oscillatory Neural Network for Image Segmentation with Biased Competition for Attention

Tapani Raiko and Harri Valpola

ZenRobotics Ltd.

www.zenrobotics.com

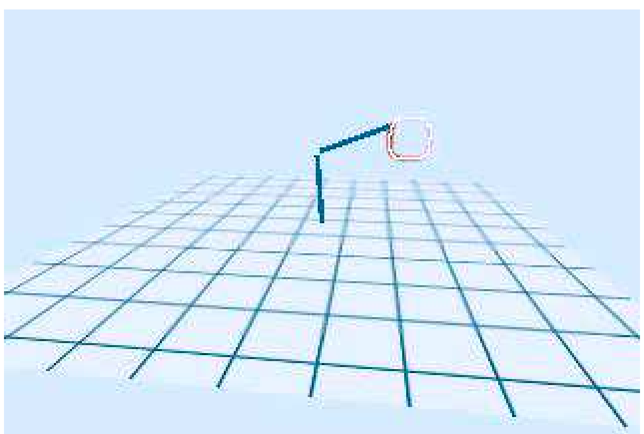
Computational neuroscience group

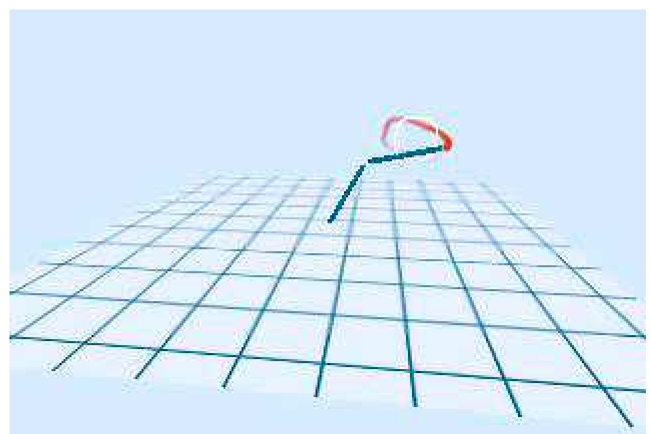
What: Figuring out how the brain works.

How: Building brains for robots = system-level modelling, implementing a whole vertebrate/mammalian brain.

Why: Because we can.

Intact cerebellum and sober





Cerebellar lesion or drunk

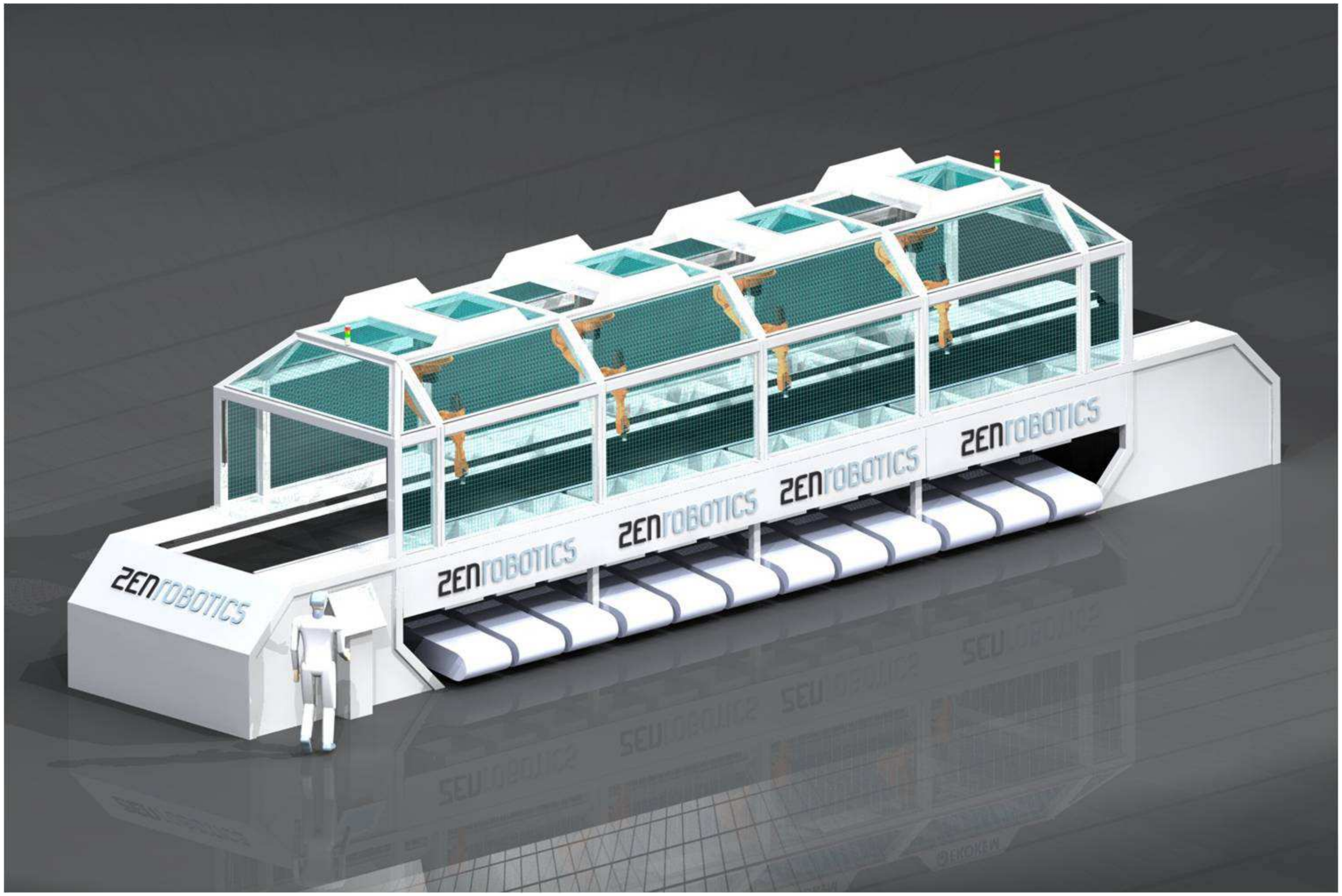
Our main strengths in research

Neuroscience ←→ **machine learning / AI / neural nets** ←→ robotics in unstructured environments

Cognitive architecture: the organisation of the whole brain

- **Cerebral cortex**
- Basal ganglia
- Cerebellum
- Hippocampus

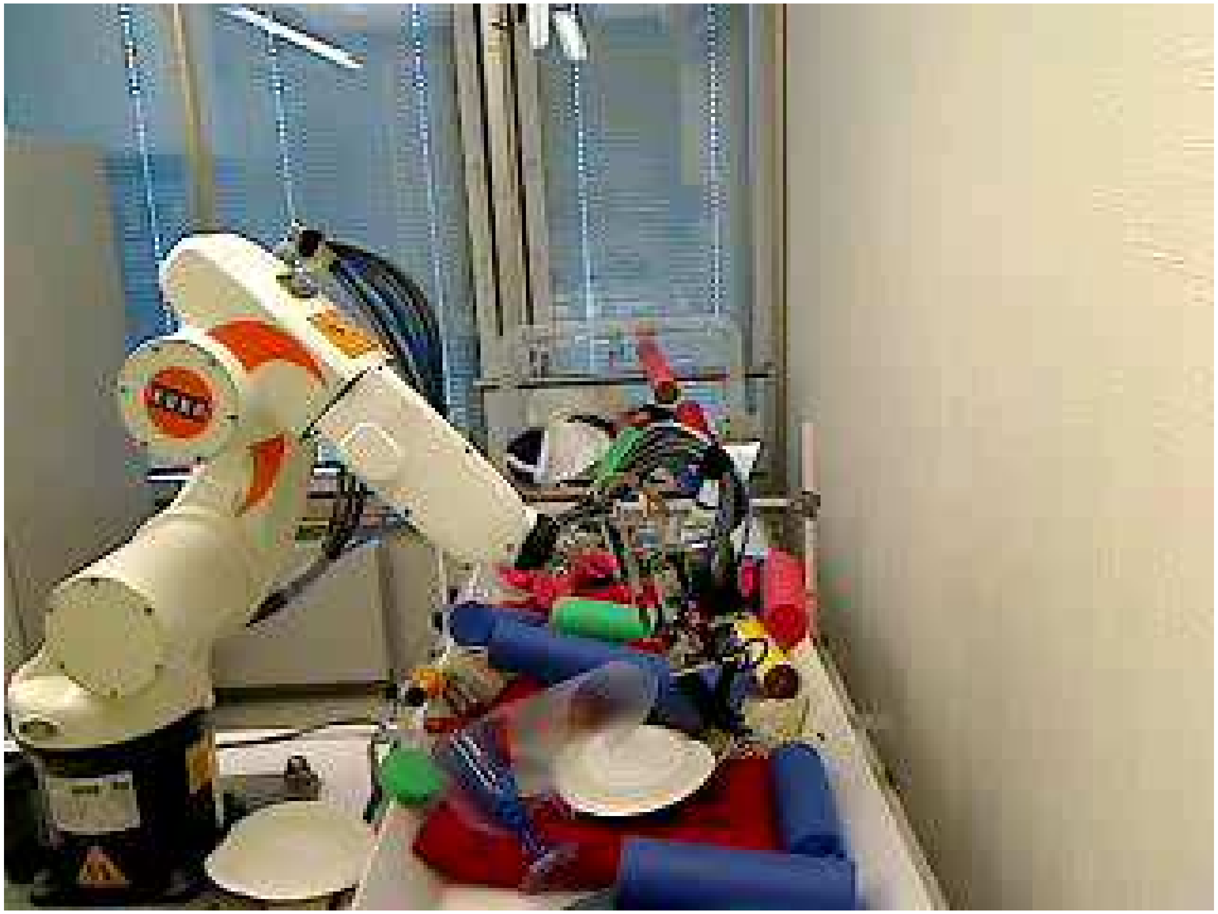
ZenRobotics Ltd.
Sorting waste with intelligent robots



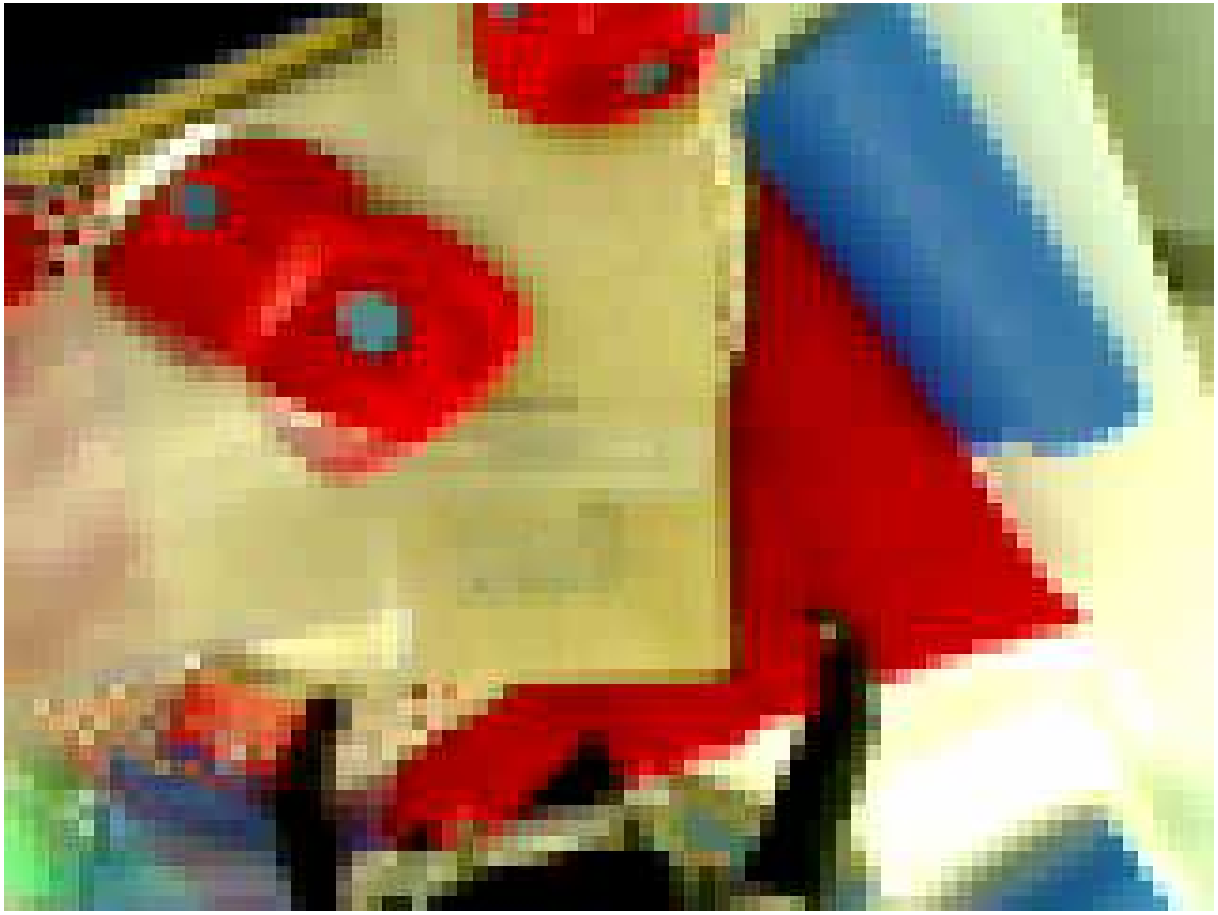
Harri Valpola, Computational Neuroscience Group, BECS,
School of Science and Technology, Aalto University, Finland

15.7.2010

ZenRobotics Ltd.
Waste recycling lab



ZenRobotics Ltd.
Robot's viewpoint



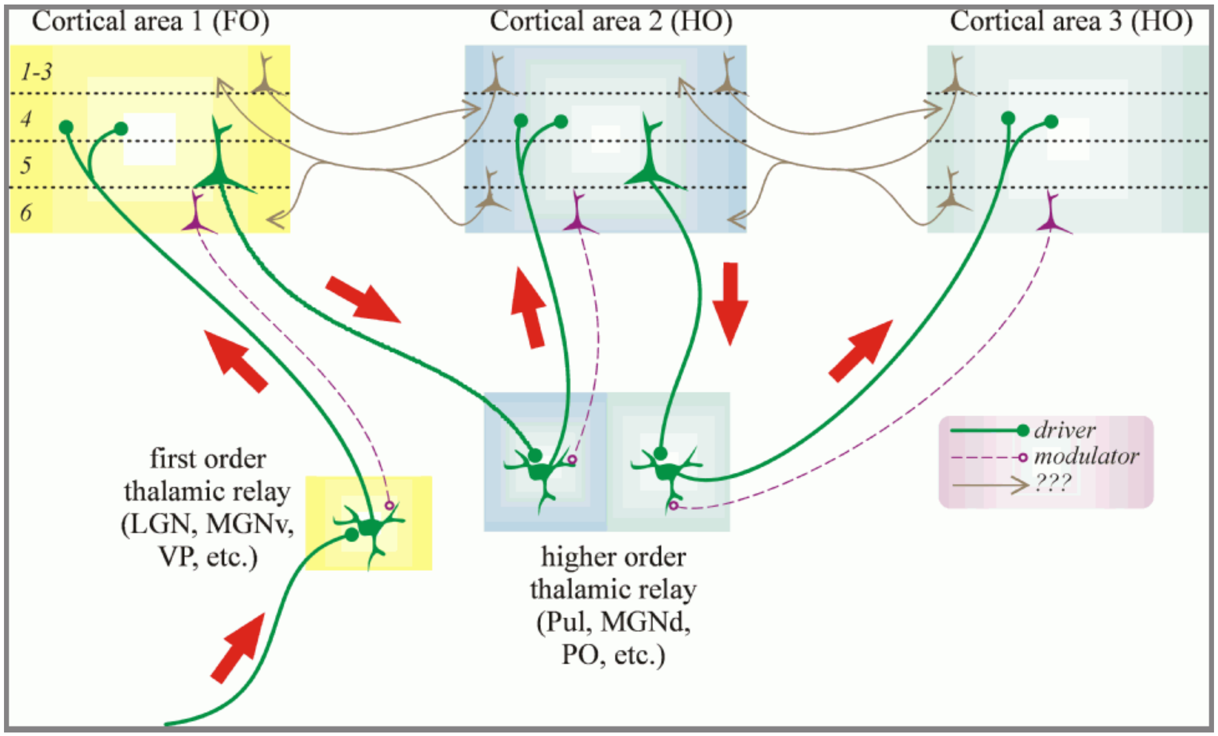
Cerebral cortex

Modelling the world (and yourself as part of it)

- Forward modelling: prediction and simulation
- Inverse modelling: figuring out which actions lead to desired consequences; planning

Different types of inputs and outputs

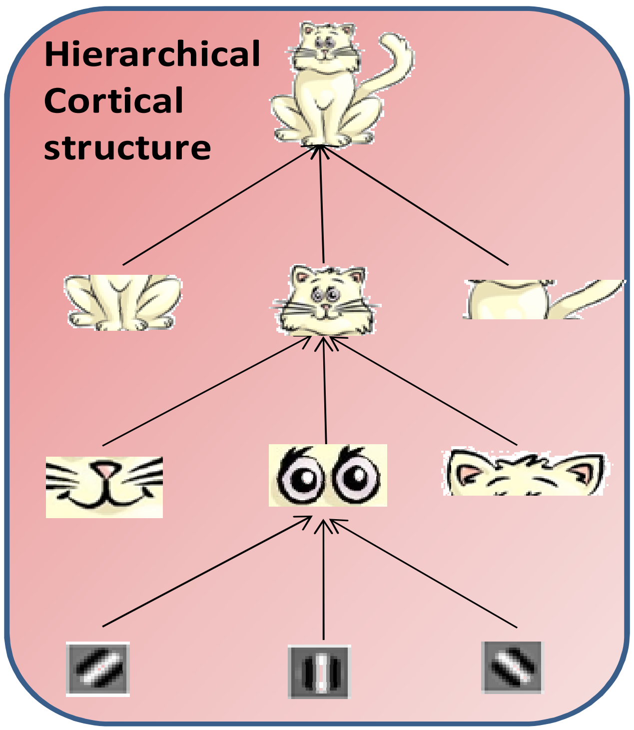
- Primary input usually from bottom-up (from the senses)
- Numerous feedback connections (order of 10 x)

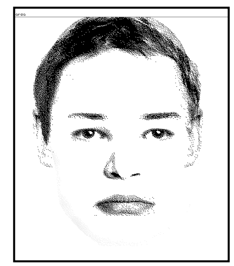


Introduction: Bayesian brain

Generative model / **Bayesian brain** hypothesis

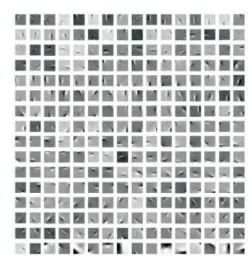
The generative model allows to **infer** the causes of sensations and to **predict** inputs





Top-down priors (context)

Perception = Probabilistic integration of information using Bayesian **inference**



Bottom-up sensory data

From the presentation of Salvador Dura, BICS 2010, 14.7.2010

Bayesian theory says:

- Decisions are based on
  1. Beliefs (measured by probability)
  2. Utilities
- The recipe:
  1. Evaluate the probabilities of all possible states of the world (probabilistic inference)
  2. Evaluate the probabilities of all outcomes for each and every potential action (probabilistic inference)
  3. Choose the action which maximises the expected utility
- This is optimal if there are no restrictions on the available computational resources

Selection of information

- But… computational resources *are* restricted →
- It is impossible to consider all the states and actions →
- It is necessary to select information in order to make decisions

Selection of information

- In practice, it has turned out to be impossible to learn complex abstractions from real data “bottom-up”
- There is too much structure → it is necessary to select which abstractions (groupings of elementary features) are meaningful
- Information will be lost
  - Example: learning the phonemes → results in inability to distinguish between foreign phonemes

Cerebral cortex

Modelling the world (and yourself as part of it)

- Forward modelling: prediction and simulation
- Inverse modelling: figuring which actions lead to desired consequences; planning

Selection of useful information

- Selective learning of features: selecting useful high-level abstractions (sensory and motor)
- Selective attention

Modelling objects and their relations

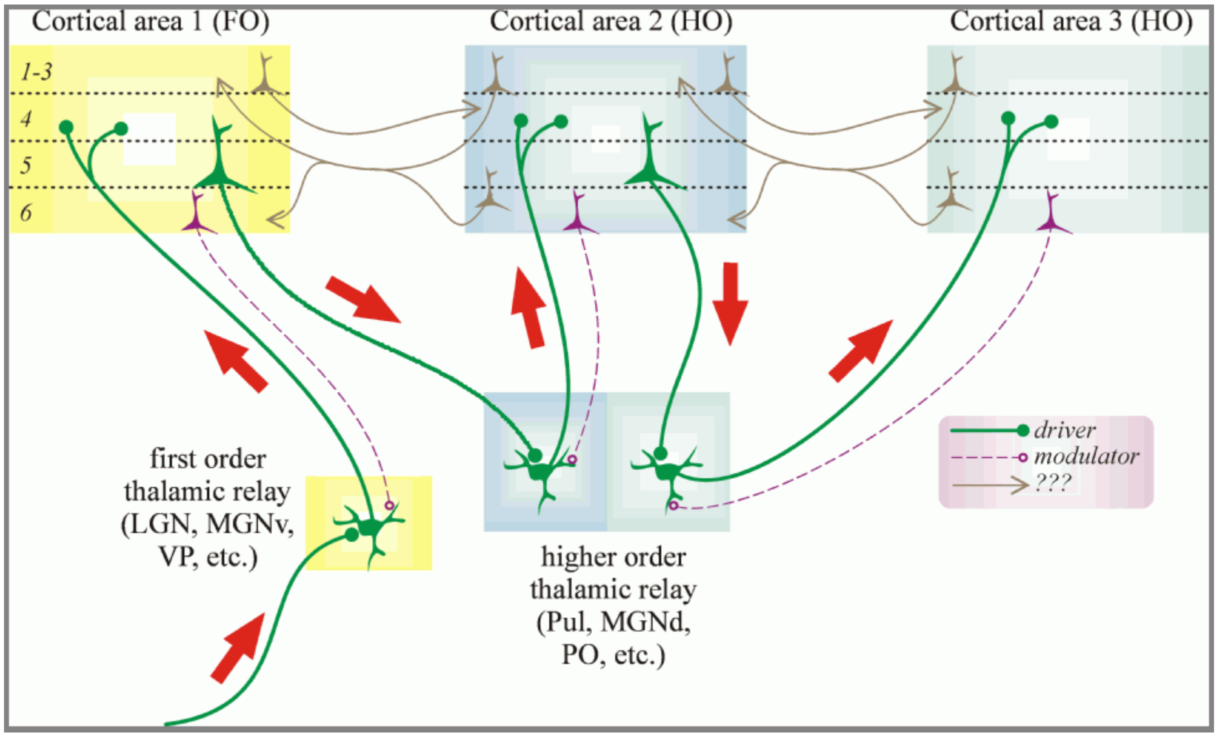
- Segmentation of objects

Key problem:
How to select useful information?

- It is necessary to select information in order to make decisions
- Selection is a type of decision, in other words:
- In order to decide we need to decide… Infinite regress!

Distributed selection on cortex

- Primary input usually from bottom-up (from the senses)
- Context (top-down or lateral) guides selection both in learning and on behavioural timescale



Invariant features

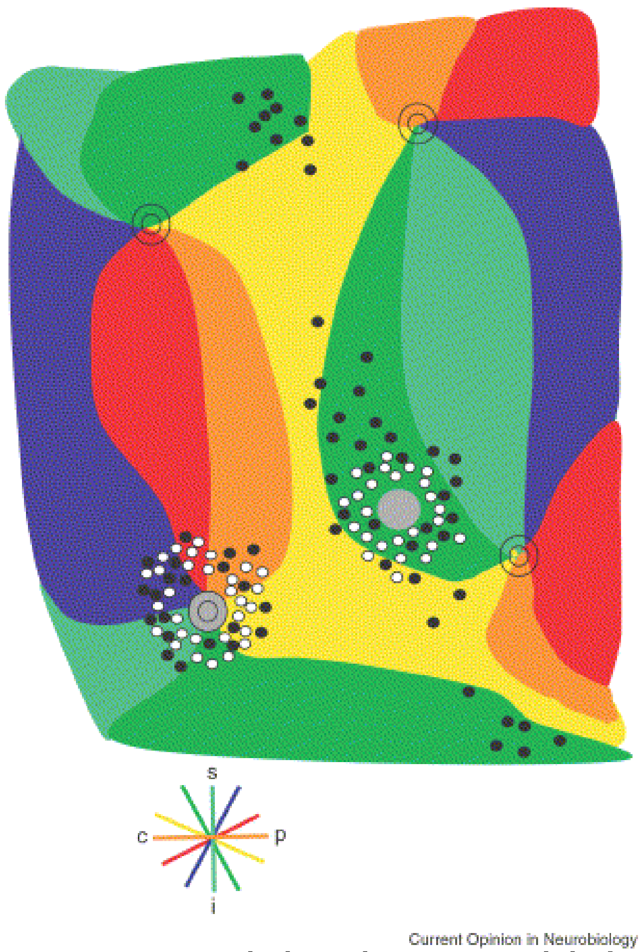
- Group simple features into complex ones in a hierarchical model
- What is the criterion?
  - Slow feature analysis: features that are activated close-by in time
  - Subspace ICA: features that are activated together
  - Canonical correlation analysis: features that are activated in the same context

Canonical correlation analysis

- A statistical technique which finds what is in common between two sets of data
- Find two projections (one from each dataset) such that their correlation is maximized
- Generalizes to several datasets, nonlinearities
- E.g., find visual features which are most relevant for motor control
  - On behavioural timescale, activations are determined mainly by visual bottom-up inputs
  - Motor context guides learning

Cortical long-range connections are specific

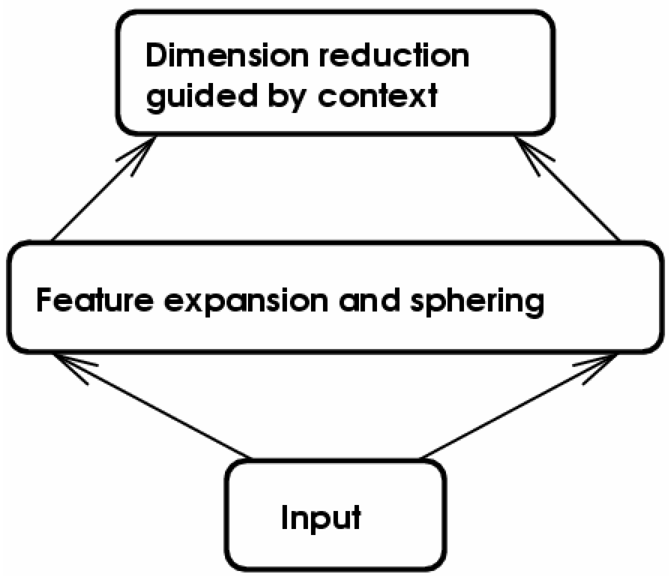
- Inhibitory connections (white dots) are local and symmetric
- Long-range excitatory connections (black dots) adapt through experience

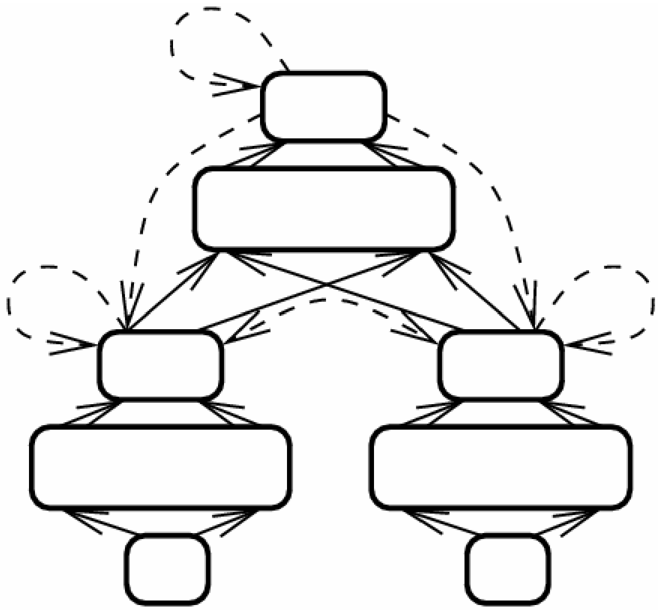


Kevan Martin, Current Opinion in Neurobiology, 2002

Context-guided learning of features

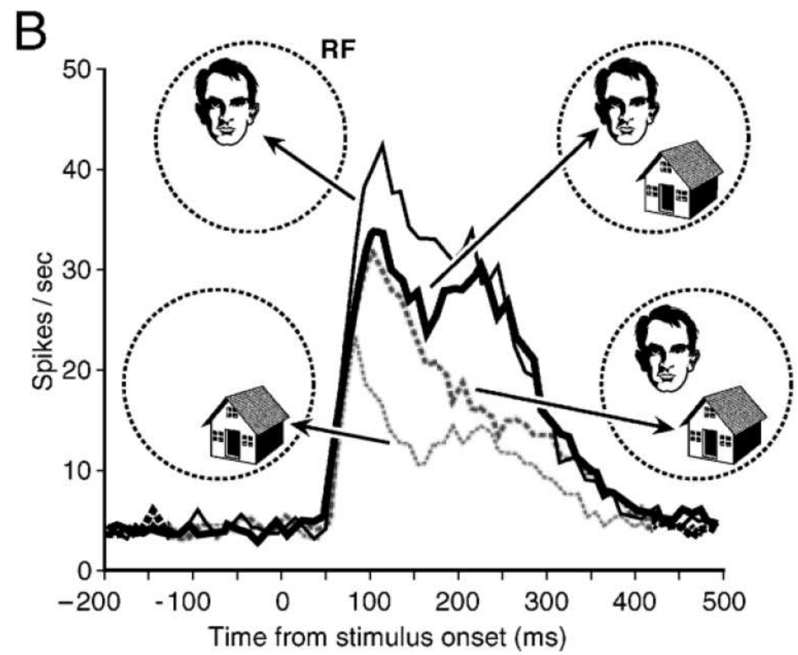
- Inputs are whitened
- Context (top-down and lateral) biases bottom-up activations
- Competitive learning → invariant features emerge (e.g., complex cells)





Attentional modulation of competition

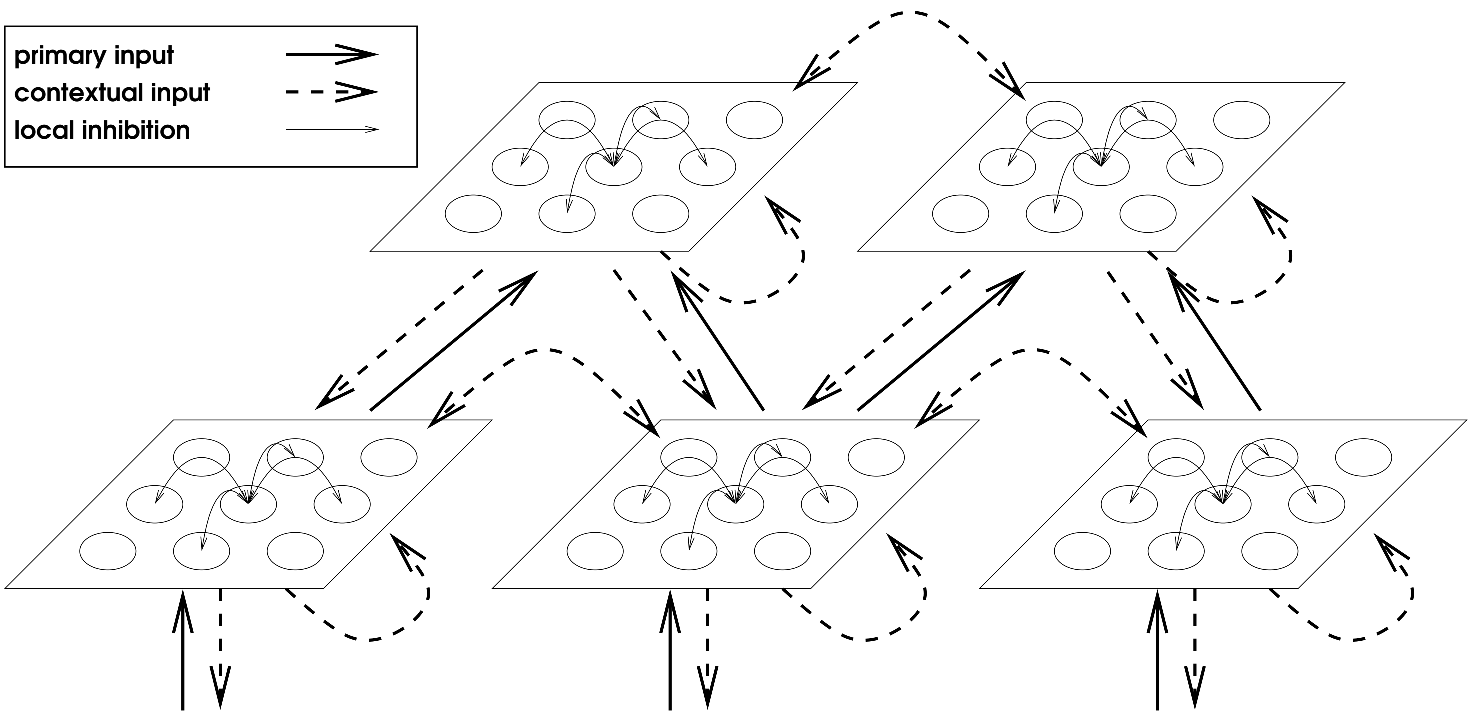
- A V4 neuron is recorded
- Weak activity for house, strong activity for face
- Intermediate activity for a combination
  - Excitation adds up but so does inhibition
- Selective attention can mask the effect of the distractor



Reynolds and Chelazzi, Annual Review of Neuroscience, 2004

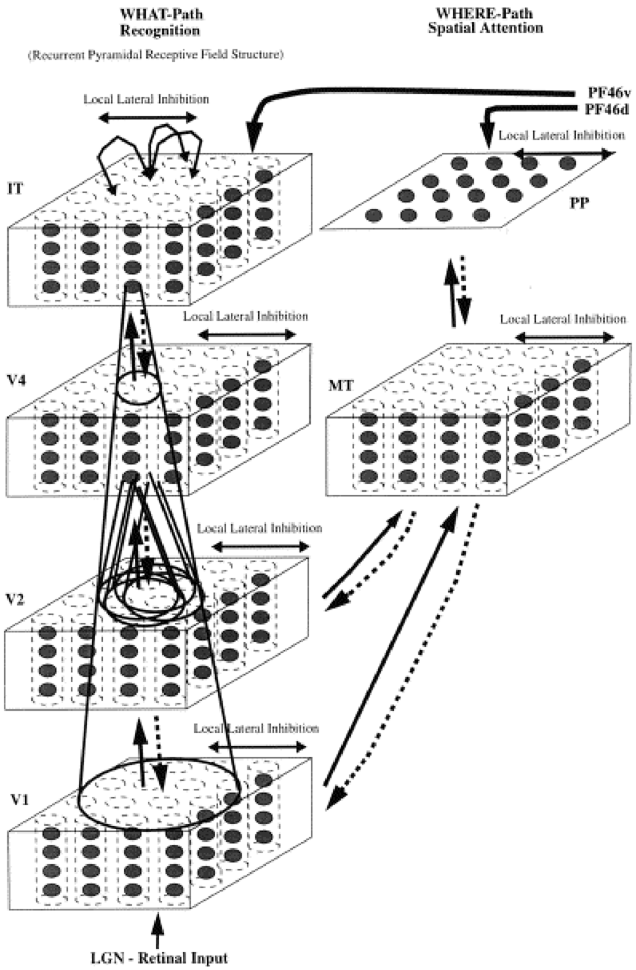
Biased competition model

- Local competition on each cortical area
- Context (top-down and lateral) biases the competition
- Selective attention emerges from the dynamics



Visual search

- Top-down biasing from working memory → implements visual search
- May look like sequential search (time increases with distractors) but the mechanism is fully parallel
- Both bottom-up and top-down phenomena related to (covert) attention can be explained



Deco and Rolls, Vision Research, 2004

Attention and learning: selection on different timescales

- Within the Bayesian framework, the only difference between perceptual inference and learning is the timescale
- In adults, perceptual learning is very strongly dependent on attention
- E.g., the same bottom-up input but different tasks → learn to perceive different aspects
- Selection in both attention and learning, only timescales differ

Biased competitive learning

- Biased competition + competitive learning (Master's thesis of Antti Yli-Krekola, 2007)
- Add adaptation (habituation, getting tired) → attention switches between objects
- Learning can be dramatically improved by switching attention (our paper at ICANN 2009)

Problems with engineering solutions to segmentation

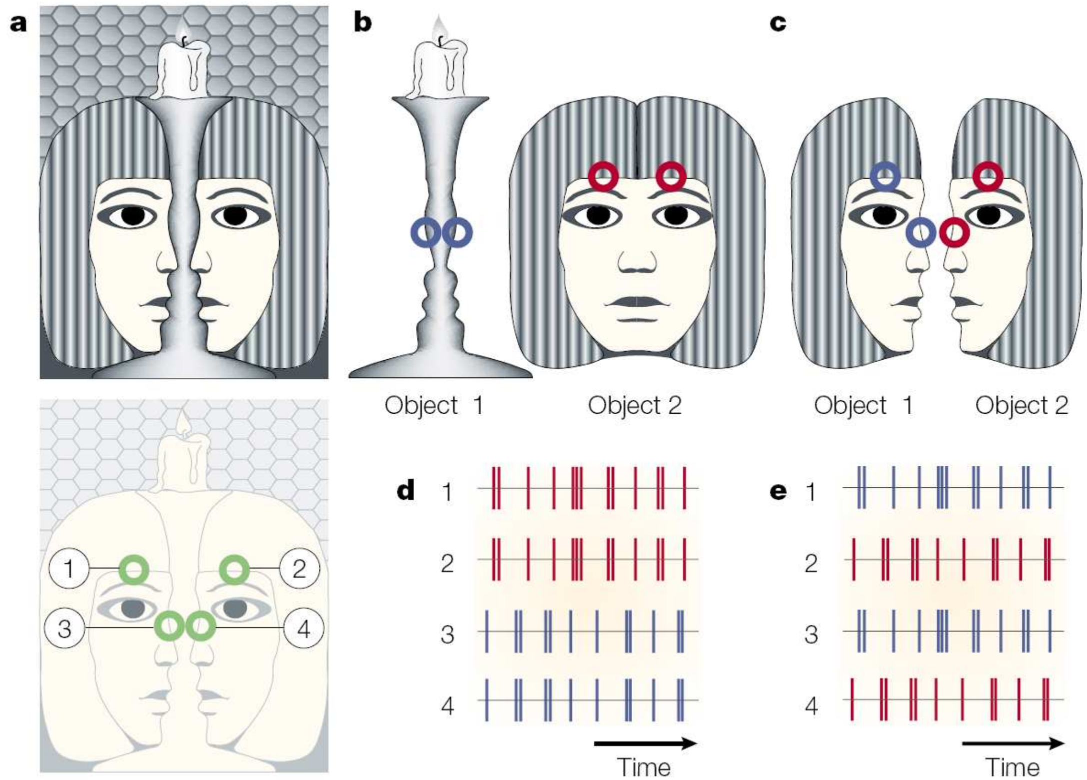
- Many engineering solutions suffer from a chicken-or-egg problem:
  - Recognition is usually successful only after segmentation
  - Segmentation is often successful only after recognition
- Iterative bottom-up / top-down message passing solves this problem in biased competition model

Segmentation still often remains poor

- Although biased competition will select one object at a time, segmentation can be poor

- From the viewpoint of an individual neuron/feature, we are effectively asking: "Do you belong to the currently active object or not?"
- An easier question would be: "Do you belong to object1 or object2 or … or something else?"
- In Bayesian terms: explaining away

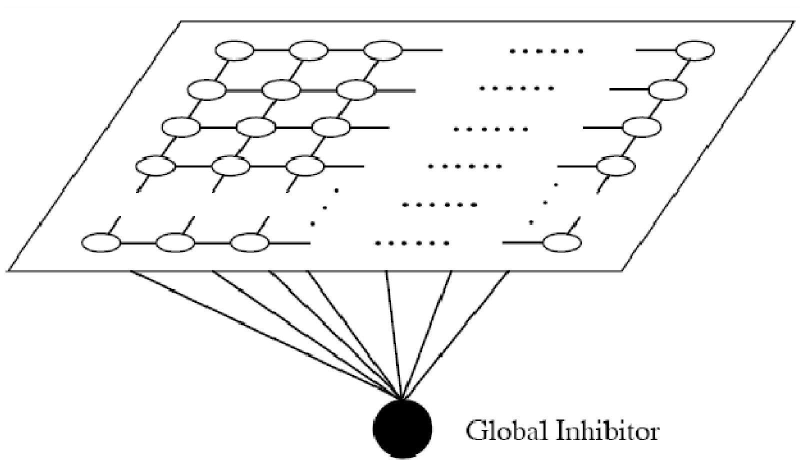
Segmentation and synchrony

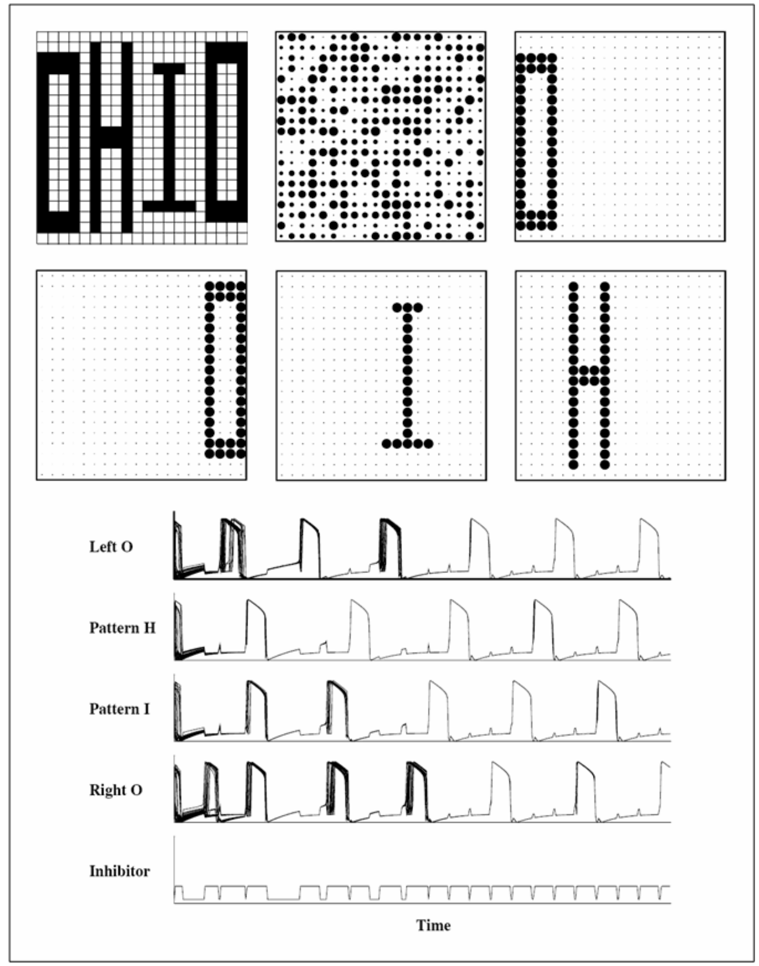


Engel, Fries and Singer, Nature Reviews Neuroscience, 2001

- Hypothesis: neurons encoding the same object synchronize

Segmentation with weakly coupled oscillators: LEGION model





http://www.scholarpedia.org/article/LEGION:_locally_excitatory_globally_inhibitory_oscillator_networks

Biased competitive learning + coupled oscillators

- Our goal in BICS 2010 paper is to study how biased competitive learning could be combined with coupled oscillators for better segmentation

Abstraction level:

- The behaviour of one area is what we design
  - try to keep it as simple as possible
- Try to come up with emergent oscillatory segmentation in a network of interconnected areas

Desired emergent properties

1. Integrates information from local patches of each object → translates into the gain of the population
2. The objects with globally highest gain emerge
3. Objects synchronise completely (not pieces of objects)
4. Objects desynchronise between each other (when their representations overlap)

Properties of a single area – part 1

Pretty much like biased competitive learning but:

- Intrinsic oscillators built from excitatory and inhibitory neurons
- Low-pass filtering is needed somewhere to build oscillators. Low-pass filter inhibition to avoid distorting the signal carried by excitatory neurons.
- Keep activations in check with gain control

Properties of a single area – part 2

- If no bottom-up input → no activation
- If bottom-up input →
  - If no support → little oscillation, little activation
  - If constant support → some oscillation, more activation
  - If oscillating support → phase-locked oscillation, strong activation
- Competing features push each others phases further away from each other

Details are not crucial, there are many ways to implement these properties

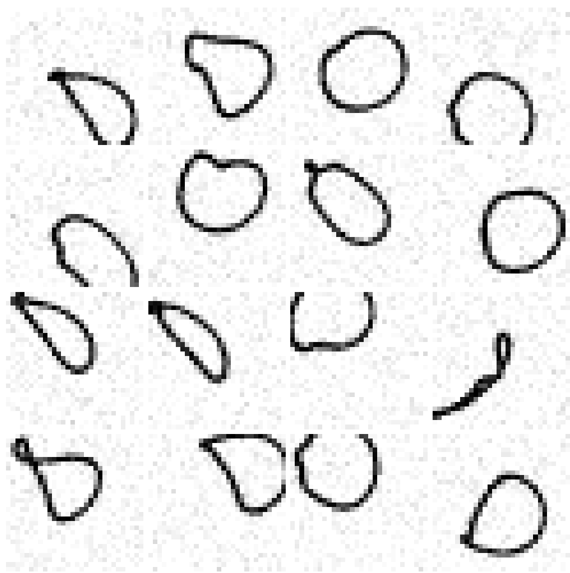
Results – part 1

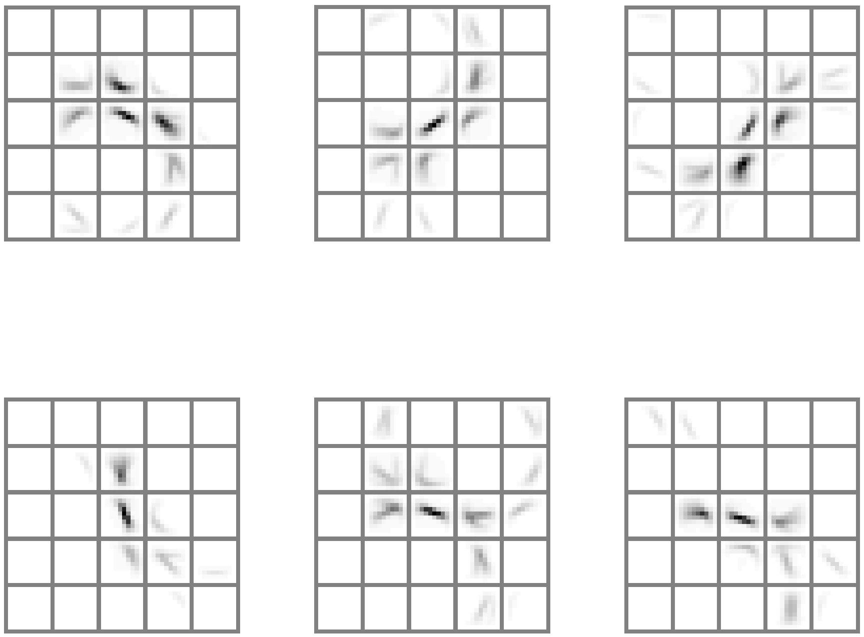
- First we checked that our implementation got the properties for a single area right

  (using externally generated bottom-up and lateral input)

- Seemed to work ok
- What happens now when we connect many areas together?

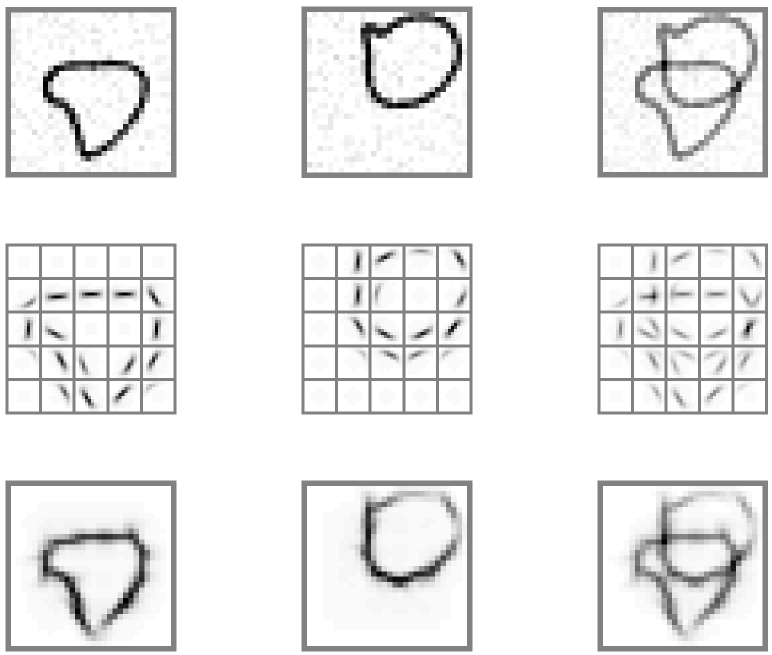
Results – part 2

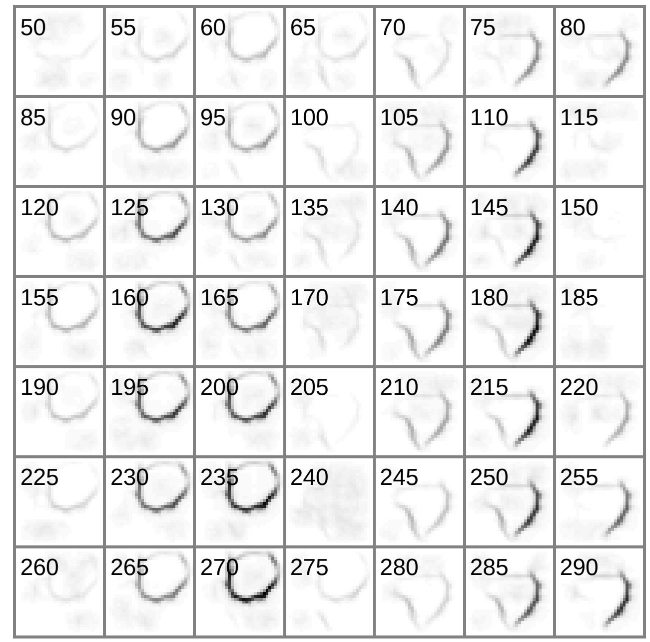


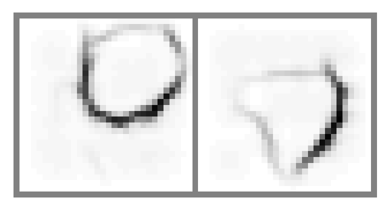


- Use simple visual inputs
- A network of 5×5 inputs
- Learn bottom-up features and their lateral connections (between excitatory neurons)

Results – part 3







Discussion

- Looks promising
- There was a problem with multiple activations from the same object within one area
  - Hierarchy and top-down support would help tie them together? Complex cells? Synchronize inputs?
- Obvious next step: test a hierarchical model

But I want to see it work in motor control:

- Is this really useful?
- Even context-guided learning and biased competition haven't been properly integrated with motor learning → test them separately first

Future / ongoing work

- Sensory and motor abstractions
  - Development of sensory abstractions guided by motor context and vice versa (akin to canonical correlation analysis)
  - Predictive power as a measure of value of information (needed for selection of information)
- Decision-making as biased competition on motor representations
- Better mappings
  - A hierarchical model of correlations at lower levels; latent variables describing the operating points of the system

Conclusions

- The experiments with artificial data have not proven that the system works in real life
- Nevertheless, looks promising
- I expect to revisit this work: once we need sophisticated segmentation, synchrony really might be a useful ingredient

Thank you!

www.zenrobotics.com