Encoding and Decoding Target Locations With Waves in the Turtle Visual Cortex

Xiuxia Du*, Bijoy K. Ghosh, *Fellow, IEEE*, and Philip Ulinski

***Abstract*—Visual stimuli elicit waves of activity that propagate across the visual cortex of turtles. An earlier study showed that these waves encode information about the positions of stimuli in visual space. This paper addresses the question of how this information can be decoded from the waves. Windowing techniques were used to temporally localize information contained in the wave. Sliding encoding windows were used to represent waves of activity as low dimensional temporal strands in an appropriate space. Expanding detection window (EDW) or sliding detection window (SDW) techniques were combined with statistical hypothesis testing to discriminate input stimuli. Detection based on an EDW was more reliable than detection based on a SDW. Detection performance improved at a very early stage of the cortical response as the length of the detection window is increased. The property of intrinsic noise was explicitly considered. Assuming that the noise is colored provided a more reliable estimate than did the assumption of a white noise in the cortical output.**

***Index Terms*—B space representation, Karhunen–Loeve (KL) decomposition, statistical hypothesis testing, turtle visual cortex.**

I. Introduction

SENSORY stimuli produce waves of activity that propagate across the cerebral cortex [1]. Such waves have been described in the primary visual [2], [3] and somatosensory cortices of mammals [4]–[6], including humans [7]. Propagating waves have also recently been studied in the visual cortex of freshwater turtles using both multielectrode and voltage sensitive dye methods [8]–[10]. Analysis of the waves of activity occurring in turtle visual cortex raised the possibility that these waves encode information about stimuli in visual space. Nenadic *et al.* [11] used a large-scale model of turtle visual cortex to simulate the responses of the cortex to stationary and moving visual inputs. The response waves can be viewed as movies consisting of a sequence of frames, each of which represents the level of depolarization seen at an array of points in the cortex. To decode the information encoded in the waves, Nenadic *et al.* [11] used a two-step Karhunen–Loeve (KL) decomposition. The first decomposition represents a wave as a linear combination of a series of spatial modes with time-varying coefficients. Thus, the wave is adequately represented (as has already been shown by Senseman and Robbins [10]) by a trajectory in a phase space called A-space. Most of the energy contained in the original wave can be captured by the decomposition coefficients corresponding to the first three principal modes. A further reduction of the dimensionality of the wave is achieved by a second KL decomposition which maps the trajectory in A-space into a point in a low-dimension space called B-space. B-space is spanned by temporal modes. Maximum likelihood estimation methods [12] were used to show that position and velocity information in the visual stimulus could be reliably estimated from the spatio-temporal dynamics of the cortical wave.

This paper continues the analysis of information encoding and decoding in cortical waves by addressing three points that were not considered by Nenadic *et al.* [11]. Firstly, Nenadic *et al.* [11] used cortical responses that lasted 1500 ms to estimate visual inputs. However, it is clear that turtles can make reliable visual discriminations in much less than 1500 ms. Turtles can be trained to discriminate targets that are presented for times ranging between 107 ms and 148 ms, depending upon the size and color of the target [13]. This raises the question of how much time is required to make accurate estimates of visual stimuli from the spatio-temporal dynamics of the waves. The problem was considered here by using windowing techniques, including a *sliding encoding window* in the wave encoding process and *expanding detection window* (EDW) and *sliding detection window* (SDW) techniques in the information decoding process, to estimate the position of stimuli in visual space. These techniques make it possible to localize the information contained in individual time segments of the cortical response. The estimation approach in [11] can be considered as a single detection window that spans from time 0 to 1 500 ms. Secondly, the model used by Nenadic *et al.* [11] was deterministic in that no noise source was included in the model. Variability was introduced by perturbing the positions of individual neurons in the model between simulations. However, the stochastic nature of transmitter release and the opening and closing of ion channels is a source of noise in the real cortex. This paper explicitly considers the effects of neural noise on estimating visual stimuli. A Gaussian white noise current was injected into the neurons in the model. The noise contained in the cortical output was shown to be colored and highly correlated even though the injected noise current was white. A detection approach that uses a colored noise model yields a better performance than does a model that assumes white noise. Thirdly, this paper examines the effect of window size on the encoding and decoding process.

Manuscript received June 27, 2004; revised July 18, 2004. This work was supported in part by the National Science Foundation (NSF) through the Collaborative Research Computational Neuroscience Program under Grant EIA-0218186. *Asterisk indicates corresponding author.*

*X. Du is with the Department of Electrical and Systems Engineering, Washington University, St. Louis, MO 63130 USA (e-mail: duxiuxia@netra.wustl.edu).

B. K. Ghosh is with the Department of Electrical and Systems Engineering, Washington University, St. Louis, MO 63130 USA.

P. Ulinski is with the Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL 60637 USA.

Digital Object Identifier 10.1109/TBME.2004.841262

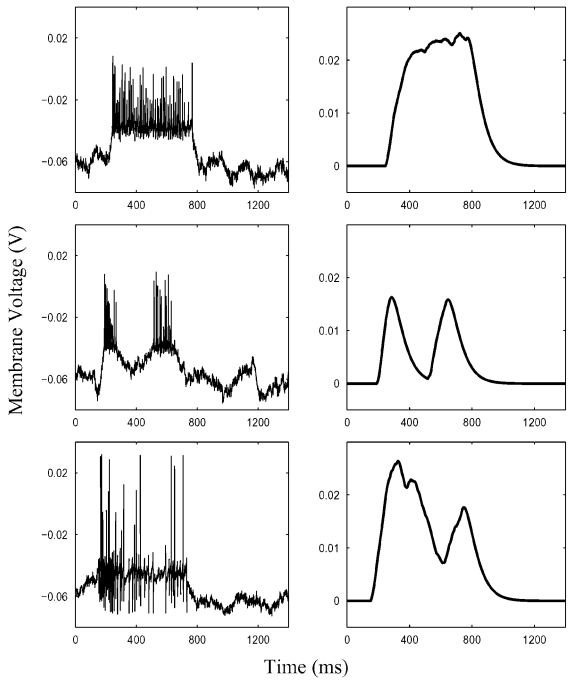

Fig. 1. Responses of model pyramidal cells. The traces in the left column show voltage traces from three model pyramidal cells. The traces in the right column show the smoothed spike rate of the same three cells.

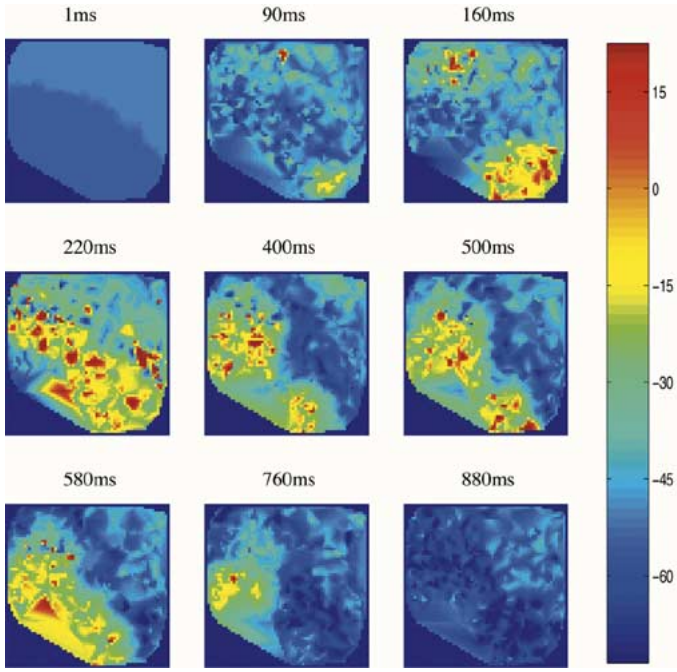

Fig. 2. Responses of the cortical model. This figure shows selected frames from the spatio-temporal response of the model in response to a left stimulus. Color scale indicates smoothed membrane potential. The values near the color bar are in millivolts. Times indicate time elapsed from stimulus presentation.

II. Large-Scale Model of Turtle Visual Cortex

This study used the large-scale model of turtle visual cortex that was described in detail by Nenadic *et al.* in [14]. In brief, the model consists of 744 neurons in the visual cortex and 201 neurons in the dorsal lateral geniculate complex. Pyramidal cells constitute the largest population of neurons in the cortex in the model (679 pyramidal neurons); the remaining cells are two types of inhibitory interneurons. Cortical neurons are distributed in a two-dimensional array that captures the spatial distribution of the three major populations of cortical neurons. Each cortical neuron is represented by a multicompartmental model based on the anatomy and physiology of that type of cortical neuron. The soma compartment of each model neuron contains voltage-gated conductances sufficient to generate action potentials. These currents are specified by Hodgkin-Huxley equations. Excitatory and inhibitory synaptic conductances establish interactions between individual neurons. Geniculate neurons are modeled as single compartments with ionic currents that generate action potentials following the injection of a stimulating current in the neuron. Generation of an action potential produces a signal that propagates across the model cortex with a velocity based on the known conduction velocity of geniculate afferents. Stimulation of points along the left, center, and right edge of the horizontal meridian of visual space was simulated by electrical activation of clusters of 20 geniculate neurons situated at the left, center and right edges, respectively, of the line of geniculate neurons. Stationary stimuli were simulated by presenting a 150 ms square current pulse to each of the geniculate neuron in a cluster. A zero-mean Gaussian noise current with a standard deviation of 4 nA was injected into each neuron. The left column in Fig. 1 shows the spike trains of three representative pyramidal neurons in response to a left stimulus. The right column shows the smoothed spike rates of the same three neurons (see the Appendix I for details). In this example, a movie of the whole cortical response to a stimulus was constructed by spatial interpolation of the voltage values between neurons at each time, t. Fig. 2 shows nine snapshots from a representative movie. A total of 100 sample waves were generated for each of the three stimulus locations. In this paper, "wave" and "movie" are used interchangeably to refer to the spatio-temporal response signal of the cortex.

III. Localizing Information Using Windowing Techniques

Nenadic *et al.* [11] showed that analysis of a movie lasting 1500 ms produced a single point in the B-space. It is reasonable to expect that the value of this point might depend upon the length of the movie analyzed. For example, very short segments of the cortical responses may not contain enough information to reliably estimate the position of a stimulus in visual space from the dynamics of the cortical response. Analysis of a short segment of the corresponding movie should, then, produce a point in B-space different from the point that results from analysis of a longer segment of the movie. In this study, the two-step decomposition developed in [11] was modified to analyze segments of the cortical response using *sliding encoding window* techniques. As shown in Fig. 3, the time axis was covered by equal-length, overlapping encoding windows and the double KL-decomposition was applied to the segment of the spike rate signal within each window. Both the starting and ending times of the windows changed while the length of the window remained constant. Each segment of the cortical wave was, thus, mapped to a point in B-space. Plotting images of successive windows produced a sequence

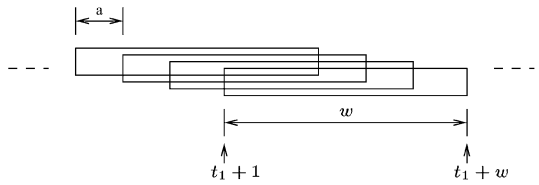

Fig. 3. Encoding window. The time axis is covered by equal-length, overlapping, sliding encoding windows. Both the starting and ending times of the windows slide over the time axis while the length of the window remains constant. a is the amount of time that the window slides and w is the width of each encoding window.

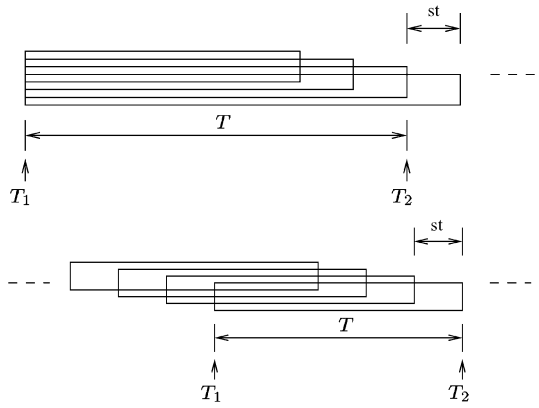

Fig. 4. Decoding windows applied over the β-strand representation of cortical waves. T_1 and T_2 are the starting and ending time of the detection windows respectively. T is the width of the detection window. The amount of time that the detection window slides is equal to st. (Top) EDW with increasing duration. The starting time remains the same at $T_1 = 1$ ms while the ending time changes. (Bottom) SDW with equal duration. Both T_1 and T_2 slide over the time axis.

of points, or β-strand (see Fig. 6). This is a vector-valued function of time and is an alternative way to represent the original movie in a space with a much lower dimension. In the decoding process, we used EDW and SDW techniques to cover the β-strands and applied detection theory to the β-strand segment within each detection window (see Fig. 4). For EDWs, the detection algorithm was applied to a series of time windows of increasing duration. The starting time remained the same at $T_1 = 1$ ms while the ending time changed. For SDWs, the detection algorithm was applied to a series of SDWs of equal length. Our motivation in considering these two cases was to address two potential mechanisms by which the cortex processes information. SDWs represent the possibility that the cortical information processing system has a transient memory mechanism of very short duration, equal to the width of the detection window. EDWs, on the other hand, represent the possibility that information accumulates in the cortex over a relatively longer time window. In particular, we analyzed the extreme case when the system remembers the entire past. We next describe the process of encoding cortical waves and decoding them from their β-strand representations.

A. Encoding Cortical Waves With β Strands

For a particular input stimulus, let $I(t,n), 0 \le t \le T, 1 \le n \le N$ denote the smoothed spike rate of the cell as a spatio-temporal response signal, where t is time and n is the index of the pyramidal neuron. $I(t,n)$ can be viewed as a matrix. The tth row represents the spike rate of each neuron at time t in response to a particular stimulus. The nth column corresponds to pyramidal neuron n. Let the length of each time window be w and let $I(t,n), t_1+1 \le t \le t_1+w, t_1 = 0, a, 2a, \ldots$ be the response signals for different time windows. Here a is the amount of time that the encoding windows slide (see Fig. 3). Let M denote the total number of cortical response movies in response to stimuli in the left, center, and right visual field. For the kth, $k = 1, 2, \ldots, M$ movie, the spatio-temporal signal in this time window can be viewed as a collection of vectors $\{I^k(t_1+1), I^k(t_1+2), \ldots, I^k(t_1+w)\}$ where $I^k(t_1+i) \in \mathrm{R}^{1\times N}, i = 1, 2, \ldots, w$. The dimensionality of the cortical response is reduced by two KL transforms into A-space and B-space, respectively. We first describe the KL transform into A-space. The covariance matrix $C_1 \in \mathrm{R}^{N\times N}$ for a family of M movies is calculated as

$$C_1 = \frac{1}{Mw}\sum_{k=1}^{M}\sum_{i=1}^{w}(I^k(t_1+i))^T(I^k(t_1+i)) \qquad (1)$$

where $(I^k(t_1+i))^T$ is the transpose of $I^k(t_1+i)$. The matrix C_1 is symmetric and positive semi-definite, so its eigenvalues are all real and nonnegative and the corresponding eigenvectors form an orthonormal basis in R^N. The eigenvectors corresponding to the largest p eigenvalues of C_1 are called the principal eigenvectors, or modes, and the pth-order successive reconstruction of the spatio-temporal signal $I^k(t) \in \mathrm{R}^{1\times N}$ is given by

$$\hat{I}^k(t_1+i) = \sum_{j=1}^{p}\alpha_j^k(t_1+i)\phi_j^T, \quad i = 1, 2, \ldots, w \qquad (2)$$

where $\phi_j \in \mathrm{R}^{N\times 1}$ is the jth principal mode, the time coefficients $\alpha_j^k(t_1+i)$ are given by $\alpha_j^k(t_1+i) = \langle I^k(t_1+i), \phi_j^T\rangle$ and $\langle\cdot,\cdot\rangle$ stands for the standard inner product. The coefficients $\alpha_j^k(t_1+i)$ of the KL-decomposition are uncorrelated in terms of j and we call $\alpha_j^k(t), t_1+1 \le t \le t_1+w, 1 \le j \le p$ the pth-order A-space representation of the movie segment within the corresponding time window for the kth movie. The vector function

$$\left[\alpha_1^k(t), \alpha_2^k(t), \ldots, \alpha_p^k(t)\right], \quad t_1+1 \le t \le t_1+w \qquad (3)$$

can be viewed as a sample function of a vector random process since the model cortex is disturbed by random noise. Statistical analysis of a random process can be facilitated if the process is further parameterized using a second KL decomposition. Let

$$\gamma_j^k = \begin{bmatrix}\alpha_j^k(t_1+1)\\ \alpha_j^k(t_1+2)\\ \vdots \\ \alpha_j^k(t_1+w)\end{bmatrix}, \quad (\xi^k)^T = \begin{bmatrix}\gamma_1^k\\ \gamma_2^k\\ \vdots\\ \gamma_p^k\end{bmatrix} \qquad (4)$$

where $j = 1, 2, \ldots, p$. Calculating the covariance matrix as in (1), we have

$$C_2 = \frac{1}{M}\sum_{k=1}^{M}(\xi^k)^T(\xi^k). \qquad (5)$$

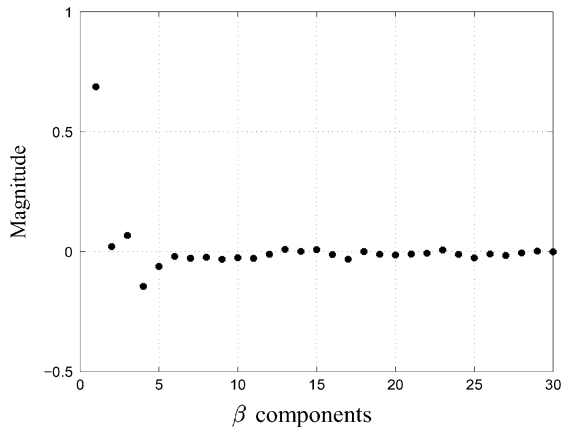

Fig. 5. Weighting coefficients for B space representation. This figure shows the first 30 (out of 300) β components of the B space representation of a particular movie within time window 211–220 ms. The components are ordered in terms of the descending eigenvalues of the correlation matrix C_2 in (5). The horizontal axis represents the indices of the β components and the vertical axis represents the value of the β components. It is clear that the first few components can represent most of the information contained in the complete β vector.

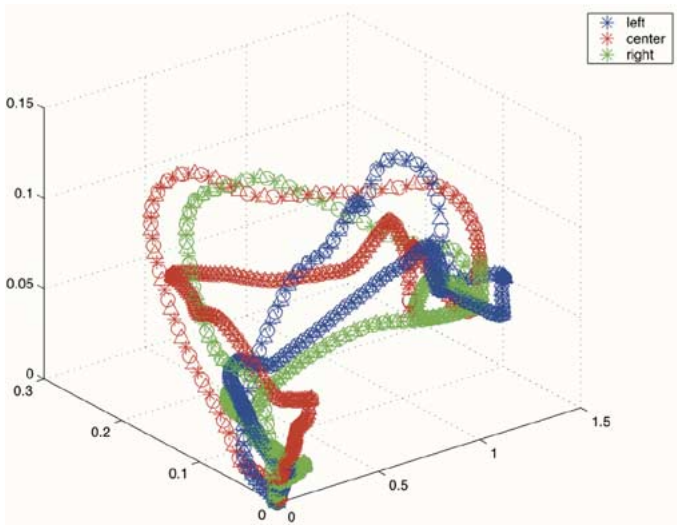

Fig. 6. Mean third-order B-space representation of cortical responses. This figure shows the mean β strands for cortical responses to left, center and right stimulus. Each strand is the statistical mean of the β-strands resulting from 100 responses to each stimulus. The axes represent the values of the components of the β vectors. Each β-strand starts from near the origin, curves around and finally returns to the origin when the cortical waves die out. Here $w = 10$. Note the temporal evolution of the distance between the β-strands.

The qth-order successive approximation of the kth vector ξ^k is given by

$$\hat{\xi}^k = \sum_{j=1}^{q} \beta_j^k \psi_j^T \tag{6}$$

where $\psi_j, j = 1, 2, \ldots, q$ are the eigenvectors corresponding to the largest q eigenvalues of the matrix C_2. The coefficients β_j^k are found by orthogonal projection of ξ^k onto the jth eigenvector $\beta_j^k = \langle \xi^k, \psi_j^T \rangle$. In our case, each β vector has 300 components and is referred to as the B-space representation of the cortical movie within the corresponding time window. Fig. 5 shows, in the descending order of the eigenvalues of C_2, the first 30 β-vector components corresponding to time window 211–220 ms of a left-input movie generated by a flash of light from the left. Clearly, a few β components capture most of the information contained in the original movie since the rest of the β components are close to zero. Repeating the above data processing procedure for all the sliding encoding windows of a movie produces a β-strand as a vector-valued function of time. We refer to this β-strand as the B-space representation of this movie. By discarding those components that are close to zero, we obtain a low dimensional representation of the original movie segment. If, for each sliding encoding window, the first q components of each β vector are used, we say that the vector consisting of these q components is the qth-order B-space representation of the movie. The statistical mean of the β-strands of the left-, center-, and right-stimuli movies can be easily obtained.

In our analysis, we used $w = 10, M = 300, p = 679, q = 10$. The values of t_1 were chosen to be $0, 2, 4, \ldots$. Fig. 6 shows the mean β-strand for 100 presentations for stimuli presented at the left, center, and right clusters of geniculate neurons.

B. Decoding From the β Strands

In this section, we show that β-strands contain information about the location of the input stimulus. Consistent with the idea of localizing information in time, we detect the location of the visual stimulus based on different segments of the β-strands. Each segment is a detection window. We now designate the mean β-strands for the left, center, and right stimuli by $s_1(t), s_2(t)$, and $s_3(t)$, respectively. Let $r(t)$ be the β-strand of an arbitrary cortical movie in response to a stimulus from an unknown location. We assume that the location of the stimulus is unknown, but restricted to either left, center or right. Finally, we denote the start and end times of the detection window as T_1 and T_2, respectively, so $T = T_1 - T_2$ is the size of the detection window.

Intuitively, discriminating the three stimuli from each other should be easiest when the three mean β-strands $s_i(t)$, $i = 1, 2, 3$ are farthest from each other. The distance between two β-strands can be calculated as:

$$d_{ij} = \sum_{t=T_1}^{T_2} \sum_{k=1}^{q} \left(s_i^k(t) - s_j^k(t) \right)^2, \quad i, j = 1, 2, 3 \tag{7}$$

where k is the index to the components of $s_i(t), i = 1, 2, 3$. Fig. 7 is a plot of this distance as a function of T_2 for the SDW technique with $T = 99$ ms. The curves represent $d_{12} (= d_{21})$, $d_{23} (= d_{32})$, and $d_{31} (= d_{13})$, respectively. All three distances reach their peaks between $T_2 = 200$–300 ms and then decline. Since the origin of the wave has a latency of approximately 20 ms after stimulus presentation, these plots suggest that the information content of cortical waves reaches a maximal value 180–280 ms after the onset of the wave. The further apart the average β-strand segments of $s_i(t), i = 1, 2, 3$ are, the easier it is to discriminate an arbitrary β-strand. Fig. 7 indicates that 180–280 ms will be the time period when the discrimination is most reliable, which is confirmed by the detection results described next. It is worth mentioning that the distance between left and right β-strands is smaller than that between center and

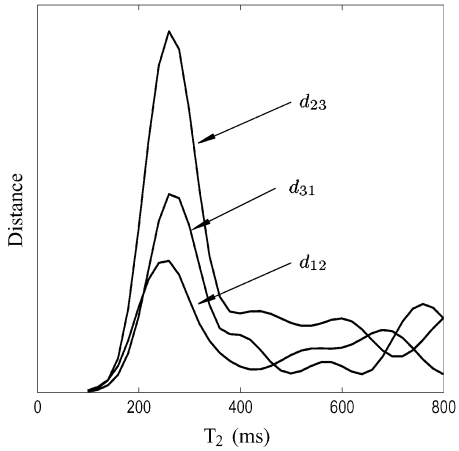

Fig. 7. Distance between the corresponding segments of two average β-strands vs. ending time, T_2. Here, the distance is calculated using (7) and T_1 and T_2 are the starting and ending time of SDWs.

right β-strands. This is an indication that the physical separation between visual stimuli might not necessarily be translated to the separation of the β-strands. How they are actually related is unknown at this moment and requires further investigation.

Replacing $s_i(t)$ by $r(t)$, the β-strand representation of an arbitrary wave (generated from an input whose location is unknown, but restricted to left, center, or right), in (7), we obtain d_1, d_2, and d_3 as the distance between $r(t)$ and $s_1(t), s_2(t), s_3(t)$ respectively as follows:

$$d_i = \sum_{t=T_1}^{T_2} \sum_{k=1}^{q} \left(r^k(t) - s_i^k(t)\right)^2 \tag{8}$$

for $i = 1, 2, 3$. Our goal is to find out which of the three β-strands s_1, s_2, s_3 is closest to the strand $r(t)$. The smallest of the three distances d_1, d_2, d_3 indicates the location of the input stimulus, i.e., if d_1 is the smallest, there is a high probability that the stimulus is from the left, and similarly for d_2, d_3 and the center, right stimuli, respectively. Fig. 8 shows d_1, d_2, and d_3 for each of the 300 movies. Identifying the correct input stimulus for a given movie is easiest when one of the distances is closest to zero and the other two are significantly larger. The reliability of the discrimination is quantified in Fig. 12(A)–(B) which shows the probability of a detection error for the EDW and the SDW technique respectively.

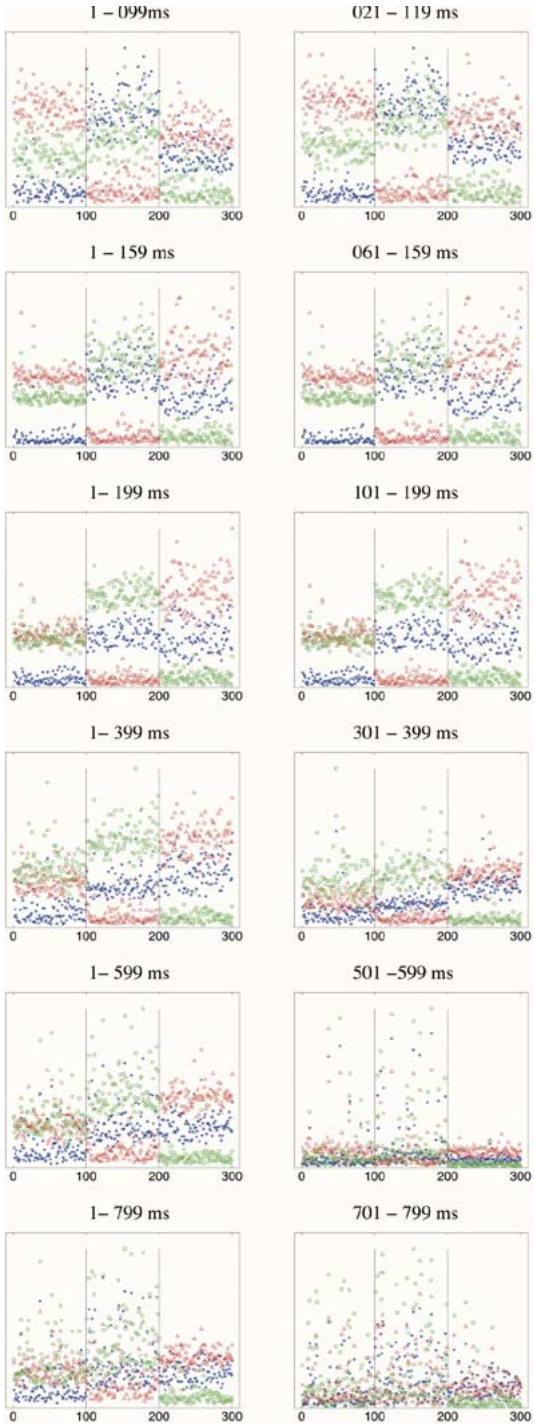

Fig. 8. Results of detection by distance. The horizontal axis is the indices of the cortical movies. Movies 1–100 correspond to a left stimulus, 101–200 to a center stimulus, and 201–300 to a right stimulus. The vertical axis shows the distance d_1, d_2, and d_3 as defined in (8) for each of the 300 movies. d_1 is shown in blue, d_2 in red and d_3 in green. The time above each figure represents the starting and ending time of the detection window. (Left) Results obtained using EDW technique. (Right) Results obtained using SDW technique.

IV. Effect of Noise

Biological noise was simulated by injecting a Gaussian white noise into the cortical neurons. With noise present, a rigorous approach is to formulate the problem of detection as a hypothesis testing problem. Let the three types of stimuli correspond to three hypotheses. H_1 denotes the hypothesis that the stimulus is from the left; H_2 denotes the hypothesis that the stimulus is from the center; H_3 denotes the hypothesis that the stimulus is from the right. Let us write

$$r(t) = s_i(t) + n(t), \quad i = 1, 2, 3 \tag{9}$$

where $n(t)$ represents the vector-valued noise process contained in the β-strand with mean 0. To investigate the effect of noise on detection performance, we solve the hypothesis testing problem assuming that the noise is colored, Gaussian and compare the results to that when the noise is assumed to be white, Gaussian.

A. Series Expansion of Sample Functions of Random Processes

The β-strand, $r(t)$, can be regarded as a sample function of a vector stochastic process. It is well known that a deterministic waveform with finite energy can be represented in terms of a series expansion. This idea can be extended to include sample functions of a random process as well. In our case, we propose to obtain a series expansion of the β-strand within a chosen detection window. This process involves finding a complete orthonormal set $\{\phi_i(t), i \in \mathbf{N}\}$ ($\mathbf{N}$ denotes the set of integers) and expanding $r(t)$ as:

$$r(t) = \text{l.i.m.}_{L\to\infty} \sum_{i=1}^{L} r_i \phi_i(t), \quad T_1 \le t \le T_2 \tag{10}$$

where $\phi_i(t)$ are vectors of the same dimension as $r(t)$. Let us denote $r^k(t)$ and $\phi_i^k(t)$ to be the kth component of the vectors $r(t)$ and $\phi_i(t)$ respectively. In (10), l.i.m. denotes "limit in the mean" which is defined as

$$\lim_{L\to\infty} E\left[\sum_{k=1}^{q}\left(r^k(t) - \sum_{i=1}^{L} r_i \phi_i^k(t)\right)^2\right] = 0, \quad T_1 \le t \le T_2 \tag{11}$$

where E is the expectation operator. The coefficients r_i, to be defined next (14), are required to be uncorrelated with each other. This is to say that, if $E[r_i] = m_i$, then we would like to have

$$E[(r_i - m_i)(r_j - m_j)] = \lambda_i \delta_{ij}. \tag{12}$$

The value of r_i^2 has a simple physical interpretation. It corresponds to the *energy* along the coordinate function, $\phi_i(t)$, in a particular sample function. It is shown in [12] that if $m_i = 0$, then λ_i is the *expected* value of the energy along $\phi_i(t)$. Clearly, $\lambda_i \ge 0$ for all i. The complete orthonormal set $\phi_i(t)$ is the solution of the integral equation:

$$\lambda_i \phi_i^k(t) = \sum_{j=1}^{q} \int_{T_1}^{T_2} K_{kj}(t,u)\phi_i^j(u)\,du \tag{13}$$

where $k = 1, 2, \ldots, q, T_1 \le t \le T_2$, and $K(t,u)$ is the covariance matrix of the noise process $n(t)$, i.e., $K_{ij}(t,u) = E[n_i(t)n_j(u)]$. Here, t and u denote time and i and j denote indices of the component of the vector noise process. In (13), λ_i is called the eigenvalue of the noise process and $\phi_i(t)$ is called the corresponding eigenfunction. Once the coordinate functions $\{\phi_i(t), i \in \mathbf{N}\}$ are obtained, one can project the sample function $r(t), T_1 \le t \le T_2$ onto $\phi_i(t)$ and obtain the coefficient r_i as

$$r_i = \sum_{k=1}^{q} \int_{T_1}^{T_2} r^k(t)\phi_i^k(t)\,dt. \tag{14}$$

Recall from Section III-A that q is the number of β components we choose for the B-space representation of the cortical movies. The νth-order representation of $r(t)$ can then be written as a vector $R = [r_1, r_2, \ldots, r_\nu]$.

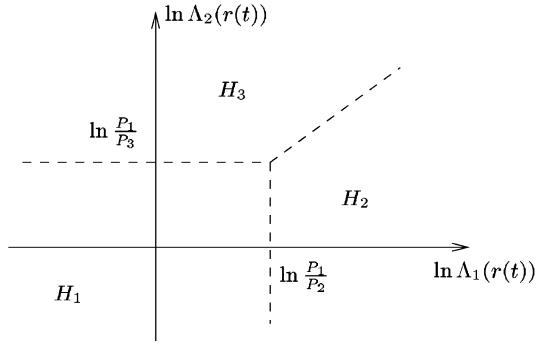

Fig. 9. Decision space divided into three regions, H_1, H_2, and H_3, in terms of the logarithm likelihood ratio (15) and (16). H_1 is the hypothesis that the visual input is from left; H_2 is the hypothesis that the visual input is from center; H_3 is the hypothesis that the visual input is from right. For any give β-strand $r(t)$, the region that the pair $\Lambda_1(r(t))$ and $\ln \Lambda_2(r(t))$ fall into in the decision space determines which hypothesis is true.

B. Hypothesis Testing

The proposed detection algorithm is based on computing conditional probability densities and choosing a decision criterion (see [12] for details). Commonly used decision criteria include the Bayes and Neyman-Pearson criteria. In this paper, we use the former for two reasons. The first is that the hypotheses are governed by probability assignments which we denote as $P_j, j = 1, 2, 3$, i.e., hypothesis H_1 occurs with probability P_1, hypothesis H_2 with probability P_2 and hypothesis H_3 with probability P_3. The second reason is that a certain cost is incurred each time an experiment is conducted. We propose to design our decision rule so that *on the average* the cost is as small as possible. It is demonstrated [12] that for a decision using the Bayes criterion, the optimum detection consists of computing the logarithm likelihood ratio and comparing it to a threshold. If we assign the cost of correct detection to be zero and that of a wrong detection to be 1, the logarithm likelihood ratio can be computed (see [12, Section 2.3]) as follows:

$$\ln \Lambda_1(r(t)) = \frac{p_{r\,|\,H_2}(R\,|\,H_2)}{p_{r\,|\,H_1}(R\,|\,H_1)} \tag{15}$$

$$\ln \Lambda_2(r(t)) = \frac{p_{r\,|\,H_3}(R\,|\,H_3)}{p_{r\,|\,H_1}(R\,|\,H_1)}. \tag{16}$$

The decision regions in the decision space are determined by the following threshold comparisons:

$$\ln \Lambda_1(r(t)) \underset{H_2 \text{ or } H_3}{\overset{H_1 \text{ or } H_3}{\gtrless}} \ln \frac{P_1}{P_2} \tag{17}$$

$$\ln \Lambda_2(r(t)) \underset{H_3 \text{ or } H_2}{\overset{H_1 \text{ or } H_2}{\gtrless}} \ln \frac{P_1}{P_3} \tag{18}$$

$$\ln \Lambda_2(r(t)) \underset{H_3 \text{ or } H_1}{\overset{H_2 \text{ or } H_1}{\gtrless}} \ln \Lambda_1(r(t)) + \ln \frac{P_2}{P_3}. \tag{19}$$

The associated decision space has been sketched in Fig. 9. For a particular strand $r(t)$, we say that the hypothesis H_1 is true, i.e., the stimulus is from the left part of the visual space, if the logarithm likelihood ratio pair falls in region H_1. Likewise, the same can be said for H_2 and H_3.

In Fig. 9, if $P_1 = P_2 = P_3 = 1/3$, the dividing line between regions H_1 and H_2 becomes the negative vertical axis, the dividing line between regions H_2 and H_3 becomes the diagonal

line which is 45 degrees counterclockwise from the positive horizontal axis, and the dividing line between regions H_3 and H_1 becomes the negative horizontal axis. In the following discussion, we assume that $P_1 = P_2 = P_3 = 1/3$. The vector noise process $n(t)$ in (9) could be either white or colored and we address these two cases next.

C. Decoding With Additive White Noise Model

If the vector noise process is white, i.e., $E[n(t)n^T(u)] = N_0 I\delta(t-u)$, where $N_0 \in \mathrm{R}$, I is the identity matrix, and $\delta(\cdot)$ is the Dirac function, the eigenfunctions of the noise process turn out to be the orthonormalization of $\{s_i(t), i = 1,2,3\}$. So, instead of solving the integral equation (13), we apply Gram-Schmidt orthogonalization procedure on $\{s_i(t), i = 1,2,3\}$ to get $\{\phi_i(t), i = 1,2,3\}$ as

$$\begin{aligned}\phi_1(t) &= s_1(t)/\mathrm{norm}(s_1(t))\\ \phi_2(t) &= \psi_2(t)/\mathrm{norm}(\psi_2(t))\\ \phi_3(t) &= \psi_3(t)/\mathrm{norm}(\psi_3(t))\end{aligned}$$

where

$$\begin{aligned}\psi_2(t) &= s_2(t) - c_1 * \phi_1(t)\\ \psi_3(t) &= s_3(t) - c_2 * \phi_1(t) - c_3 * \phi_2(t)\\ c_1 &= \mathrm{IP}(s_2(t), \phi_1(t))\\ c_2 &= \frac{\begin{vmatrix}\mathrm{IP}(s_3(t),\phi_1(t)) & \mathrm{IP}(\phi_2(t),\phi_1(t))\\ \mathrm{IP}(s_3(t),\phi_2(t)) & \mathrm{IP}(\phi_2(t),\phi_2(t))\end{vmatrix}}{\begin{vmatrix}\mathrm{IP}(\phi_1(t),\phi_1(t)) & \mathrm{IP}(\phi_2(t),\phi_1(t))\\ \mathrm{IP}(\phi_1(t),\phi_2(t)) & \mathrm{IP}(\phi_2(t),\phi_2(t))\end{vmatrix}}\\ c_3 &= \frac{\begin{vmatrix}\mathrm{IP}(\phi_1(t),\phi_1(t)) & \mathrm{IP}(s_3(t),\phi_1(t))\\ \mathrm{IP}(\phi_1(t),\phi_2(t)) & \mathrm{IP}(s_3(t),\phi_2(t))\end{vmatrix}}{\begin{vmatrix}\mathrm{IP}(\phi_1(t),\phi_1(t)) & \mathrm{IP}(\phi_2(t),\phi_1(t))\\ \mathrm{IP}(\phi_1(t),\phi_2(t)) & \mathrm{IP}(\phi_2(t),\phi_2(t))\end{vmatrix}}\end{aligned}$$

and $\mathrm{IP}(\cdot,\cdot)$ and $\mathrm{norm}(\cdot,\cdot)$ are defined, respectively, as

$$\begin{aligned}\mathrm{IP}(a(t),b(t)) &= \sum_{k=1}^{q}\int_{T_1}^{T_2} a^k(t)b^k(t)\,dt\\ \mathrm{norm}(a(t)) &= \sqrt{\mathrm{IP}(a(t),a(t))}.\end{aligned}$$

The remaining $\phi_i(t)$ consist of an arbitrary orthonormal set whose members are orthogonal to $\phi_1(t)$, $\phi_2(t)$, $\phi_3(t)$ and are chosen so that the entire set is complete. We then project $r(t)$ onto this set of orthonormal coordinate functions to generate coefficients r_i as in (14). All of the r_i except r_1, r_2, r_3 do not depend on which hypothesis is true and are statistically independent of r_1, r_2, r_3. The mean values of r_1, r_2, r_3 depend on which hypothesis is true: $E[r_i \,|\, H_j] = m_{ij}, i, j = 1,2,3$. Note also that the coefficients r_1, r_2, r_3 are uncorrelated with each other. Based on the Gaussian assumption, the logarithm likelihood ratio (15) and (16) can be calculated as

$$\begin{aligned}\ln\Lambda_1(r(t)) &= \sum_{i=1}^{3}\frac{1}{N_0}\left(r_i m_{i2} - \frac{1}{2}m_{i2}^2 - r_i m_{i1} + \frac{1}{2}m_{i1}^2\right)\\ \ln\Lambda_2(r(t)) &= \sum_{i=1}^{3}\frac{1}{N_0}\left(r_i m_{i3} - \frac{1}{2}m_{i3}^2 - r_i m_{i1} + \frac{1}{2}m_{i1}^2\right).\end{aligned}$$

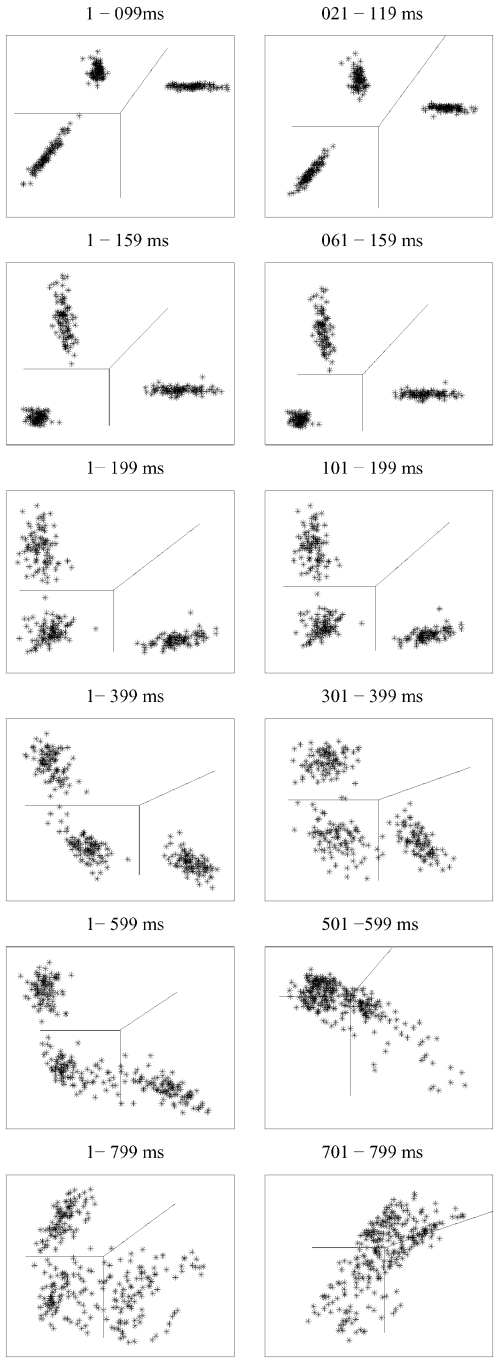

Fig. 10. Detection results in the decision space by using white noise model. Each point in the decision space is defined by the logarithm likelihood ratio (15) and (16) and represents a cortical wave in response to an input of unknown location. The region that the point falls in the decision space determines the input location. The time above each figure represents the starting and ending time of the detection window. (Left) Detection results with EDW. (Right) Detection results with SDW.

Fig. 10 shows the decision spaces for estimates with the assumption that $n(t)$ is white. In each figure, each point in the decision space represents a given cortical wave in response to an unknown stimulus. Ideally, any point corresponding to a left, center, or right stimulus should fall in the region of H_1, H_2, or H_3, respectively. Any point that does not fall in its corresponding region in the decision space produces a detection error.

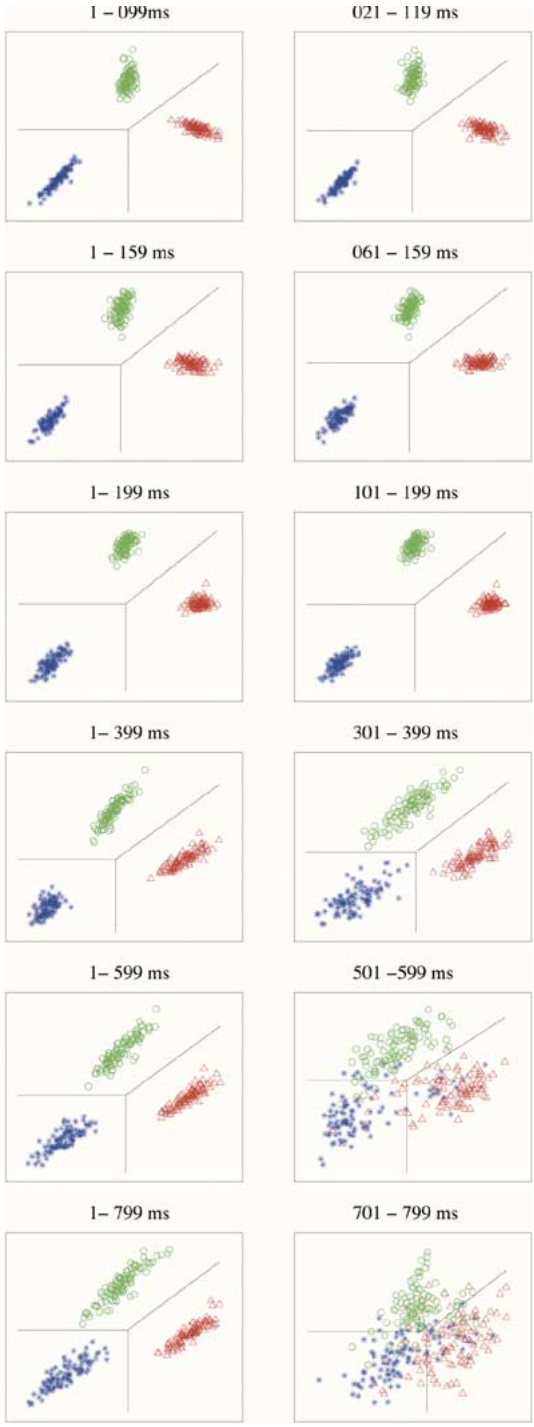

Fig. 11. Detection results in the decision space by using colored noise model. Each point in the decision space is defined by the logarithm likelihood ratio (15) and (16) and represents a cortical wave in response to an input of unknown location. The region that the point falls in the decision space determines the input location. The time above each figure represents the starting and ending time of the detection window. (Left) Detection results with EDW. (Right) Detection results with SDW.

Performing the detection over a continuum of detection windows and summing the total detection error for each detection window yields the relationship between the probability of detection error and detection window as shown in Fig. 12(C)–(D).

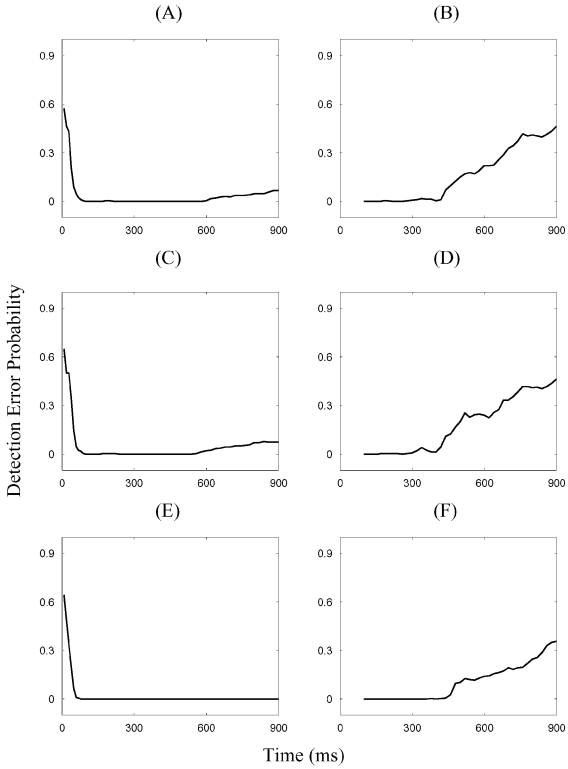

Fig. 12. Detection error probability vs. ending time T_2 of detection windows. $T = 99$ ms for the SDW technique. (A) Detection error probability of EDW by detection-by-distance. (B) Detection error probability of SDW by detection-by-distance. (C) Detection error probability of EDW by detection-by-white-noise-model. (D) Detection error probability of SDW by detection-by-white-noise-model. (E) Detection error probability of EDW by detection-by-colored-noise-model. (F) Detection error probability of SDW by detection-by-colored-noise-model.

D. Decoding With Additive Colored Noise Model

If the noise process is colored, i.e., the matrix $K(t,u) = E[n(t)n^T(u)]$ has nonzero off-diagonal elements, we solve the integral equation (13) for λ_i and ϕ_i, project $r(t)$ onto $\phi_i(t)$, and compute the logarithm likelihood ratio (15) and (16) as follows:

$$\ln \Lambda_1(r(t)) = \sum_{i=1}^{\nu} \frac{1}{\lambda_i} \left(r_i m_{i2} - \frac{1}{2} m_{i2}^2 - r_i m_{i1} + \frac{1}{2} m_{i1}^2 \right)$$

$$\ln \Lambda_2(r(t)) = \sum_{i=1}^{\nu} \frac{1}{\lambda_i} \left(r_i m_{i3} - \frac{1}{2} m_{i3}^2 - r_i m_{i1} + \frac{1}{2} m_{i1}^2 \right)$$

where ν is the number of series expansion coefficients that are chosen to represent $r(t)$. Fig. 11 shows the decision spaces for estimates obtained using the EDW technique (left column) and the SDW technique (right column) with $T = 99$ ms. Interpretation of each figure is the same as that of the figure in Fig. 10. The detection performance is quantified in Fig. 12(E)–(F).

V. Detection Performance Analysis

Fig. 12 shows that both the EDW and SDW techniques permit reliable estimates of stimulus position for the first 400 ms of the movie, regardless of the approaches used. However, a major difference between the two techniques is that the reliability of the estimates decreases significantly for time windows after 400 ms when the SDW techniques is used, but continues to be relatively

high after 400 ms when the EDW technique is used. For the detection-by-distance approach, Fig. 12(A)–(B) shows that the detection error probability is close to 0 up to 600 ms with the EDW technique. It increases after 600 ms, but only to a value of 0.07 at 900 ms. The detection error probability is close to 0 for time up to 400 ms with the SDW technique, but then increases dramatically to more than 0.4 at 900 ms. This difference is shown more dramatically in Fig. 12(E) where the analysis with the EDW technique and the assumption of colored noise indicates that estimates that are correct with nearly 1.0 probability can be obtained even when the wave has almost returned to baseline at 900 ms. For the detection-by-white-noise-model, Fig. 12(C)–(D) indicates that the probability of detection error is low for the first 550 ms using EDW technique and for the first 400 ms using SDW technique, but then increases. However, the probability is only 0.07 even toward 900 ms of the response using the EDW technique.

For the EDW technique, the detection error starts with a nonzero value. It decreases with the ending time of the detection window and reaches zero at about 100, 100, and 60 ms using the detection-by-distance, the detection-by-white-noise-model, and the detection-by-colored-noise-model approach, respectively. This error-free detection breaks down at about 600 and 580 ms for detection-by-distance and detection-by-white-noise-model and increases from then on. For detection-by-colored-noise detection, the error-free detection breaks down after 900 ms.

It appears that the waves can be used to reliably discriminate stimuli positioned at opposite edges of the visual fields using windowing technique and hypothesis testing. It is striking that the detection performance with the EDW technique is very reliable at the very late stage of the cortical response. However, it is important to note that the discrimination task used in this study is not very demanding and probably does not challenge the information coding capabilities of the waves. It should also be noted that the results reported here and by Nenadic *et al.* [11] show that information is encoded in the cortical waves, but do not demonstrate that the turtle actually uses this information for behavioral tasks.

VI. Effect of Window Size

Finally, we examine the effect of the size of the encoding and decoding window on detection performance. The detection results we have seen in Fig. 12 were obtained using an encoding window size, $w = 10$. Detection with other encoding window sizes is shown in Fig. 13(A)–(D) where detection is performed using the EDW technique with the colored noise model. Detection based on encoding window sizes $w = 4, 30, 50$ is compared with that based on $w = 10$. For all the analyzes, $p = 679, q = 10$ are used. The figure shows that, for the discrimination task described in this paper, a larger encoding window yields a better detection performance at the early stage of the cortical response, probably because of the larger information content in a larger detection window. After the wave is developed, the detection performance is not very different with different w sizes. This is probably because the discrimination task used in this study was not very difficult. The effect of the size of the SDW shown in Fig. 14 lies in the difference in the time when the detection can

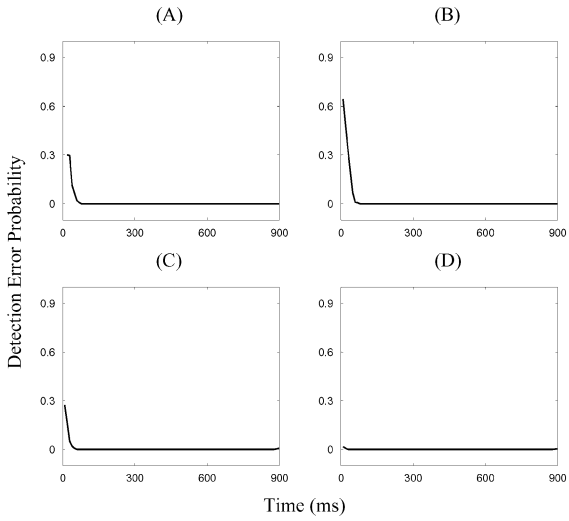

Fig. 13. Relationship between the detection error probability and size of encoding window. Detection was performed using EDW technique with the colored noise model. The horizontal axis denotes ending time T_2 (ms) of the expanding detection windows and vertical axis denotes the detection error probability. (A). $w = 4$; (B). $w = 10$; (C). $w = 30$; (D). $w = 50$.

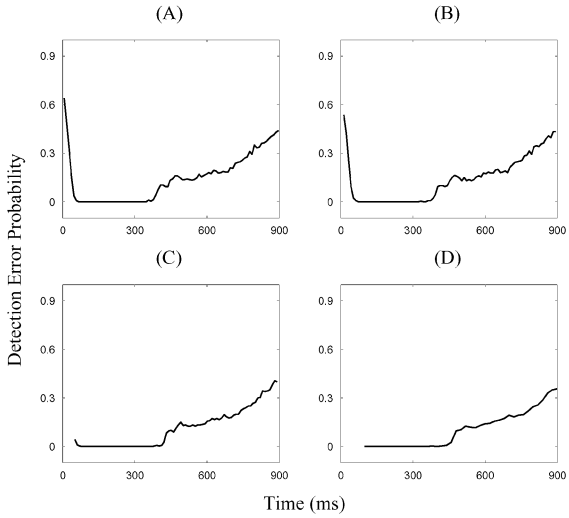

Fig. 14. Relationship between the detection error probability and size of SDW. All detections were performed using colored noise model and $w = 10$. The horizontal axis denotes ending time T_2 (ms) of the SDWs and vertical axis denotes the detection error probability. (A). $T = 6$; (B). $T = 12$; (C). $T = 50$; (D). $T = 99$.

start and the time when the error-free detection breaks down. A smaller detection window makes it possible to start the detection earlier, but the error-free detection also ends earlier. Other than this difference, the overall detection performance is similar in terms of the detection performance dynamics and the actual values of the errors.

Figs. 13 and 14 suggest that, for the discrimination task addressed in this paper, neither the encoding nor the decoding process requires a long time integration once the wave is developed. This might imply that neuronal data processing is relatively local in time and that a relatively short-term memory is adequate for the turtle to handle simple inputs. Larger encoding and decoding windows might give better detection performance for more challenging tasks. However, (1), (5),

and (13) suggest that, for larger encoding and decoding windows, the calculation intensity is also heavier. Thus, there is a trade-off between the detection performance and the computation intensity. Depending on the performance criteria, we can choose an "optimal" encoding window and decoding window (if the SDW is used).

It is worth mentioning that for encoding window size $w = 4$, the detection can not be performed at the very early stage of the cortical response [see Fig. 13(A)]. This is due to a λ (eigenvector of the noise process) which is close to zero while the cortex has not been excited yet.

VII. Discussion

A. Evolving Information Encoding in Cortical Responses

The classical approach to characterizing the coding properties of neurons in visual cortex involves measuring the firing rates of individual neurons over relatively long time windows. Notably, Hubel and Wiesel [15] have constructed orientation tuning curves for neurons in the primary visual cortex of cats by plotting the firing rates of neurons as a function of the orientations of bars of light presented in the visual field of the cats. Studies such as this demonstrated that cortical neurons code information about sensory stimuli but provided no information about the dynamics of the encoding process. Several laboratories have now examined the dynamics of the responses of individual neurons in the primary or secondary visual cortices of cats or monkeys using reverse correlation methods [16]–[23]. Other researchers have studied the dynamics of responses using models of cortical networks [24], [25]. All of these studies have shown that the responses of cortical neurons develop with time and, depending upon the cortical area or species being studied, reach their peaks somewhere between 50 and 150 ms following stimulus presentation. Sharon and Grinvald [3] extended these results to populations of neurons using voltage sensitive dye methods in the secondary visual cortex of cats. The extent to which the selectivity of neurons changes with time remains controversial. Some research has reported that the selectivity of neurons to stimulus orientation changes with time while others have found little or no change in selectivity throughout the course of the response (see [26] for a recent discussion of this controversy). However, there is a strong consensus that the response develops with time.

The simulations reported in this study are consistent with the findings in mammalian visual cortex in showing that the cortical responses to activation of different clusters of geniculate neurons in our large scale model of turtle visual cortex develop with time. The β-strands corresponding to stimuli present at the left, right or center edge of visual space are initially close to each other in B-space but are maximally separated at about 180 ms following stimulus presentation. They then approach each other as time proceeds and are again close by about 900 ms. The time courses in the model [14] and real cortices [10] are observed to be similar in terms of both wave origination and propagation. In the model, the wave begins to form in the rostral pole of the cortex with a latency of about 20 ms following stimulation of the geniculate neurons. It propagates across the cortex, reaching the caudal pole by about 200 ms after stimulus onset. It then decays slowly back to the baseline activity of the cortex by about 900 ms following stimulus onset.

The detection results shown in Figs. 8, 10, and 11 and the detection performance shown in Fig. 12 also indicate that the information content of the cortical waves is a dynamic variable whose temporal evolution is demonstrated in the detection error, the distance between and the spread of the clusters in the decision space. Fig. 7 shows that this information content reaches a maximal value between 180–280 ms following the origin of the wave and the most reliable estimates of stimulus position occur while the wave is propagating across the cortex. Fig. 12 shows that there is clearly an advantage to the detection system during the time period that the β-strands are maximally separated and less of an advantage before and after. Furthermore, even when the detection error is very low, the distance between and the spread of the clusters in the decision space vary. This is another indication of the temporal evolution of the information content in the cortical waves. Each cluster spreads in different directions, but whether or not this has any significance is unknown at this moment.

The fact that detection, based on very early stage of the cortical response (0 to about 80 ms), is unreliable is due to a lack of sufficient information in the chosen time windows. The asymptotic increase of detection error in the late stages of the cortical response can be attributed to an increase in the signal-to-noise ratio of the response signal. After about 600 ms, the cortex is dominated by noise. An information-theoretic analysis of the cortical waves has been performed [27], wherein the Kullback-Leibler distance [28] was calculated between waves generated by two different stimuli. Plots of the cumulative Kullback-Leibler distance as a function of time have shown that the information content of the waves increased rapidly from 100 to 200 ms following stimulus onset and then slowly reached a plateau over the subsequent 600 ms.

B. Effect of Noise on Decoding

Visual cortex has intrinsic noise sources due to the stochastic properties of chemical synapses and voltage-gated channels (See [29] for a discussion of noise in neural systems). Intrinsic noise was introduced in the model used in the present study by injecting a Gaussian white noise current into individual geniculate neurons in the model. Analysis of the power spectra of the β-strands indicated that the cortical responses contained a colored noise. This is not surprising given that the kinetics of the voltage-gated channels can be expected to act as nonlinear filters and the dendritic integration is strongly nonlinear as well.

A recent study [30] has focused on the possibility that cortical neurons exhibit a form of stochastic resonance whose role is to amplify the signal-to-noise ratio of cortical responses to weak input stimuli. The analysis presented here has addressed the issue of whether or not the additive noise eventually improves the reliability of estimates. The main conclusion of our analysis is that by assuming the additive noise is white, the decoding algorithm is at best similar to a decoding algorithm that does not assume any noise model (detection-by-distance approach). However, the algorithm improves considerably by assuming that the noise source is additive and colored. We do not make any explicit claim that any of these noise models are actually being used by the animal.

It is worth mentioning that the cortex model used in this paper has a high noise level. The cortex shows a high level of spiking

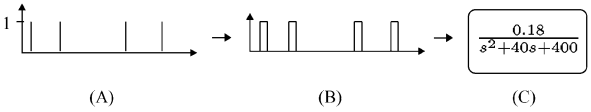

Fig. 15. Block diagram of spike train filtering.

activity when only a noise input is present and increasing the noise level even further would be unrealistic. It is almost certain that the detection error probability will decrease when the noise level is decreased.

APPENDIX I
FILTERING OF SPIKE TRAINS

Each spike train of the cortical cell was preprocessed before the double KL decomposition and detection operations were performed. Fig. 15 illustrates this preprocessing procedure as follows:

A) The membrane voltage of each cell was thresholded and a spike was considered to have occurred when the voltage exceeded the threshold. The threshold was set at -0.034 and -0.04 volts for lateral and medial pyramidal cells, respectively. Each occurrence of a spike was represented as a 1.

B) Each spike was then represented as a unit pulse of width 0.5 ms. Since the refractory period of spike generation is 10 ms, adjacent unit pulses do not overlap with each other.

C) The train of unit pulses was passed through a second-order filter. The output from the filter can be considered as equivalent to the spike rate signal of the corresponding cell.

REFERENCES

[1] G. B. Ermentrout and D. Kleinfeld, "Traveling electrical waves in cortex: Insights from phase dynamics and speculation on computational role," *Neuron.*, vol. 29, pp. 33–44, Jan. 2001.

[2] N. P. Cottaris, "Temporal dynamics of chromatic tuning in macaque primary visual cortex," *Nature*, vol. 395, pp. 896–900, 1998.

[3] D. Sharon and A. Grinvald, "Dynamics and constancy in cortical spatiotemporal patterns of orientation processing," *Science*, vol. 295, pp. 512–515, 2002.

[4] D. Derdikman, R. Hildeheim, E. Ahissar, A. Arieli, and A. Grinvald, "Imaging spatiotemporal dynamics of surround inhibition in the barrels somatosensory cortex," *J. Neurosci.*, vol. 23, pp. 3100–3105, 2003.

[5] A. A. Ghazanfar and M. A. L. Nicolelis, "Spatiotemporal properties of layer V neurons of the rat primary somatosensory cortex," *Cerebral Cortex*, vol. 9, pp. 348–361, 1999.

[6] C. C. H. Petersen, A. Grinvald, and B. Sakmann, "Spatiotemporal dynamics of sensory responses in layer $2/3$ of rat barrel cortex measured *in vivo* by voltage-sensitive dye imaging combined with whole-cell voltage recordings and neuron. reconstructions," *J. Neurosci.*, vol. 23, pp. 1298–1309, 2003.

[7] H. R. Wilson, R. Blake, and S. H. Lee, "Dynamics of travelling waves in visual perception," *Nature*, vol. 412, pp. 907–910, 2001.

[8] J. C. Prechtl, L. B. Cohen, P. P. Mitra, B. Pesaran, and D. Kleinfeld, "Visual stimuli induce waves of electrical activity in turtle visual cortex," *Proc. Nat. Acad. Sci.*, vol. 94, pp. 7621–7626, 1997.

[9] D. M. Senseman, "Correspondence between visually evoked voltage sensitive dye signals and activity recorded in cortical pyramidal cells with intracellular microelectrodes," *Vis. Neurosci.*, vol. 13, pp. 963–977, 1996.

[10] D. M. Senseman and K. A. Robbins, "Modal behavior of cortical neural networks during visual processing," *J. Neurosci.*, vol. 19, no. RC3, pp. 1–7, 1999.

[11] Z. Nenadic, B. K. Ghosh, and P. Ulinski, "Modeling and estimation problems in the turtle visual cortex," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 8, pp. 753–762, Aug. 2002.

[12] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968, ch. 2, pp. 23–46.

[13] A. M. Granada and C. A. Dvorak, "Vision in Turtles: The visual system in vertebrates," in *Handbook of Sensory Physiology*, F. Crescitelli, Ed, Berlin, Germany: Springer-Verlag, 1977, vol. VII/5, pp. 451–495.

[14] Z. Nenadic, B. K. Ghosh, and P. Ulinski, "Propagating waves in visual cortex: A large scale model of turtle visual cortex," *J. Computational Neurosci.*, vol. 14, pp. 161–184, 2003.

[15] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol.*, vol. 160, pp. 106–154, 1962.

[16] S. Celebrine, S. Thorpe, Y. Trotter, and M. Imbert, "Dynamics of orientation coding in area V1 of the awake primate," *Vis. Neurosci.*, vol. 5, pp. 811–825, 1993.

[17] M. Volgushev, T. R. Vidyasagar, and X. Pei, "Dynamics of the orientation tuning of postsynaptic potentials in the cat visual cortex," *Vis. Neurosci.*, vol. 12, pp. 621–628, 1995.

[18] I. A. Shevelev, U. T. Eysel, N. A. Lazareva, and G. A. Sharaev, "The contribution of intracortical inhibition to dynamics of orientation tuning in cat striate cortex neurons," *Neurosci.*, vol. 84, pp. 11–23, 1998.

[19] D. C. Gillespie, I. Lampl, J. S. Anderson, and D. Ferster, "Dynamics of the orientation-tuned membrane potential response in cat primary visual cortex," *Nature Neurosci.*, vol. 4, pp. 1014–1019, 2001.

[20] C. E. Bredfeldt and D. L. Ringach, "Dynamics of spatial frequency tuning in macaque V1," *J. Neurosci.*, vol. 22, pp. 1976–1984, 2002.

[21] G. Felsen, Y.-S. Shen, H. Yao, G. Spor, C. Li, and Y. Dan, "Dynamic modification of cortical orientation tuning mediated by recurrent connections," *Neuron.*, vol. 36, pp. 945–954, 2002.

[22] J. A. Mazer, W. E. Vinje, J. McDermott, P. H. Schiller, and J. L. Gallant, "Spatial frequency and orientation tuning dynamics in area V1," *Proc. Nat. Acad. Sci.*, vol. 99, pp. 1645–1650, 2002.

[23] D. L. Ringach, M. J. Hawken, and R. Shapley, "Dynamics of orientation tuning in macaque V1: The role of global and tuned suppression," *J. Neurophysiol.*, pp. 342–352, Jul. 2003.

[24] D. McLaughlin, R. Shapley, M. Shelley, and D. J. Wielaard, "A neuronal network model of macaque primary visual cortex (V1): Orientation selectivity and dynamics in the input layer 4C?," in *Proc. Nat. Acad. Sci.*, vol. 97, 2000, pp. 8087–8092.

[25] M. Pugh, D. L. Ringach, R. Shapley, and M. J. Shelley, "Computational modeling of orientation tuning dynamics in monkey primary visual cortex," *J. Comput. Neurosci.*, vol. 8, pp. 143–159, 2000.

[26] D. L. Ringach, C. E. Bredfeldt, R. W. Shapley, and M. J. Hawken, "Suppression of neural responses to nonoptimal stimuli correlates with tuning selectivity in macque V1," *J. Neurophysiol.*, vol. 87, pp. 1018–1027, 2002.

[27] X. Du and B. K. Ghosh, "Information-theoretic analysis of turtle cortical waves," in *Proc. 42nd IEEE Conf. Decision and Control*, 2003, pp. 6423–6428.

[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[29] W. Gerstner and W. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[30] M. Rudolph and A. Destexhe, "Do neocortical pryamidal neurons display stochastic resonance?," *J. Comput. Neurosci.*, vol. 11, pp. 19–42, 2001.

Xiuxia Du received the B.S. and M.S. degree in electrical engineering from Hefei University of Technology, China, and worked on electric power generation using renewable energy at the Institute of Electrical Engineering, Chinese Academy of Sciences. She is currently working towards the D.Sc. degree in the Department of Electrical and Systems Engineering at Washington University, St. Louis, MO.

Her research interests include information-theoretic analysis and linear and nonlinear dynamics of complex biological systems and quantitative biology.

Bijoy K. Ghosh (S'78–M'79–SM'90–F'00) received the B.Tech. and M.Tech. degrees in electrical and electronics engineering in India in 1977 and 1979, respectively. In 1983, he received the Ph.D. degree in engineering from the Decision and Control Group of the Division and Applied Sciences at the Harvard University, Cambridge, MA.

Since 1983, he has been a faculty member in the Department of Electrical and Systems Engineering at Washington University where he is currently a Professor and directs the center for BioCybernetics and Intelligent Systems. His research interests are in multivariable control theory, machine vision, robotics, biological control systems, and systems biology. In 1988, he received the American Automatic Control Council's Donald P. Eckman award in recognition of his outstanding contributions in the field of Automatic Control. In 1993, he had been an UNDP consultant under the TOKTEN program and visited the Indian Institute of Technology, Kharagpur, India. In 1997, he received the Japan Society for the Promotion of Science Invitation Fellowship for research in Japan and visited Tokyo Denki University and Mechanical Engineering Laboratory, Tsukuba City, Japan. He has also held short term visiting positions at Osaka University, Japan in 1992, Tokyo Institute of Technology, Japan in 1995, Padova University, Italy in 2001 and Institut Mittag-Leffler, Sweden in 2003. He is a permanent visiting Professor at the Tokyo Denki University and in the spring of 2001 he visited the Electrical Engineering Department at Yale University, New Haven, CT.

In 2000, he became a Fellow of the IEEE for fundamental contributions in systems theory with applications to robust control, vision and multisensor fusion. He is presently an elected member of the Board of Governors of the IEEE Control Systems Society and is a committee chair for the technical committee on BioSystems and Control.

Philip S. Ulinski received the Ph.D. degree in zoology from Michigan State University, East Lansing, in 1968.

He is currently Professor and Chairman of the Committee on Computational Neuroscience at the University of Chicago, Chicago, IL. His research interests include the functional organization of the cerebral cortex and computational models of neural circuits.