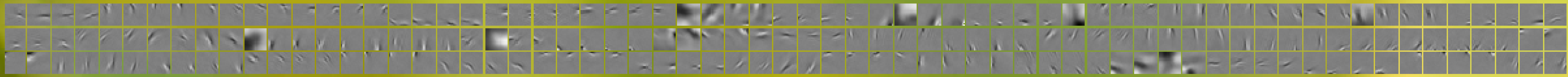
From Natural Images to Visual Neurons

Urs Köster
Helsinki Institute for Information Technology, Department of Computer Science University of Helsinki, Finland koster@cs.helsinki.fi



Intro

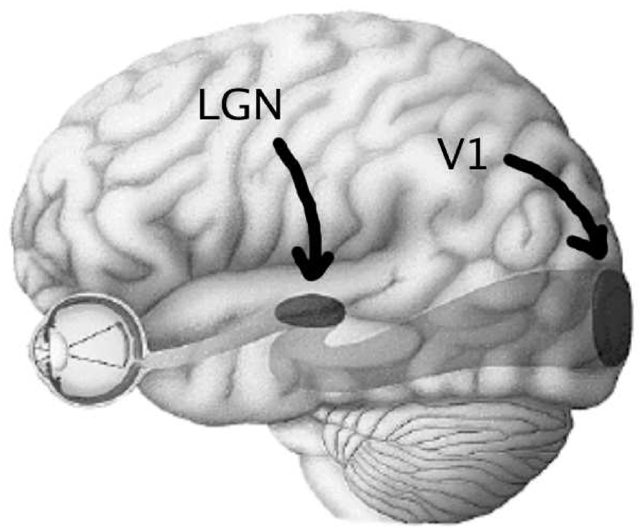
The visual cortex is the most "accessable" area of the brain, and has been studied extensively. Still it is also very complex and does remarkable things - it is far from being understood. It has been studied in countless ways, three of which obvious when we look at it as a black box receiving some input, doing someting with it, and generating output:

- starting at goals or output (e.g. face detection) - computer vision
- starting at the system, recording neurons - "biological" neuroscience
- starting at the properties of the inputs - statistical modelling

I work on the third. The information contained in the light hitting the retina has to be coded in an efficient way to make best use of neural resources. This gets us to **information theory** and therefore statistics. By having a model of the environment, coded into the neural network, the amount of data that needs to be transported and processed can be greatly reduced.
As an additional advantage, the network can use it's knowledge about the statistics of the environment to make intelligent guesses (**Bayesian Inference**) when it is missing information.
In ANN (artificial neural network) research, **unsupervised learning** is used to generate statistical models like that. The ANN does not map inputs to given outputs, but forms an **internal representation** of the statistical structure.
This internal representation can be linked to the **receptive fields** (RF's) of neurons (Hubel and Wiesel described visual RF's in 1963). Each neuron corresponds to some meaningful feature in the image. For early vision (V1, V2) the features should be related to the properties of the input. My goal is to explain the properties of visual RF's in terms of image statistics. So far, experimental evidence has preceded models, but since the interpretation of physiological experiments past V1 is so hard, this might change soon!

Neural processing

The visual processing pathway starts at the **retina**, where adaptation and lateral inhibition normalize contrast and brightness of the signal. Then the **lateral geniculate nucleus**, LGN with its center-surround receptive fields effectively decorrelates the data.

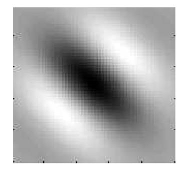


In **primary visual cortex**, cells can be classified into two important types:

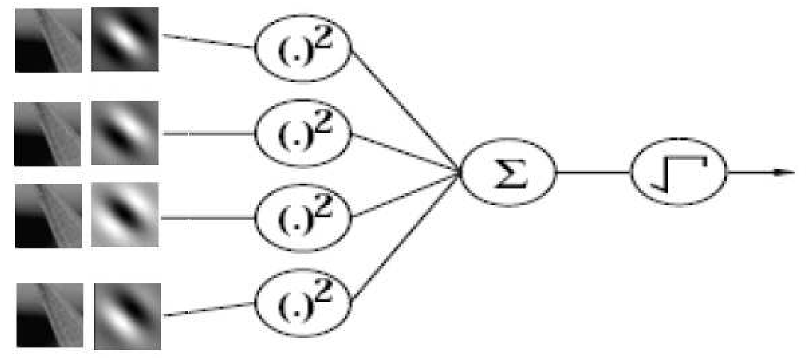
- **Simple cells**, which can be modelled as linear filters, which take the shape of Gabor wavelets, being oriented, bandpass and localized.
- **Complex cells**, which can be modelled by pooling a number of simple cells followed by a nonlinearity. They show somewhat nonlinear responses, e.g. phase invariance.

Simple and Complex Cells

This is what the receptive fields look like:



Simple Cell RF: This cell responds to a diagonal bar or grating stimulus.



Complex Cell RF: Pooling over different phases, this cell reacts to a bar or grating even if it is moving along and the phase is changing.

Modelling

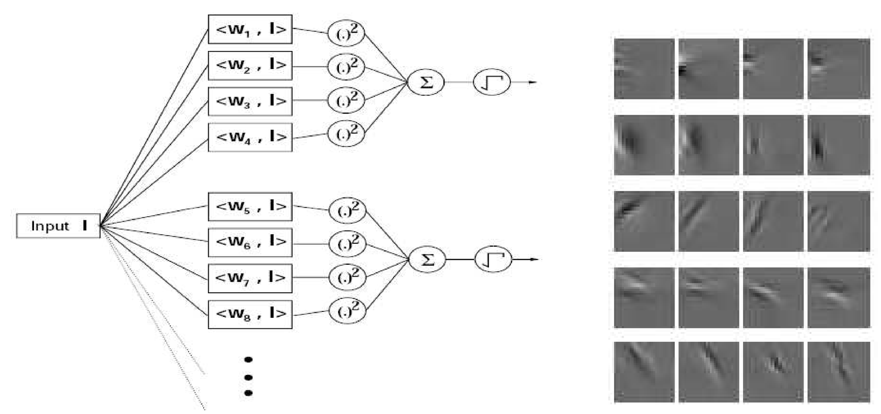
To extract the statistical structure of images to understand the neural processing, we can use **Independent Component Analysis (ICA)** It finds features in the data that provide a basis for reconstructing images, and has useful advantage:

- Only few neurons are active at any one time
- The code is robust and energy-efficient
- It is matched to the properties of images
- Similar to JPEG image compression

This is also called **Sparse Coding**.
ICA models image patches x as a linear superposition of features a, where the activations s are independent, so $x = As$.
ISA extends this with a second layer. The features are grouped into *subspaces* where dependencies are allowed. One subspace is $u = \left(\sum_i^n |s_i|^d\right)^a$, the nonlinearity corresponds to an L_p-norm. To estimate these models, adjust features and parameters to **maximize likelihood** of the model.

Independent Subspace Analysis

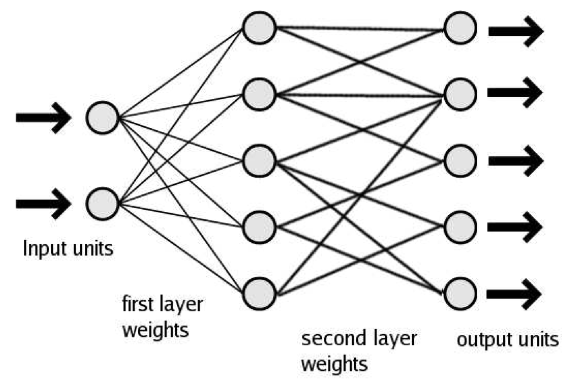
A possible explanation for complex cell properties is given by statistical optimality in the framework of *Independent Subspace Analysis* (ISA). The diagram shows how groups of simple cells, each receiving the same input, are pooled to give complex cells. The receptive fields shown are derived from ISA models and illustrate how simple cells with similar orientation and position are pooled to give phase-invariant complex cells.



I did some work on extending ISA (Independent Subspace Analysis, [4]) to L_p-ISA, that allows learning the optimal pooling for subspaces which are spherical under the L_p-norm. This is achieved by maximizing the likelihood of the model forcing a subgaussian probability distribution.

Score Matching

ISA is quite restricted, a full **two-layer model** would be more powerful. Estimating these is very hard in general. Our novel model, **Score Matching** uses a mathematical trick to make this estimation possible.



Currently I am working on such a two layer model, which is flexible enough to model V1 in more detail than ever. Eventually, this might be extended to more layers, so the hierarchy of V1 up to V2 and possibly beyond can be modelled.

Acknowledgments: Urs Köster is supported by the Alfried Krupp von Bohlen und Halbach-Foundation.