Bioloģiski iedvesmota mašīnmācīšanās datorredzes uzdevumu pildīšanai

Georgijs Kanonirs
Vadītājs: asoc. prof. Jānis Zuters

Mērķi

- Izveidot mašīnmacīšanas algoritmu, kas balstās uz "atmiņas-paredzēšanas" intelekta teorijas [1] un neokorteksa [2] uzbūves/darbības principiem
- Pielietojums – objektu detektēšana un atpazīšana

[1] http://en.wikipedia.org/wiki/Memory-prediction_framework

[2] http://en.wikipedia.org/wiki/Neocortex

Kāpēc?

Cilvēka smadzenes ir labākais rīks kognitīvo darbību pildīšanai

* http://www.salon.com/2013/07/18/inside_george_zimmermans_brain/

Kāpēc?

Cilvēka smadzenes ir labākais rīks kognitīvo darbību pildīšanai

mašīnmācīšanas metodes ir tas labāka alternatīva

Dati

Pielietojuma sfēras zināšanas

parametrizēts adaptīvais modelis

Klassifikācija, prognozēšana u.c.

* http://www.salon.com/2013/07/18/inside_george_zimmermans_brain/

Atmiņas-paredzēšanas teorija (1)

Jeff Hawkins
2004

- Smadzenes nav dators, bet atmiņas sistēma, kas saglabā pieredzi tādā veidā, kas atspoguļo patieso pasaules struktūru, atceroties notikumu secības un to attiecības.
- Paredzēšana ir intelektuālo darbību pamatmehānisms, kā arī intelekta kritērijs (nevis uzvedība).
- Cilvēka neokorteksс ir atbildīgs par gandrīz visiem augsta līmeņa funkcijam, t.sk. redzi, dzirdi, valodu, kustību plānošanu u.c.

http://en.wikipedia.org/wiki/On_Intelligence

Atmiņas-paredzēšanas teorija (2)

Jeff Hawkins
2004

- Atmiņas-paredzēšanas teorijas apkopojums:

  - Neokortekss izmanto vienotu algoritmu dažādām modalitātēm;

  - Neokortkesam ir hierarhiskā struktūra;

  - Neokorteksss tiek konstruēts no skaitļošanas moduļiem, kas ir zināmi ar nosaukumu „kanoniskā kortikālā ķēde“;

  - Neokorteksa pamatfunkcija ir pasaules modeļa veidošana (no telpiskiem un laiciskiem šabloniem). Šis modelis tiek izmantots, lai veikt prognozes par ieejas datiem;

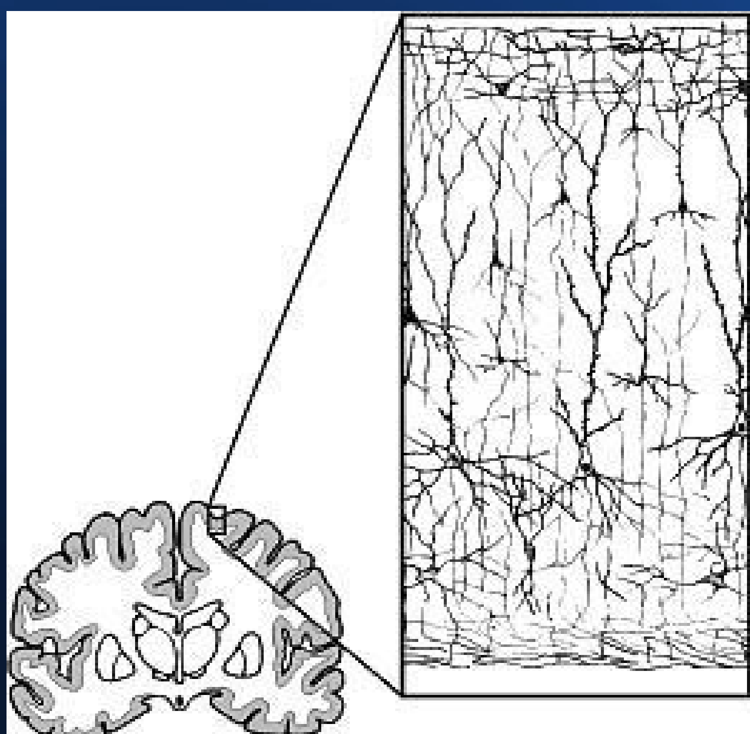
  - Neokorteksss veido pasaules modeli nekontrolētā veidā („unsupervised learning“);

  - Informācija tiek nodota uz augšu un uz leju hierarhijā, lai atpazītu un atrisinātu neskaidrības, kā arī tiek pavairota uz priekšu laikā, lai prognozētu nākamo ievadu.

Vienots algoritms (1)

- Regulārā struktūra

  1978 – V.B. Mountcastle „An Organizing Principle for Cerebral Function: The Unit Model and the Distributed System“



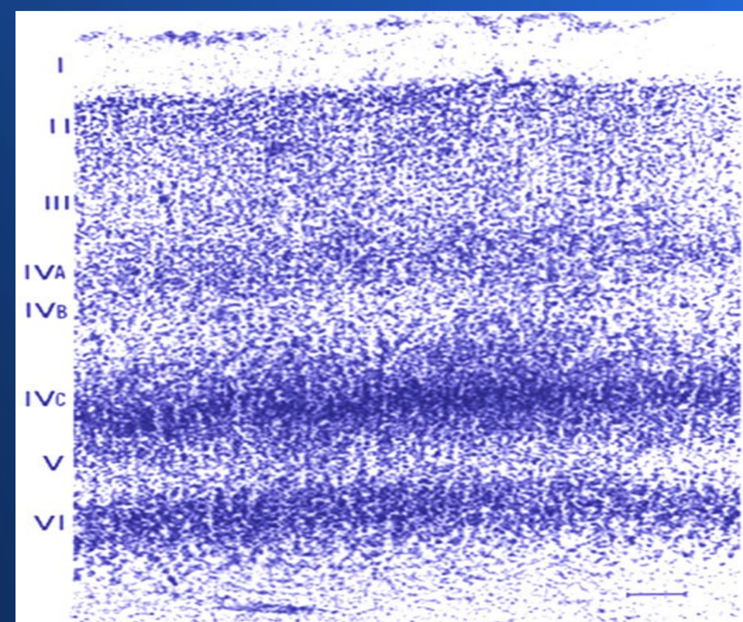


Figure 13. Nissl stain of the visual cortex reveals the different layers I through VI quite clearly.

* „How to make computers work like the brain“, G. Dileep, 2010 (presentation)

Vienots algoritms (2)

- Maņu aizvietošanas eksperimenti



1. http://www.stepbystep.com/eyemusic-device-converts-images-to-music-55746/
2. http://www.nei.nih.gov/news/briefs/weihenmayer.asp
3. http://www.creativeapplications.net/environment/feelspace/

No Free Lunch

- Neviens mācīšanās algoritms nav labāks par citiem visu problēmu risināšanai (Wolpert, 1995)

=> Algoritma pārākums ir atkarīgs no pieņēmumiem par problēmu

No Free Lunch

- Neviens mācīšanās algoritms nav labāks par citiem visu problēmu risināšanai (Wolpert, 1995)

=> Algoritma pārākums ir atkarīgs no pieņēmumiem par problēmu

Datiem no dažādām modalitātēm pamatā ir viena un tā paša statistiska struktūra

Jautājums

Kāda ir pamatpieņēmumu kopa, kas ir

- pietiekami specifiska, lai mācīšanos padarītu iespējamu saprātīgā laikā
- pietiekami vispārīga, lai tā būtu piemērota plašākam problēmu lokam

High-level properties of neocortex

- Hierarchical representation

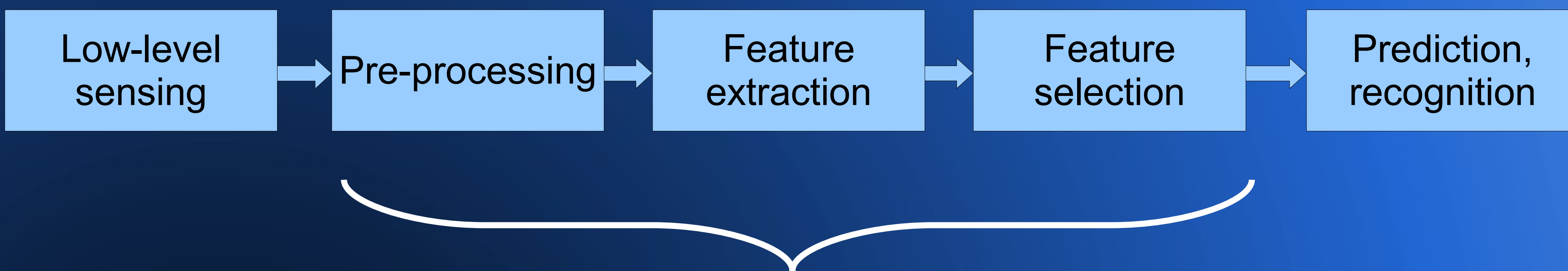
  - spatial and temporal
- Sparse distributed representation
- Using time as superviser
- Ability to make predictions
- Feed-forward and feed-back connections
- Sensori-motor integration

The pipeline of machine visual perception (1)

Low-level sensing → Pre-processing → Feature extraction → Feature selection → Prediction, recognition

- Most critical for accuracy
- Most time-consuming in dev. cycle
- Often hand-crafted in practice (SIFT, SURF, HOG etc.)

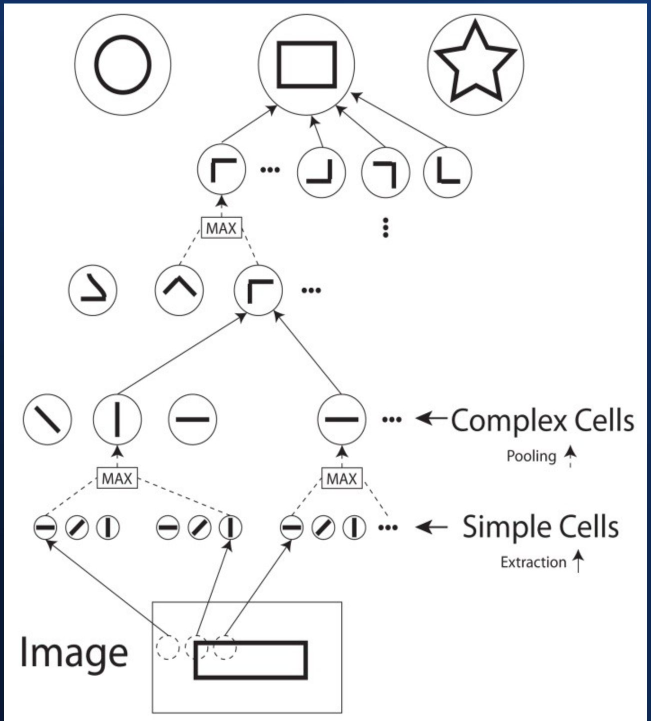
The pipeline of machine visual perception (2)



Instead of design features, let's design feature learning mechanisms

State-of-Art: Deep learning

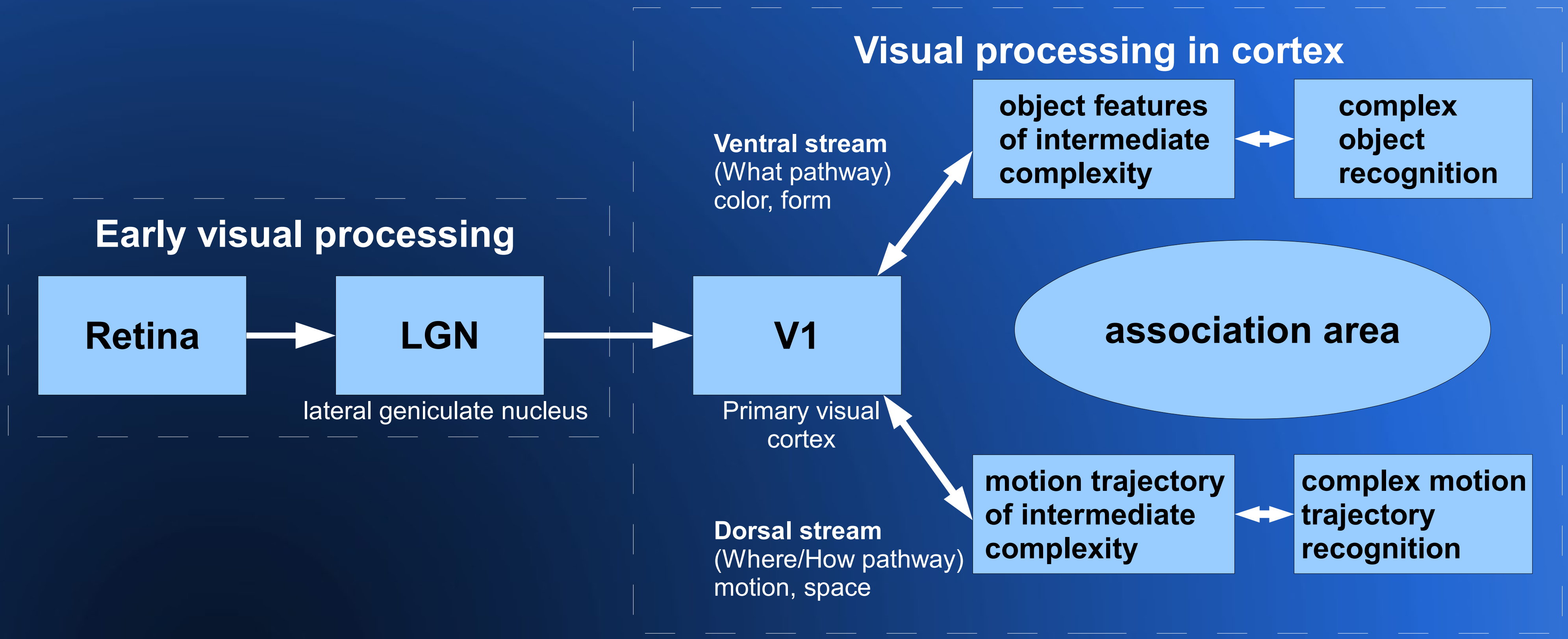
- **DEFINITION.** Deep learning is a set of algorithms in machine learning that attempt to learn in multiple levels of representation, corresponding to different levels of abstraction.



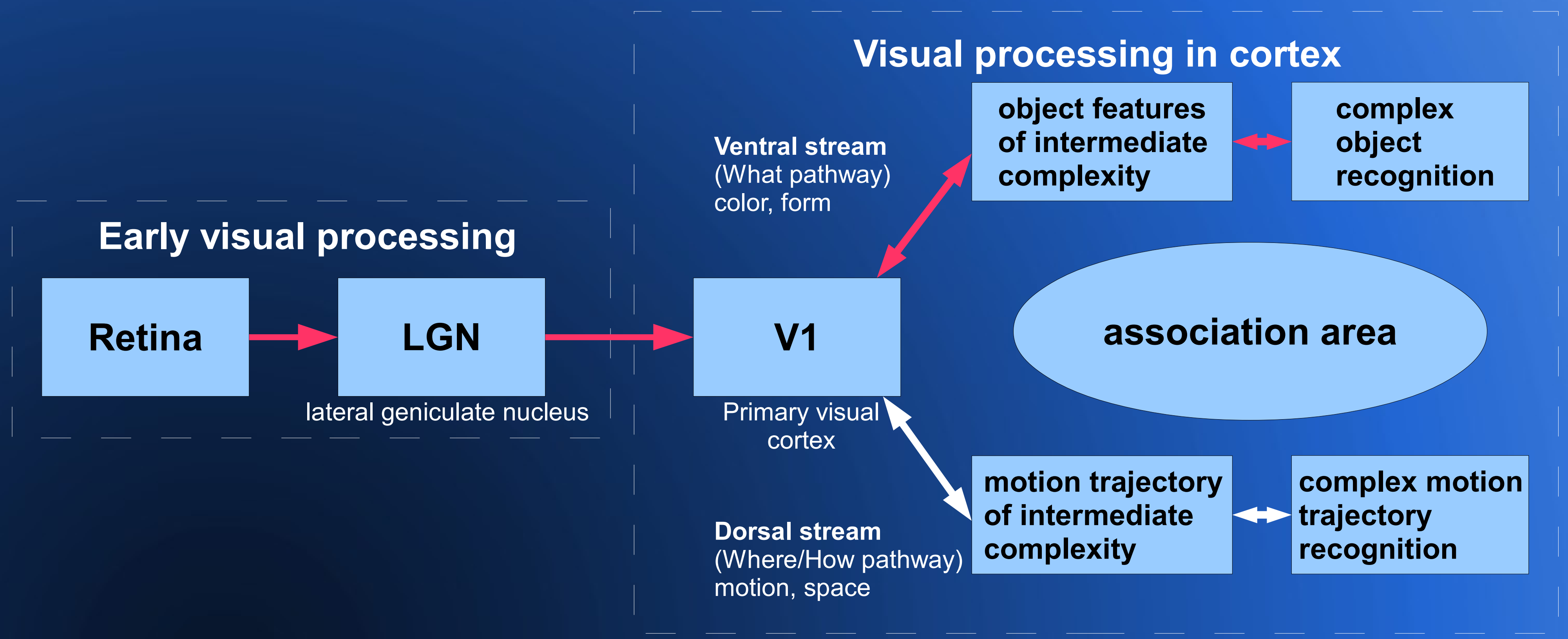
Algorithms

- Convolutional neural networks, HMAX
- Deep belief network (DBN) based on
  - restricted Boltzmann machines (RBM)
  - autoencoders
- Hierarchical Temporal Memory (HTM)
- Deep LCA (locally competitive algorithm)
- etc.

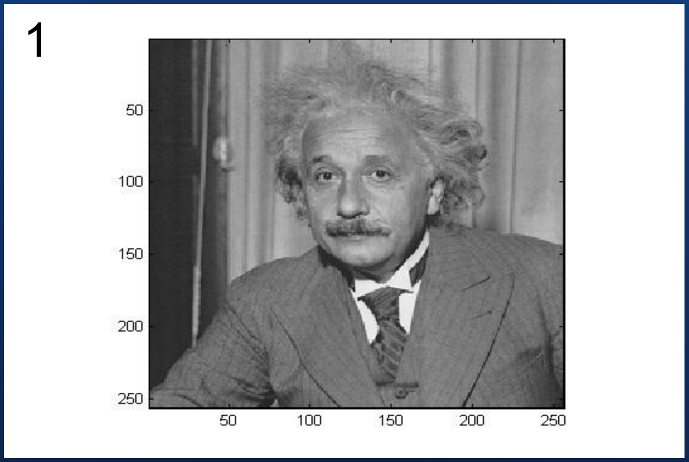
Visual information flow



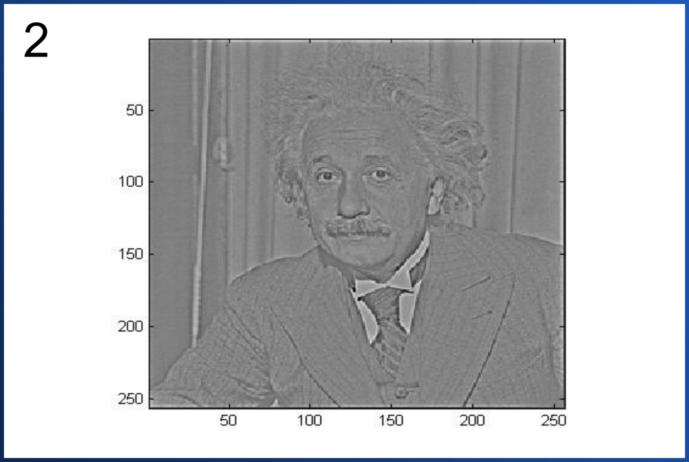
Visual information flow



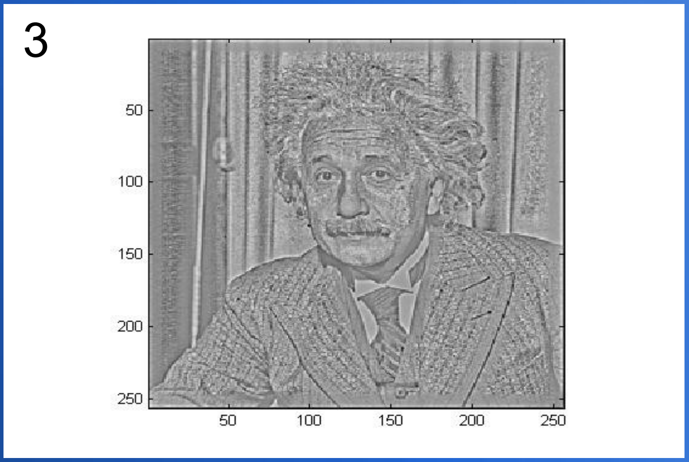
Early visual processing



Original image



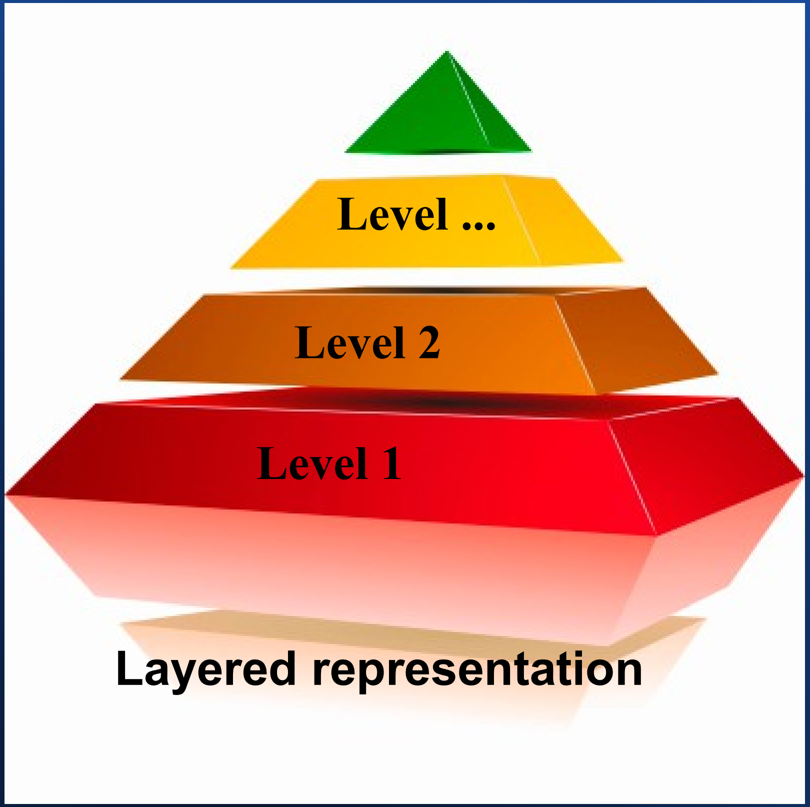
After whitening
(removes pairwise, linear correlations)

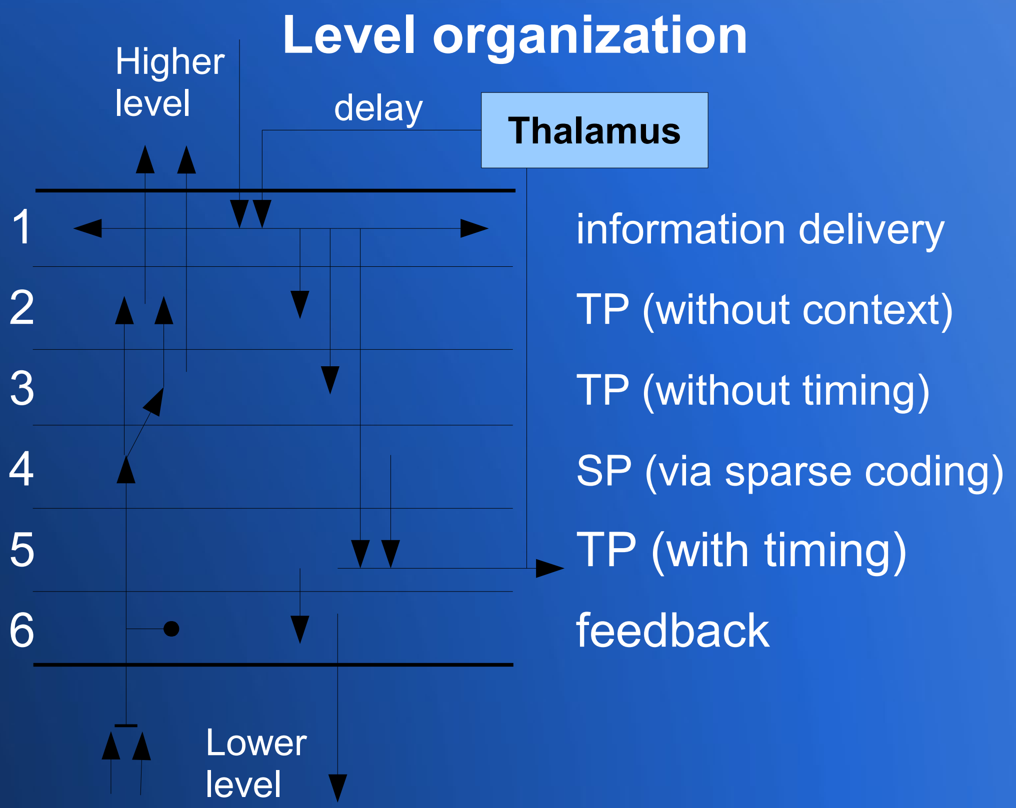


After local contrast normalization
(brings out image details)

http://redwood.berkeley.edu/bruno/npb261b/lab2/lab2.html
http://redwood.psych.cornell.edu/papers/graham-chandler-field-2006.pdf

Architecture





- Invariant spatio-temporal representation is formed on each layer
- Hierarchical Bayesian inference is used for layer interconnection

1. TP – temporal pooling
2. SP – spatial pooling

Efficient spatio-temporal coding

- **Hierarchical structure**
  - Goal: To capture hierarchical model of the world
  - Using spatial and temporal pooling
- **Sparse representation**
  - Goal: To build hierarchical representation efficiently
  - Using sparse coding via lateral inhibition
- **Sequence memory**
  - Goal: To make invariant representation
  - Using memory and prediction
- **Topographic maps**
  - Goal: To make connections shorter
  - Using exicitation and inhibition connections

Sparse code

- The sparse code is a kind of neural code in which each item is encoded by the strong activation of a relatively small set of neurons. For each item to be encoded, this is a different subset of all available neurons.
- Given a potentially large set of input patterns, sparse coding algorithms attempt to automatically find a small number of representative patterns which, when combined in the right proportions, reproduce the original input patterns.

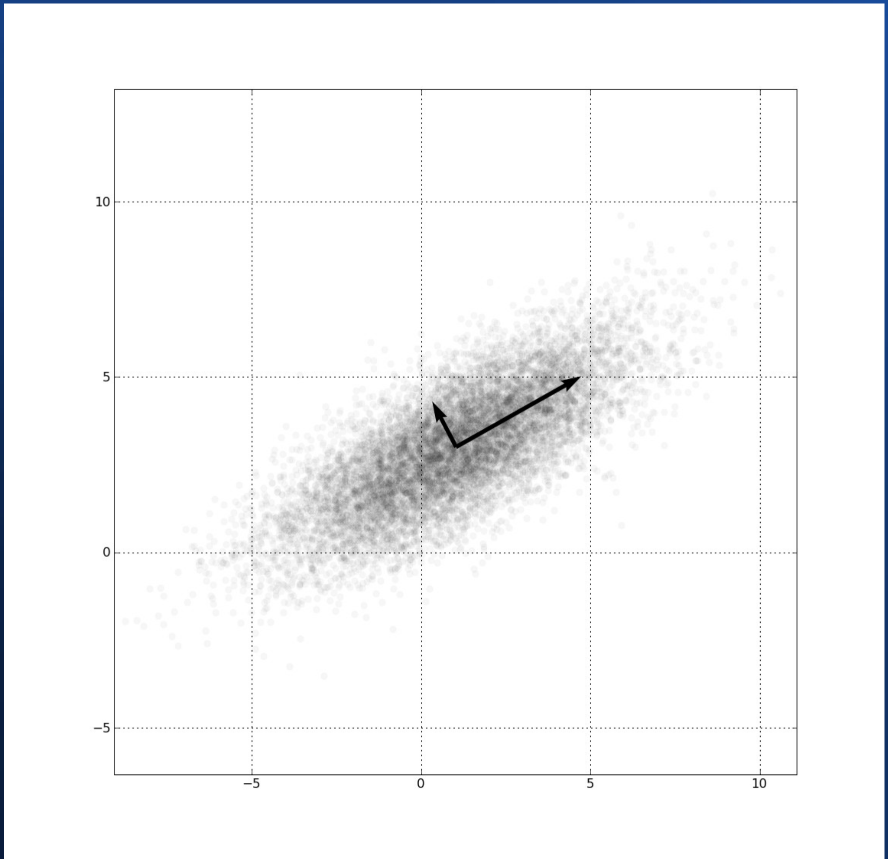
http://en.wikipedia.org/wiki/Sparse_coding

Principal Component Analysis (1)

- Principal component analysis (PCA) is a mathematical procedure that uses orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

http://en.wikipedia.org/wiki/Principal_component_analysis

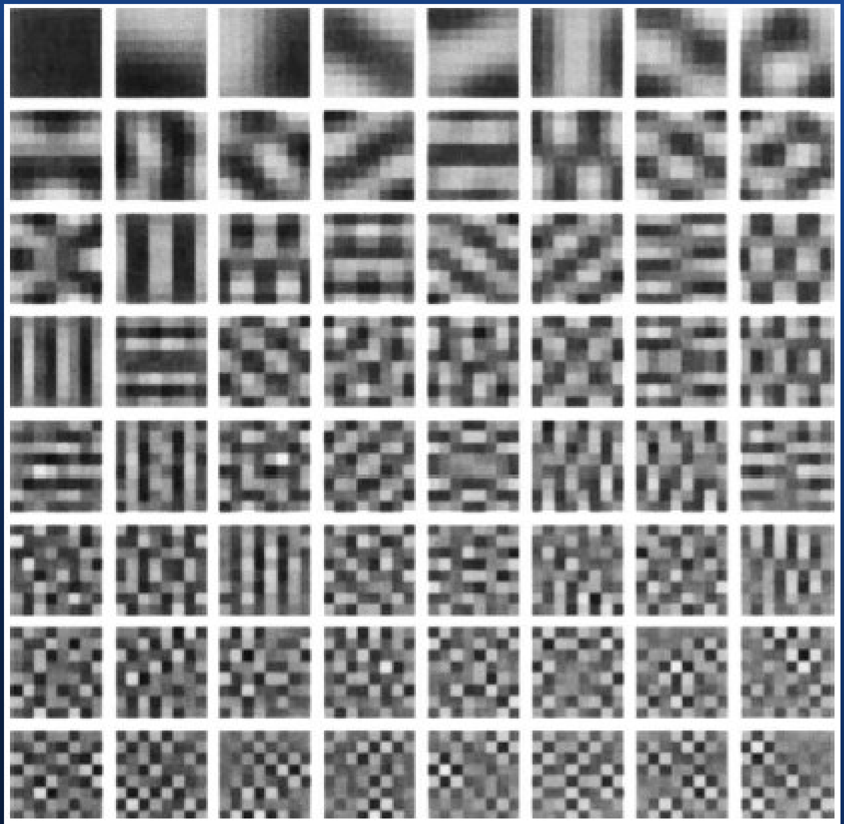
Principal Component Analysis (2)



- Goal is to find a set of mutually orthogonal basis functions that capture the directions of maximum variance in the data and for which the coefficients are pairwise decorrelated

- PCA is appropriate for capturing the structure of data that are well described by a gaussian cloud, or in which the linear pairwise correlations are the most important form of statistical dependence in the data

Principal Component Analysis (3)



← Principal components calculated on 8 x 8 image patches extracted from natural scenes

The receptive fields are not localized, however, and the vast majority do not at all resemble any known cortical receptive fields

„Emergence of simple-cell receptive field properties by learning a sparse code for natural images“, B.A. Olshausen, D.J. Field

Hubel & Wiesel: simple cells

A simple cell in the primary visual cortex is a cell that responds primarily to oriented edges and gratings (bars of particular orientations). These cells were discovered by Torsten Wiesel and David Hubel in the late 1950s.

- D. H. Hubel and T. N. Wiesel Receptive Fields of Single Neurones in the Cat's Striate Cortex J. Physiol. pp. 574-591 (148) 1959
- D. H. Hubel and T. N. Wiesel Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex J. Physiol. 160 pp. 106-154 1962

Introducing sparsity

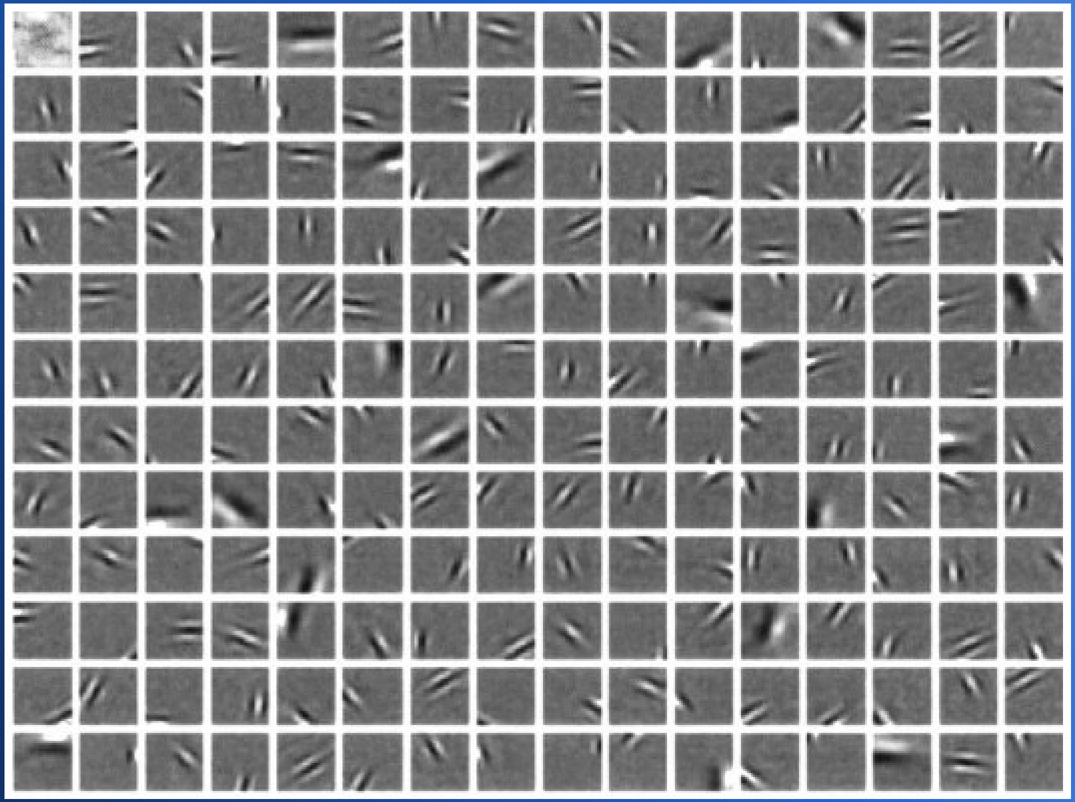
$$I(x,y) = \sum_i a_i \phi_i(x,y) \tag{1}$$

$$E = -[\text{preserve information}] - \lambda[\text{sparseness of } a_i] \tag{2}$$

$$[\text{preserve information}] = -\sum_{x,y}\left[I(x,y) - \sum_i a_i \phi_i(x,y)\right]^2 \tag{3}$$

The search for a sparse code as an optimization problem

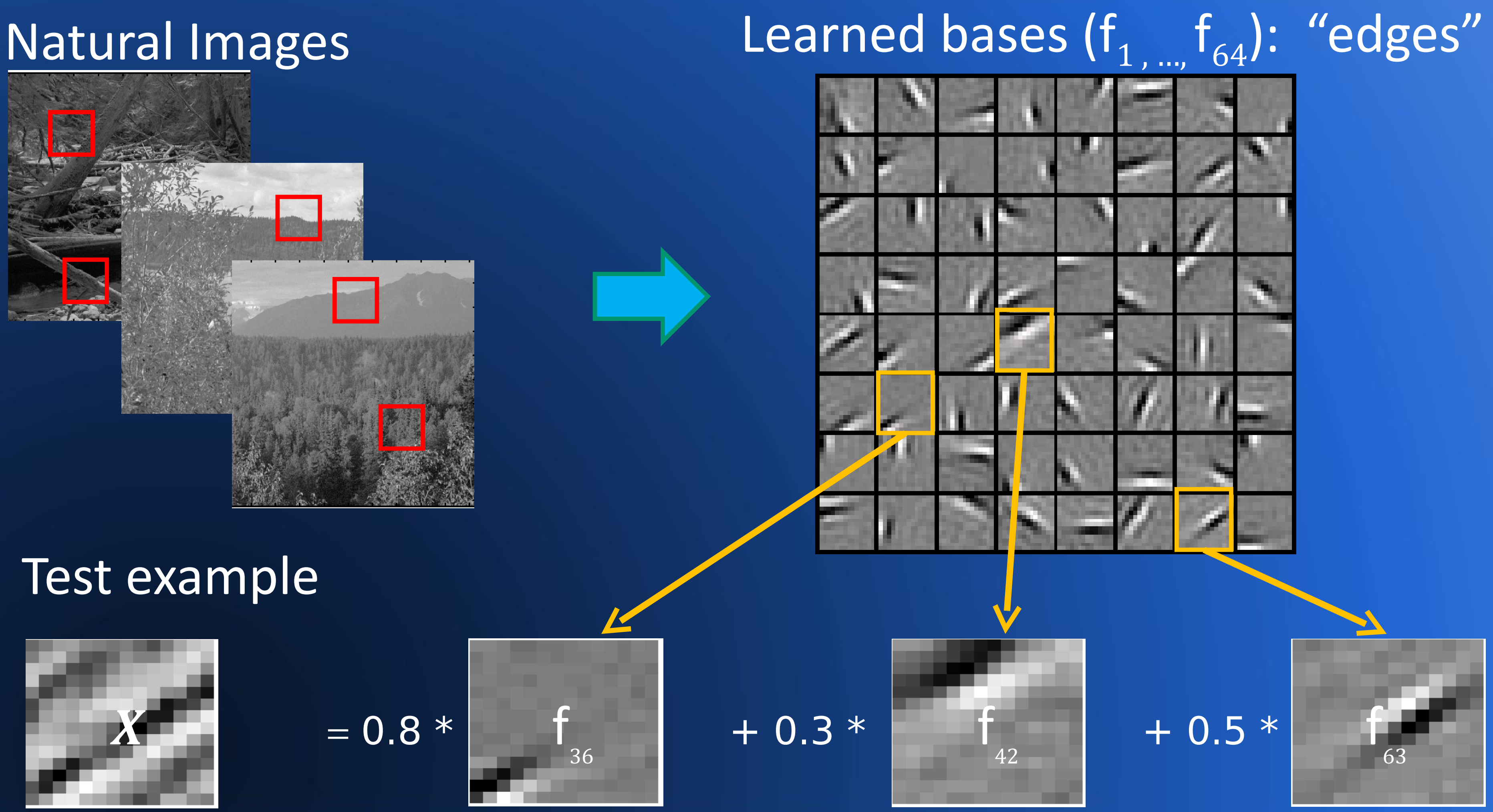
„Emergence of simple-cell receptive field properties by learning a sparse code for natural images“, B.A. Olshausen, D.J. Field



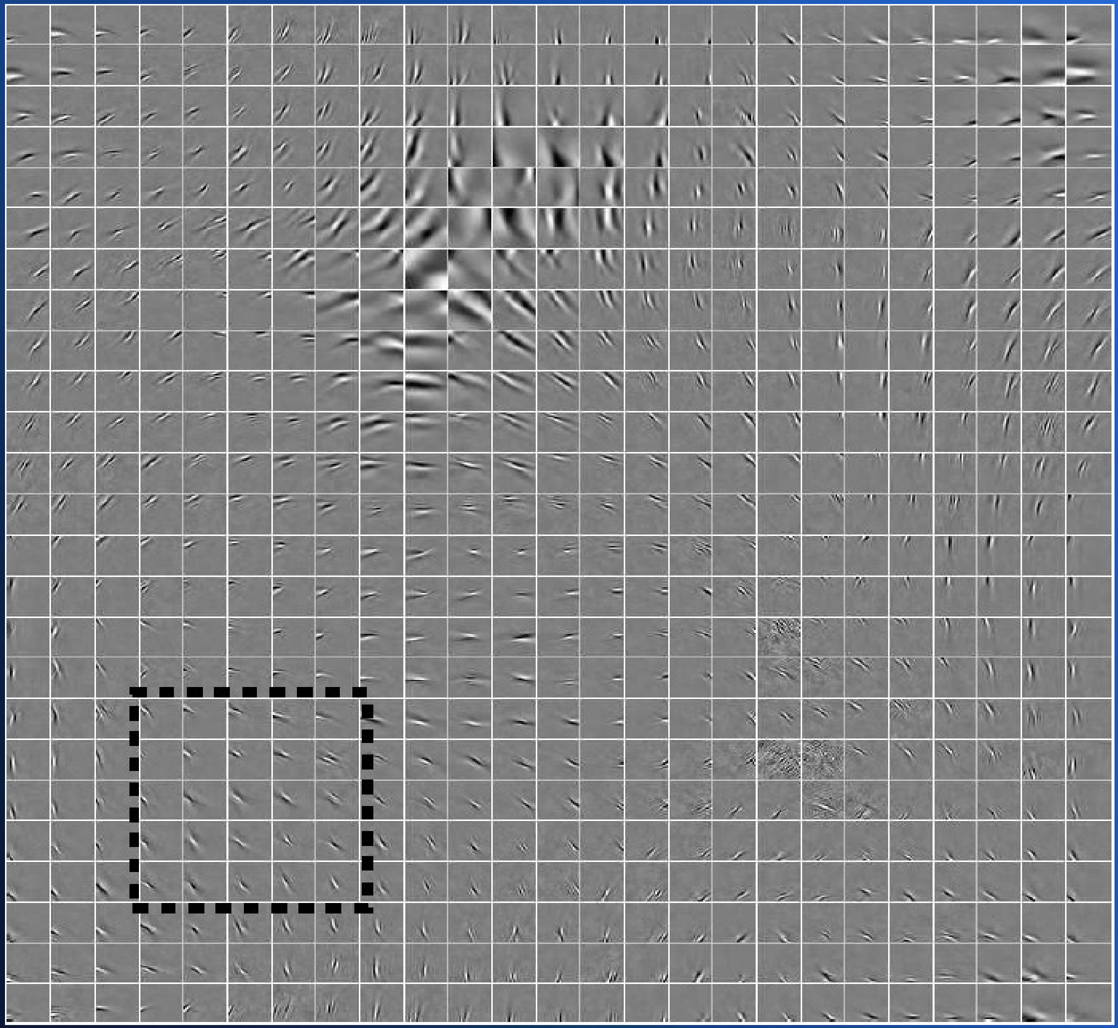
The result of training the system on 16 x 16 image patches extracted from natural scenes

Sparse coding software by B. Olshausen - Sparsenet @ http://redwood.berkeley.edu/bruno/sparsenet/

Sparse coding illustration

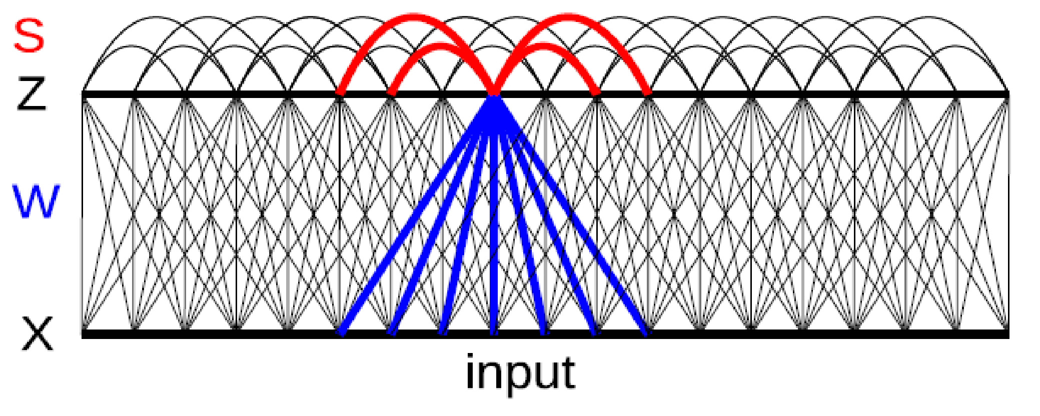


Topographic maps



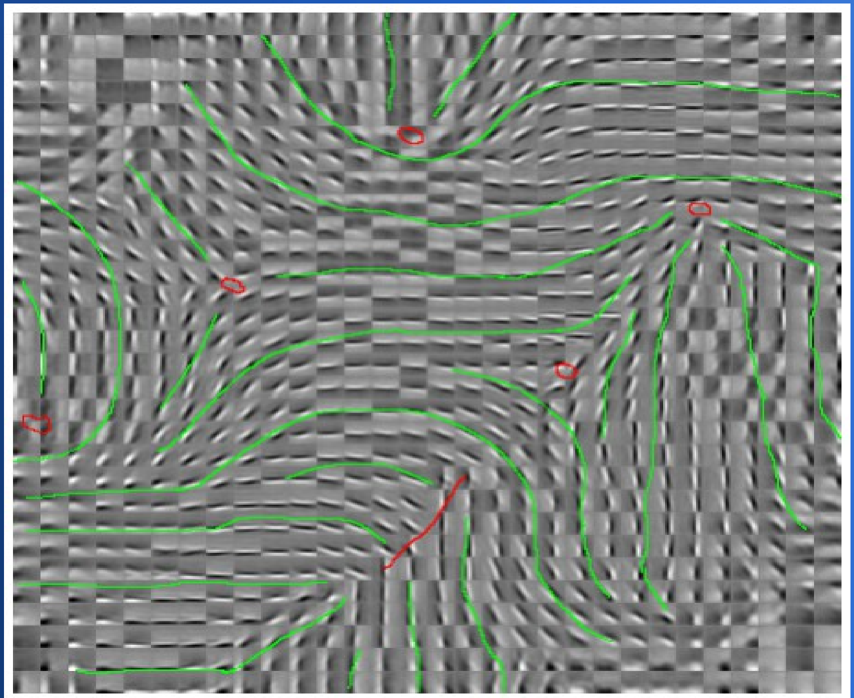
„A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images“
Aapo Hyvarinen, Patrik O. Hoyer

Structured sparse coding via lateral inhibition



It is sometimes appropriate to enforce more structure on Z than just sparsity.

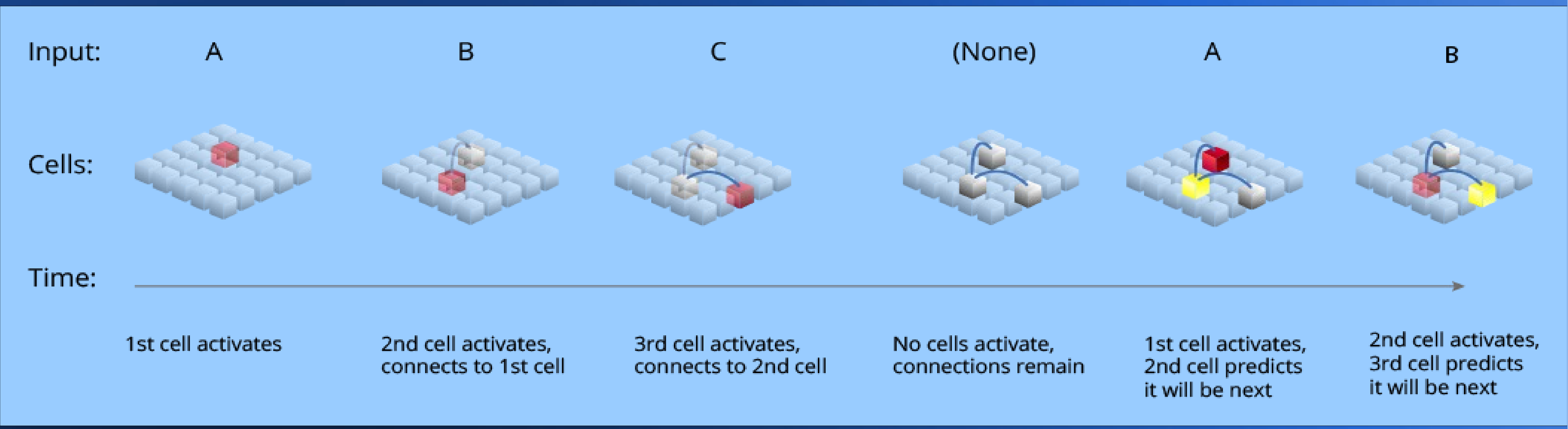
The result is a topographic map with pinwheel patterns similar to those observed in the visual cortex.



- „Structured sparse coding via lateral inhibition“, K. Gregor, A. Szlam, Y. LeCun

- http://www.cs.nyu.edu/~yann/research/sparse/index.html

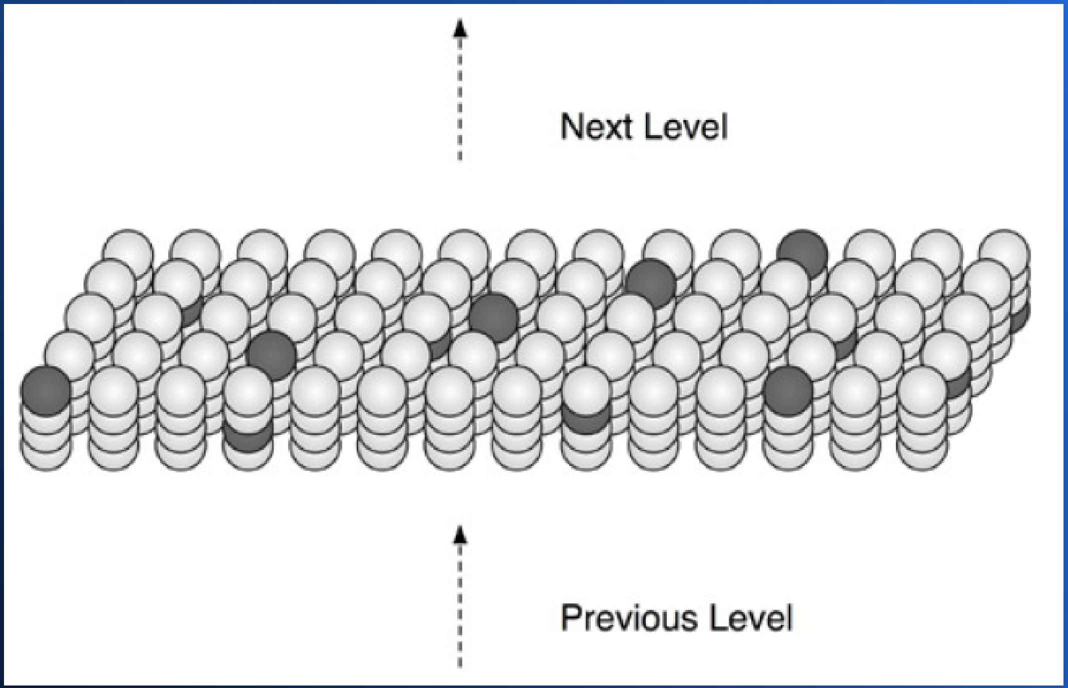
First-order sequence memory using HTM CLA



http://www.groksolutions.com/technology.html#

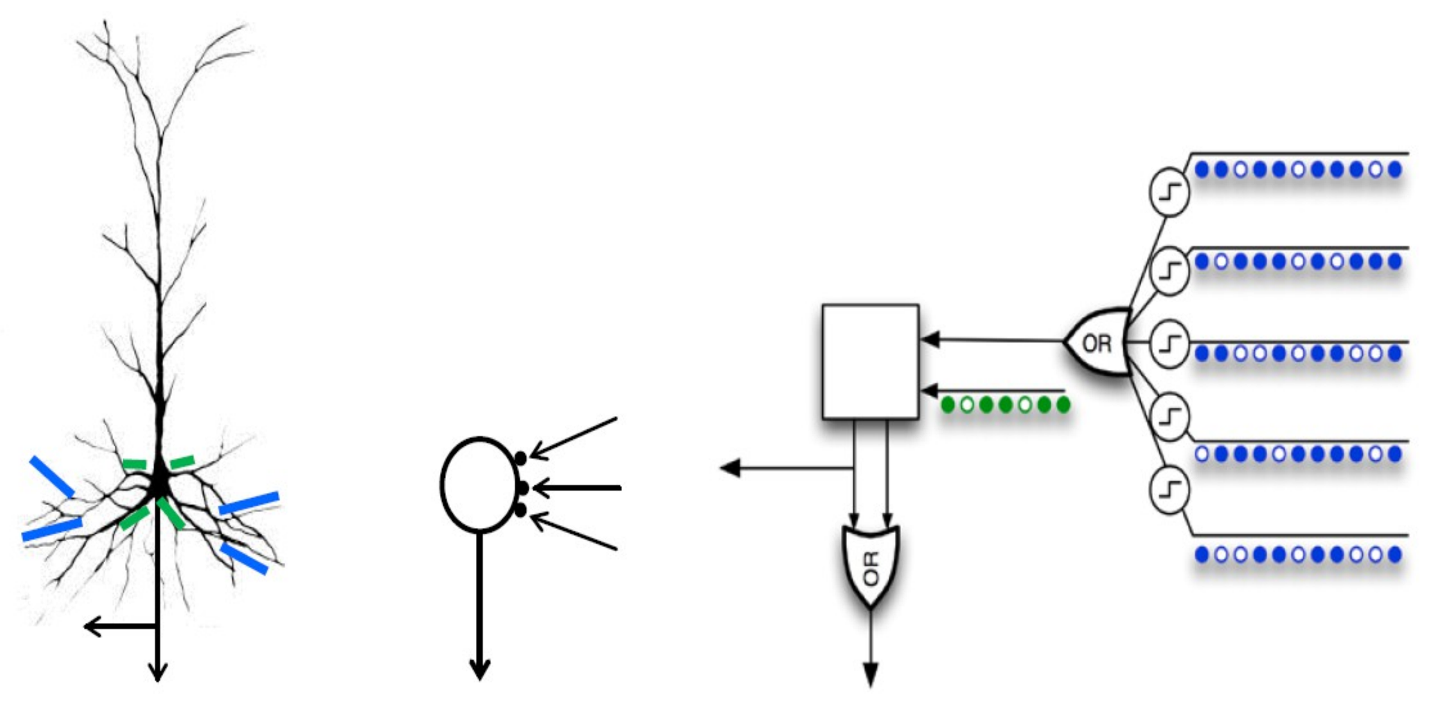
Hierarchical Temporal Memory: cortical learning algorithm

Variable-order sequence memory using HTM CLA



Use multiple cells per column to capture temporal context

Neurons

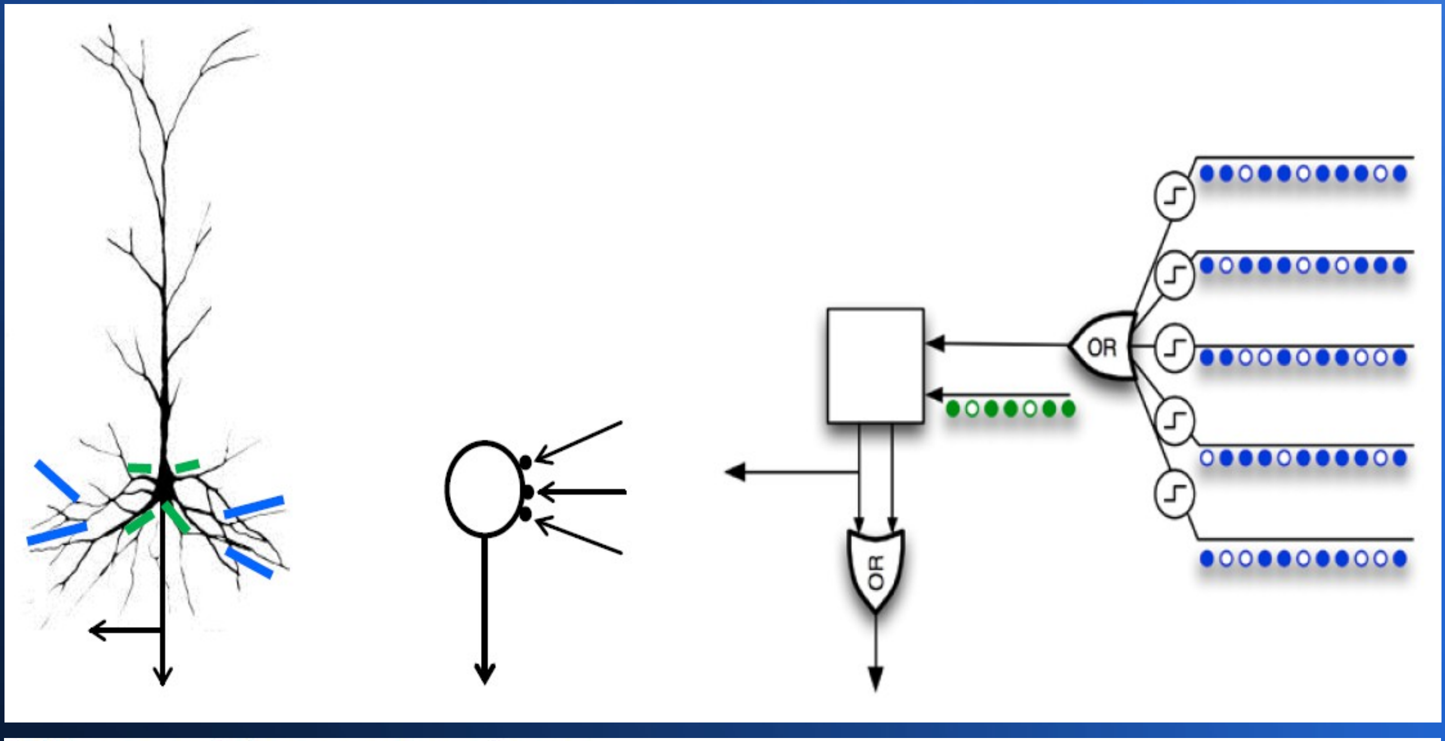


Real neuron

ANN neuron
(computation)

HTM neuron
(coincidence detection)

Neurons



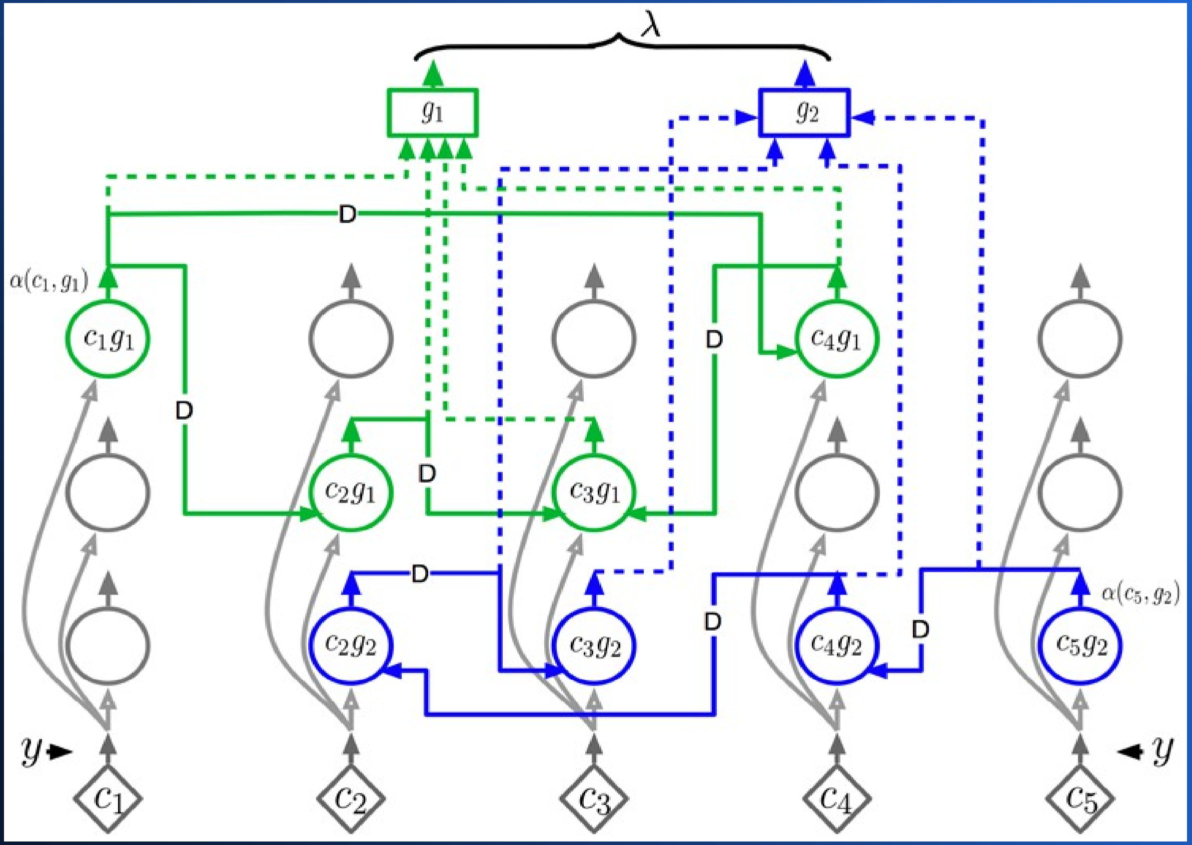
The cable equation

$$\lambda^2 \frac{\partial^2 V_m(x,t)}{\partial x^2} = \tau_m \frac{\partial V_m(x,t)}{\partial t} + \left(V_m(x,t) - E_m\right) - r_m I_{inject}(x,t)$$

W. Rall model - http://www.scholarpedia.org/article/Rall_model

Sequence formation

- Define sequences using graph partitioning algorithm

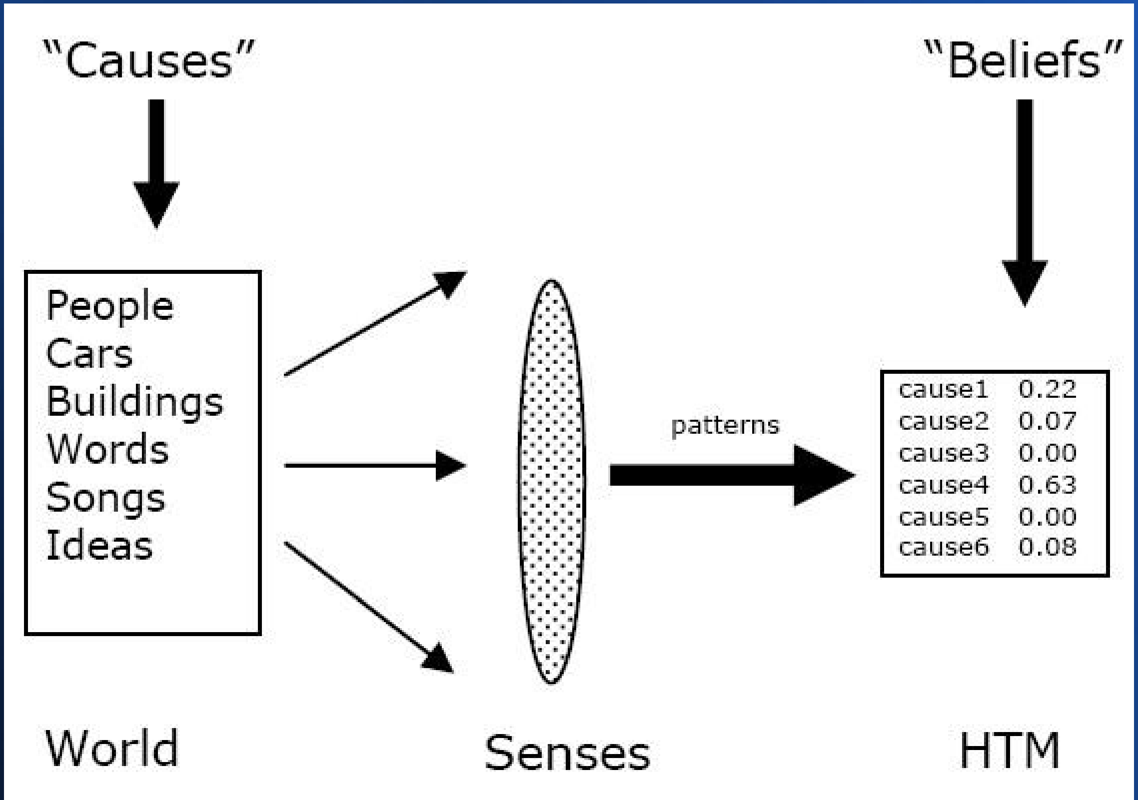


* „Towards a Mathematical Theory of Cortical Micro-circuits“, G. Dileep, J . Hawkins

Spatio-temporal pooling

1. **Spatial pooling**
2. **Temporal pooling with timing**
   - Used for behaviour
   - Based on (1)
3. **Temporal pooling without timing**
   - Used for invariant form representation development
   - Based on (1) and (2)
4. **Temporal pooling without context**
   - Used for invariant motion representation development
   - Based on (1) and (3)

How it works?



* Numenta HTM Concepts

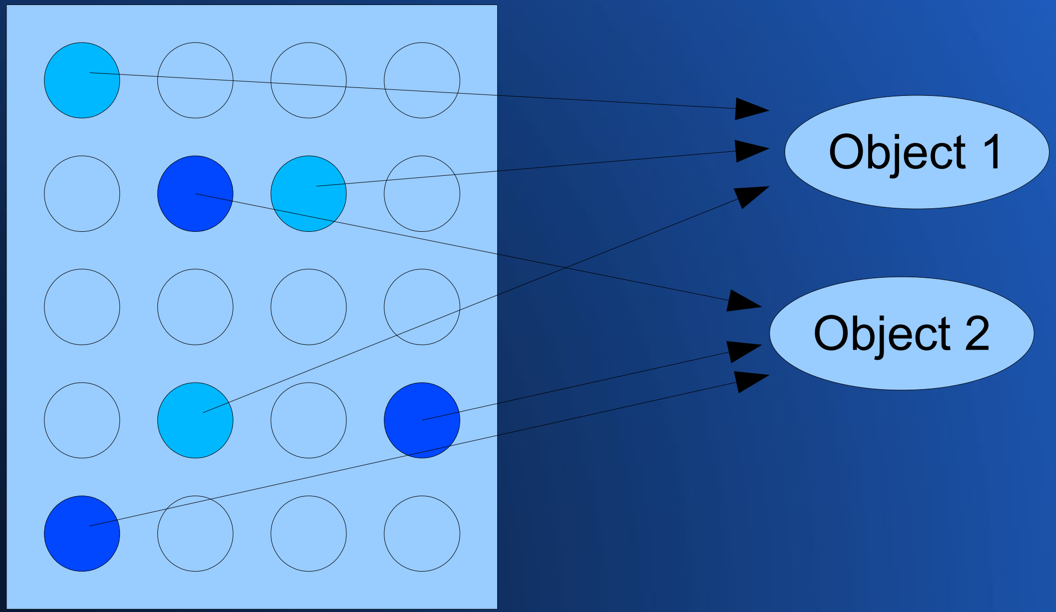
- Calculate spatio-temporal pattern probabilities

- Perform inference (probabilistic influence flow)

- Recognize objects

Recognition

- Make/use <Random variables>-to-<object> mapping using supervised learning



Kopsavilkums

- Lai buvēt īsti intelektuālās mašīnas ir nepieciešams ņemt vērā neokorteksa uzbūves un darbības principus
- „Atmiņas-paredzēšanas“ intelekta teorija ir labākais sākumpunkts
- Vienotā algoritma mērķis ir invariantas reprezentācijas veidošana, pielietojot efektīvas kodēšanas principus

Nākotnes plāni

- Kopējas arhitektūras veidošana
    - "Spatio-temporal pooler" veidošana
    - Integrācija pēc PGM principa
    - Reprezentācijas dekodēšana
- Pielietošana konkrētam uzdevumam
    - Vidējas sarežģītības pakāpes objektu atpazīšana

Paldies!