

THE SPATIAL BASIS OF NEURAL REPRESENTATION

Toru Yanagawa^{1,2}, Fumihiko Taya^{1,3}, and Ken Mogi¹

1 Sony Computer Science Laboratories,

Takanawa Muse Bldg. 3-14-13, Higashigotanda, Shinagawa-ku, Tokyo, 141-0022 Japan

2 Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama, 226-8503, JAPAN

3 Department of Medicine, Osaka University 2-1 Yamadaoka, Suita, Osaka, 565-0871 JAPAN

ABSTRACT

The spatial organization of cortical neural activities is one of the most important element linking neural network models and natural neural networks. The columnar organization indicated by several works[1] suggest a strong spatial constraint on neural network functions. Recently, attention has grown on the correlation between a vigorous firing of a particular neuron and the "supporting" neural firings in the cortical vicinity. Here, we investigate the relation between a vigorous firing of a particular neuron and the surrounding neural activities using a generic neural network architecture. We analyze the the spatial firing pattern in the vicinity of a particular neuron when the neuron is firing vigorously (preferred cortical state, PCS). We find that the spatial heterogeneity of the PCS is more pronounced in the both extreme ends of activity scale, and not so much in the intermediate activity region. This result, taken together with an analysis of the reaction of the network to a dynamic stimulus, put a strong constraint on the computational capabilities of a neural network. We discuss the implications of our result for the spatial constraints imposed on the neural information processing.

1. INTRODUCTION

A typical cortical neuron is connected via 10^4 synapses to as many neurons. Heterogeneity arises from the spatial distribution of the synapses, and the different types of neurotransmitters involved. This dense and heterogeneous connection is the physical foundation for the remarkable computational capability of the brain.

The dense synaptic connectivity puts certain constraints on the capability of the cortical neurons. One of the constraints is that the cortical neurons cannot fire independently of each other. While the elements in a Turing machine[2] can in principle take arbitrary values independently of other elements, the activities of cortical neurons is highly constrained by the synaptic connectivity. When the synaptic connectivity is given in terms of the physical distance in the cortical surface (as is often the case), it is implied that the state of the neurons which are close to each other cannot take arbitrary values. The coherence in the cortical neural activities are revealed by local field potentials and single unit recordings [3],[4].

A recent study [5] investigated the relation between single unit activity and real-time optical imaging in areas 17 and 18 of the visual cortex of the anesthetized cats. The orientation selective cells investigated fired vigorously when the surrounding neurons fired in a pattern (preferred cortical state, PCS) similar to that observed in the evoked state, when the single unit was stimulated with drifting grating of optimal orientation.

The fact that a neuron is likely to fire in the PCS even when spontaneously firing puts a potentially very strong constraint on the computational capabilities of the cortical neurons.

2. MODEL

We investigated the mutual dependence of the cortical neurons. We constructed a generic neural network (Fig.1) where the excitatory and inhibitory synaptic connections are given with the following rules. The excitatory synapses are distributed with a Gaussian distribution in the range of $0 < d < 4\sigma$, where d is the distance from the presynaptic neuron and σ is the standard deviation of the Gaussian. The inhibitory synapses are distributed uniformly within the range of $0 < d < 2\sigma$. The excitatory neurons project to both excitatory and inhibitory neurons. The inhibitory neurons project only to excitatory neurons. The ratio of the excitatory and inhibitory neurons is set to be 10:3.

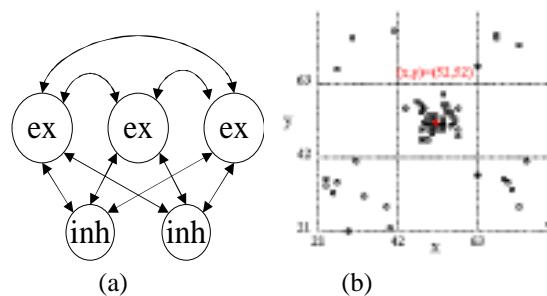


Fig. 1: The Architecture of the Network. A single layer network is represented as a composite of excitatory and inhibitory neuron layers. The synaptic connections are distributed with a gaussian distribution for the excitatory neurons, The inhibitory synapses are distributed in a smaller range. There are additional long range excitatory connections between the alternating columns.

The dynamics of the network is given by

$$u_{i,ex} = \frac{\bar{N}_{ex-ex}}{N_{i,ex-ex}} \sum_j w_{ij} x_{j,ex} - \frac{\bar{N}_{ex-inh}}{N_{i,ex-inh}} \sum_j w_{ij} x_{j,inh}$$

$$x_{i,ex} = \frac{1}{1 + e^{-\beta(u_{i,ex} - h)}}$$

$$u_{i,inh} = \frac{\bar{N}_{inh-ex}}{N_{i,inh-ex}} \sum_j w_{ij} x_{j,ex}$$

$$x_{i,inh} = \frac{1}{1 + e^{-\beta(u_{i,inh} - h)}}$$

where w_{ij} is the synaptic weight, $u_{i,ex}$ and $u_{i,inh}$, $x_{i,ex}$ and $x_{i,inh}$ are the membrane potentials and mean firing frequencies for the excitatory and inhibitory neurons respectively. $N_{i,ex-ex}$, $N_{i,ex-inh}$, $N_{i,inh-ex}$ are the numbers of synaptic connections, \bar{N}_{ex-ex} , \bar{N}_{ex-inh} , \bar{N}_{inh-ex} are mean numbers of synaptic connections. Note that the membrane potentials are linear sums of the input coupled with the synaptic weight, except for the term which normalizes for the synaptic number variation. This normalization term rectifies the quantitative heterogeneity introduced by the number of synaptic connections, leaving only the spatial heterogeneity.

To introduce a generic heterogeneity, the network is subdivided into two kinds of columns, where the columns are distributed in an alternating checkerboard pattern. There are 126 x 126 neurons in the excitatory layer, which are divided into columns of 21 x 21 dimension.

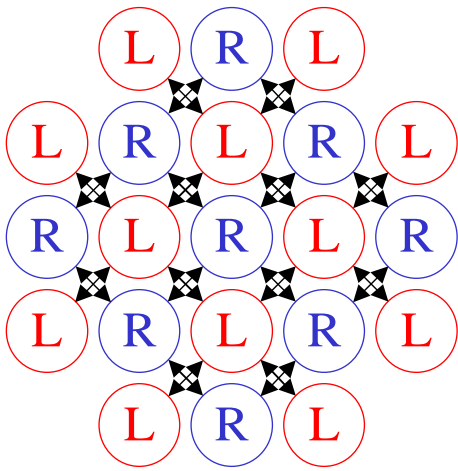


Fig.2 Generic spatial heterogeneity in the network. Here, a generic spatial heterogeneity is introduced into the network by dividing the cortical surface into two kinds of alternating columns. Long range excitatory connections are formed between the neurons in neighboring columns of the same kind. The two kinds of columns are marked as R and L for convenience.

The sigmoid function for excitatory and inhibitory neurons are given different slopes (Fig.3)

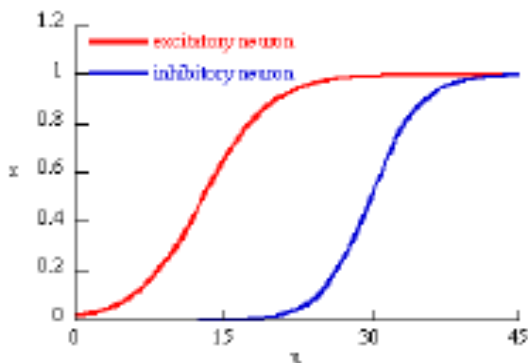


Fig.3 The Sigmoid Function for the excitatory and inhibitory neurons.

In the simulations, we have assumed the periodic boundary condition. Although the actual cortical state is naturally does not obey the periodic boundary condition, this arrangement avoids the border effect and is expected to represent a homogeneous cortical condition.

3. RESULTS

When a random initial activity pattern was given, the network soon went into a steady state, which was characterized by quasi-periodic fluctuation. An arbitrarily chosen excitatory neuron was monitored. Although the PCS is originally defined only for the vigorous firing state[5], we generalized the concept here to preferred states for arbitrary levels of activity. This generalization can be justified on the ground that it is possible to ask what kind of cortical state is preferred for any level of neural activity, although a particular functional significance might not be assigned to the intermediate activation level.

We monitored the activity of a single arbitrary neuron and analyzed its PCS. Figure 4 shows the typical time course of the neural activity. Here, the neuron was taken to be the one at (52, 52).

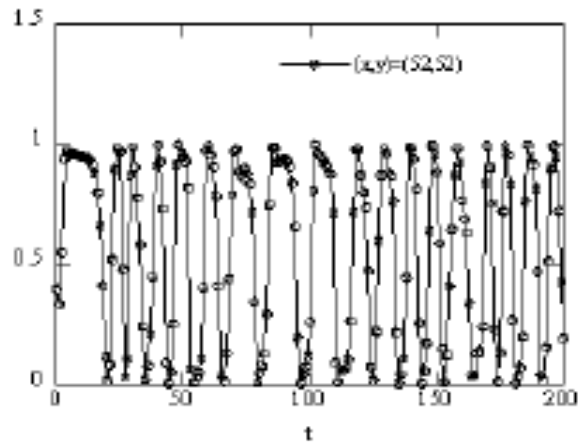


Fig.4 A typical time course of neural activity.

Fig.5 shows the spatial distribution of the neural activities (preferred cortical state, PCS) when the neuron at (52, 52) was firing vigorously ($x>0.75$) or silently ($x<0.25$), or intermediately ($0.5<x<0.75$ or $0.25<x<0.5$). The activities are shown only for the excitatory neurons in a color scale between 0-1. As is demonstrated in figures 5(a) and 5(d), There is a clear spatial heterogeneity of the PCS in the case of vigorous firing ($x>0.75$) and low firing ($x<0.25$). This heterogeneity clearly reflects the generic synaptic heterogeneity introduced as in Fig.2. On the other hand, there is no clear indication of a spatial heterogeneity when the neural activity is in the intermediate range (Fig 5(b) and Fig.5(c)).

In the second set of simulations, we gave the network an external input, specifically to stimulate a particular neuron. Fig.6 shows the spatial distribution of neural activity when the neuron (52, 52) was stimulated. The average is taken for the time when the neuron exhibited an activity of >0.75 . There is a recruitment of the neighboring columns of the same kind, as is seen in the spontaneous activity of Fig.5(a), although the activity is less pronounced. There is in addition an over-all

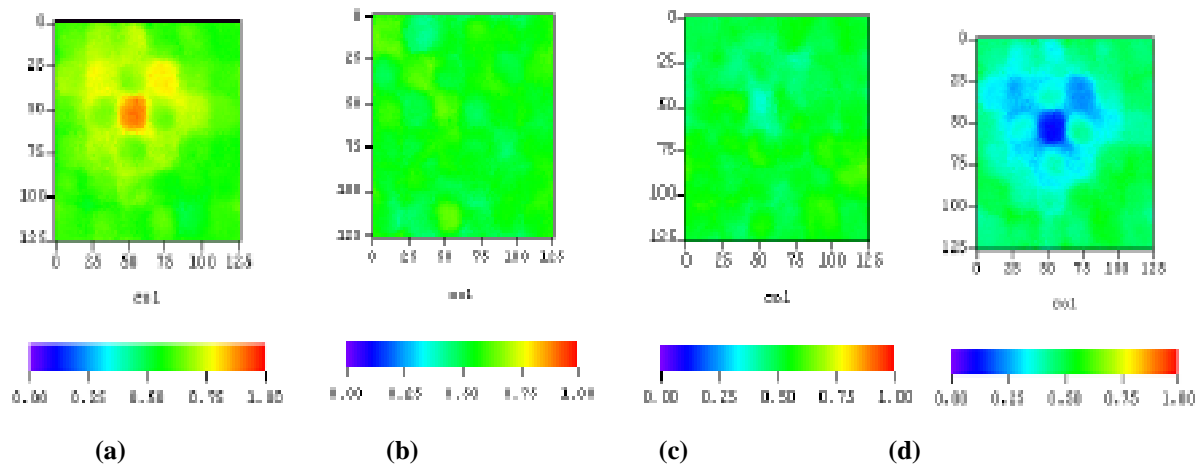


Fig.5 The preferred cortical states for different level of activation for (a) $0.75 < x$, (b) $0.5 < x < 0.75$, (c) $0.25 < x < 0.5$, (d) $x < 0.25$. The activity was measured for the neuron at (52, 52).

inhibition of the surrounding area. Thus, the PCS is induced by an external stimulation, with a possibly higher activity contrast. This result is consistent with the finding of [5].

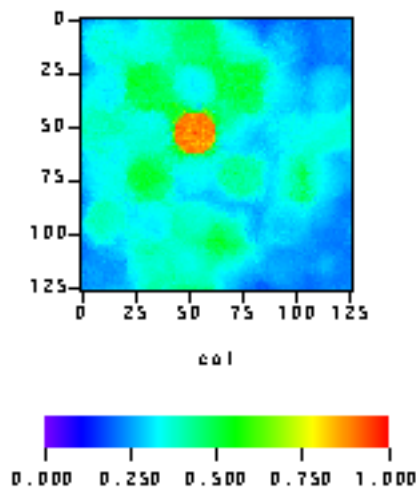


Fig.6 The preferred cortical state in the case of external stimulation. The neuron at (52, 52) was evoked to high level of activity by external stimulus. The preferred cortical state ($x > 0.75$) shows a much pronounced surround inhibition.

The strong spatial constraint as expressed in the PCS has some implications for the computational capability of the network. Specifically, it is interesting to ask how the network is able to respond to a dynamically changing stimulus, when the sequence of inputs might lead to a incompatible PCS. If the input was given as $I(x, y, t)$, where (x,y) are spatial coordinates and t is the time, there is in general a compatibility problem between $I(x, y, t)$ and $I(x, y, t + \Delta t)$. Inconsistencies between the PCS for these successive inputs might lead to the inability of the network to dynamically adapt to the input.

In order to investigate the dynamical property of the network, we gave the network a moving a stimulus. Specifically, a focused moving stimulus was given

sequentially to one of the alternating columns. In order to estimate the ability of the system to adapt dynamically to the moving stimulus, the measure $x_R - x_L$ (the difference in between the neural activities in the R type and L type columns, averaged over the neighboring columns) were calculated. This particular measure can be interpreted, for example, as one describing the competition between two eyes in binocular rivalry [6], [9-11].

Fig.7 shows the histogram of the duration time $t_{R > L}$, which is defined as the length of the period when $x_R - x_L$ is positive. T1 represents the condition where the input is given to a particular column for 1 time unit, before moving on to the next column of the same kind. In the T7 condition, the same stimulus is repeated for 7 time units for the same column before moving on to the next column. Namely, the stimulus moves 7 times faster in the T1 condition. The stimulus is moved in a specified band (the y coordinate is constant) and the data is taken for the neurons located at 2 columns apart from the movement track.

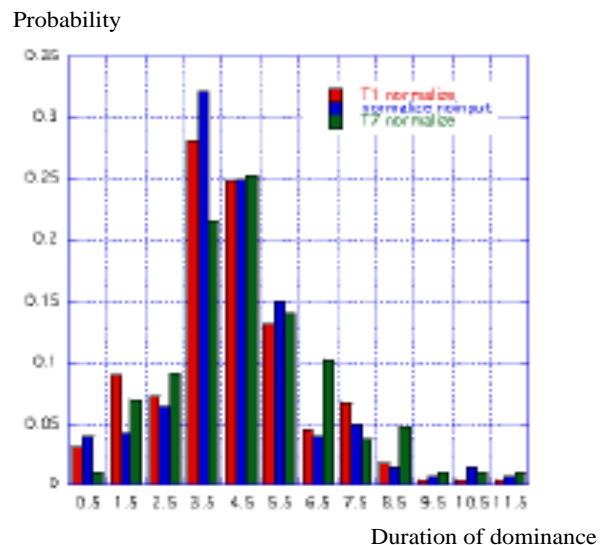


Fig.7 Probability distribution of the duration time of dominance.

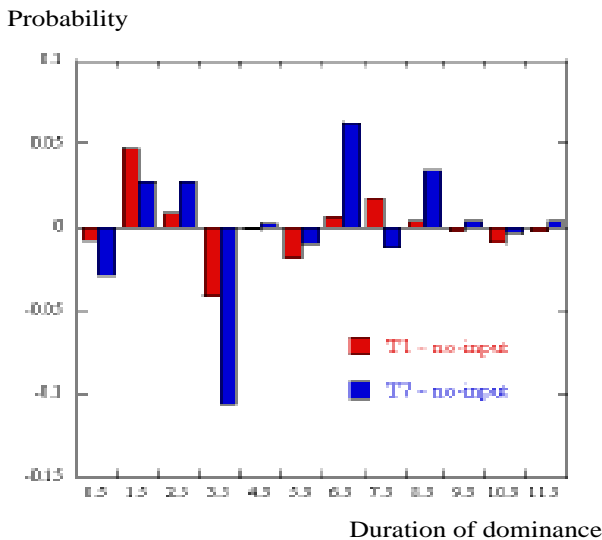


Fig.8. Shift of the duration time

Fig.8 represents the shift of the duration time in the moving stimulus conditions compared with the spontaneous firing state. As seen in Fig.7, the spontaneous fluctuation of the network leads to a dominance duration time distribution which peaks at 3-4 seconds. When the moving stimulus is given, the peak is shifted towards the duration time specific to the moving stimulus. In the T1 condition, the evoked duration time is around 1 units, whereas in the T7 condition it is around 7 units. We can estimate the ability of the system to adapt to these conditions by the shift of the duration time.

As can be seen from Figs.7 and 8, the duration time peak does not shift completely to the one specified by the external input. The system adapts to the moving stimulus within the constraint specified by the inherent fluctuation of neural activities.

4. DISCUSSIONS

The selectivity of neural response to external stimuli has been employed in the analysis of biological and artificial neural networks [7-8]. Selectivity in this sense is a mapping between a certain feature of the stimuli (such as the orientation of the bar, movement in a certain direction, a particular wavelength, etc.) and a subset of the (in general) spatio-temporal neural activity pattern. The concept of stimulus selectivity is certainly helpful in understanding the brain, which is at its outset a pattern matching machine.

However, even if we take the view that stimulus selectivity is important in the analysis of the brain and the construction of the artificial neural network, it is the internal connection between the neurons in the network which makes the selectivity possible. It is therefore necessary to "internalize" the stimulus selectivity, by spelling out the necessary condition for a particular neuron to fire in terms of the properties of the internal state.

The PCS is in this sense an internalization of the stimulus selectivity. In [5], it was observed that the PCS in evoked state is the same as the PCS in the spontaneously firing state. This result suggests that the

neural network does not go into the selectively firing state *de novo*. The selectively firing state supported by the PCS is just one of the possible transient states in the spontaneous firing state. The appropriate stimulus input just increases the probability for the network to stay in the PCS, which in turn induces the activity of the neuron with that selectivity.

When a particular stimulus induces a neuron to fire, it does so by inducing the neural network to go into the PCS. In other words, an external stimulus is able to evoke the firing of a particular neuron only when it successfully recruits the PCS, which is already present in the spontaneous firing state. Since the PCS is the result of the particular spatial pattern of the synaptic connection in the network, the spatial constraint as expressed in the PCS plays an important role in such a computation.

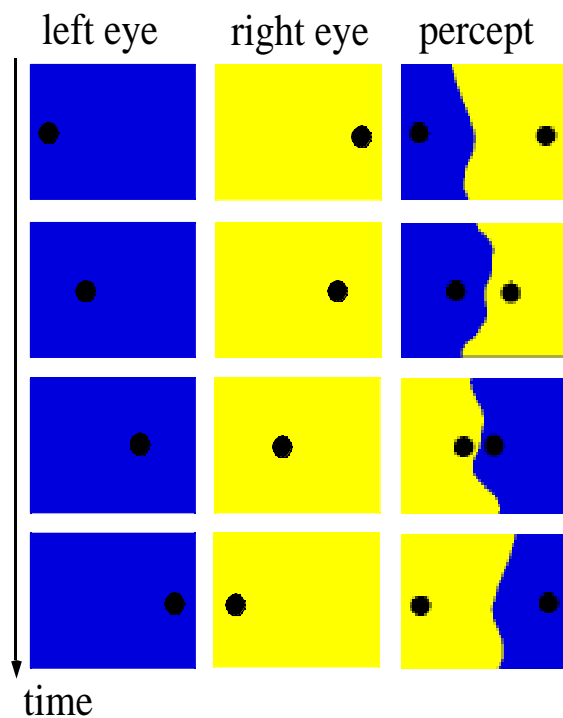


Fig.9 The dynamical adaptability of the visual system in binocular rivalry. The salient features (moving circles) are always present in the visual percept, even if that means the highly complicated modulation of the ocular dominance pattern. Adapted from [11].

The neural network somehow has to adapt to the changing stimulus environment within such a spatial constraint. In binocular rivalry [6]. [9-11], the visual system is able to adapt to the dynamical change in the environment, so that, for example, the salient features are always present in the conscious percept, even if that means that the ocular dominance pattern should be very flexibly modulated (Fig.9). In order to realize such a dynamical adaptability, the neural network has somehow to overcome the inconsistency in the PCS for the neurons selective for the successive stimulus inputs. In biological system from cell motility[12] to the nervous system[13], spontaneous fluctuation is known to correlate with the ability of the system to adapt dynamically to the changing environment. Our

simulation results suggest that the adaptability of the system might correlate with the inherent fluctuation in the spontaneous state, which induces the system to go through transitions between possibly multiple PCSs.

One of the mysteries of the property of neural firing in the brain is the functional significance of the spontaneous activity and the variability of neural activity in the evoked state[14]. In a conventional digital computer, the level of variability observed in cortical neural activities would not be permissive, as it would lead to frequent computational errors. Apart from a role in the stochastic resonance [14], spontaneous variability of neural activities has not been considered in the context of positive functionality. Our result suggests that spontaneous fluctuation of neural firing might contribute to the functionality of the system by providing the conditions necessary for the system to dynamically adapt to the environment.

5. CONCLUSIONS

In this paper, we investigated the properties of the preferred cortical state (PCS) in the spontaneously firing state by building a generic neural network model. We introduced a generic spatial heterogeneity by assuming two types of columns to be distributed in a checkerboard pattern. The PCS for the vigorously firing state as well as the silent state was found to be highly heterogeneous, consistent with the synaptic heterogeneity introduced in the system. On the other hand, the PCS for the intermediately firing state is less heterogeneous in space. This property of PCS sets a spatial constraint on the computational ability of the network, e.g., in terms of the consistency relations between the PCS for different neurons with different stimulus selectivity.

When the activity of a particular neuron was evoked by external stimulus, the PCS in this condition resembled that in the spontaneous firing state. This result is consistent with the hypothesis that in an evoked state, the neural network does not build the response *de novo* but picks up one of the PCS already present in the spontaneous state.

The ability of the neural network to adapt dynamically to the changing environment is expected to be highly constrained by the multiple PCSs and the mutual consistency relationship between them. When our generic neural network was provided with a moving stimulus, it was found that the dynamical ability of the network correlates with the inherent fluctuation of neural firing in the system. It is possible that the neural network adapts to the changing environment by modulating its inherent fluctuation. If this conjecture is the case, a neural network would not be able to adapt to a change in the environment without a certain level of spontaneous variability of the activity of the its neurons. This argument might provide *raison d'être* for the existence of the observed variability of neural firing in the central nervous system.

Acknowledgments. We thank Yuichi Mitsudo, Kei Omata, Ayako Onzo, Takayasu Sekine, Tamami Sudo, Arnaud Vaudour and Mario Tokoro for helpful discussions.

6. REFERENCES

- [1] Fujita, I., Tanaka, K., Ito, M. and Cheng, K. Columns for visual features of objects in inferotemporal cortex. *Nature* 360,343-346 (1992)
- [2] Turing, A. Proc. Lond. Math. Soc. (2) 42, 230-265 (1936-7); correction *ibid.* 43, 544-546 (1937).
- [3] Destexhe, A., Contreas, D., Steriade, M. J. *Neurosci.* 19, 4595 (1999)
- [4] Lampl, I., Reichova, I, Ferster, D. *Neuron* 22, 361 (1999)
- [5] Tsodyks, M., Kenet, T., Grinvald, A. & Arieli, A. *Science* 286, 1943-1946 (1999)
- [6] Taya, F. and Mogi, K. *Consciousness and Cognition* 9, 53. (2000)
- [7] Hubel, D.H. and Wiesel, T.N. *Journal of Physiology*, 165,559--568, (1963).
- [8] von der Heydt, R., Peterhans, E. and Baumgartner, G. *Science* 224:1260-1262. (1984).
- [9] Logothetis, N.K. et al. *Science*. 245,761-3.(1989)
- [10] Leopold D.A. et al. *Nature* 379, 549-53.(1996)
- [11] Taya, F. and Mogi, K. Submitted. (2002)
- [12] Ishijima A., Kojima H., Funatsu T., Tokunaga M., Higuchi H., Tanaka H., Yanagida T. *Cell* 92,161-171. (1998)
- [13] Levin, J.E. and Miller, J.P. *Nature*,380:165-168. (1996).
- [14] Britten K.H., Shadlen M.N., Newsome W.T., Movshon J.A. *Visual Neurosci* 10:1157-1169. (1993)