

# Phenomenological model for the adaptation of shape-selective neurons in area IT

Martin A. Giese<sup>1</sup>, Pradeep Kuravi<sup>2</sup>, and Rufin Vogels<sup>2</sup>

<sup>1</sup>Section Computational Sensomotrics, Dept. of Cognitive Neurology,  
CIN&HIH, University Clinic Tübingen  
Ottfried-Müller Str. 25, 72076 Tübingen, Germany

<sup>2</sup>Lab. Neuro en Psychofysiologie, Dept. of Neuroscience, KU Leuven  
Campus Gasthuisberg, O&N 2, Herestraat 49, BE-3000 Leuven, Belgium  
Email: [martin.giese@uni-tuebingen.de](mailto:martin.giese@uni-tuebingen.de)

**Abstract.** Shape-selective neurons in inferotemporal cortex show adaptation if the same shape stimulus is shown repeatedly. Recent electrophysiological experiments have provided critical data that constrain possible underlying neural mechanisms. We propose a neural model that accounts in a unifying manner for a number of these critical observations. The reproduction of the experimental phenomenology seems to require a combination of input fatigue and firing rate fatigue mechanisms, and the adaptive processes need to be largely independent of the duration of the adapting stimulus. The proposed model realizes these constraints by combining a set of physiologically-inspired mechanisms.

**Keywords:** object recognition, adaptation, inferotemporal cortex, fatigue, neural field

## 1 Introduction

Shape-selective neurons in inferotemporal cortex (area IT) show adaptation for repeated stimulus presentation [1]. This phenomenon has been of strong interest in neuroscience and functional imaging [2], since it might be contributing to high-level after-effects [3], might be related to the observation of repetition suppression effects in functional imaging [4], and efficient coding [5]. Various theories about the origin of adaptation effects have been proposed [6] and different models for such adaptation effects have been put developed [7, 8, 4, 9]. However, the precise underlying neural processes remain largely unknown. Recent electrophysiological experiments provide strong constraints for the possible underlying neural mechanisms and their computational properties. Based on a collection of such experiments, we propose a neural model that accounts simultaneously for all of them, exploring a variety of possible neural adaptation mechanisms. We found that accounting for this data requires a combination of adaptation processes that act on different resolution levels in feature space. Our solution combines *input fatigue* and *firing-rate fatigue* [10]. In addition, we assume that the relevant adaptation processes act largely independently of the duration of the

adapting stimulus, while they are sensitive to the number of its repetitions. The proposed neural model combines a synaptic gain-control mechanism and a saturating firing-rate-dependent adaptation mechanism that shifts the threshold of the neurons in order to account for the data. We assume that the shape-selective neurons are recurrently connected and embedded in a neural field, resulting in a competition between different views and recognized shapes.

In the following, we briefly describe the model and the assumed adaptation mechanisms. Subsequently, we present simulations of a set of critical experiments that provide strong constraints for the underlying mechanisms, followed by a discussion.

## 2 Neural model

Consistent with earlier work, we model IT neurons by radial basis function units that are selective for a learned shapes. We assume that these neurons are embedded within in a recurrent neural network, which can be approximated in the mean-field limit by a dynamic neural field that results in competition between neurons different recognized shapes and views. In other work such recurrent models have been successfully exploited in order to account for object and action recognition (e.g. [11–13]). This recurrent network of shape-selective neurons is augmented by adaptation mechanisms. We have tested a variety of possible mechanisms and report here only a combination that accounts for the critical data sets discussed below. In addition, the model contains a mechanism for spike-rate adaptation that is important to reproduce the signal shape of the post-stimulus time histograms (PSTHs) of IT neurons. The following sections give a more detailed description of the different model components.

### 2.1 Recurrent network (field) of shape-selective neurons

We assume that the individual shape-selective IT neurons obtain their input from a previous layer that encodes input features. We model this layer in an idealized way as two-dimensional neural field that represents input features with a well-defined metric for feature similarity. The dimensionality of this space could be chosen differently, as long as there is a defined similarity metric between the feature vectors. Let  $v(\mathbf{y}, t)$  define the activity of the neurons in this input layer. We assume that this activity is always non-negative.

Consistent with electrophysiological data, we assume that the shape-selective IT neurons encode shapes in a view-specific manner, and we specify by the vector  $\mathbf{x} = [\phi, \theta]$  the encoded shape and view. In our implementation we assumed a two-dimensional space for this representation, one dimension encoding view angle  $\phi$ , and the other the location  $\theta$  along a one-dimensional shape continuum. The dimensionality of the shape space is not critical for the model, and one could assume higher-dimensional shape spaces if necessary. The embedding of neurons in such metric shape spaces simplifies the treatment of pattern similarity, which is a critical variable that was manipulated in the experiments. We assume that

$u(\mathbf{x}, t)$  is the average activity of the neuron (ensembles) whose shape and view selectivity is given by the vector  $\mathbf{x}$ .

The shape-selective neurons are modeled by radial basis function (RBF) units that receive their inputs through a linear weight kernel  $m(\mathbf{x}, \mathbf{y})$  from the input layer. This kernel specifies the strength of the synaptic connections from the input layer to the shape-selective IT neurons. We assume that, without adaptation, this kernel has a Gaussian characteristics, thus defining Gaussian RBFs. One of the assumed adaptation processes acts on the values of this weight kernel. This makes this kernel time-dependent.

The recurrent network of shape-selective neurons is modeled by a dynamic neural field of [14] that receives input from the input layer through the synaptic weight kernel  $m$ . The recurrent interactions in the field are specified by the interaction kernel  $w$ , resulting in the dynamical equation:

$$\begin{aligned} \tau \frac{d}{dt} u(\mathbf{x}, t) = & -u(\mathbf{x}, t) + \int w(\mathbf{x} - \mathbf{x}') H(u(\mathbf{x}', t)) d\mathbf{x}' - h \\ & + \underbrace{\int m(\mathbf{x}, \mathbf{y}, t) v(\mathbf{y}, t) d\mathbf{y}}_{s(\mathbf{x}, t)} \\ & - F(a(\mathbf{x}, t)) + k_c c(\mathbf{x}, t) \end{aligned} \quad (1)$$

In this equation  $H(x)$  is the Heaviside function, thus  $H(x) = 1$  for  $x > 0$  and  $H(x) = 0$  otherwise. The positive constants  $\tau$  and  $h$  define the time scale and the resting potential of the field.

## 2.2 Firing-rate fatigue adaptation

The first adaptation mechanism is based on *firing rate fatigue*, i.e. an increase of the neuron thresholds after they have been continuously firing. This adaptation process is modeled by an adaptation variable  $a(\mathbf{x}, t)$  that increases the effective threshold of the neurons. The dynamics of this variable is determined by a differential equation that is applied separately to each neuron (point in the neural field):

$$\tau_a \frac{d}{dt} a(\mathbf{x}, t) = -a(\mathbf{x}, t) + H(u(\mathbf{x}, t)) \quad (2)$$

The time constant  $\tau_a$  of the adaptation process was chosen to be 1200 ms. The adaptation variable couples into the field dynamics through a saturating nonlinear function  $F(a) = k_a \min(a, a_{\max})$ , with  $k_a > 0$  and  $a_{\max} > 0$ . This nonlinearity bounds the adaptation effects for longer adaptor durations and turns out to be essential in order to account for the result of the experiment with variable adaptor durations (see below).

## 2.3 Input fatigue adaptation

A second adaptation process is acting on the synaptic strength of the input signals of the field, which is specified by the function  $m$ . An alternative interpretation is that this process captures adaptive changes in previous hierarchy layers

of the shape recognition pathway. We assume that the strength of the synaptic connection between neurons at position  $\mathbf{y}$  in the input layer and position  $\mathbf{x}$  in the IT layer is decaying, when the input layer neuron has been activated. In addition, extensive simulation work shows that it has to be assumed that the main effect of the input fatigue adaptation emerges when the input signals decays after a sufficiently long activation period. A highly simplified mathematical model for this is a process that depends on the thresholded negative derivative of the input neuron activation. In our model we captured this by assuming a second adaptation variable  $b$  that follows the dynamical equation:

$$\tau_b \frac{d}{dt} b(\mathbf{y}, t) = -b(\mathbf{y}, t) + \left[ -\frac{\partial}{\partial t} v(\mathbf{y}, t) \right]_+ \quad (3)$$

Here  $v(\mathbf{y}, t)$  signifies the activity of the input neuron at position  $\mathbf{y}$  (with the linear threshold function  $[a]_+ = \max(a, 0)$ ). As the time constant  $\tau_b$  of this adaptation process we chosed 1440 ms. We assume that the adaptation variable  $b$  modulates the strength of the synaptic input weights of the neurons by reducing their gain according to the relationship:

$$m(\mathbf{x}, \mathbf{y}, t) = m(\mathbf{x}, \mathbf{y}) \cdot \left( \frac{1}{b(\mathbf{y}, t)/c_b + 1} \right) \quad (4)$$

The function  $m(\mathbf{x}, \mathbf{y})$  was chosen as two-dimensional gaussian filter kernel. The positive constant  $c_b$  determines a threshold level for the input fatigue adaptation process.

## 2.4 Spike rate adaptation

In order to reproduce the signal shape of the Peristimulus Time Histogram (PSTH) of typical IT neurons, we added another very fast transient process that models spike rate adaptation. This process acts on a much faster timescale than the other discussed adaptive processes. The effect of spike rate adaptation is that the neurons show a short overshoot of activity after stimulus onset that quickly decays. This phenomenon was modeled by adding a transient component to the effective input signal of the IT neurons that decays with a very fast time constant  $\tau_c$ , which was about 9.6 ms. The spike rate adaptation is modeled by third adaptation state variable  $c(\mathbf{x}, t)$  that obeys the dynamic equation

$$\tau_c \frac{d}{dt} c(\mathbf{x}, t) = -c(\mathbf{x}, t) + \left[ \frac{\partial}{\partial t} s(\mathbf{x}, t) \right]_+, \quad (5)$$

where  $s(\mathbf{x}, t)$  is the effective input signal of the IT neuron at position  $\mathbf{x}$  in the neural field. (See also equation (2).) The spike rate adaptation process has only a small effect on the simulation results related to adaptation but is important to reproduce the shapes of the neural responses.

### 3 Simulation results

The model provides a unifying account for several critical experiments that are discussed in the following in comparison with the simulation results.

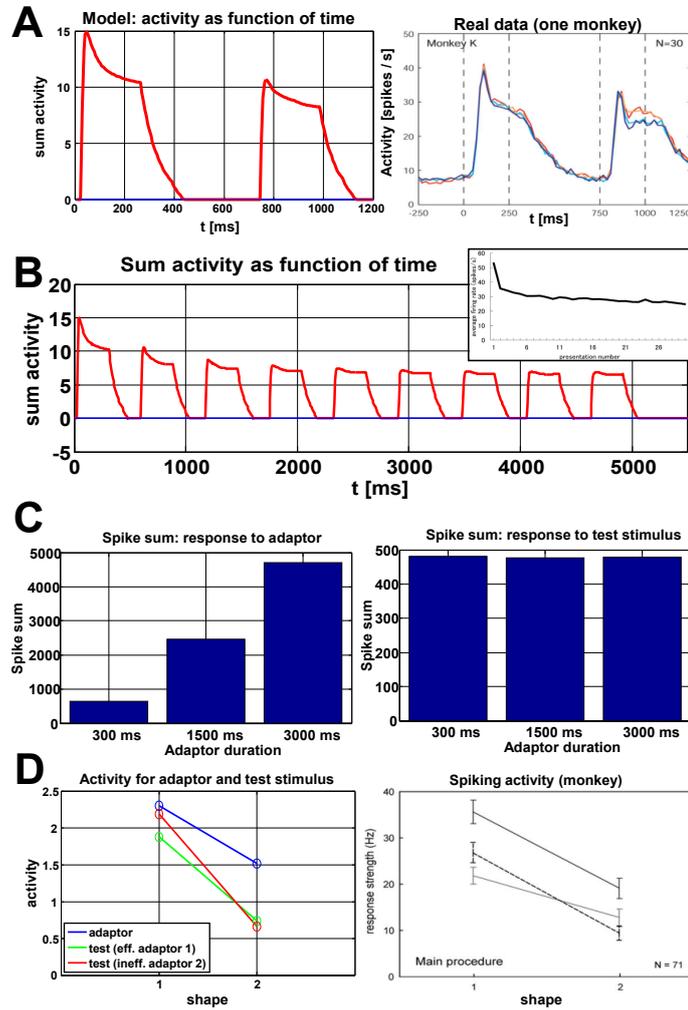
Figure 1 A shows a comparison between the PSTHs from a single stimulus repetition from a typical IT neuron [15] (right panel) in comparison with the simulation result (left panel). Due to the spike rate adaptation process, the model reproduces the signal overshoots after stimulus onset that is present for many IT neurons.

Figure 1 B shows a simulation of the responses for many repetitions of the same shape stimulus, which is optimally stimulating the tested model neuron. The timing parameters match the ones by [1]. Consistent with the experiment the adaptation effect saturates largely after 5 stimulus repetitions. In the real experimental data there occurs a further slight decay that even continues after more than 10 stimulus repetitions. This effect is not captured by our model.

Figure 1 C shows the simulation of an experiment (not yet published) that varied the duration of the adaptor stimulus (showing a shape that stimulated the neuron maximally). For testing stimuli with a fixed duration of 300 ms were shown after a fixed inter-stimulus interval. Quite unexpectedly, the duration of the adaptor stimulus almost did not affect the observed adaptation effects (equal response to test stimulus in all three condition) for a wide range of adaptor durations. The same behavior was observed for real IT neurons in the monkey (Unpublished data). This experimental result is highly constraining for models and could not be reproduced by many model variants, e.g. ones that included adaptation mechanisms that responded to the tonic behavior of the adaptor stimuli without saturation. It can be reproduced by adaptation mechanisms that are dependent on the decay of neural activity.

Finally, figure 1 D shows a simulation of another highly constraining experimental result by [10]. The neurons were stimulated with an effective stimulus (shape 1), that is a shape that elicited a maximum response in the neuron, and with an ineffective stimulus (shape 2), which elicited still a selective response, but a relatively weak one. Presented as an adaptor, obviously, the effective stimulus elicits a higher response than the ineffective stimulus, in the real data as well as for the model. If testing is done with an effective stimulus, evidently, the adaptation effect is larger for an effective adaptor stimulus than for an ineffective adaptor. An interesting situation emerges, however, when the model is tested with ineffective stimuli, and adapted with the effective or the ineffective stimulus. In this case, a statistical interaction occurs where the adaptation effect for an ineffective adaptor stimulus is larger than the one induced by an effective adaptor.

Consistent with the analysis in [10], simulations with many different versions of the model confirm that this type of interaction cannot be obtained with models that do not contain an input fatigue mechanism. The key for obtaining the strong adaptation effect for ineffective test stimuli by adaptation with ineffective stimuli is that the adaptation occurs highly locally in the pattern space of the input neu-



**Fig. 1.** Simulation results: **A** Simulated PSTH (left) in comparison with PSTH from typical IT neuron for single stimulus repetition [15]. **B** Simulated decay of stimulus responses for many repetitions of the same stimulus in comparison with real data (inset) from [1]. **C** Total responses strength (number of spikes integrated over time) for adaptors with different durations (300, 1500, and 3000 ms) (left), and following test stimuli with a fixed duration (300 ms). A similar behavior was observed in real data from area IT (unpublished). **D** Responses for adaptor stimulus (line) for effective and ineffective stimulus, and responses for test stimuli after adaptation either with an effective or an ineffective adaptor stimulus. Model responses (left) compared to corresponding physiological data from [10].

rons. For firing rate fatigue the spatial precision of the induced adaptation effect is bounded by the spatial low-pass characteristics of the feed-forward kernel  $m$  and the lateral interaction kernel  $w$ . The width of the activation peak in the field determines the shape tuning, and thus the difference between the responses to the effective and the ineffective adaptor. Adapting a neuron with an ineffective stimulus will thus induce smaller activity than adaptation with an effective stimulus, resulting in a lower adaptation, which also remains visible when the neuron is tested with an inefficient stimulus. In contrast, if adaptation happens at the input or synaptic level, if one assumes that the tuning in the input level is highly localized in the  $\mathbf{y}$  parameter-space, adaptation and testing with a inefficient stimulus leads to a strong adaptation effect since both, adaptation and test stimulus activate the same neurons in the input layer. At the same time, small adaptation emerges at the input level if adaptation and test stimulus are different, which explains the interaction effect. Detailed simulations show that both, input fatigue and firing rate fatigue are necessary to reproduce the results from [10].

## 4 Conclusions

We have presented a phenomenological model that reproduces simultaneously a number of critical experimental results on adaptation effects in neurons in area IT. Testing many different variants of the model, we found that an account for these results necessitates and input fatigue as well as a firing rate fatigue process. A second constraint from the data is the absence of an influence of the adaptor duration on the strength of the adaptation effect. In order to reproduce this result, we assumed a fast nonlinear saturation of the firing rate fatigue, and a dependence of the input fatigue on the decays of the synaptic input signals. With many other tested mechanisms, including models with transient-dependent firing-rate fatigue or saturating input fatigue mechanisms, we were not able to reproduce the data.

Since the model is qualitative and makes a lot of ad hoc assumptions future work will have to verify the proposed mechanisms, ideally by deriving predictions from the model that can be tested physiologically by causal manipulations of the proposed levels (input synapses, output firing rates, recurrent interactions between shape-selective IT neurons). In addition, the model has to be linked closer to specific biophysical mechanisms and details, such as dynamic processes dependent on specific channel types.

## Acknowledgements

We thanks L. Fedorov for interesting comments. Funded by the EC FP7 ABC PITN-GA-011-290011, HBP FP7-604102; Koroibot FP7-611909, H2020 ICT-644727 CogImon; BMBF FKZ: 01GQ1002A, and DFG GZ: GI 305/4-1 and KA 1258/15-1.

## References

1. Sawamura, H., Orban, G.A., Vogels, R.: Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the fMRI adaptation paradigm. *Neuron* 49(2), 307-318 (2006)
2. Krekelberg, B., Boynton, G.M. and van Wezel, R.J.A. : Adaptation: from single cells to BOLD signals. *Trends in Cognitive Sciences*, 12(6), 230–236 (2006)
3. Leopold, D.A., O’Toole, A.J., Vetter, T., Blanz, V.: Prototype-referenced shape encoding revealed by high-level aftereffects. *Nat Neurosci.* 4(1), 89–94 (2006)
4. Grill-Spector, K., Henson, R., Martin, A.: Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci.* 10(1), 14–23 (2006)
5. Wark, B., Lundstrom, B.N., Fairhall, A.: Sensory adaptation. *Curr Opin Neurobiol.* 17(4), 423–429 (2007)
6. Vogels, R.: Sources of adaptation of inferior temporal cortical responses. *Cortex* pii: S0010-9452(15)00334-2. (2015)
7. Sohal VS(1), Hasselmo ME.: A model for experience-dependent changes in the responses of inferotemporal neurons. *Network* 11(3),169–190 (2000)
8. Friston K.: A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci.* 360(1456), 815–836 (2005)
9. Wark, B., Fairhall, A., Rieke, F.: Timescales of inference in visual adaptation. *Neuron* 61(5), 750-761 (2009)
10. De Baene, W., Vogels R.: Effects of adaptation on the stimulus selectivity of macaque inferior temporal spiking activity and local field potentials. *Cereb Cortex.* 20(9):2145-2165 (2010)
11. Riesenhuber, M., Poggio, T.: Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2(11), 1019–1025 (1999)
12. Giese, M.A., Poggio, T.: Neural Mechanisms for the recognition of biological movements and action. *Nat. Neurosci.* 4, 179–192 (2003)
13. Giese, M.A.: Skeleton Model for the Neurodynamics of Visual Action Representations. In: ICANN 2014. 24th International Conference on Artificial Neural Networks, Hamburg, Germany, September 15-19, Proceedings, pp 707–714. Springer International Publishing (2014)
14. Amari, S.: Dynamics of pattern formation in lateral inhibition type neural fields. *Biol. Cybernetics* 27, 77–87 (1977)
15. Kaliukhovich, D.A., Vogels, R.: Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. *Cereb Cortex* 21(7), 1547-1558 (2011)