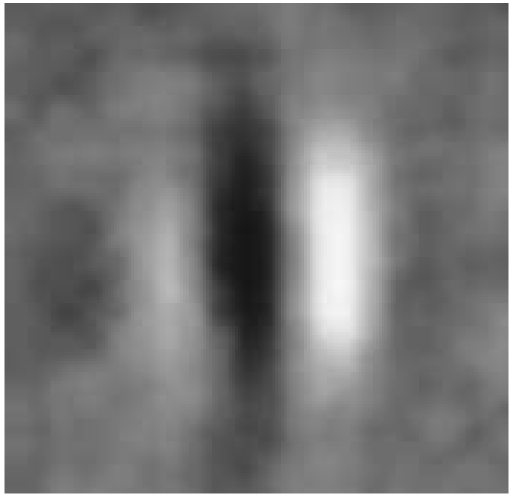
Natural image statistics and cortical visual processing

Aapo Hyvärinen
Dept of Mathematics and Statistics & Dept of Computer Science
University of Helsinki

Ecological approach to receptive fields

- Why are the receptive fields in visual cortex the way they are?



Classical theories of receptive fields

- Edge detection
- Joint localization in space and frequency
- Texture classification
- But: these give only vague predictions.

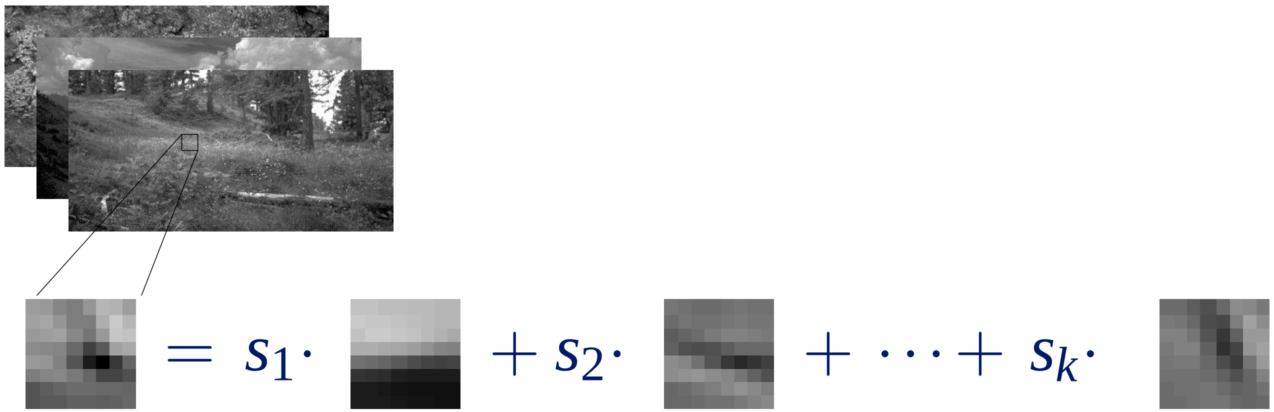
Statistical-ecological approach

- What is important in a real environment?
- Natural images have statistical regularities.
- Can we "explain" receptive fields by basic statistical properties of natural images?
- Emergence : a lot of precise predictions from only a couple statistical assumptions.

Outline of this talk:

- Statistical models that account for some properties of the (primary) visual cortex.
  - simple cells
  - complex cells
  - topography (spatial organization)
- Multi-layer approach can predict properties beyond V1.

Linear statistical models of images



- Denote by $I(x,y)$ the gray-scale values of pixels.
- Model as a linear sum of basis vectors:

$$I(x,y) = \sum_i A_i(x,y) s_i \tag{1}$$

- What are the "best" basis vectors for natural images?

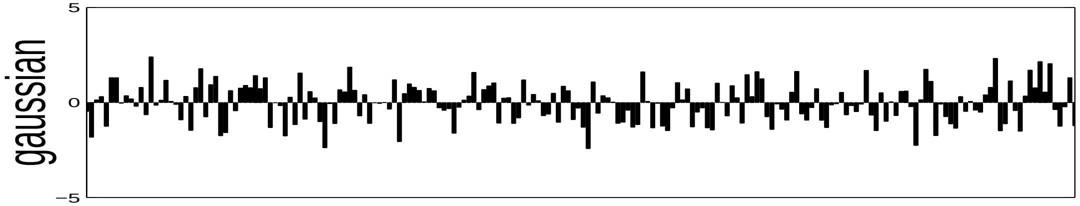
Independent Component Analysis (Jutten and Hérault, 1991)

- In ICA, we assume that
  - The s_i are mutually statistically independent
  - The s_i are nongaussian, e.g. sparse
  - For simplicity: the number of basis vectors equals the number of pixels
- Then, the actual basis vectors can be estimated, if the data is actually generated using the linear model (Comon, 1994).
- Thus we get the best basis vectors from one statistical viewpoint.
- Inverting the system: $s_i = \sum_{x,y} W_i(x,y) I(x,y)$, we see that the s_i are linear filter (simple cell) outputs.

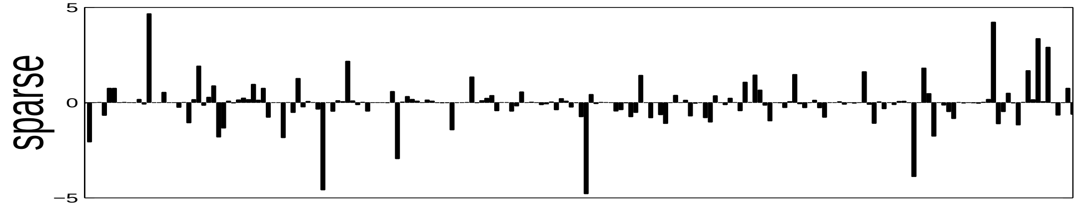
Sparseness

- A form of nongaussianity often encountered in natural signals
- A random variable is "active" only rarely

gaussian:



sparse:



- Outputs of linear filters are usually sparse when input is natural images.

Sparse coding and ICA

- Sparse coding: Find linear representation

$$I(x,y) = \sum_i A_i(x,y) s_i \tag{2}$$

  so that the s_i are as sparse as possible.

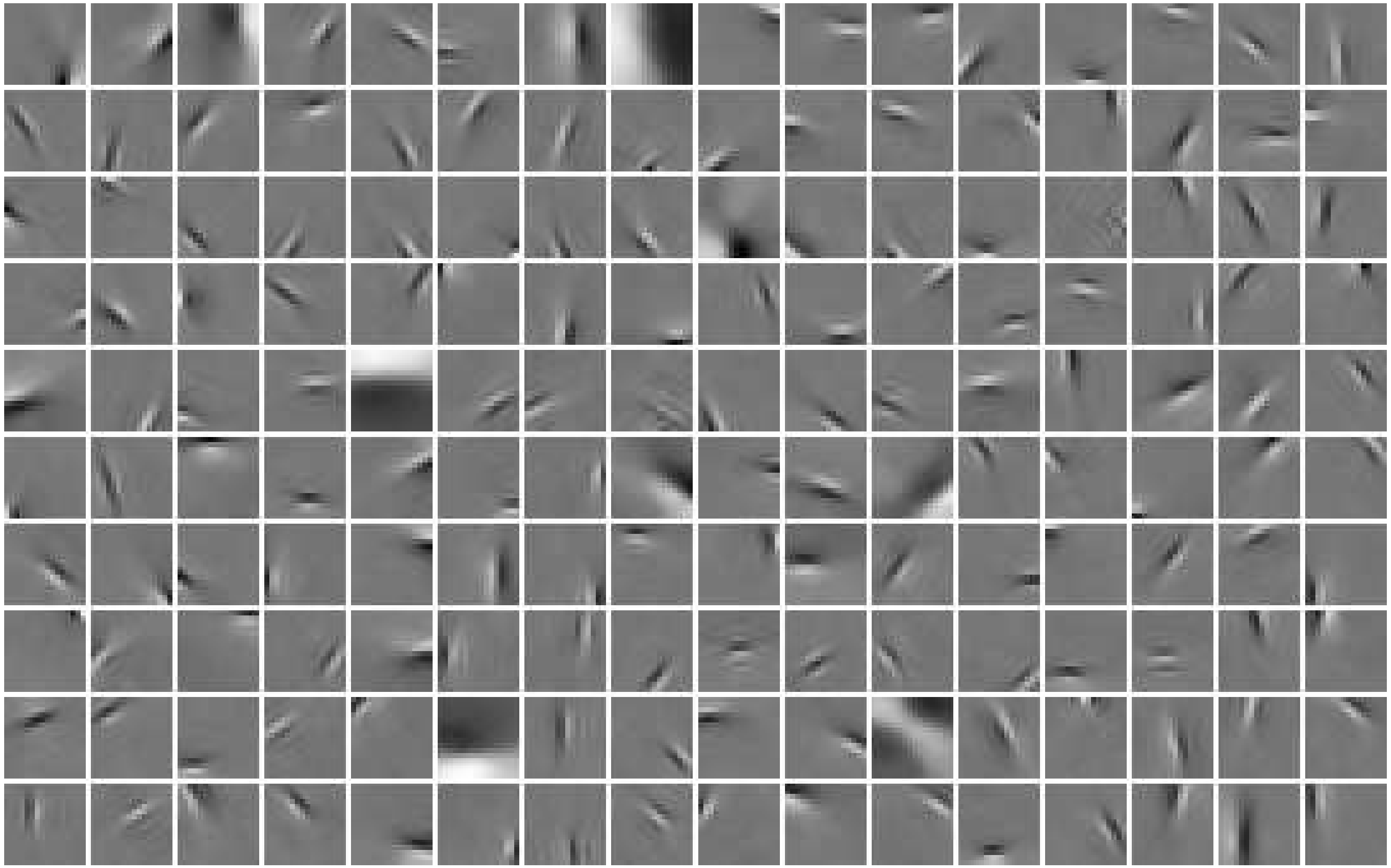
- Important property: a given data point is represented using only a limited number of “active” (clearly non-zero) components s_i.

- In contrast to PCA, active components change from image patch to patch.

- Deep result: For images, ICA is sparse coding.

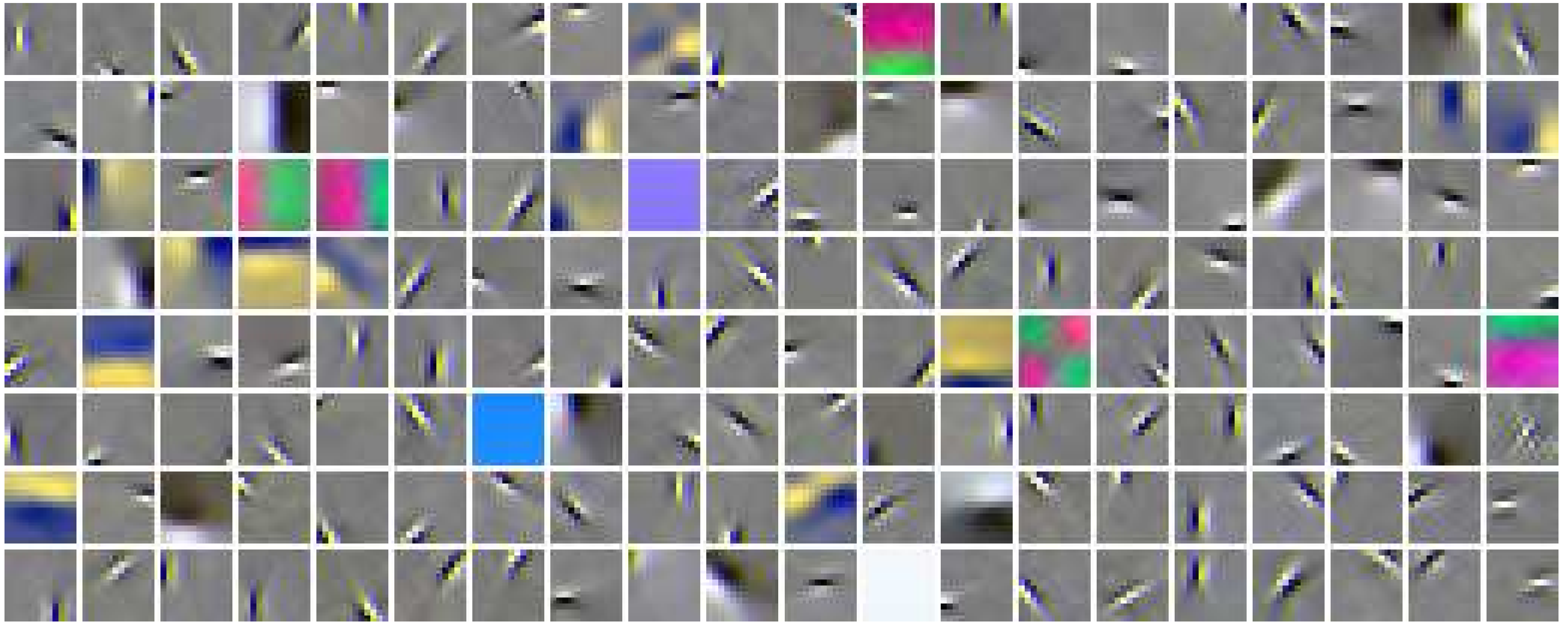
ICA / sparse coding of natural images

(Olshausen and Field, 1996; Bell and Sejnowski, 1997)



ICA of natural images with colour

(Hoyer and Hyvärinen, 2000)



Model II: Independent subspace analysis

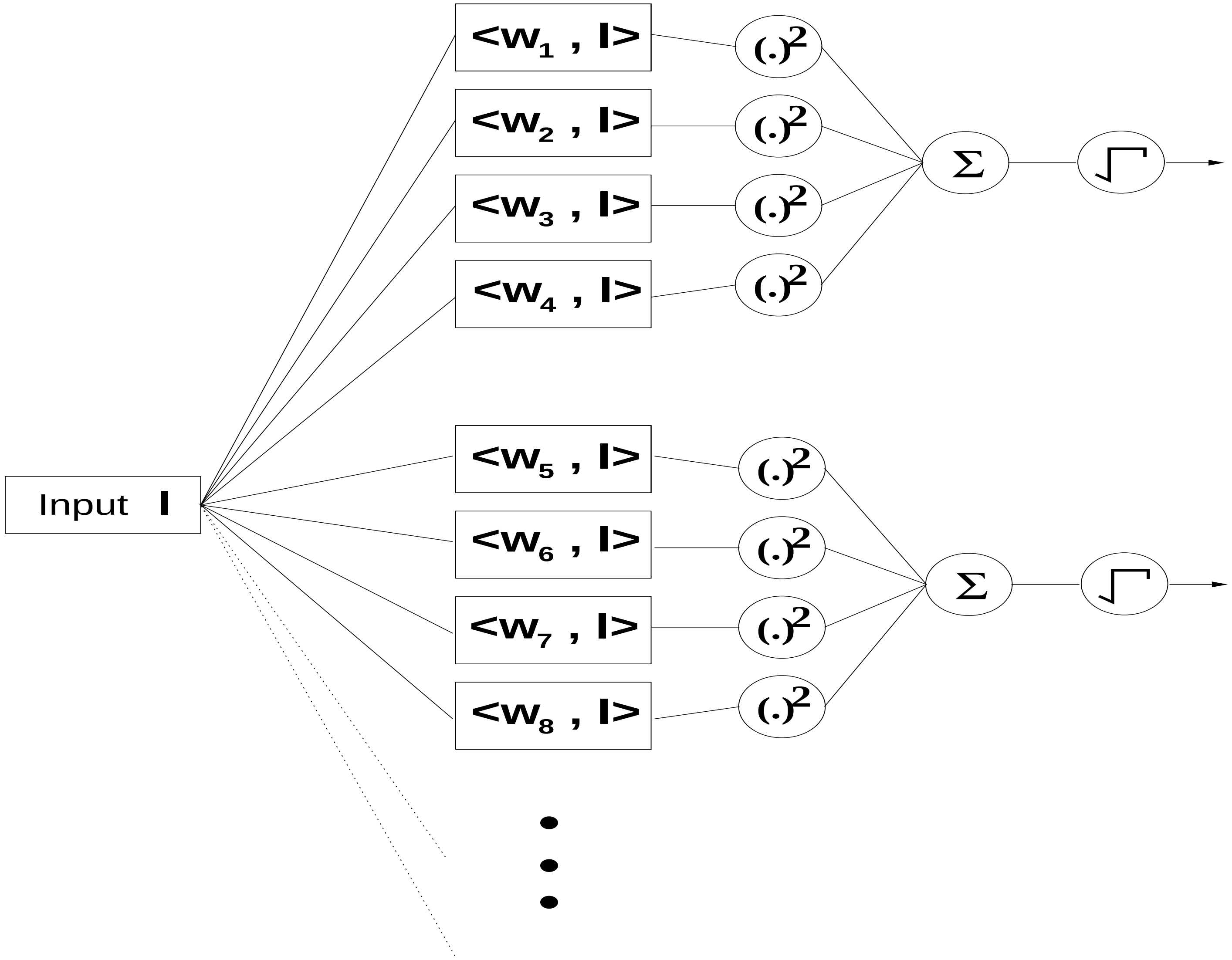
- Components estimated from natural images are not really independent.
- The statistical structure much more complicated (of course!).
- In fact, independent components cannot be found for most kinds of data: There are not enough free parameters.
- Next, we model some dependencies of simple cell (linear filter) outputs.
- This leads to a model of complex cell receptive fields: insensitivity to phase of input.

Independent subspaces

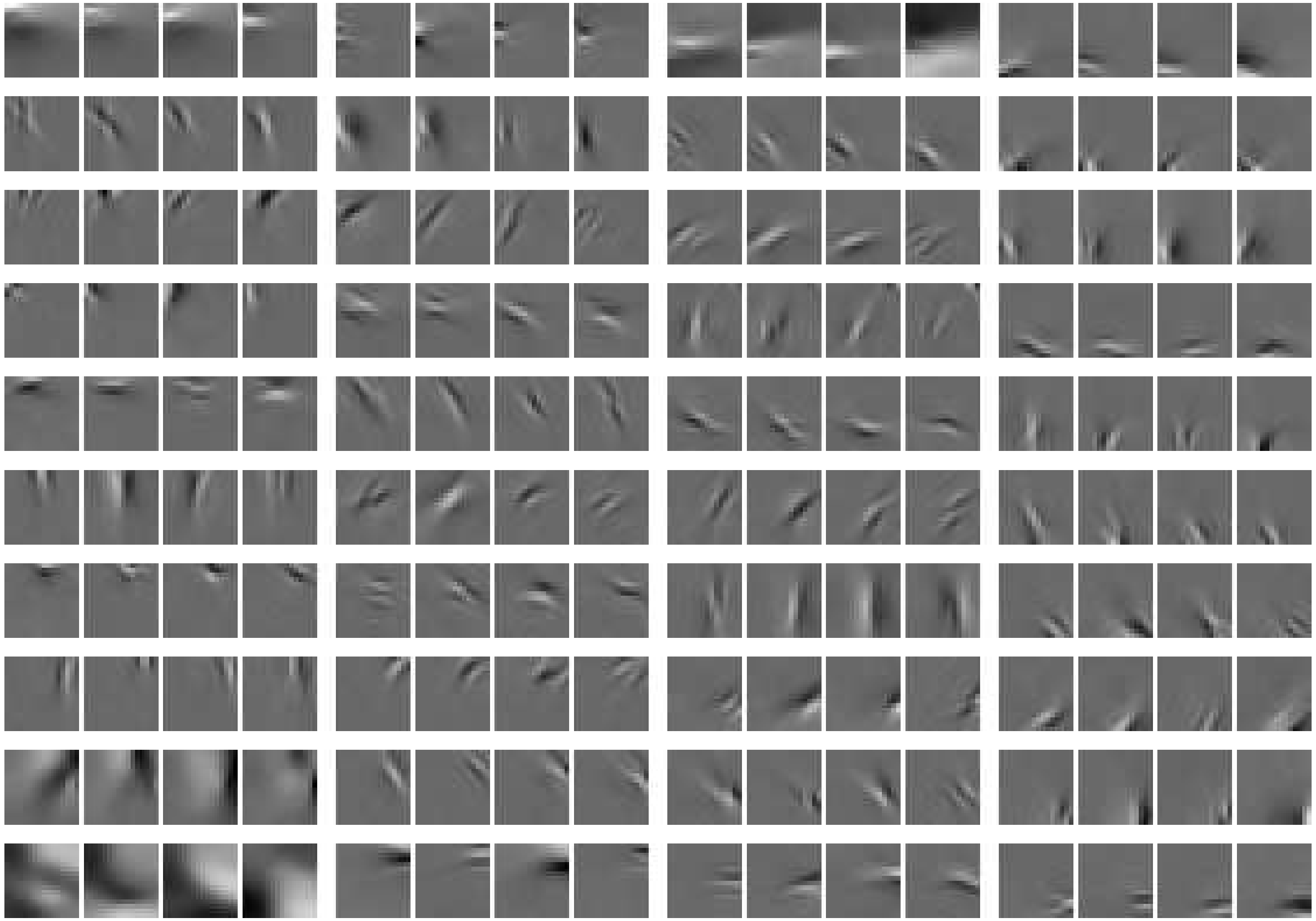
(Hyvärinen and Hoyer, 2000)

- A very basic approach to modelling dependencies.
- Assumption: the s_i can be divided into groups or subspaces, such that
  - the s_i in the same group are dependent on each other
  - dependencies between different groups are not allowed.
- We also need to specify the distributions inside the groups
  $\Rightarrow$ Inspiration from energy pooling models.

Energy pooling inside groups



Independent subspaces of natural image patches



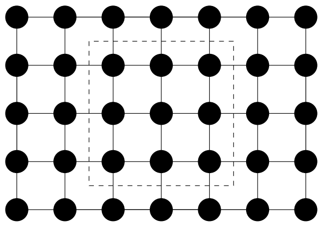
Each group of 4 basis vectors corresponds to one complex cell.

Model III:
Spatial organization (topography) in V1

- Receptive field properties mostly change continuosly when moving on the cortical surface.
- Retinotopy: localization changes smoothly.
- Orientation changes smoothly except in “pinwheels” (Bonhoeffer and Grinvald, 1991; Blasdel, 1992).
- There are low-frequency regions, possibly co-incident with CO blobs (Tootell et al 1988; Edwards et al, 1995).
- Phase changes randomly (DeAngelis et al, 1999).
- Original inspiration for the Kohonen Map (Kohonen, 1982).

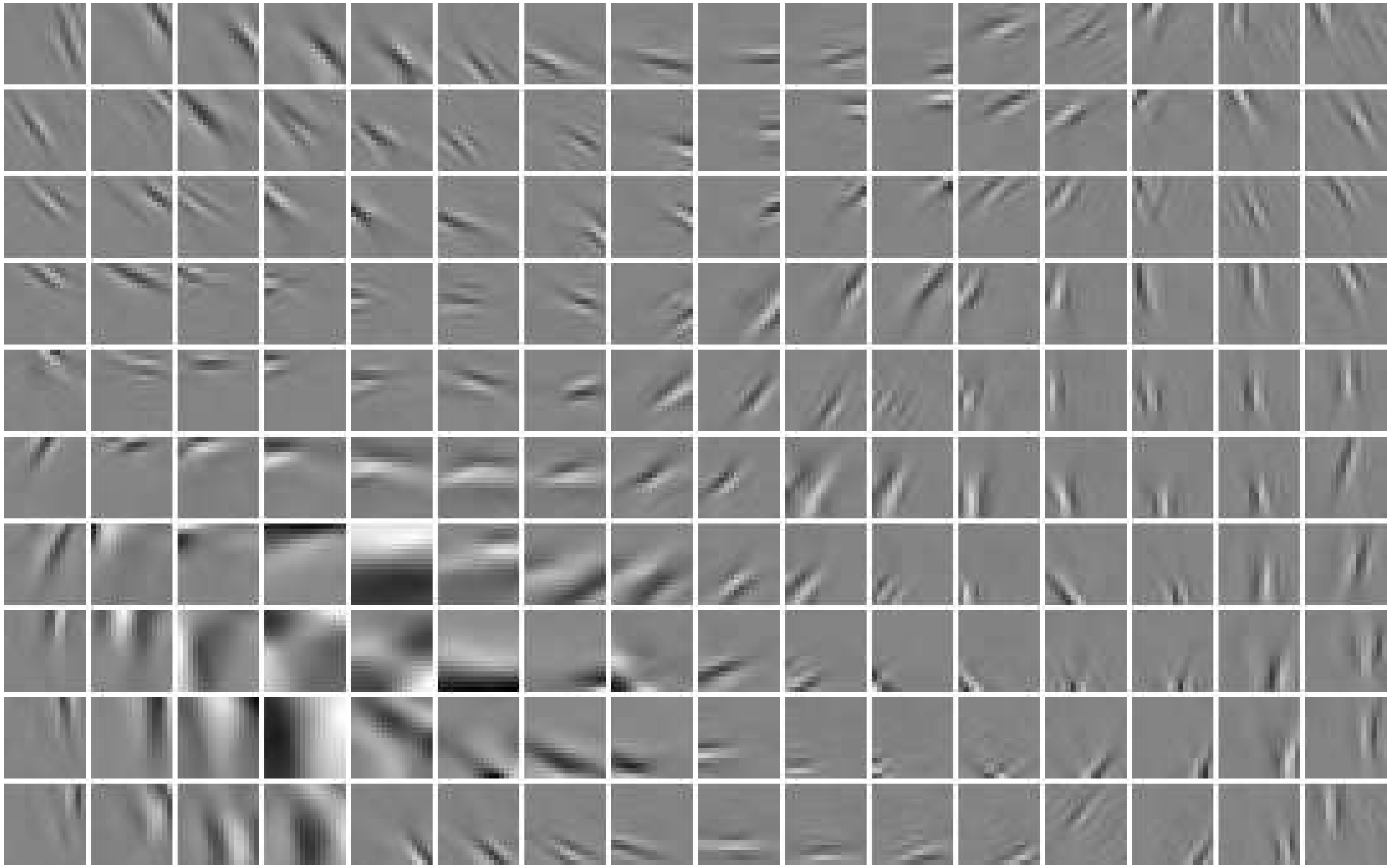
Topographic ICA (Hyvärinen and Hoyer, 2001)

- Cells (components) are arranged on a two-dimensional lattice



- Again, simple cell outputs are sparse, but not independent.
- Statistical dependency of components follows topography.

Topographic ICA on natural image patches



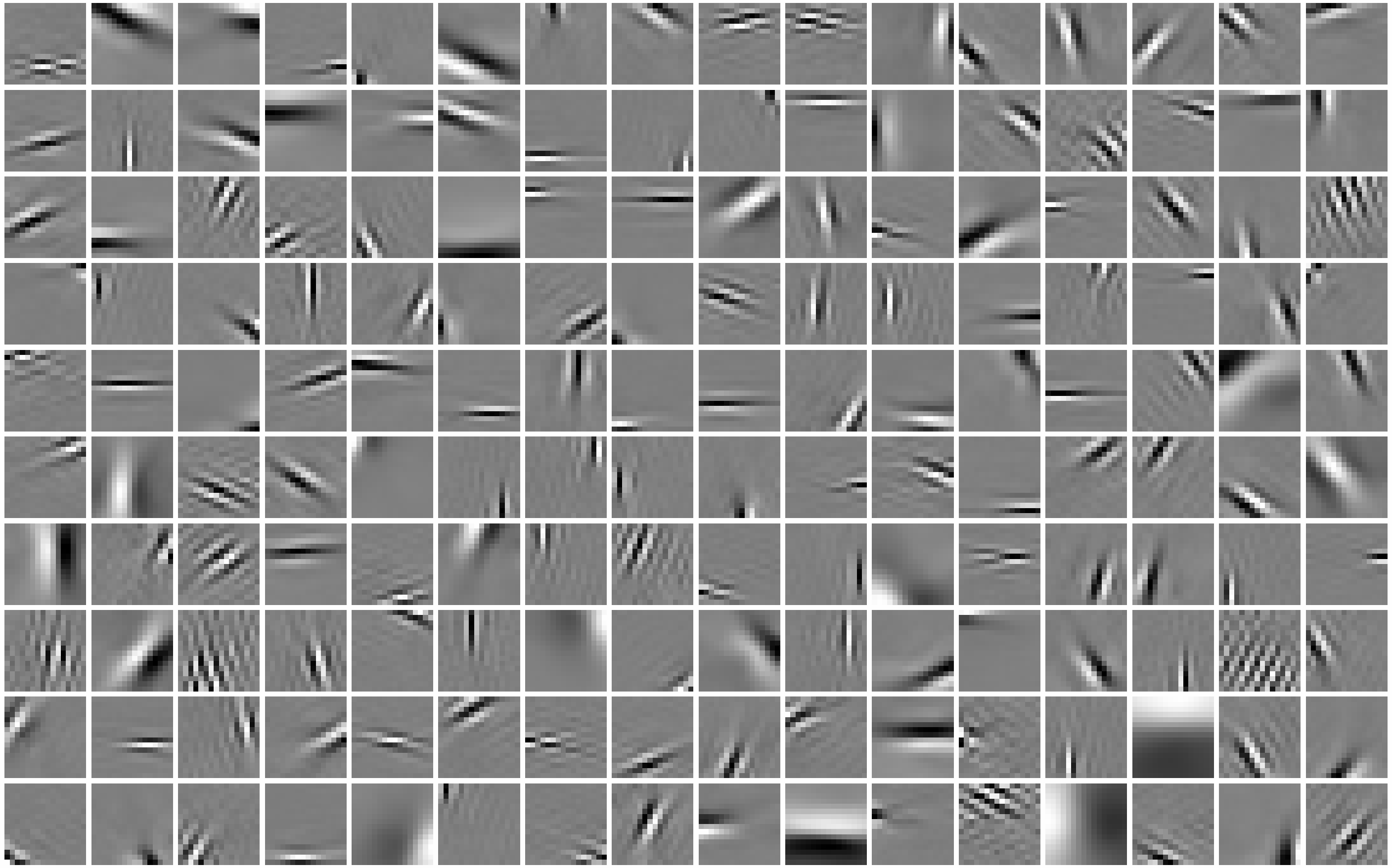
Basic vectors (simple cell RF's) with spatial organization

Model IV: Temporal coherence of simple cell outputs

- In image sequences (video) we can look at the temporal correlations.
- An alternative to sparseness.
- Look at the dependencies of $s_i(t)$ and a lagged version $s_i(t-\Delta t)$.
- Using linear correlations gives only Fourier-like receptive fields.
- We propose: Maximize correlation between $s_i^2(t)$ and $s_i^2(t-\Delta t)$.

Temporal coherence results on natural images

(Hurri and Hyvärinen, 2003)



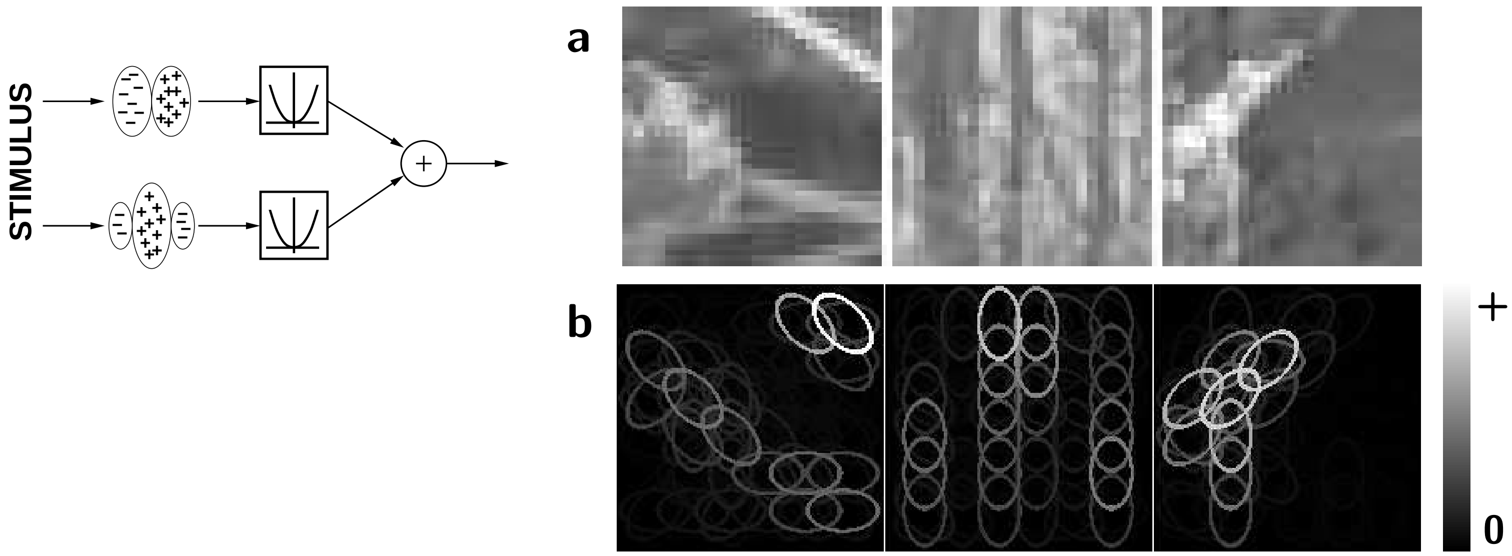
Spatial basis vectors estimated from image sequences.

Beyond the primary visual cortex

- What the next stage of processing be like?
- To predict this, we can perform ICA on complex cell outputs

Model V: ICA on complex cell outputs

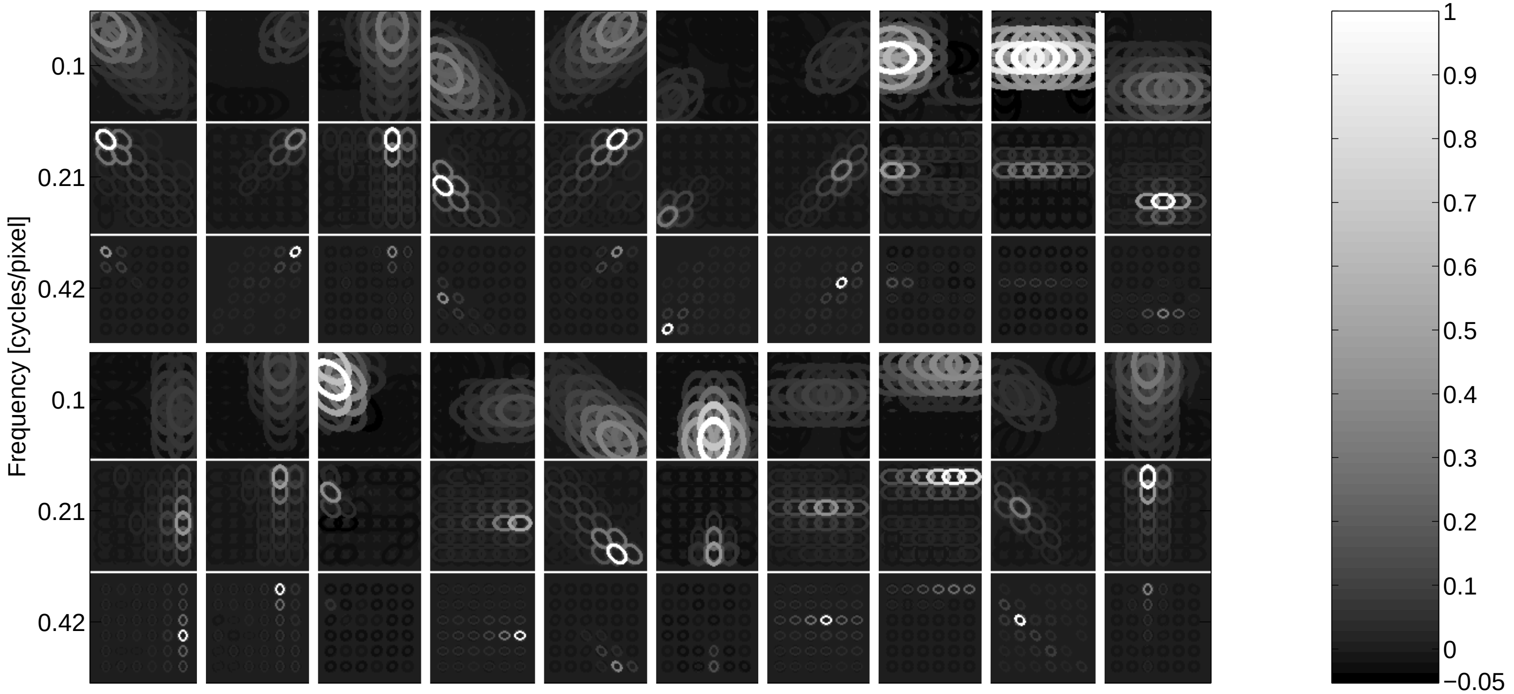
- Compute complex cell outputs for natural images



- Do ICA on this complex cell output data.

ICA on complex cell outputs

(Hyvärinen, Gutmann, Hoyer, 2005)

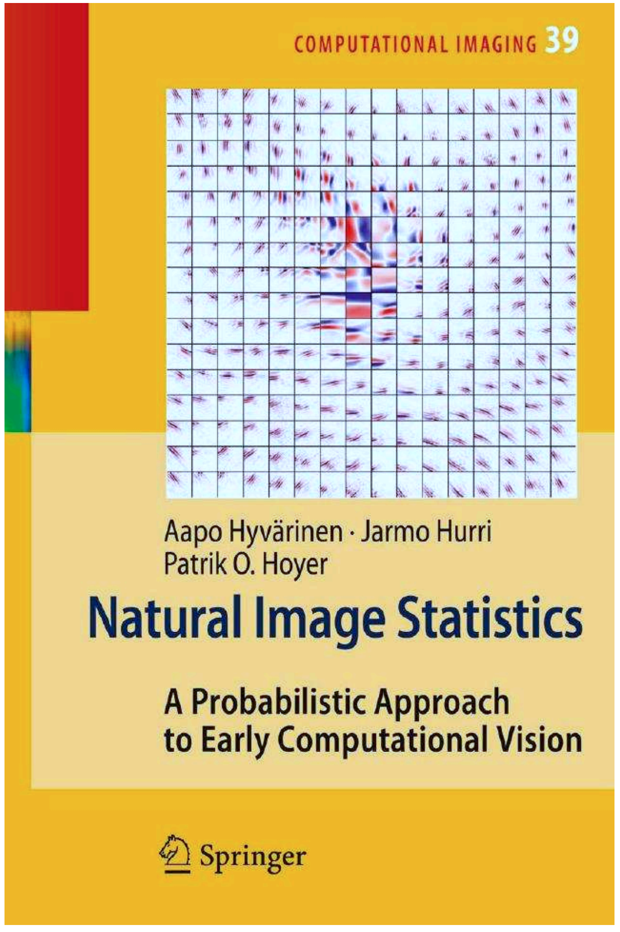


Each higher-order cell corresponds to 3 frequency displays

Emergence of contours and pooling over frequencies

- Elongated contour units
- Classic view emphasizes separate frequency channels: here we have pooling of frequency channels
- An example of predictive modelling

For more information:



Conclusion

- Properties of visual neurons can be quantitatively modelled by statistical properties of natural images.
- Simple cell receptive fields can be learned by maximizing independence / sparseness.
- By modelling dependencies between simple cell ouputs we can model complex cells and topography.
- Instead of sparseness, temporal coherence can be used.
- Modelling complex cell outputs yields frequency-pooling contour coding units.
- Many more models can be built and properties predicted using this approach.