Intrinsic Plasticity in a Generative Model of V1

Cornelius Weber & Jochen Rietzsch
Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, Frankfurt, Email: c.weber@ffas.uni-frankfurt.de

Overview

Generative, sparse-coding models explain the learning of localized edge detector neurons in primary visual cortex V1. Current models, which we use as a basis, work on two time scales:

1. The intrinsic neurons’ activations adapt to a visual stimulus;
   - the weights adapt to the set of stimuli.

Here we distinguish a separate mechanism that may act on a medium time scale: the adaptation of a neuron’s excitability, "intrinsic plasticity" (IP). Hereby a neuron maintains firing rate homeostasis in a dynamic environment.

We set a desired exponentially distributed firing rate, which represents spike coding, and adapt neuronal transfer parameters a and b to maintain the desired rate distribution.

With natural image input, localized edge detectors emerge.

The novel time scale of the intrinsic plasticity parameters allows to explore visual aftereffects such as the tilt aftereffect (TAE). In this effect, after a weaver adapts to a grid of a certain orientation, grids of a nearby orientation will be perceived as tilted away from the adapted orientation.

Our results show that adapting the neurons’ gains but not the threshold parameter leads to quantitatively similar aftereffects as found psychophysically, without the use of horizontal connectivity. This happens, because neurons coding for the adapting stimulus attenuate their gain while others increase it.

In contrast, long-term exposure to gratings of, e.g., 45° orientation yields enhanced representations of that orientation, while after training with natural images, the model also reproduces enhanced representations of vertical and horizontal edges, as experimentally observed.

Model

Adaptations of a and b:

\[
\hat{z}_i = z_i \exp(\alpha_{ai} + \beta_{bi})
\]

with \(\sum a_{ij} w_{ij} + 1\) for each neuron i.

IP parameters a and b:

\[
\Delta a_{ij} = \alpha_{ai} w_{ij} z_i
\]

\[
\Delta b_{ij} = \beta_{bi} w_{ij} z_i
\]

transfert function

prior probability distribution imposed on the activations

Adaptation of \(a_i\) and \(b_i\):

\[
\Delta a_i = a_i (\hat{z}_i - a_i) z_i
\]

\[
\Delta b_i = b_i (\hat{z}_i - b_i) z_i
\]

with \(\sum z_i = 1\) for each input neuron i.

The adaptation of the IP parameters a and b alters the firing rate of the neuron (eq. 1).

The adaptation of the IP parameters a and b affects the gain of every orientation (eq. 2).

The orientation- (gamma) frequency- (beta) distribution of the receptive fields of all orientations and a broad frequency range are accessed.

Adaptation times from the frame rate to the time scale of the model TAE does not saturate.

The TAE due to gain increase or decrease. If it is adapted only by the coarse (i.e., low-frequency) input from the adapting stimulus, then also the orientations of the other stimulus interfere.

Adaptation times from the frame rate to the time scale of the model TAE does not saturate.

Discussion

- One model has realistic assumptions, learns realistic V1 neuron receptive fields, and accounts for adaptation on a time scale of time of seconds, including the TAE.
- It has relatively few parameters and is easy to use.
- It is information theoretically bounded, yet has local learning rules.
- The IP adaptation works also with Okazawa’s generative network algorithm (not shown). The weight normalization used here does away with Okazawa’s weight decay parameter.
- IP parameter adaptation could alternatively be regarded to happen on a time scale of 50 hours, to account for Turagawa’s experiments (not then an explanation of the TAE).
- Gain adaptation accounts for the TAE. Unlike the gain parameters, the neuron’s threshold parameters are all very similar. They may be modifiable.

Bednar & Miikkulainen’s model explains the TAE with rapidly strengthened inhibitory horizontal V1 connections. Activity-induced inhibition is hard to distinguish from rate adaptation. With horizontal weights, any model may also replicate the indirect TAE.

Acknowledgments

We acknowledge financial support by the European Union through projects FP6-2005-015803 (“Daisy”), MEXT-CT-2006-042484 (“PLICON”) and by the Herta Foundation.

The way trained from natural stimulate has a slight preference for cardinal orientation. Better, after additional training with stimuli of non-cardinal orientations better, the orientation preference can be fitted by Okazawa.

The distribution of the IP parameter a is over the receptive field size (mean).

The distribution of the IP parameter b is over the receptive field size (mean).

The centers of the receptive fields in the model retain their respective distributions.

The orientation sensitivity (gamma) frequency distribution of the receptive fields of different orientations and a broad frequency range are accessed.

The orientation sensitivity (gamma) frequency distribution of the receptive fields of all orientations and a broad frequency range are accessed.

The distribution of the IP parameter a is over the receptive field size (mean).

The distribution of the IP parameter b is over the receptive field size (mean).

The centers of the receptive fields in the model retain their respective distributions.

The orientation sensitivity (gamma) frequency distribution of the receptive fields of all orientations and a broad frequency range are accessed.

The orientation sensitivity (gamma) frequency distribution of the receptive fields of all orientations and a broad frequency range are accessed.

The orientation sensitivity (gamma) frequency distribution of the receptive fields of all orientations and a broad frequency range are accessed.

The orientation sensitivity (gamma) frequency distribution of the receptive fields of all orientations and a broad frequency range are accessed.