Slow Feature Analysis (SFA): Towards a generative model interpretation

Rich Turner (turner@gatsby.ucl.ac.uk)

Gatsby Unit, 16/08/2005

Motivation for Slow Features Analysis (SFA)

- Output from low-level sensors **varies quickly** as they respond to localised features
- Higher-level descriptions (eg. perception) varies on a **much longer time scale**
- Use temporal slowness to extract the higher-level descriptions
- cf. ICA where statistical independence of the features is the guiding principle, but which is often applied to temporal settings where correlations might also be sufficient to achieve blind source separation.

SFA

- Given a time-series of multidimensional observations: $\mathbf{y}_{1:T}$
- find a set of transformations $\hat{x}_{i,t} = g_i(\mathbf{y}_t)$ such that:

$$\Delta_i = \left\langle \frac{d\hat{x}_{i,t}}{dt}^2 \right\rangle \quad (1)$$

is minimal under the constraints:

$$\langle \hat{x}_{i,t} \rangle = 0 \quad (2)$$

$$\langle \hat{x}_{i,t} \hat{x}_{j,t} \rangle = \delta_{i,j} \quad (3)$$

SFA Intuition

- 1: minimise the the temporal variation of the output signal
- 2: mathematical covenience
- 3: given 2 avoids
  - trivial solution $\hat{x}_{i,t} = const$
  - each extracted feature being identical $\hat{x}_{i,t} = \hat{x}_{j,t}$

Implementing SFA

- It would be nice to find the non-linear functions automatically, but **calculus of variations** is too tough. Instead, use the usual trick of:
  - **Expanding the observations by passing them through a family of non-linearities** eg. all polynomials up to and including degree 2
  - **apply linear SFA to the expanded input** $\hat{x}_{i,t} = \mathbf{w}_i' \mathbf{y}_t$

- choosing the expanded obsevations to have zero mean, we integrate the remaining constraints with the objective function (Hinton, 1992):

$$\operatorname{argmin} \frac{\mathbf{w}_j' \langle \dot{\mathbf{y}} \dot{\mathbf{y}}' \rangle_t \mathbf{w}_j}{\mathbf{w}_j' \langle \mathbf{y} \mathbf{y}' \rangle_t \mathbf{w}_j} = \operatorname{argmin} \frac{\mathbf{w}_j' A \mathbf{w}_j}{\mathbf{w}_j' B \mathbf{w}_j} \tag{4}$$

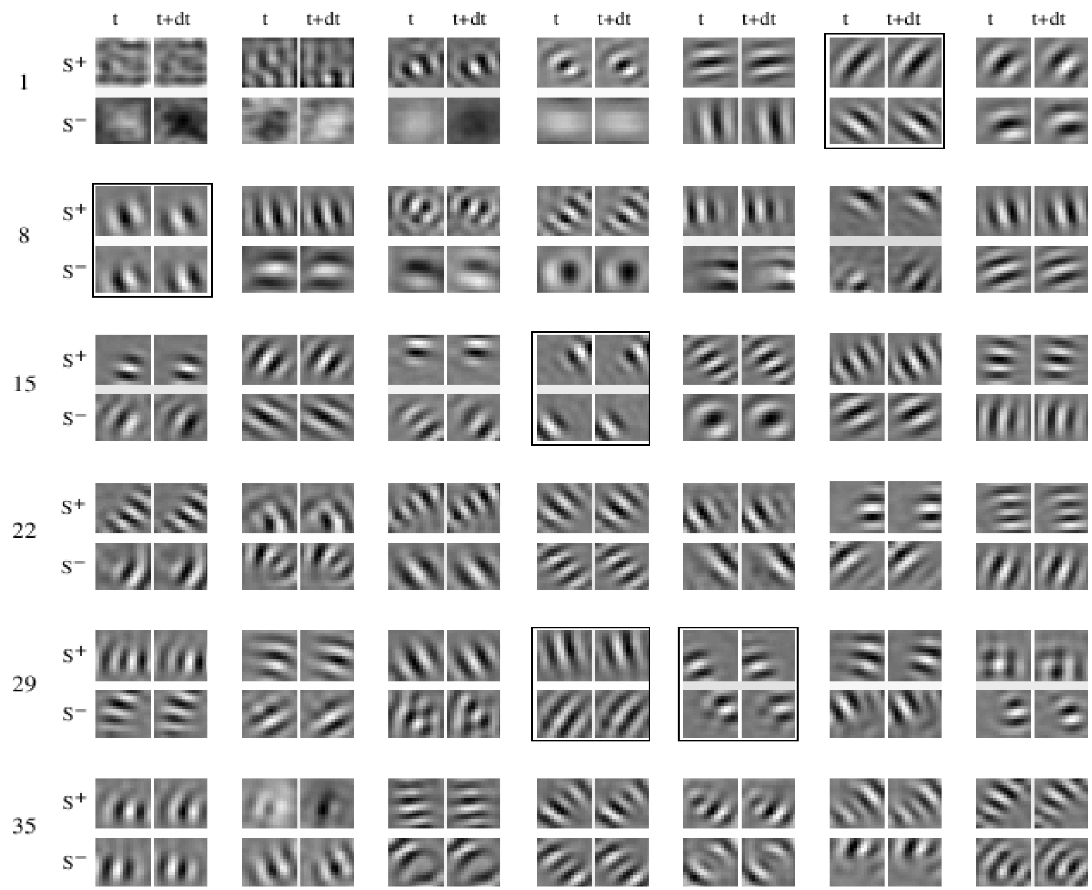
- The solutions are given by the **generalised eigenvalue** problem:

$$AW = \Lambda BW \tag{5}$$

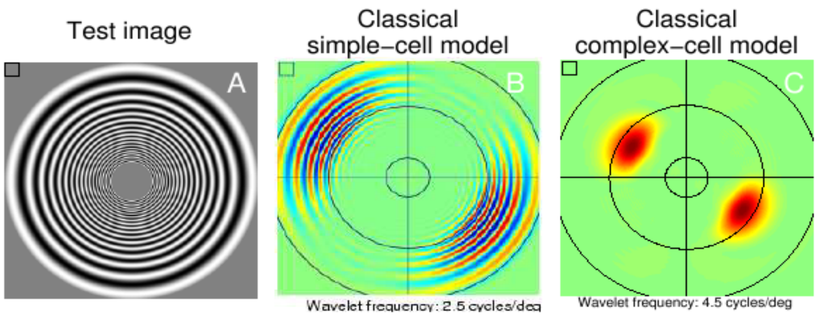
Methods

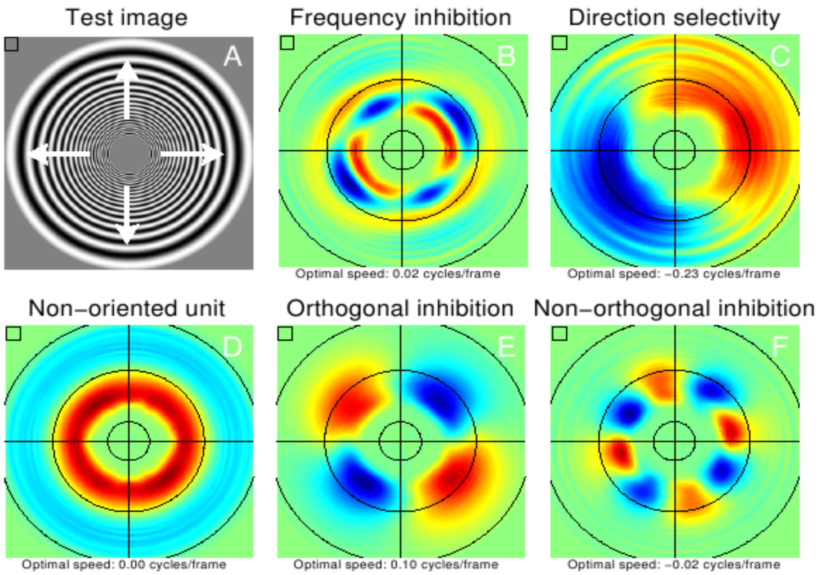
- take images of natural scenes
- contruct a time series by **translating, rotating, and zooming** a window
- augment the input (at one time step) with two consecutive frames so a **full spatial-temporal field** can be recovered by SFA

Results: best stimuli



Results: spatial frequency / orientation tuning





Towards a probabilistic interpretation

- Interpret the $\hat{x}_{i,t}$ as the **posterior mean** of a **latent variable** denoted $x_{i,t}$
- Relax the condition $\langle \hat{x}_{i,t} \hat{x}_{j,t} \rangle = \delta_{i,j}$ to apply to the average posterior value of $x_{i,t} x_{j,t}$
- Now let's think about the forward model which is implicit in SFA
- We'll make all the distributions Gaussian (which is consistent with the square in the cost function $\left\langle \frac{d\hat{x}_{i,t}}{dt}^2 \right\rangle$)
- We know that all the solutions lie in a white subspace (from the above constraints) and that within this subspace, slowness is used to find a further subspace

- MOVE TO WHITE BOARD - DRAW OBSERVATION SPACE DIAGRAM AND THE GRAPHICAL MODEL COMPLETE WITH PROBABILITY DISTRIBUTIONS

Some references

Slow Feature Analysis: A Theoretical Analysis of Optimal Free Responses
Wiskott, Neural Computation 15(9):2147-2177 (2003)

Slow Feature Analysis: Unsupervised Learning of invariances
Wiskott, Sejnowski, Neural Computation 14(4) 2002

Extracting Slow Subspaces from Natural Videos Leads to Complex cells
Kayser, Einhausser, Dummer, Konig, Kording

Is Slowness a learning principle of the visual cortex
Wiskott, Berkes, Zoology 106 2003

Independent Slow Feature Analysis and Non-linear Blind Source Separation
Blaschke and Wiskott, (unsure where!)