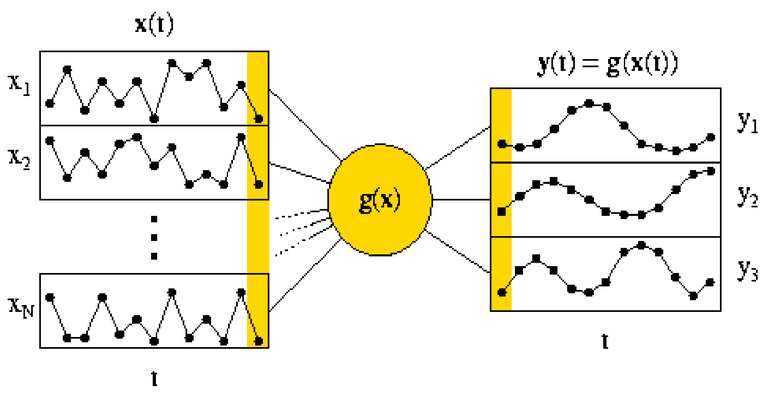
Slowness Leads to Place Cells

M. Franzius, H. Sprekeler, L. Wiskott
Institute of Theoretical Biology, HU Berlin

1. Abstract

We present a model for the self-organized formation of hippocampal place cells based on unsupervised learning on natural visual stimuli. Our model consists of a hierarchy of Slow Feature Analysis (SFA) modules, which were recently shown to be a good model for the early visual system (Berkes & Wiskott, Journal of Vision 5(6):579). The system extracts a distributed representation of position, which is transcoded into a place field representation by sparse coding (ICA). We introduce a mathematical framework for determining the solutions of SFA, which accurately predicts the distributed representation of computer simulations.

2. Slow Feature Analysis



Learning Task:

Given a function space $\mathcal{F}$ and a vectorial input signal $\mathbf{x}(t)$, find a set of *instantaneous* input-output functions $g_j(\mathbf{x}) \in \mathcal{F}$, whose output signals $y_j(t) = g_j(\mathbf{x}(t))$ minimize

$$\Delta(y_j(t)) = \langle \dot{y}_j(t)^2 \rangle_t$$

under the constraints

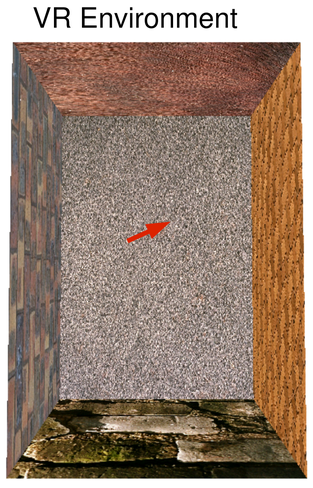
$$\begin{aligned} \langle y_j(t) \rangle_t &= 0 && \text{(zero mean)} \\ \langle y_j(t)^2 \rangle_t &= 1 && \text{(unit variance)} \\ \forall i < j \,:\, \langle y_i(t) y_j(t) \rangle_t &= 0 && \text{(decorrelation \& order)} \end{aligned}$$

Wiskott & Sejnowski, Neural Computation 14:715

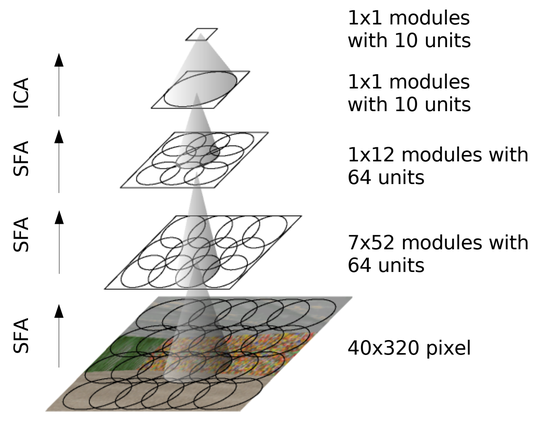
3. Inputs and Processing

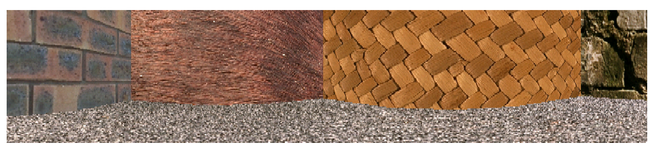
A rat's visual input is simulated by a random walk in a virtual reality system (openGL, VisTex textures), consisting of random translations and rotations. The system's input of colored pixel images (40x320°, 1px/°) is processed by hierarchical layers of SFA modules, each learning a set of quadratic functions, followed by one step of linear Independent Component Analysis (ICA) as a sparse coding mechanism.

Top View of Rat in VR Environment



Hierarchical Processing





Corresponding Input View

4. Theoretical Analysis

Assumptions:

- No restriction of the function space $\mathcal{F}$
- Isotropic and homogeneous movement of the rat

➔ **Objective function** for the response $g_j(\mathbf{r}, \phi)$ of the input-output functions at position $\mathbf{r}$ and head direction ϕ:

$$\Psi(g_j) = v^2 \langle (\partial_{\mathbf{r}} g_j)^2 \rangle_{\mathbf{r}} + \omega^2 \langle (\partial_\phi g_j)^2 \rangle_{\mathbf{r}} - \sum_{i \le j} \lambda_{ij} \langle g_i g_j \rangle_{\mathbf{r}}$$

Using variational calculus, we show that the optimal solutions fulfill the eigenvalue equation

$$-v^2 \partial_{\mathbf{r}}^2 g_j - \omega^2 \partial_\phi^2 g_j = \Delta_j g_j$$

with the boundary condition

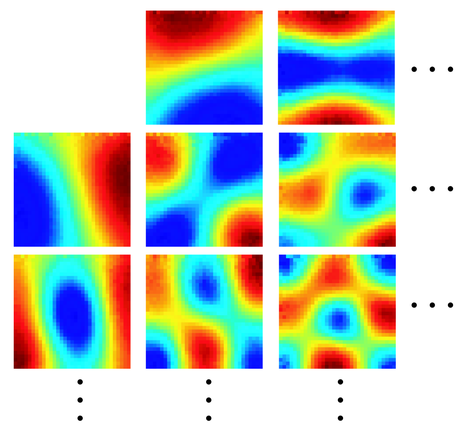
$$\partial_{\mathbf{r}} g_j = 0.$$

For fast rotation and a quadratic room of size L, the slowest solutions are given by

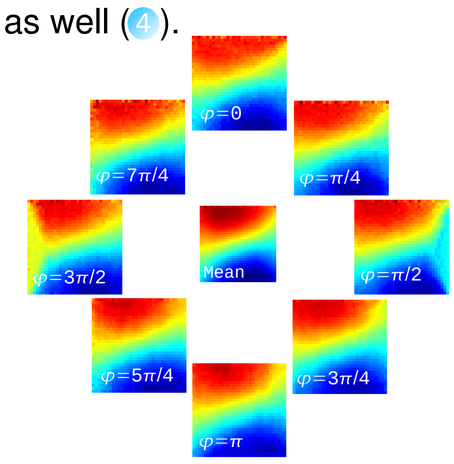
$$g_{nm}(x, y) = 2 \cos(\frac{\pi m}{L} x) \cos(\frac{\pi n}{L} y)$$

5. Simulation Results

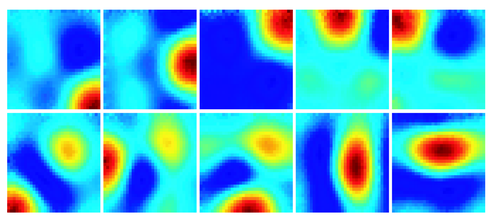
The system learns orientation invariant outputs (2) in the highest SFA layer. The set of all SFA units corresponds to the theoretical prediction (1). Sparse coding in the last step transforms the distributed spatial code into localized representations similar to Place Fields (3; Muller, Neuron 17:813). These units are orientation invariant as well (4).



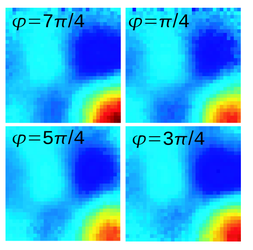
1 SFA Output Units: Spatial Selectivity



2 Single SFA Output Unit: Orientation Invariance



3 ICA Output Units: Spatial Selectivity



4 ICA Output Units: Orientation Invariance

6. Summary

We present a model for the formation of hippocampal place cells based on the unsupervised learning principles of slowness (SFA) and sparse coding (ICA). The architecture of the parameter-free model is inspired by the hierarchical organization of cortex. Our studies show that the model is capable of extracting high level cognitive information such as an animal's position from its complex visual perception, while becoming completely invariant to the animal's orientation. We have developed a general mathematical framework which allows exact analytical predictions of the system's behavior. Further studies show that the model performance is highly robust to noise and architectural details. An intermediate representation displays spatial periodicity similar to the recently discovered Grid Cells (Hafting et. al., Nature 436:801).