Stability of Visual Features and Learning Disparity Selective Complex Cells

Selim Onat[1,2], Christoph Kayser[2], Peter König[1,2]
[1]Neurobiopsychology Department,Institute of Cognitive Science, University of Osnabrück,Germany
[2]Institute of Neuroinformatic, ETH,Zurich,Switzerland

Learning from 3D Natural Images

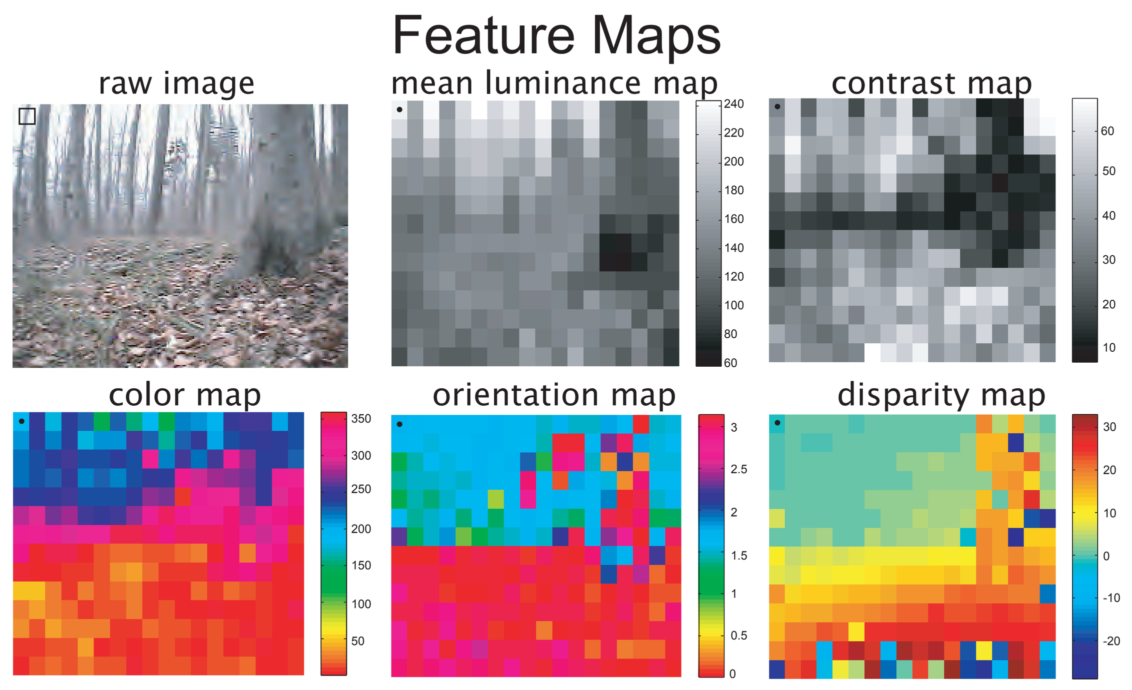
The principles governing the organisation of receptive field (RF) properties are an essential question. This question is investigable by using an unsupervised learning scheme where the activities of neurons are optimized with respect to a goal function. By using such an approach, it has been shown that RFs like those of simple and complex cells can be learnt by representing natural images as stable as possible. In this work, we are extending this approach to stereoscopic images. The extraction of disparity feature in a way similar to real neurons constitute a challenge for this approach. In the present work, we are producing optimally stable representations of stereoscopic natural

Recording Stereoscopic Natural Movies

weight = 71 gr
interocular distance = 4.8 cm
27556 non compressed RGB images (752x583x24 at 25 Hz)

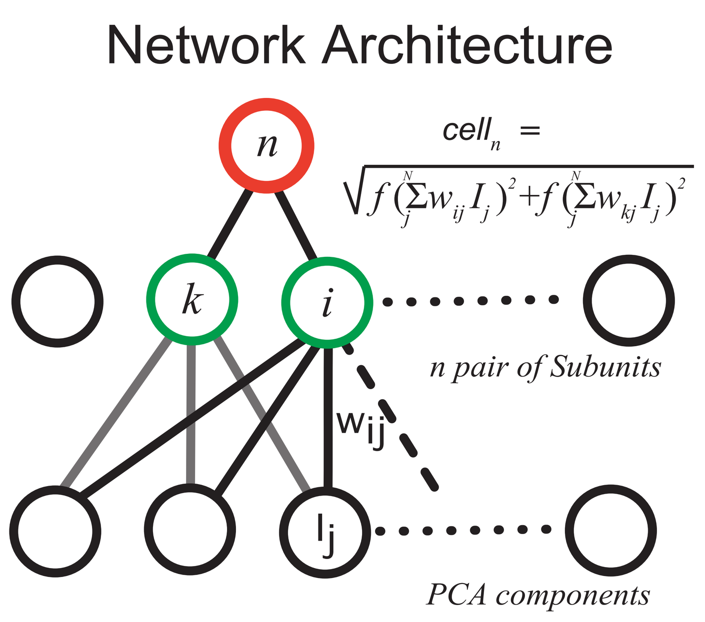
Cat-carried cameras make it possible to capture spatio-temporal course of visual input without introducing antropomorphic biases on the temporal aspects.

Spatio-Temporal Correlations of Visual Features



Inside patches of 40x40 pixel, different visual features are extracted. This allows us to quantitatively analyze the temporal correlation of mean luminance, color, orientation, contrast and disparity over time and space.

Extraction of Stable Features from Natural Movies



Stability Goal Function

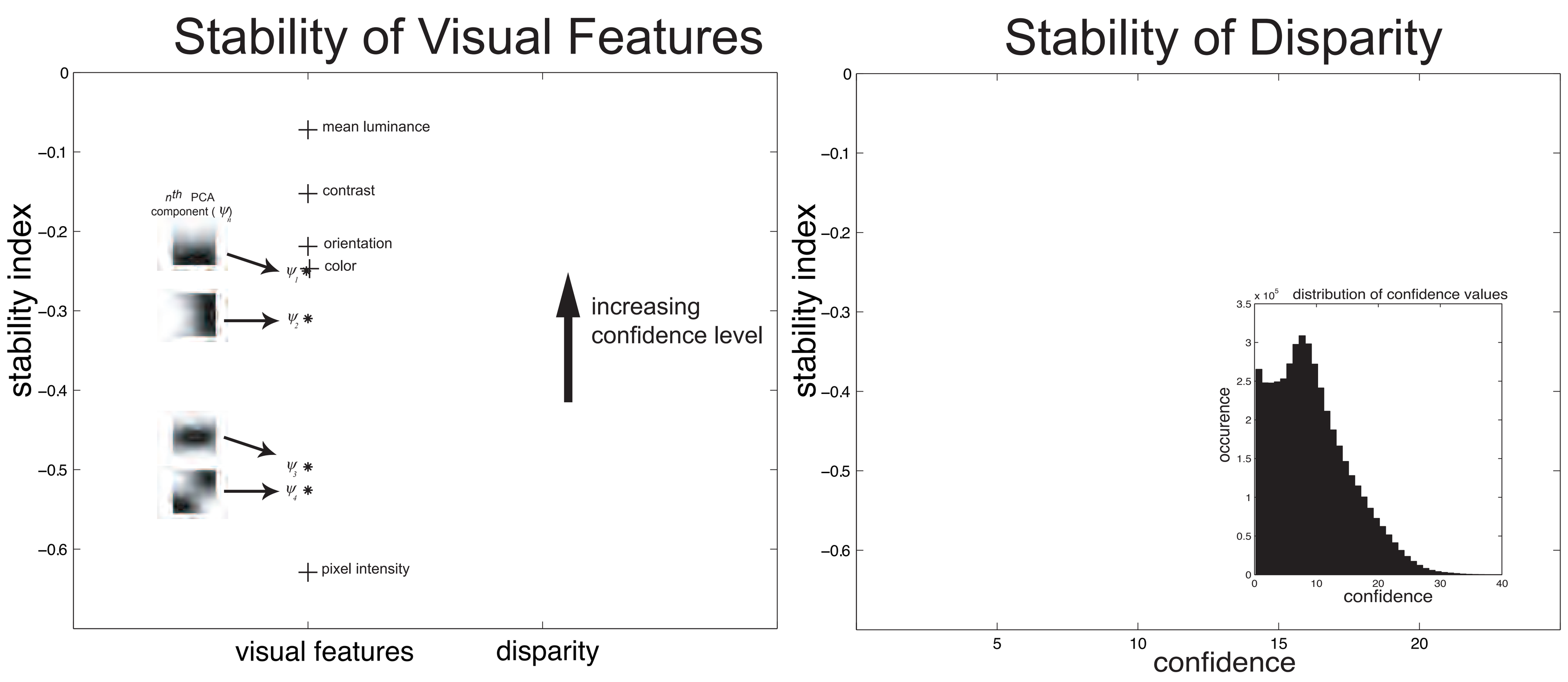
$$\Psi_{stability} = -\frac{1}{N}\sum_i^N \frac{\left\langle \left[A_i(t) - A_i(t+\Delta t) \right]^2 \right\rangle_{stimuli}}{\sigma_i(t)\cdot\sigma_i(t+\Delta t)}$$

$$\Psi_{decorr} = -\frac{1}{(N-1)^2}\left\langle \sum_{i,j}^N CC(A_i A_j) \right\rangle_{stimuli}$$

$$\Psi_{total} = \Psi_{stability} + \Psi_{decorr}$$

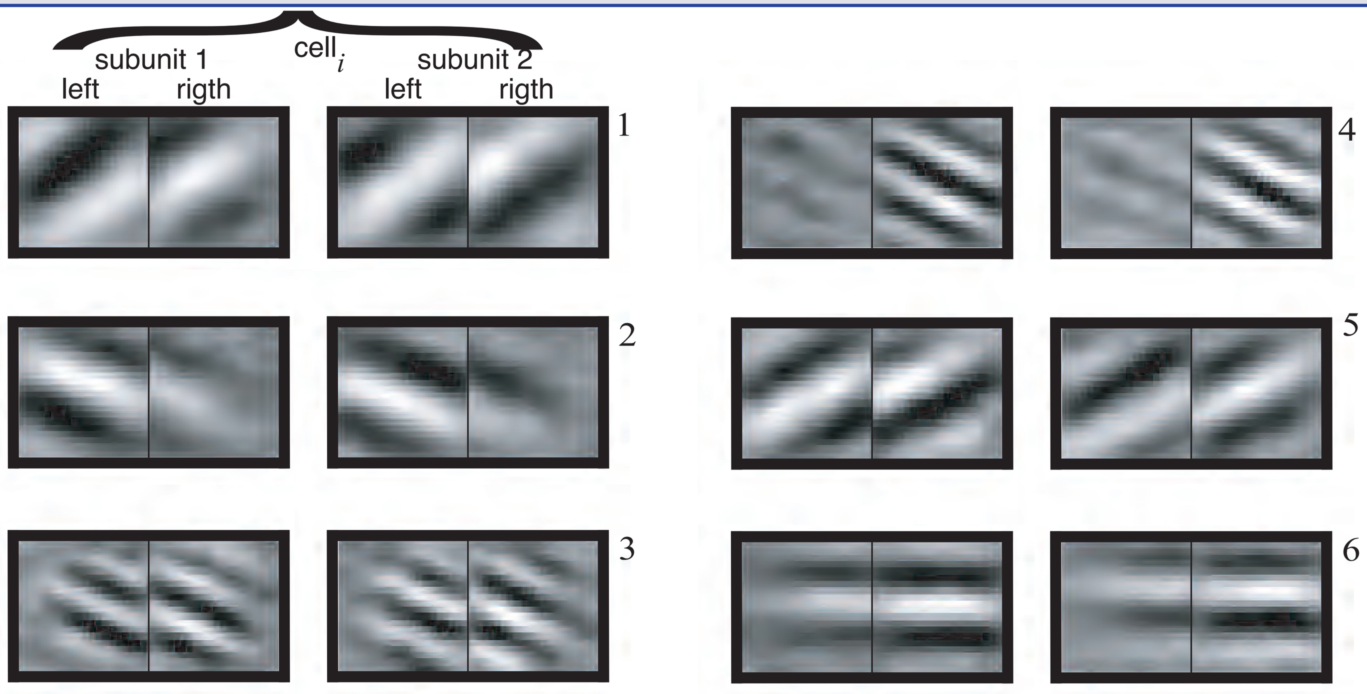
The stimulus set used in the simulations contained 50000 of 40x40 pixels image patches per time t. Additionnaly, we applied a selection criterion on the patches based on their binocular similitude giving 2 different stimulus sets: In confidence condition, the patches were selected according to a confidence measure yielding correlated binocular input. And in random condition, the confidence criterion was relaxed making possible to have uncorrelated binocular input.

Disparity is the SLOWEST Visual Feature



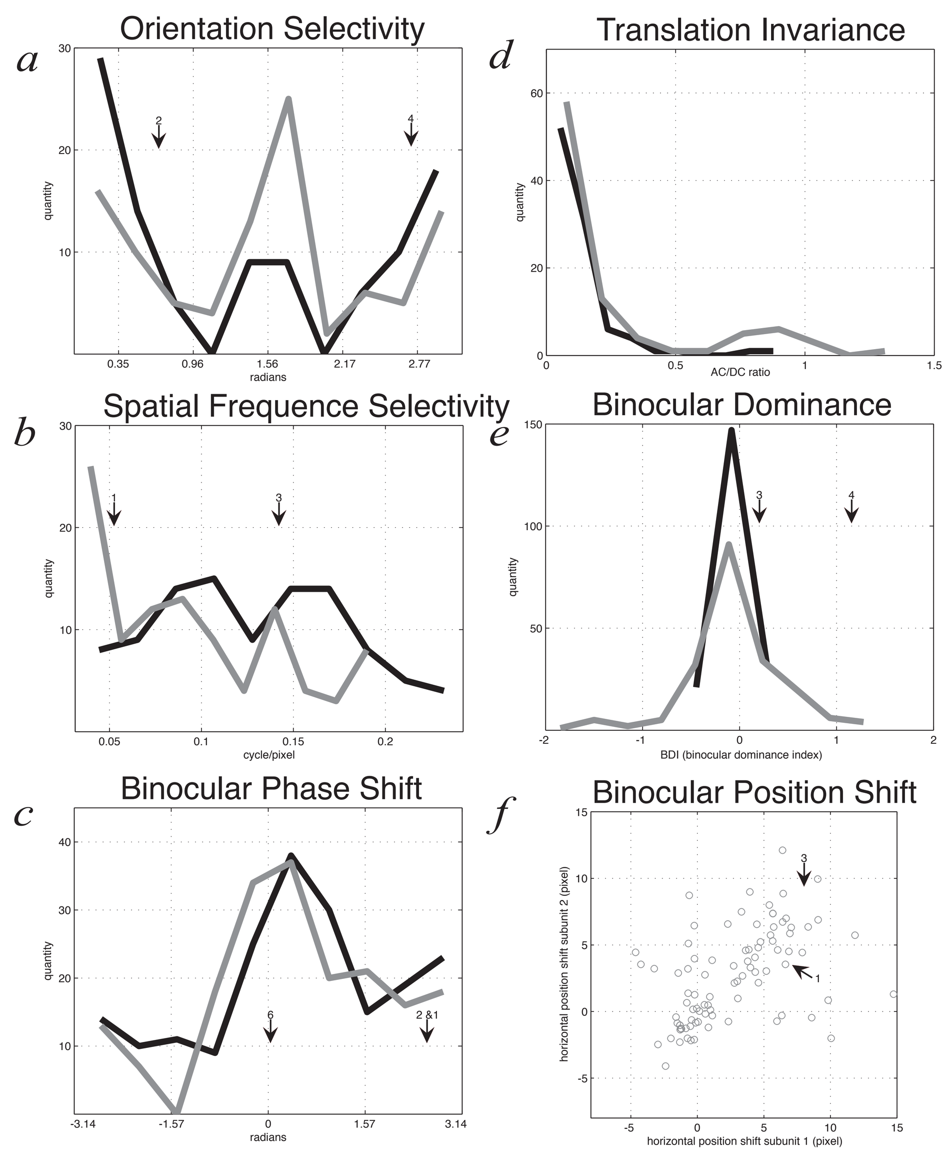
Stability Index is measured according to the stability goal function introduced in the previous paragraph. The disparity feature is the most slowly changing feature, followed by color and orientation. This means that disparity information can be extracted by a network exploiting

Stable Representations are similar to Binocular Complex Cells



Optimally stable representations of stereoscopic natural movies share similar properties with disparity selective complex cells: RFs of subunits are selective for a given orientation and spatial frequency. A given pair of subunits as well as different eyes of a given subunit exhibit similar preferences. Left and right receptive fields of subunits of cell 1 encode disparities with a shift in the phase of their monocular receptive fields. On the other side, the cell 3 encode disparities by positional shifts of the receptive fields. Cells 2, 4 and 6 are binocularly unbalanced. Cells seems to be selective for horizontal disparities independent of the orientation

Receptive Fields Cover the Stimulus Space.



(*a,b*) RFs are selective for different spatial frequencies and orientations. (*c*) Phase differences between left and right eye are predominantly distributed between 0 and π reflecting the disparity distribution of the input statistics due to the arrangement of the cameras. (*d*) Most of the neurons trained with both stimuli sets are invariant to the position of their preferred stimuli although confidence conditions produces better results. (*e*) Random condition produces more binocularly unbalanced receptive fields characterized by the width of the distribution. (*f*) shows

Optimally stable representations of stereoscopic natural movies are similar in many respects to disparity selective complex cells. We claim that it is possible to understand the properties of binocular complex cells in terms of input statistics.