A Method for Evaluating Tuning Functions of Single Neurons based on Mutual Information Maximization

Lukas Brostek[*], Thomas Eggert[†], Seiji Ono[**], Michael J. Mustari[**], Ulrich Büttner[*] and Stefan Glasauer[*]

[*]*Clinical Neurosciences and Bernstein Center for Computational Neuroscience, Ludwig-Maximilians-Universität, Munich, Germany*
[†]*Department of Neurology, Ludwig-Maximilians-Universität, Munich, Germany*
[**]*Regional Primate Center, University of Washington, Seattle, WA, USA*

Abstract. We introduce a novel approach for evaluation of neuronal tuning functions, which can be expressed by the conditional probability of observing a spike given any combination of independent variables. This probability can be estimated out of experimentally available data. By maximizing the mutual information between the probability distribution of the spike occurrence and that of the variables, the dependence of the spike on the input variables is maximized as well. We used this method to analyze the dependence of neuronal activity in cortical area MSTd on signals related to movement of the eye and retinal image movement.

Keywords: Neuroscience, Mutual Information, Neuronal Tuning, Neuronal Latency
PACS: 87.19.lo, 87.19.L-

INTRODUCTION

Neurons code the information they transmit using a binary code. The so called action potentials, or spikes, are the only form of neuronal membrane potential fluctuation that can propagate over long distances. It is usually assumed that the information is conveyed in the rate of spiking activity [1, 2].

Neuronal tuning functions are typically defined by the functional relation between the rate of spiking activity and uni- or multivariate independent variables. Tuning functions have provided a first-order description of virtually every sensory system, from orientation colums in the vertebrate visual cortex up to wind-detecting neurons in the cricket cercal system [2, 3, 4].

A difficulty in the analysis of neuronal data is the estimation of proper latency values between any of the variables and the neuronal activity. Appropriate estimation of neuronal latencies is important, since the choice of these latency values has great influence on the tuning function [5, 6].

A common approach to the problem of latency estimation is minimizing the residual error in regression analysis using a linear, quadratic, or any other model [7, 8, 9]. To overcome the limitations of model-based system identification we developed an information based approach for evaluating the dependence of neuronal activity of single cells on combinations of one or multiple independent variables.

The proposed approach is similar to a method used for the registration of medical images: the relative position and orientation of two different images is adjusted by transforming one of the images until the mutual information between both intensity distributions is maximized [10, 11]. Analogous to the spatial alignment used in this methods, our method performs temporal alignment of two random variables by maximizing the mutual information.

THE MAXIMUM MUTUAL INFORMATION METHOD

Basically, our approach consists of two components: first, a method for the determination of a neuronal tuning function, and second, an information-theoretic technique for estimating neuronal latencies and selecting those variables that show the greatest dependence on the neuronal activity.

Tuning function determination

A neuronal tuning function describes the rate of spiking activity in a neuron depending on one or multiple independent variables. This dependence is ideally expressed by the conditional probability $p_{S|V}(s|v)$ of observing a spike given any combination of the variables. By multiplying with the sampling rate, this probability translates directly into an expectation value of the rate of spiking activity. Using Bayes' theorem, $p_{S|V}(s|v)$ can be expressed as the quotient of the joint probability mass function $p_{V,S}(v,s)$ divided by $p_V(v)$:

$$p_{S|V}(s|v) = \frac{p_{V,S}(v,s)}{p_V(v)},$$

where $p_V(v)$ is the marginal probability mass function of observing any combination of variables. The normalization on $p_V(v)$ allows the estimation of the tuning function in unbalanced designs (unequal number of observations across independent variables).

Estimates of $p_V(v)$ and $p_{V,S}(v,s)$ can be attained by histogramming the experimental data. Note, that the joint probability mass function $p_{V,S}(v,s)$ critically depends on the assumed neuronal latency. For optimal bin width estimation we adapted an algorithm proposed by Knuth [12]. According to this, the optimal bin width is defined by the Bayesian estimate of the number of segments of a piecewise constant probability function that is limited to a fixed interval. For smoothing the histogramms, we used a symmetrical Gaussian low-pass filter with a standard deviation of two bin widths. Bins containing a number of values less than 0.5 percent of total spike count were omitted in the analysis.

The amount of data needed for histogramming increases exponentially with the number of dimensions. However, the duration one single neuron can be recorded is restricted due to experimental and physiological constraints. Hence, there is a limitation in the number of variables this method for tuning function determination can be applied on.

Mutual information maximization

Applied to the two random variables V and S from previous section, the mutual information $I(V;S)$ can be stated as

$$I(V;S) = H(S) - H(S|V).$$

with $H(S)$ being the entropy of S and $H(S|V)$ the conditional entropy of S given V, also referred to as noise entropy. These are defined by

$$H(S) = -\sum_s p_S(s) \log p_S(s)$$

$$H(S|V) = -\sum_v p_V(v) \sum_s p_{S|V}(s|v) \log p_{S|V}(s|v),$$

where $p_S(s)$ denotes the probability mass function of spike occurence. The conditional probability $p_{S|V}(s|v)$ denotes the tuning function, determined by the method mentioned in the previous section. This probability and herewith $H(S|V)$ depend on both, the choice of variables analyzed, and the choice of latencies between these variables and the neuronal activity.

The proposed approach chooses the latency in such a way, that the dependence of the neuronal activity on the independent variables is maximized by maximizing the mutual information between V and S. As $H(S)$ is defined by the neuronal activity alone, the maximization is achieved by minimizing the noise entropy $H(S|V)$.

We define the mutual information rate

$$IR(V;S) = \frac{I(V;S)}{H(S)}.$$

This measure specifies the percentage of information about S that can be gathered by knowledge of V.

Due to the limitation in the number of variables mentioned in previous section, in practice the dimension of the tuning function will not exceed values of two or three. To investigate the dependence of a spike on a higher number of independent variables, we determined the tuning functions of a single neuron for any pairwise selection V_k of those variables. For each of these pairs neuronal latencies of both variables were estimated by maximizing the mutual information $I(V_k;S)$. As $I(V_k;S)$ quantifies the dependence of the spike on the selected pair of variables, those two variables that are most related to the spiking activity can be determined by comparing the maximal mutual information rates of the two-dimensional tuning functions.

APPLICATION

The data presented in this paper consists of a 400 s long extracellular recording in cortical area MSTd from a behaving monkey (Macaca mulatta, 5-7 kg), born in captivity at the Yerkes National Primate Research Center (Atlanta, GA). Experimental procedures

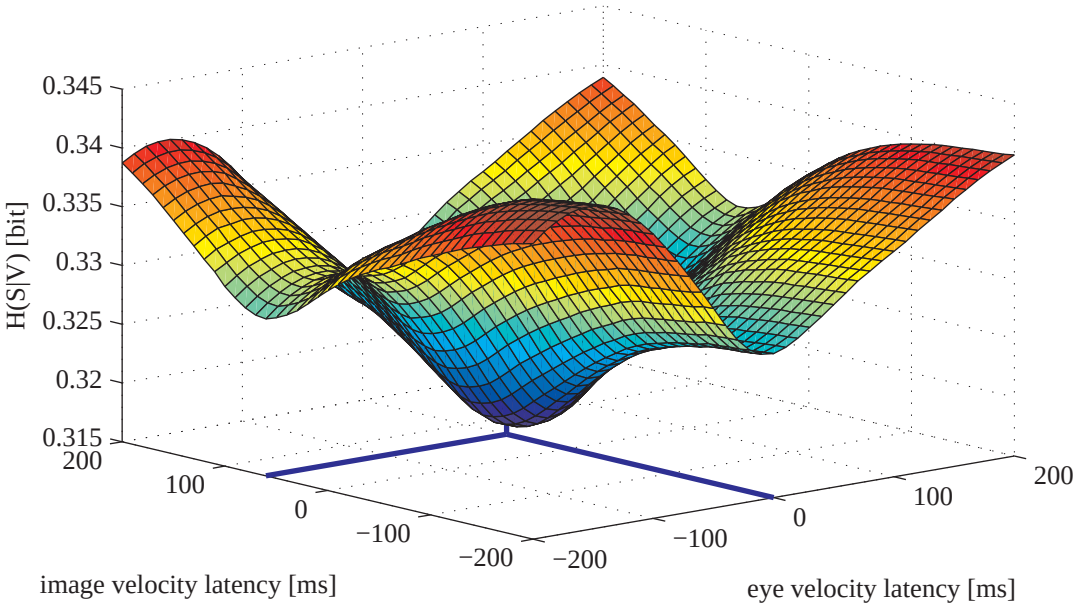

FIGURE 1. Noise entropy $H(S|V)$ against various latencies of image and eye velocity.

are explained in detail in [13]. During the experiments the monkey was seated in a primate chair with his head fixed in the horizontal stereotaxic plane in a completely dark room. Single unit activity was recorded with customized epoxy-coated tungsten microelectrodes. Using a hardware window discriminator a total number of 26015 action potentials was detected and sampled at 1 kHz. Eye movements were detected with standard electro-magnetic methods using scleral search coils [14]. The recorded eye position traces were filtered with a Gaussian low-pass (cutoff frequency 10 Hz) and three-point differentiated to obtain the velocity traces. Saccades were detected and removed with a slow-phase estimation algorithm as described in [15].

The stimulus consisted of a moving large field (35° x 35°) random dot pattern. For optimal coverage of the value range, we used quasi random motion with a flat frequency spectrum (white noise) with maximal eccentricity of 25° and velocity up to $100°/s$. During presentation of the visual motion the monkey had to fixate a small target spot at the center of gaze, though large-field stimulation always produces slight optokinetic eye movements.

We related the neuronal activity to variables, supposed to be potentially coded in MSTd during visual stimulation [16, 17, 18]. The retinal variables consisted of image velocity and acceleration, whereas the group of extraretinal variables contained eye position, velocity and acceleration. Data were acquired only for those movement directions that were previously identified to be the preferred direction of the neuron. This means, the direction which elicits maximal spiking activity for a moving large field stimulus in the analyzed neuron.

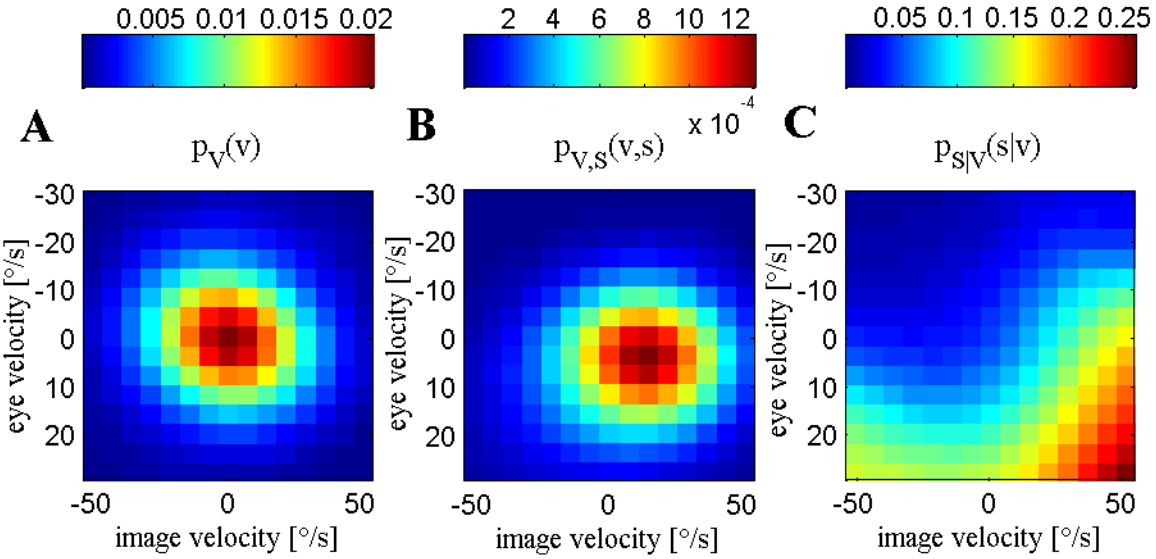

FIGURE 2. Tuning function determination. Dividing $p_{V,S}(v,s)$ (B) by $p_V(v)$ (A) yields the conditional probability $p_{S|V}(s|v)$ of observing a spike given any combination of the variables (C).

RESULTS

Figure 1 shows the noise entropy $H(S|V)$ against various latencies of the variables retinal image velocity and eye velocity. Both variables were shifted in the range between -200 and +200 ms relative to spiking activity, with negative and positive delay meaning backwards and forwards shifts, respectively. For the image velocity latency of +60 ms and eye velocity latency of 0 ms $H(S|V)$ had a minimum of 0.3268 bits. With $H(S) = 0.3456$ bits, the mutual information $I(V,S)$ accounted for 0.0189 bits for these estimates of neuronal latency. The mutual information rate $IR(V,S)$ was 5.46 %, meaning that this portion of information contained in the spiking activity was the maximum that could be explained by the information of that variable pair.

Figure 2 demonstrates the determination of the neuronal tuning function for the variable pair image velocity & eye velocity by application of Bayes' rule. The estimated probability mass function $p_V(v)$ of the occurrence of combinations of the independent variables image velocity and eye velocity is plotted in Fig. 2A. Figure 2B shows the estimated joint probability mass function $p_{V,S}(v,s)$ of coincident variable and spike occurrence. Note that both variables were shifted relative to the neuronal activity according to the estimated neuronal latencies. Dividing $p_{V,S}(v,s)$ by $p_V(v)$ yields the conditional probability $p_{S|V}(s|v)$ of observing a spike given any combination of the variables (Fig. 2C).

In the same way the neuronal tuning functions were determined for all variable combinations (Fig. 3). Neuronal activity in area MSTd is non-linearly related to combinations of the considered eye movement and retinal image movement variables. In the analyzed

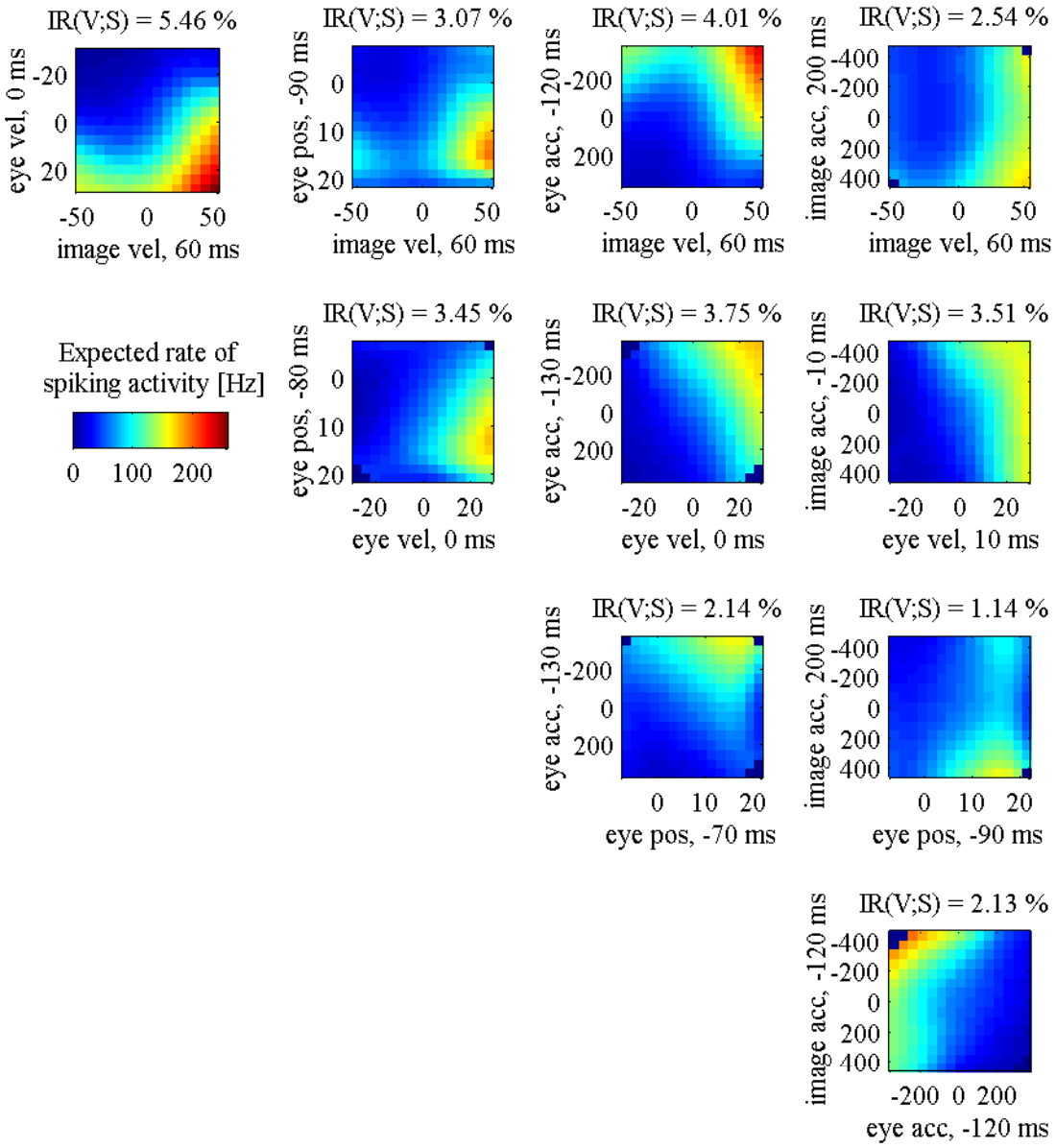

FIGURE 3. Two-dimensional tuning functions for all pairs of analyzed variables. Here, colors indicate the expected rate of spiking activity, which results by multiplying the conditional probability $p_{SV}(s|v)$ with the sample rate of 1 kHz. Each axis is labeled by respective variable and the estimated latency in regard to the neuronal activity.

neuron the mutual information rate $IR(V,S)$ for the variable combination image velocity & eye velocity was larger then any other combination. Hence, this combination was most related to spiking activity. The estimated latencies agree well with results based on other approaches [16, 19]. The latencies of all variables depended only little on the combination, except that of image acceleration. This is related to the low dependence of spiking activity on image acceleration, also apparent in the respective tuning functions.

CONCLUSIONS

The proposed method for tuning function determination allows the identification of any neuronal tuning function. It can be applied in unbalanced designs and allows quantification of any possible dependence of the neuronal activity on the independent variables. However, the dimension of the tuning function is limited by the length of the neuronal recording.

Analyzing the mutual information is the adequate tool for evaluating tuning functions defined in this probabilistic framework. This method is independent of model assumptions. Maximizing the mutual information allows estimation of neuronal latency and comparison of the coherence between spiking activity and different variable combinations. Since neuronal tuning functions can be versatile and highly non-linear, the proposed method is especially suitable for analyzing these.

ACKNOWLEDGMENTS

This work was supported by the Bernstein Center for Computational Neuroscience Grant BMBF 011GQ0440 and NIH Grants EY013308, RR00166.

REFERENCES

1. E. D. Adrian, *J. Physiol.* **61** (1926).
2. P. Dayan, and L. F. Abbott, *Theoretical Neuroscience*, MIT Press, 2001.
3. F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code*, MIT Press, 1999.
4. D. A. Butts, and M. S. Goldman, *PLoS Biol* **4**, 639–646 (2006).
5. J. Seal, D. Commenges, R. Salamon, and B. Bioulac, *Brain Research* **278**, 382–386 (1983).
6. H. S. Friedman, and C. E. Priebe, *J. Neurosc. Methods* **83**, 185–194 (1998).
7. S. Ono, V. E. Das, J. R. Economides, and M. J. Mustari, *J. Neurophysiol.* **93**, 108–116 (2004).
8. U. J. Ilg, S. Schumann, and P. Thier, *Neuron* **43**, 145–151 (2004).
9. M. C.-K. Wu, S. V. David, and J. L. Gallant, *Annu. Rev. Neurosci.* **29**, 477–505 (2006).
10. A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, *Information Processing in Medical Imaging* **3**, 263–274 (1995).
11. W. M. Wells, and P. Viola, *Med. Image Anal.* **1**, 35–51 (1996).
12. K. H. Knuth, *ArXiv Physics e-prints* (2006), `arXiv:physics/0605197`.
13. S. Ono, L. Brostek, U. Nuding, S. Glasauer, U. Büttner, and M. J. Mustari, *J. Neurophysiol.* **103** (2010).
14. A. F. Fuchs, and D. A. Robinson, *J. Appl. Physiol.* **21**, 1068–1070 (1966).
15. J. Ladda, T. Eggert, S. Glasauer, and A. Straube, *Exp. Brain Res.* **182**, 343–356 (2007).
16. W. T. Newsome, R. H. Wurtz, and H. Komatsu, *J. Neurophysiol.* **60** (1988).

17. F. Bremmer, U. J. Ilg, A. Thiele, C. Distler, and K.-P. Hoffmann, *J. Neurophysiol.* **77**, 944–961 (1997).
18. S. B. Hamed, W. Page, C. Duffy, and A. Pouget, *J. Neurophysiol.* **90**, 549–558 (2003).
19. K. Kawano, M. Shidara, Y. Watanabe, and S. Yamane, *J. Neurophysiol.* **71**, 2305–2324 (1994).