INVITED PAPER

Bio-Inspired Networks of Visual Sensors, Neurons, and Oscillators

Images collected by controlled camera networks can be processed to simultaneously locate the position and track the movement of target objects.

By Bijoy K. Ghosh, *Fellow IEEE*, Ashoka D. Polpitiya, *Member IEEE*, and Wenxue Wang, *Member IEEE*

ABSTRACT | Animals routinely rely on their eyes to localize fixed and moving targets. Such a localization process might include prediction of future target location, recalling a sequence of previously visited places or, for the motor control circuit, actuating a successful movement. Typically, target localization is carried out by fusing images from two eyes, in the case of binocular vision, wherein the challenge is to have the images calibrated before fusion. In the field of machine vision, a typical problem of interest is to localize the position and orientation of a network of mobile cameras (sensor network) that are distributed in space and are simultaneously tracking a target. Inspired by the animal visual circuit, we study the problem of binocular image fusion for the purpose of localizing an unknown target in space. Guided by the dynamics of "eye rotation," we introduce control strategies that could be used to build machines with multiple sensors. In particular, we address the problem of how a group of visual sensors can be optimally controlled in a formation. We also address how images from multiple sensors are encoded using a set of basis functions, choosing a "larger than minimum" number of basis functions so that the resulting code that represents the image is sparse. We address the problem of how a sparsely encoded visual data stream is internally represented by a pattern of neural activity. In addition to the control mechanism, the synaptic interaction between cells is also subjected to "adaptation" that enables the activity waves to respond with greater sensitivity to visual input. We study how the rat hippocampal place cells are used to form a cognitive map of the environment so that the animal's location can be determined from its place cell activity. Finally, we study the problem of "decoding" location of moving targets from the neural activity wave in the cortex.

KEYWORDS | Cortical waves; eye movement; formation sensing; gaze control; Hebbian and anti-Hebbian adaptation; Kuramoto model; Listing's law; localization; neural network; oscillator network; place cells; sensor network; sparse coding; theta phase precession

I. INTRODUCTION

In recent years, wireless sensor networks have made important advances in many scientific disciplines such as manufacturing, agriculture, construction, and transportation. They include such wide applications as measuring traffic on city roads, evaluating people's interest in an art gallery, monitoring the perimeter of an oil refinery, or safeguarding the remote water reservoirs of the nation. There are of course many technical challenges associated with a successful deployment of such a network. The challenges include acquisition of sensory information, sharing of sensory data between nodes over a wireless network, and synchronization of the data processing cycle between a node and its neighbor, to name a few. There are many interesting review articles in recent literature that describe the general problems associated with sensor networking (see [21], [22], [31], [32], [75], [97], and the references therein).

As charge coupled devices (CCDs) and complementary metal–oxide–semiconductor (CMOS) image/video cameras become smaller, less expensive, and more power

Manuscript received September 6, 2005; revised August 31, 2006. This work was supported in part by the National Science Foundation (NSF) under Grant EIA-0218186 and in part by the NSF under Grant ECS-0323693.

B. K. Ghosh and **W. Wang** are with the Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409-1042 USA (e-mail: bijoy.ghosh@ttu.edu; wenxue.wang@ttu.edu).

A. D. Polpitiya is with Pacific Northwest National Laboratory, Richland, WA 99352 USA (e-mail: ashoka.polpitiya@pnl.gov).

Digital Object Identifier: 10.1109/JPROC.2006.887320

efficient, camera networks will become increasingly more powerful and widespread. A critical first step in any camera network deployment is "localization." The camera nodes must determine their positions and orientations in three-dimensional (3-D) space. Most sensor network localization techniques use acoustic delays and radio-frequency intensities to create a set of pairwise distances in order to localize [19]. Our goal in this paper is to emphasize visual sensing, localizing, and target tracking problems associated with building a visual sensor network. The greatest of the challenges associated with a network of camera sensors is solving a "simultaneous localization and tracking (SLAT)" problem. Manually measuring the locations and orientations of all cameras in the network is a very tedious and time-consuming task. If a moving object is seen in a camera's field of view and a few moments later, the same object is observed by another camera, knowing the trajectory of the object, we could infer the relative position of the two cameras. Unfortunately, without an independent localization system like GPS, the object's trajectory remains unknown. Various approaches to the SLAT problem have been addressed in the literature (see [19], [35], and [48]).

The approach in this paper is derived from biology and relies on the fact that as cameras move, their orientations remain constrained. Their orientations are chosen to be a function of the gaze direction, satisfying the well-known Listing's law [38]. In this way, camera sensors are actuated in such a way that the orientation ambiguity is resolved.

Inspired from biology, we also study how images are to be acquired, followed by how they are encoded and decoded. In so doing, we introduce other networks that are well studied in biology and in nonlinear dynamical system science [2], [13], [30], [53]. In neuroscience, these networks arise as networks of neurons that are built as models of the "visual cortex," "optic tectum," and "hippocampus," to name a few. In nonlinear system science, these networks arise as networks of "coupled oscillators" with a certain guaranteed region of stability.

We introduce a sensing paradigm (see Fig. 1) derived from animal vision with the eventual goal of constructing a biologically inspired sensor network. Such a network captures salient features of the target in order to make predictions about its future location and recall event sequences from the past. The sensors in the network are expected to be able to attend to a specific visual target by simultaneously coordinating their relative orientations as a function of their respective positions and gaze directions. The images captured from multiple sensors are fused and encoded prior to any further computations. Multisensor image fusion is carried out using a set of basis functions that are specifically chosen to be "overcomplete" (i.e., the number of basis functions are more than what would be minimally required) so that the number of nonzero coefficients obtained, in the basis function expansion, is small (sparse codes).

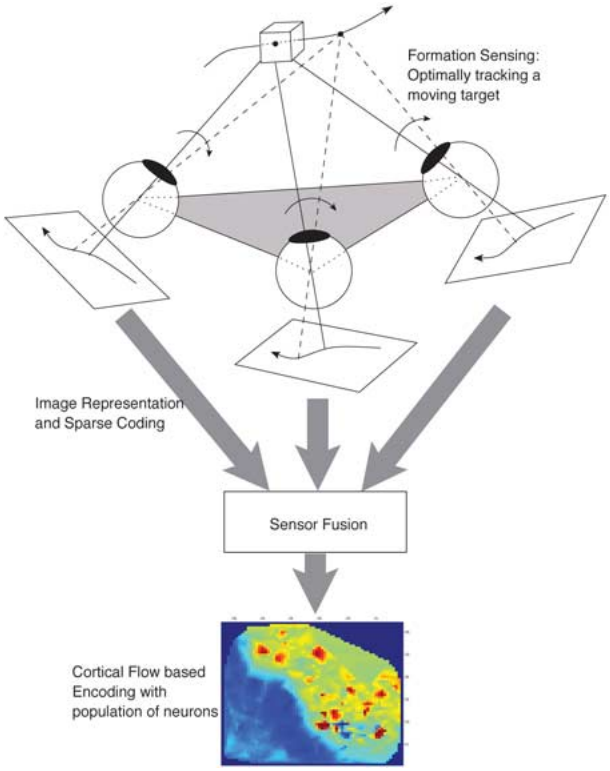

Fig. 1. ***The figure shows a triplet of visual sensors that simultaneously capture images of a moving target. The rotational movement of each sensor satisfies Listing's constraint and the gaze directions are constrained to pass through a moving point in space. Captured images are fused and encoded using a population of neurons that produces a traveling wave of activity.***

In principle, the sparse codes obtained can be further processed in a variety of ways depending upon the objective of the sensor network. In this paper, we consider the following two objectives inspired from biology.

- Prediction of a target location, utilizing the cortical circuit.
- Recalling a sequence of prior places visited, utilizing the hippocampal place cells.

Inspired by the turtle's visual cortex, we design a neural population network that generates waves of activity as a result of a visual stimulus from the retina. We hypothesize that the activity waves encode features from the visual scene. It can be easily verified by carrying out the decoding process using alternative algorithms discussed in [29], [63], [64], and [95]. The activity wave sustained by the neural population is controlled by a feedforward and a feedback loop—sketched in Fig. 6—and various adaptation mechanisms, one of which is discussed in detail using *Hebbian and anti-Hebbian rules*.

The problems we survey are pertinent in visual multisensor fusion. They include *geometric optimal control problems in formation sensing* [28], [47], [89], [96]; *image*

representation and sparse coding with over-complete basis functions [34], [68], [69]; *cortical flow based encoding with population of neurons* [63]; *Hebbian and anti-Hebbian adaptation* of the neural population [24], [39], [41]; and finally *decoding and prediction from the cortical activity waves* [29], [63], [64], [95]. The decoding problem has been addressed in the literature using three distinct choices of algorithms. The first algorithm utilizes *statistical hypothesis testing* [91] from the activity patterns in the cortex to detect the location of the target [29], [63], [64] (see Fig. 10). The second method utilizes *linear time invariant filters* [37] to reconstruct features from the visual scene [95]. These two methods have not been elaborated in this paper. The third procedure, which we describe in this paper, is how activity patterns are discriminated using a *coupled network of oscillators* (see [44], [46], and [56]) with multiple equilibria.

At the very outset, we would like to emphasize that the problem of image acquisition is tightly connected to the problem of "sensor actuation." In the case of visual sensing, mobility has many advantages and improves sensing performance in the following three ways [51]: first, a mobile sensor can cover a large volume of space. Second, mobile network components can reposition themselves to overcome obstacles and other anisotropies in the medium. Finally, cameras can move in order to sample the event in space where the highest resolution coverage is required. In order to construct an example of a mobile visual sensor network, Cyclops [79] can be coupled with a sensor network node such as the Berkeley Mote [21]. The combination would represent a wireless vision network node. Together with a Pan/Tilt module, this provides an actuation enabled wireless vision node that has already been successfully applied for habitat monitoring, supporting diverse domains in monitoring bird nesting, reptile populations, and plant phenology.

Specifically, we are interested in the visual systems of animals in so far as how a moving target is "acquired" and "kept in view" by active eye movement. Animals locate a target through a well-developed "overt" attention mechanism which consists of a combination of eye and head movement (in this paper we only describe the eye movement). The sole purpose of this movement is to keep a target in view. For a functional sensor network, it is not enough just to be able to physically view a target. Rather, it is important to internally represent events from the visual scene via a process of "covert" internal representation. In the case of binocular vision, the physical movements of the two eyes are synchronized in such a way that corresponding retinal images are aligned. For example, an interesting target is always imaged on the fovea. Additionally, for a given target location, the images have a unique orientation. Thus, assuming no head movement, fixed objects look the same visually (i.e., without any orientation ambiguity) no matter when you see them. This is achieved because eye movement is always constrained by Listing's law [38], which we shall discuss in detail later.

The process of image alignment is not perfect, and it has been noted by Anderson and Van Essen [3] that "misalignments can be several minutes of arc under optimal conditions and an order of magnitude larger under more realistic conditions." In order to alleviate the misalignment problem, shifter circuits have been proposed [3] that would dynamically shift the relative alignments of the input images without losing the local spatial relationship. We do not explore the alignment compensation problem any further in this paper (see [66] and [67] for details). In the turtle's visual circuit, images on the retina cannot be used subsequently in visual discrimination or control signal generation for a suitable motor action, unless they are encoded and transmitted to the visual cortex whose functional role is described in considerable detail in this paper. This encoding process should require a small amount of neural resource, should be easily amenable to sensor fusion, and should be easily accessible for subsequent signal processing and many other post-sensory actions. The first part of this encoding process is addressed by utilizing a sparse, over-complete set of basis functions to represent natural scenes. The algorithm (originally proposed in [34], [68], and [69]) chooses basis functions in such a way that a set of natural images is represented by a small number of basis functions. Images from various sensors are fused utilizing principal component analysis. A set of basis functions is generated that can represent a pair of images by sparse codes "a_1," "a_2" (see Fig. 2) simultaneously with respect to a dominant set of bases. The codes "a_1" and "a_2" are fused to obtain a single vector "a" over a suitable window of time which is the required output of the sensor fusion process.

We study how sparsely encoded and fused retinal images are represented via a wave-generating pattern of neural activity (see Figs. 1 and 2) by the visual cortex. The cortical pattern of neural activity is in the form of a traveling wave (see Fig. 7). Animals actively internalize spatiotemporal cues from the visual environment and generate patterns of activity in the brain. The activity patterns, in turn, are used for recognition, associative recalling, and prediction. We illustrate this process of cortical encoding using neural circuits of turtle vision, and show how visual images from two eyes are fused into a pattern of activity waves. Typically these waves last longer than the visual event itself—for example, a flash of light incident on the turtle retina for 100 ms would result in a cortical wave that would last for about 800 ms. The wave-generating mechanism is easily amenable to sensor fusion and can be readily used as an input to a decoder, which we describe in detail in this paper.

A wave-generating population of neurons that is to be attached to the output of a visual sensor must generate waves continually. In effect, this means that waves must be generated and killed, and then generated again in a

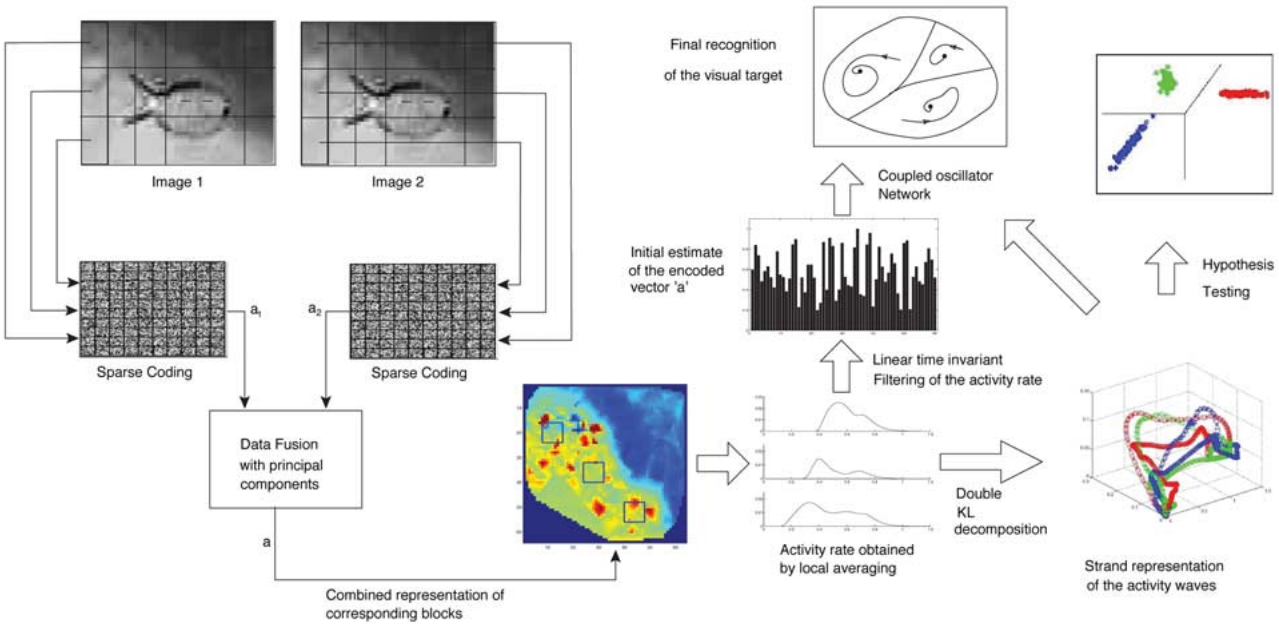

Fig. 2. ***The figure shows that a pair of images from each of the two eyes are sparse coded, fused, and subsequently used as an input to the visual cortex in generating waves. Activity rates of neurons in the cortex are used in target recognition. Three alternative schemes are proposed: In the first scheme, activity rates are represented as a temporal strand using principal component analysis. Targets are recognized by discriminating the strands statistically. In the second scheme, activity rates are used as an input to a network with multiple equilibria, after preprocessing through a linear filter. Each target is assumed to correspond uniquely to an equilibrium. A target is recognized by observing the equilibrium point that the network converges to. In the third scheme, the strands are used as an input to the network with multiple equilibrium and the details are analogous to scheme two.***

repetitive fashion. This is achieved via a process of Hebbian and anti-Hebbian adaptation [24], [39], [41], which we now describe. Adaptation is intrinsic to any population of interacting neurons, and it is designed to either reinforce or weaken the strength of interaction (synaptic strength). Hebb's basic postulate is that when the axon of neuron A is near enough to excite neuron B and repeatedly takes part in firing it, some growth process or metabolic change takes place in one or both of the neurons such that A's efficiency as one of the cells firing B increases. It follows that in Hebbian adaptation, the synaptic strength between two neurons increases in proportion to the pre- and post-synaptic firing rates. Likewise, in anti-Hebbian adaptation, the synaptic strength is reduced. In this paper, we show that a suitable combination of these two adaptation rules is precisely what we need for repeated generation of traveling-wave activity in the cortex model.

Wave propagation is believed to play a critical role in biological pattern formation. In network-based computation, it has already found important application in amorphous computing [1]. Vast numbers of unreliable microsensors, actuators, and communication devices interconnected in unknown ways would apply paradigm from cellular cooperation in biological organisms. It would be possible to assemble a system incorporating a multitude of information processing units at virtually no cost. These units would consist of logic circuits, microsensors, actuators, and communication devices, all in the same chip, producing particles that could be mixed with bulk materials such as paints, gels, and concrete. Coating buildings and bridges with "smart paint," it would be possible to sense and report traffic, wind loads, or structural integrity. A smart paint coating on a wall could sense vibrations, monitor the premises for intruders, and cancel noise. The particles would organize themselves by interacting locally and would produce global behavior, such as generation of a traveling wave. An initial anchor particle, chosen by a cue from the environment, would broadcast a message to all of its neighbors. These neighbors propagate the message to their neighbors, and so on, creating a diffusion wave that would spread throughout the system. The amorphous computing environment could play the role of a visual cortex and could be used to experiment with synthetic biology.

We now comment on the hippocampal place cells and how they are used to memorize and recall events encoded by the cortex. Studies on rat hippocampal cells by O'Keefe in the 1970s have shown the existence of place cells [99]. Place cells are cells that fire selectively in correlation to a rat's position while navigating a known environment. The position in the environment that each place cell

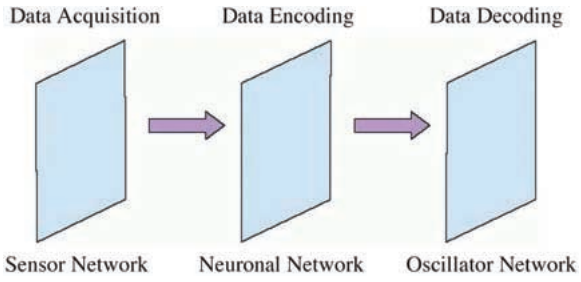

Fig. 3. ***A schematic diagram showing the role of network. The sensor network of cameras acquire target data from the 3-D world. The acquired data from sensors are encoded by the neuronal network in the form of traveling waves or spatially localized activity. Finally the neuronal activity is decoded by an interconnection of coupled oscillators in a network such as the Kuramoto network.***

corresponds to is called its place field. Together, the place cell/place field relationship forms a cognitive map of the environment, in that the rat's location can be determined from its place cell activity [17], [18]. Further studies of monkeys by Ono and Nishijo have shown that monkey hippocampal neurons also exhibit place cell properties that are sensitive to spatial location within a known environment [85]. In addition to the place cells, the hippocampus also sustains a slow oscillation in the theta (4–8 Hz) frequency range. This has been observed in all mammals including humans. This oscillation is often referred to as the "theta rhythm" and each period is referred to as a "theta cycle." Firing of place cells and observation of the theta rhythm has led to the proposal that information in the hippocampus is contained in the phase of this rhythm and not by the cell's firing rate [57]. In this paper, we shall elaborate how the phase coding is used in the hippocampus to memorize and recall a sequence of events.

Our final interest is in the task of decoding. Typically if we have point targets localized at different corners of visual space, a specific sensing task should be able to discriminate between events that are separated in space from the associated activity waves generated by the cortical circuit. One strategy described in this paper is to use a synchronized set of coupled oscillators in a network. The oscillator parameters have been adjusted to have multiple stable equilibria. The network is initialized based on an initial estimate of the target in visual space. Subsequently the target is discriminated on the basis of the equilibria where the network eventually settles down.

In summary, this paper highlights three important areas of visual sensing (see Fig. 3). The first is "active acquisition of visual images in a sensor network" and we describe the problem of formation sensing. The second is the problem of "sparse coding and data fusion" followed by an internal representation using a network of neurons. Such a network can synthesize a wave generator, and we describe in detail the cortical circuit of freshwater turtles (see Fig. 6). The network can also synthesize a sequence of interacting place cells mimicking the hippocampal circuit of rats (see Fig. 12). We also show how adaptation plays an important role in sustained generation of waves in the cortex and binding of place cells in the hippocampus. Third, we discuss the problem of decoding from the firing pattern of the family of neurons in the cortex. Image recognition using nonlinear dynamic models, especially the "Kuramoto model," is proposed and described in detail.

II. EYE MOVEMENT SATISFYING LISTING'S LAWS

In this section, we are interested in studying a group of visual sensors that move in a formation. We shall use the phrase "formation-sensing" for this purpose. Underlying motivation to consider this problem comes from the biomechanics of eye rotation [82], wherein the problem is to simultaneously control the gaze directions of a pair of eyes. Simply speaking, an animal has to move its eyes to keep a target in view, and do so without introducing any orientation ambiguity. When different parts of a visual space are viewed in a sequence using a visual sensor, every direction, in principal, can be viewed with an ambiguous set of orientations. Physiological evidence supports the view that orientation ambiguity is not present in human eye movement; this is possible because eye movements obey a law known as *Listing's law* (see [38] and [90]). It states that "eye rotation takes place such that the axes of rotation always reside in a fixed plane called Listing's Plane."

If eye movement is constrained to satisfy Listing's law, every gaze direction corresponds to a specific orientation of the eye. This is a desirable feature for image acquisition using one or multiple sensors. If sensors move while satisfying Listing's constraint, multiple images acquired by a given sensor, using a specific gaze direction, do not have to be corrected for orientation ambiguity. The sensor localization problem reduces to ascertaining only the position, orientation being predetermined by the gaze. Of course, the price we pay is in implementing the controller, which would require messages to be passed between the sensors. This leads to an extra cost of communication. In this section and the next, we consider a formation of mobile sensors that rotate satisfying Listing's law. The problems we consider are described as follows.

Problem I (Optimal Gaze Control Problem): Consider a sensor that is allowed to rotate in 3-D following Listing's law. How would one steer the sensor optimally from an initial gaze direction to the desired final direction?

Problem II (Binocular Formation Sensing): Consider a pair of sensors allowed to rotate in 3-D, each constrained by Listing's law. The sensors are also allowed to translate with

respect to each other. The sensors have to simultaneously maintain a gaze direction that passes through a moving point target. How would one steer the sensor pair optimally from an initial position of the target to a final position?

The formation sensing problem can be posed likewise using any number of sensors. The binocular problem is illustrated in this paper.

A. Background of the Eye Movement Problem

Since as early as 1845, modeling the eye plant in order to generate various eye movements has been studied closely by Listing, Donders, and Helmholtz (see for example [82]). The eyes rotate with three degrees of freedom, making the problem interesting yet simple, when compared to other somewhat more complex human movements. The human eye is spherical in shape and has $SO(3)$ as configuration space where $SO(3)$ is the space of all 3×3 rotation matrices. However, from the point of view of target tracking, only the gaze direction vector is important. The precise orientation of the eye is not significant as long as it remains invariant for a given fixed gaze direction. Said differently, the orientation of an eye is ambiguous up to a full circle of rotation matrices for a specific gaze direction. One needs to resolve the ambiguity while tracking the gaze prescribed by the moving target. Listing's law describes precisely how this rotational ambiguity is resolved. It states that all eye rotations occur about axes orthogonal to the primary gaze direction. If we were to take the (x_1, x_2, x_3) axes such that the x_3 axis was aligned with the normal gaze direction, then Listing's law amounts to a statement that all eye rotations take place about axes that lie in the plane given by $x_3 = 0$. What complicates the rotational movement is that the axis of rotation, while staying inside this plane, may change in time. In order to solve optimal control problems, it becomes imperative to be able to parameterize the subset of all eye orientations that satisfy Listing's law.

B. Understanding Rotational Motion Through Quaternion

A quaternion is a 4-tuple of real numbers (q_0, q_1, q_2, q_3) which can be added and scalar multiplied as a vector. Additionally, two elements in the set of quaternions can be added and multiplied which gives quaternions. The set of quaternions is indeed a noncommutative division ring (see [55] for details). Writing $q = (q_0, q_1, q_2, q_3) = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$ and $s = (s_0, s_1, s_2, s_3) = s_0 + s_1\mathbf{i} + s_2\mathbf{j} + s_3\mathbf{k}$, one can define

$$q + s = (q_0 + s_0, q_1 + s_1, q_2 + s_2, q_3 + s_3)$$

and

$$q.s = w \tag{1}$$

where w can be easily written down using the rule $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$ as follows:

$$\begin{aligned} w = {} & (q_0 s_0 - q_1 s_1 - q_2 s_2 - q_3 s_3) \\ & + (q_0 s_1 + q_1 s_0 + q_2 s_3 - q_3 s_2)\mathbf{i} \\ & + (q_0 s_2 - q_1 s_3 + q_2 s_0 + q_3 s_1)\mathbf{j} \\ & + (q_0 s_3 + q_1 s_2 - q_2 s_1 + q_3 s_0)\mathbf{k}. \end{aligned}$$

For the purpose of this discussion, we are interested in unit quaternions, i.e., when $\sum_{i=0}^{3} q_i^2 = 1$, and denote the space of unit quaternions by S^3. Our main point is the following: *Every unit quaternion (q_0, q_1, q_2, q_3) can be viewed as an element of $SO(3)$, i.e., a rotation matrix, which rotates a vector $v \in \mathbb{R}^3$ by an angle ϕ about axis (q_1, q_2, q_3) where $\cos(\phi/2) = q_0$.*

The map between S^3 and $SO(3)$ is given as follows:

$$\Psi : S^3 \to SO(3) \tag{2}$$

where

$$(q_0 \quad q_1 \quad q_2 \quad q_3)^T \longmapsto \begin{pmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_1q_2+q_0q_3) & q_0^2+q_2^2-q_1^2-q_3^2 & 2(q_2q_3-q_0q_1) \\ 2(q_1q_3-q_0q_2) & 2(q_2q_3+q_0q_1) & q_0^2+q_3^2-q_1^2-q_2^2 \end{pmatrix}.$$

If we denote $q = (q_0, q_1, q_2, q_3)$, $\bar{v} = (0, v)$, $\bar{u} = (0, Rv)$, one can show that $\bar{u} = q.\bar{v}.q^{-1}$, i.e., Rv is the vector part of $q.\bar{v}.q^{-1}$ obtained purely using quaternion multiplication (1).

C. Parameterizing the Space List

Let us denote by ***List***, the subset of $SO(3)$ which contains all the rotation matrices with axes of rotation in the plane $x_3 = 0$. It would follow that ***List*** is parameterized by all unit quaternion of the form $(q_0, q_1, q_2, 0)$ where $q_0^2 + q_1^2 + q_2^2 = 1$. One can pick the angles (θ, ϕ) as the coordinates for ***List*** as follows:

$$\begin{aligned} & \rho : [0, \pi] \times [-\pi, \pi] \longrightarrow \textbf{\textit{List}} \\ & (\theta, \phi) \longmapsto \left(\cos\frac{\phi}{2}, \cos\theta \sin\frac{\phi}{2}, \sin\theta \sin\frac{\phi}{2}, 0 \right) \end{aligned} \tag{3}$$

where ϕ is the angle of rotation and $(\cos\theta, \sin\theta, 0)$ is the axis of rotation in Listing's plane $x_3 = 0$. Then we have the following correspondence by the map (2) between S^3 and $SO(3)$ in the ***List*** [see equation (4) at the bottom of the next page].

An element of ***List***, parameterized by (θ, ϕ) rotates the vector $(0, 0, 1)^T$ to the vector $(\sin\theta \sin\phi,$

$-\cos\theta\sin\phi, \cos\phi)^T$, which is an element on the unit sphere S^2. Thus, if v is a unit vector in $\mathbb{R}^3$, there exists unique (θ^*, ϕ^*) such that

$$v = (\sin\theta^*\sin\phi^*, -\cos\theta^*\sin\phi^*, \cos\phi^*)^T.$$

It would follow that a unique element of ***List***, parameterized by (θ^*, ϕ^*) would rotate the vector $(0, 0, 1)^T$ to the vector v. Hence, the orientation ambiguity is resolved.

D. Riemannian Metric on *List*

Assume that the eye is a perfect sphere and its moment of inertia is equal to $I_{3\times 3}$. The configuration space $SO(3)$ is equipped with a left invariant Riemannian metric [12] described at the identity element of $SO(3)$ as follows:

$$\langle \Omega(e_i), \Omega(e_j)\rangle_I = \delta_{i,j}$$

where

$$\Omega(e_k) = \begin{pmatrix} 0 & \delta_{3,k} & -\delta_{2,k} \\ -\delta_{3,k} & 0 & \delta_{1,k} \\ \delta_{2,k} & -\delta_{1,k} & 0 \end{pmatrix}$$

and $\delta_{l,m}$ denotes the Kronecker delta function, i.e., $\delta_{l,m} = 1$ if $l = m$ and $\delta_{l,m} = 0$ if $l \neq m$, and where e_i are standard basis vectors. Via the two mappings (2), (3), the Riemannian metric on $SO(3)$ induces a Riemannian metric on ***List***. If we define

$$g_{11} = \left\langle \frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta} \right\rangle, \quad g_{22} = \left\langle \frac{\partial}{\partial\phi}, \frac{\partial}{\partial\phi} \right\rangle, \quad g_{12} = \left\langle \frac{\partial}{\partial\theta}, \frac{\partial}{\partial\phi} \right\rangle$$

one can show [73] by explicit calculation that

$$g_{11} = \sin^2\frac{\phi}{2}, \quad g_{22} = \frac{1}{4}, \quad g_{12} = g_{21} = 0$$

and the corresponding Riemannian metric on ***List*** is given by

$$g = \sin^2\frac{\phi}{2}d\theta^2 + \frac{1}{4}d\phi^2.$$

It is straightforward, but somewhat tedious to use the Riemannian metric and compute the corresponding geodesic equation on ***List*** given by

$$\begin{aligned} \ddot{\theta} + \cot\frac{\phi}{2}\dot{\theta}\dot{\phi} &= 0, \\ \ddot{\phi} - \sin\phi\dot{\theta}^2 &= 0. \end{aligned} \tag{5}$$

E. Eye Rotation Along an Optimal Trajectory

If we consider a potential function in the form $V(\theta, \phi)$, we write down the Lagrangian as

$$L(\theta, \phi, \dot{\theta}, \dot{\phi}) = \frac{1}{2}(\dot{\theta} \;\; \dot{\phi})\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}\begin{pmatrix} \dot{\theta} \\ \dot{\phi} \end{pmatrix} - V(\theta, \phi).$$

The Euler–Lagrangian equation [4]

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \tau_{q_i} \tag{6}$$

where q^i corresponds to the ith coordinate and τ_{q_i} are components of an external generalized force, gives rise to the following two equations of motion:

$$\begin{aligned} \ddot{\theta} + \cot\frac{\phi}{2}\dot{\theta}\dot{\phi} + \operatorname{cosec}^2\frac{\phi}{2}\frac{\partial V}{\partial\theta} &= \operatorname{cosec}^2\frac{\phi}{2}\tau_\theta, \\ \ddot{\phi} - \sin\phi\dot{\theta}^2 + 4\frac{\partial V}{\partial\phi} &= 4\tau_\phi \end{aligned} \tag{7}$$

where τ_θ and τ_ϕ are the generalized forces. As an illustration, let us assume that $V(\theta, \phi) = \sin^2(\phi/2)$ and that we wish to control the state $(\theta, \dot{\theta}, \phi, \dot{\phi})$ from $(\theta_0, 0, \phi_0, 0)$ to $(\theta_1, 0, \phi_1, 0)$ in T units of time while minimizing the energy

$$\int_0^T \tau_\theta^2(t) + \tau_\phi^2(t)dt. \tag{8}$$

Our choice of the potential function is motivated by the fact that all axes of rotation on Listing's plane are equally

$$\begin{pmatrix} \cos\frac{\phi}{2} & \cos\theta\sin\frac{\phi}{2} & \sin\theta\sin\frac{\phi}{2} & 0 \end{pmatrix}^T \longmapsto \begin{pmatrix} \cos^2\frac{\phi}{2} + \cos 2\theta\sin^2\frac{\phi}{2} & \sin 2\theta\sin\frac{\phi}{2} & \sin\theta\sin\phi \\ \sin 2\theta\sin\frac{\phi}{2} & \cos^2\frac{\phi}{2} - \cos 2\theta\sin^2\frac{\phi}{2} & -\cos\theta\sin\phi \\ -\sin\theta\sin\phi & \cos\theta\sin\phi & \cos\phi \end{pmatrix} \tag{4}$$

preferable. On the other hand, the potential is 0 for zero angle of rotation and rises monotonically when the angle of rotation increases in magnitude to π. The choice of a quadratic cost function on the generalized forces is somewhat arbitrary. A realistic cost function may be obtained assuming that eye rotation is actuated by three pairs of muscles. Assuming a linear model of these six muscles using a spring and damper, one can write the generalized forces as follows:

$$\tau_\theta = \sum_{i=1}^{6} [F_i + C_i] \frac{\partial l_i}{\partial \theta}$$

and

$$\tau_\phi = \sum_{i=1}^{6} [F_i + C_i] \frac{\partial l_i}{\partial \phi}$$

where F_i is the active force on the ith muscle (that is neurally actuated and can therefore be controlled), l_i is the instantaneous length of the ith muscle, and C_i is the passive force on the ith muscle (that is generated by the spring and damper and cannot be controlled). One can minimize an alternative cost function given by

$$\int_0^T \sum_{i=1}^{6} F_i^2 \, dt. \tag{9}$$

In this paper, we do not follow this alternative route (see [60] for a general introduction to muscle mechanics and eye rotation problems).

If we denote $z = (z_1, z_2, z_3, z_4) = (\theta, \dot{\theta}, \phi, \dot{\phi})$ to be the state vector, then the equations of motion in (7) become

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} z_2 \\ -z_2 z_4 \cot(z_3/2) \\ z_4 \\ z_2^2 \sin(z_3) - \frac{1}{2}\sin(z_3) \end{bmatrix} + \begin{bmatrix} 0 \\ \csc^2(z_3/2) \\ 0 \\ 0 \end{bmatrix} \tau_\theta + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 4 \end{bmatrix} \tau_\phi. \tag{10}$$

The Hamiltonian for the system is

$$\mathcal{H}(z, \lambda) = \lambda \cdot \dot{z} - \frac{1}{2}\left(\tau_\theta^2 + \tau_\phi^2\right)$$

where $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is the costate vector and the Hamilton's equations in this case are given by

$$\dot{z} = \frac{\partial \mathcal{H}}{\partial z}, \quad \dot{\lambda} = -\frac{\partial \mathcal{H}}{\partial \lambda}.$$

The optimal control can be found using the maximum principle as

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial \tau_\theta} = 0 \Rightarrow \tau_\theta = \frac{\lambda_2}{\sin^2(z_3/2)} \\ \frac{\partial \mathcal{H}}{\partial \tau_\phi} = 0 \Rightarrow \tau_\phi = 4\lambda_4. \end{aligned} \tag{11}$$

A necessary condition for the optimal path in the Listing space is given by the following system with state variables z and costate variables λ:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \dot{z}_4 \\ \dot{\lambda}_1 \\ \dot{\lambda}_2 \\ \dot{\lambda}_3 \\ \dot{\lambda}_4 \end{bmatrix} = \begin{bmatrix} z_2 \\ -z_2 z_4 \cot\left(\frac{z_3}{2}\right) + \tau_\theta^* \operatorname{cosec}^2\left(\frac{z_3}{2}\right) \\ z_4 \\ z_2^2 \sin(z_3) - \frac{1}{2}\sin(z_3) + 4\tau_\phi^* \\ 0 \\ -\lambda_1 + \lambda_2 z_4 \cot\left(\frac{z_3}{2}\right) - 2\lambda_4 z_2 \sin(z_3) \\ \Delta \\ \lambda_2 z_2 \cot\left(\frac{z_3}{2}\right) - \lambda_3 \end{bmatrix} \tag{12}$$

where $\Delta = -(1/2)\lambda_2 z_2 z_4 \operatorname{cosec}^2(z_3/2) - \lambda_4 z_2^2 \cos(z_3) + (1/2)\lambda_4 \cos(z_3) + \lambda_2 \cot(z_3/2) \csc^2(z_3/2)\tau_\theta^*$. The boundary conditions are given by $z(0) = (\theta_0, 0, \phi_0, 0)$ and $z(T) = (\theta_1, 0, \phi_1, 0)$. For a comparison between the optimal path and the geodesic path, see [74].

III. SENSING WITH A TWO EYE FORMATION—BINOCULAR VISION

In this section, we generalize the problem of optimally orienting a single eye to a pair of eyes analogous to what is shown in Fig. 1. We assume that each eye satisfies Listing's constraint, i.e., the axis of rotation always lies in a plane. For simplicity, let us assume that Listing's plane for each of the two eyes is the plane $x_3 = 0$, which is perpendicular to the primary gaze direction. The two eyes are assumed to be separated by a vector $a = (a_1, a_2, 0)$. The gaze directions of the two eyes are constrained to remain coplanar at all times during their movement, i.e., the gaze directions pass through a point at all times although this point may move with the target.

If we assume that vector a lies in the (x_1, x_2) plane, the configuration space for the "*Two Eye Formation*" is a subset of

$$\boldsymbol{List1} \times \boldsymbol{List2} \times \mathbb{R}^2 \tag{13}$$

with coordinates given by $((\theta_1, \phi_1), (\theta_2, \phi_2), (a_1, a_2))$. Note that the members of ***List*1** and ***List*2** are unit quaternions with elements having an axis of rotation in Listing's plane. The coplanarity condition on the gaze direction gives rise to the following additional constraint on the vector (a_1, a_2):

$$\begin{aligned} a_1 &= \lambda(\sin\theta_2 \tan\phi_2 - \sin\theta_1 \tan\phi_1) \\ a_2 &= \lambda(\cos\theta_2 \tan\phi_2 - \cos\theta_1 \tan\phi_1). \end{aligned} \tag{14}$$

We conclude that the required configuration space is a five-dimensional (5-D) manifold $M = \boldsymbol{List1} \times \boldsymbol{List2} \times \mathbb{R}$ with coordinates $(\theta_1, \phi_1, \theta_2, \phi_2, \lambda)$. We have the map

$$\begin{aligned} \rho : [0,\pi]\times[-\pi,\pi]\times[0,\pi]\times[-\pi,\pi]\times\mathbb{R} &\longrightarrow \mathrm{S}^3 \times \mathrm{S}^3 \times \mathbb{R}^2 \\ (\theta_1, \phi_1, \theta_2, \phi_2, \lambda)^T &\longmapsto \\ \begin{pmatrix} \cos\frac{\phi_1}{2} \\ \cos\theta_1 \sin\frac{\phi_1}{2} \\ \sin\theta_1 \sin\frac{\phi_1}{2} \\ 0 \end{pmatrix}, & \begin{pmatrix} \cos\frac{\phi_2}{2} \\ \cos\theta_2 \sin\frac{\phi_2}{2} \\ \sin\theta_2 \sin\frac{\phi_2}{2} \\ 0 \end{pmatrix}, \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \end{aligned} \tag{15}$$

where a_1 and a_2 are given by (14).

The configuration space M parameterizes the two eye complex that can rotate and separate such that the rotations satisfy Listing's constraint with respect to the axes of rotation. The separation vector is assumed to lie in a plane and the gaze vectors are assumed to be coplanar. In order to compute the optimal movement of the two eye complex, we would need to compute the Riemannian metric. Derivation of the optimal trajectory from there onwards is straightforward and analogous to that of (12). Let $G = (g_{ij})$, where g_{ij} is the ijth element of the corresponding Riemannian metric. One can show by explicit calculation that

$$g_{12} = g_{21} = g_{34} = g_{43} = 0$$

and a few of the other nonzero elements are given by

$$\begin{aligned} g_{11} &= \sin^2\frac{\phi_1}{2} + \lambda^2 \tan^2\phi_1, \\ g_{13} &= g_{31} = -\lambda^2 \tan\phi_1 \tan\phi_2 \cos(\theta_1 - \theta_2), \\ g_{14} &= g_{41} = \lambda^2 \tan\phi_1 \sec^2\phi_2 \sin(\theta_1 - \theta_2), \\ g_{15} &= g_{51} = \lambda \tan\phi_1 \tan\phi_2 \sin(\theta_1 - \theta_2), \\ g_{22} &= \frac{1}{4} + \lambda^2 \sec^4\phi_1, \\ g_{23} &= g_{32} = \lambda^2 \tan\phi_2 \sec^2\phi_1 \sin(\theta_2 - \theta_1), \\ &\text{etc.} \end{aligned}$$

The most important fact about the matrix G is that it is not diagonal, which illustrates the complexity of the two eye formation sensing problem. Interestingly, we note that G is not explicitly a function of θ_1 and θ_2 but $\theta_1 - \theta_2$ instead. This suggests different possible strategies for simplifying the G matrix as illustrated in the next subsection.

A. Different Subcases of the Binocular Formation Sensing

The 5-D configuration space and the associated Riemannian metric simplifies considerably under various different restrictive assumptions, which we shall now discuss.

Subcase A: If we assume that the axes of rotation of the two eyes remain parallel all the time, this is equivalent to constraining $\theta_1 = \theta_2$. The configuration submanifold M_1 for this case is given by the coordinates $(\theta, \phi_1, \phi_2, \lambda)$ and we have a map

$$\rho_1 : [0,\pi] \times [-\pi,\pi] \times [-\pi,\pi] \times \mathbb{R} \longrightarrow \mathrm{S}^3 \times \mathrm{S}^3 \times \mathbb{R}^2$$

which can be defined from (15) by restricting $\theta_1 = \theta_2$. The elements of Riemannian metric are given by

$$G = \begin{pmatrix} g_{11} & 0 & 0 & 0 \\ 0 & g_{22} & g_{23} & g_{24} \\ 0 & g_{23} & g_{33} & g_{34} \\ 0 & g_{24} & g_{34} & g_{44} \end{pmatrix}$$

where

$$\begin{aligned} g_{11} &= \sin^2\frac{\phi_1}{2} + \sin^2\frac{\phi_2}{2} + \lambda^2(\tan\phi_2 - \tan\phi_1)^2 \\ g_{22} &= \frac{1}{4} + \lambda^2 \sec^4\phi_1, \\ g_{33} &= \frac{1}{4} + \lambda^2 \sec^4\phi_2 \\ g_{44} &= (\tan\phi_2 - \tan\phi_1)^2, \\ g_{23} &= -\lambda^2 \sec^2\phi_1 \sec^2\phi_2 \\ g_{24} &= -\lambda \sec^2\phi_1(\tan\phi_2 - \tan\phi_1) \\ g_{34} &= \lambda \sec^2\phi_2(\tan\phi_2 - \tan\phi_1). \end{aligned}$$

Subcase B: If we assume that the axes of rotation of the two eyes remain parallel all the time and the distance between the center of the two eyes remains constant $(= 1)$, this is equivalent to constraining $\theta_1 = \theta_2 = \theta$ and $a_1^2 + a_2^2 = 1$. The configuration submanifold M_2 for this

case is given by the coordinates (θ, ϕ_1, ϕ_2) and we have a map

$$\rho_2 : [0, \pi] \times [-\pi, \pi] \times [-\pi, \pi] \longrightarrow S^3 \times S^3 \times \mathbb{R}^2$$

which can be defined from (15) by restricting $\theta_1 = \theta_2$ and restricting λ such that $a_1^2 + a_2^2 = 1$. Explicitly, the map is given as follows:

$$(\theta, \phi_1, \phi_2)^T \longmapsto \begin{pmatrix} \cos\frac{\phi_1}{2} \\ \cos\theta \sin\frac{\phi_1}{2} \\ \sin\theta \sin\frac{\phi_1}{2} \\ 0 \end{pmatrix}, \begin{pmatrix} \cos\frac{\phi_2}{2} \\ \cos\theta \sin\frac{\phi_2}{2} \\ \sin\theta \sin\frac{\phi_2}{2} \\ 0 \end{pmatrix}, \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix}.$$

The elements of Riemannian metric are given by

$$G = \begin{pmatrix} g_{11} & 0 & 0 \\ 0 & g_{22} & 0 \\ 0 & 0 & g_{33} \end{pmatrix}$$

where $g_{11} = 1 + \sin^2(\phi_1/2) + \sin^2(\phi_2/2)$ and $g_{22} = g_{33} = 1/4$. The Riemannian metric on M_2 is given by

$$g = \left(1 + \sin^2\frac{\phi_1}{2} + \sin^2\frac{\phi_2}{2}\right) d\theta^2 + \frac{1}{4}\left(d\phi_1^2 + d\phi_2^2\right)$$

and the corresponding geodesic equations are given by

$$\ddot{\theta} + \frac{1}{2}\frac{\sin\phi_1\dot{\phi}_1 + \sin\phi_2\dot{\phi}_2}{1 + \sin^2\frac{\phi_1}{2} + \sin^2\frac{\phi_2}{2}}\dot{\theta} = 0,$$
$$\ddot{\phi}_1 = \sin\phi_1\dot{\theta}^2\dot{\phi}_1, \quad \ddot{\phi}_2 = \sin\phi_2\dot{\theta}^2\dot{\phi}_2. \tag{16}$$

If we assume a potential function given by $V(\theta, \phi_1, \phi_2) = \sin^2(\phi_1/2) + \sin^2(\phi_2/2)$, then Euler Lagrange's equation (6) gives rise to the following set of equations of motion:

$$\begin{aligned}
\dot{z}_1 &= z_2, \\
\dot{z}_2 &= -\frac{1}{2}\frac{z_4\sin z_3 + z_6\sin z_5}{1 + \sin^2\frac{z_3}{2} + \sin^2\frac{z_5}{2}} z_2 \\
&\quad + \frac{1}{1 + \sin^2\frac{z_3}{2} + \sin^2\frac{z_5}{2}}\tau_\theta, \\
\dot{z}_3 &= z_4, \\
\dot{z}_4 &= z_2^2\sin z_3 - 2\sin z_3 + 4\tau_{\phi_1}, \\
\dot{z}_5 &= z_6, \\
\dot{z}_6 &= z_2^2\sin z_5 - 2\sin z_5 + 4\tau_{\phi_2}
\end{aligned}$$

where $(z_1, z_2, z_3, z_4, z_5, z_6) = (\theta, \dot{\theta}, \phi_1, \dot{\phi}_1, \phi_2, \dot{\phi}_2)$. If we wish to control the state $(\theta, \dot{\theta}, \phi_1, \dot{\phi}_1, \phi_2, \dot{\phi}_2)$ from $(\theta_0, 0, \phi_0^1, 0, \phi_0^2, 0)$ to $(\theta_1, 0, \phi_1^1, 0, \phi_1^2, 0)$ in T units of time while minimizing the energy

$$\int_0^T \tau_\theta^2(t) + \tau_{\phi_1}^2(t) + \tau_{\phi_2}^2(t)dt \tag{17}$$

we consider the Hamiltonian for the system given by

$$\mathcal{H}(z, \lambda) = \lambda \cdot \dot{z} - \frac{1}{2}\left(\tau_\theta^2 + \tau_{\phi_1}^2 + \tau_{\phi_2}^2\right)$$

where λ as before is the costate vector. The optimal path can be computed using Hamilton's equation

$$\dot{z} = \frac{\partial\mathcal{H}}{\partial z}, \quad \dot{\lambda} = -\frac{\partial\mathcal{H}}{\partial\lambda}$$

analogous to what has been shown in Section II. The optimal control can also be computed likewise using the maximum principle (11).

To end this section, we would like to reiterate some of the main points. We introduce and analyze optimal control problems in binocular sensing, wherein we consider a pair of visual sensors that can rotate and translate with the constraint that the gaze directions of the cameras always remain coplanar. The constrained space is parameterized as a Riemannian manifold and we illustrate the computation of the optimal trajectory which minimizes a suitably chosen cost function (17). Most importantly, when eye rotations take place using rotation matrices that are constrained within the associated configuration space (13), orientation ambiguity for a given gaze direction is resolved. These techniques can, in principle, be generalized to more than a pair of sensors although this fact is not illustrated in this paper.

Eye rotation satisfying List's constraints resolves the orientation ambiguity at a given gaze direction.

Remark: In closing this section, we would like to remark that the problem of optimally orienting visual sensors to track a moving target in space while satisfying Listing's constraint has been described in detail in [74]. In this

paper, we simplify the problem of sensor localization to only estimating the position of visual sensors in space. The orientation of the sensors is already constrained by Listing's constraint, implementation of which requires an additional price of internode communication. Many other sensor localization and camera calibration techniques are quite standard in the machine vision literature and we would like to refer to [33], [40], [58], [92], and [101]. For a geometric introduction to the constrained optimal trajectory tracking problem, see Hussein and Bloch [47].

IV. IMAGE REPRESENTATION AND SPARSE CODING

As has been noted in the introduction, the spatiotemporal sequence of images on the retina (or the image plane of a camera) is not very useful for subsequent processing unless it has been represented, perhaps with respect to a set of basis functions. Such a representation typically reduces the data set, highlights the more relevant parts of the visual field, and renders the data amenable for sensor fusion. In mammalian vision, for example, retinal images are internally represented by the "visual cortex" and in this paper we would like to study this representation. The problem we raise in this section is described as follows.

Problem III (Image Representation Problem): Assuming that we have a stream of images incident on the retina, what would constitute a good basis function representation for subsequent encoding of this spatiotemporal stream?

For the general problem of sensor design and fusion, the answer to Problem III will vary. We restrict our discussion to the visual circuit of mammalian vision and restrict further to the visual circuit in freshwater turtles. The mammalian visual cortex has evolved over millions of years to effectively cope with images of the natural environment. It is reasonable to think that the cortex has discovered efficient coding strategies for representing these images.

Our basic starting point in this section is to note the principle of redundancy reduction proposed by Barlow [8], [9], which states that a useful goal of sensory coding is to transform the input in such a manner that reduces the redundancy due to complex statistical dependencies among elements of the input stream. A reasonable goal of the visual system, then, is to extract the statistical dependencies so that images may be explained in terms of a collection of independent events. The hope is that such a strategy will recover an explicit representation of the underlying independent entities that give rise to the image, which would be useful to the survival of the organism. This principle has been successfully applied by Atick and colleagues, over the last decade [5]–[7], towards understanding the response properties of retinal ganglion cells using pairwise correlations among image pixels. As was observed by Field [34], Olshausen and Field [68], natural images have oriented lines and edges that have statistical dependencies of higher order than pairwise correlations.

In order to address the problem of reducing such higher order forms of redundancies, Olshausen and Field [69] have proposed to represent images described in terms of a linear superposition of basis functions. These functions are adopted so as to best account for the image structure in terms of a collection of statistically independent events. It is conjectured that the probability distribution of these events is such that a given image is represented by a small number (sparse) of basis functions chosen out of a large set (over-complete). Such a strategy, called "sparse over-complete image coding," has been detailed in [69]. In this section, we examine more closely the consequences of utilizing such a code in which the number of basis functions is greater than the dimensionality of the input. As a result, sparsification weeds out those basis functions not needed to describe a given image structure.

A. The Image Model and Representation With Sparse Codes

Assume that an image patch, $I(x)$, is described as a linear superposition of a set of basis functions, $\phi_i(x)$, with amplitudes a_i such that

$$I(x) = \sum_i a_i \phi_i(x) + \nu(x) \tag{18}$$

where x denotes spatial position within the patch and the variable ν represents independent and identically distributed Gaussian noise. The basis functions are trained on the set of images by adapting the probabilistic model to the statistics of the images. Said differently, we wish to match the distribution, $P(I|\phi)$ of images arising from the above model, to the actual distribution $P^*(I)$ of images observed in nature. In order to calculate the probability of images arising from the model, we need to specify the prior probability distribution of the coefficients $P(a)$ as well as the probability of an image arising from a certain state of the coefficients in the model $P(I|a, \phi)$. It follows that

$$P(I|\phi) = \int P(I|a, \phi)P(a)da.$$

The probability $P(I|a, \phi)$ of an image arising from a particular choice of coefficients essentially expresses our model of the level of noise. In view of (18), we obtain

$$P(I|a, \phi) = \frac{1}{Z_{\sigma_N}} e^{-\frac{|I-a\phi|^2}{2\sigma_N^2}}$$

where $|I - a\phi|^2$ denotes the sum

$$\sum_x \left[I(x) - \sum_i a_i \phi_i(x) \right]^2$$

σ_N^2 is the variance of the noise, and Z_{σ_N} is a normalizing constant. Since the basis set is over-complete, there will be infinitely different choices of the coefficient vector "a" for any given image; our problem is to choose one which is sparse. We achieve this by choosing a suitable prior distribution.

Assume first of all that the different components a_i of the parameter vector a are statistically independent, so that we have

$$P(a) = \prod_i P(a_i).$$

As was noted by Olshausen and Field [69], "the notion of sparseness is incorporated by shaping the probability distribution of each a_i to be unimodal and peaked at zero with heavy tails." This would ensure that most coefficients a_i would be close to zero. As an example, one can choose

$$P(a_i) = \frac{1}{Z_\beta} e^{-\beta S(a_i)}$$

where the function S determines the shape of the distribution, β is a parameter that controls its steepness, and Z_β is a normalizing constant. Choosing $\beta = 1$ and $S(x) = log(1 + x^2)$ corresponds to specifying a Cauchy distribution for the prior, which has the desired sparse shape.

In order to obtain the desired sparse code, one resorts to minimizing the following cost function:

$$E = \sum_x \left[I(x) - \sum_i a_i \phi_i(x) \right]^2 + \lambda \sum_i S(a_i).$$

The cost function E has two components. The first term computes the reconstruction error and the second term incurs a penalty for utilizing too many basis functions. The cost function E is minimized in two separate phases, one nested inside the other. In the inner phase, E is minimized with respect to the a_i, for each image, holding the ϕ_i fixed. In the outer phase (i.e., on a long timescale and over many image presentations), E is minimized with respect to the ϕ_i. For details on this optimization, we once again refer to [69].

B. Image Compression With Principal Components

In our discussion so far, we have obtained "sparse codes" as coefficient vectors of the form $\{a_i, i = 1, 2, \ldots\}$ that can encode an image $I(x)$ with respect to a set of basis functions $\phi_i(x)$. Instead of obtaining sparse codes for the entire image, we subdivide an image into blocks along the horizontal direction and each block is further subdivided into patches. In order to get codes of low dimension, it is natural to encode each image patch using sparse codes and subsequently concatenate the codes within each block for further data compression using principal component analysis. The details are sketched in Fig. 4. The sparse coefficient vectors for the patches within each block of the image are arranged into a single vector A^j, where $j = 1, 2, \cdots, J$ and J is the number of blocks. So the image can be sparsely represented by vectors in the form

$$A = [A^1, A^2, \cdots, A^J].$$

These vectors are further compressed using principal component analysis and the image can be encoded by a few significant principal components. The codes obtained from the principal components could be transmitted and decoded over a wireless network for any of the stated purposes discussed in the introduction. In the next two sections, we follow the biological track and use the codes as input to the visual cortex or to the hippocampus. By doing so, we introduce a network of neurons or oscillators and study how network activity is altered as a result of changes in the observed visual scene.

Remark: We would like to remark that the general strategy of sensor fusion using code concatenation makes

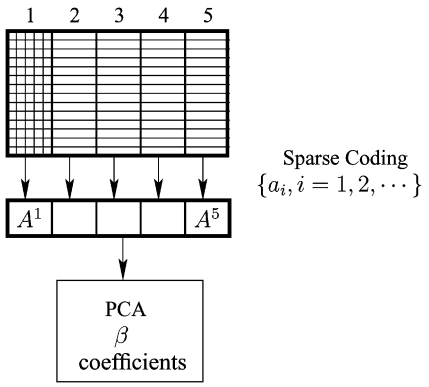

Fig. 4. *The image is subdivided into blocks along the horizontal direction and each block is further subdivided into patches. The sparse vectors of the jth block of the image are stacked into a single long vector and compressed using principal components (PCA) to generate β coefficients for the wave generator.*

sense only because the orientation ambiguity of the camera sensors has already been resolved. The codes representing multiple images are functions of the target pose and the sensor location.

V. PATTERN GENERATION WITH A POPULATION OF NEURONS—CORTICAL COMPUTATION

Sensory information is typically encoded in animals via synchronized activity of a population of neurons. For example, mammals have a cerebral cortex that embodies several topographically organized representations of visual space. Extracellular recordings show that neurons in a restricted region of the visual cortex are activated when a visual stimulus is presented to a restricted region of the visual space, the classical receptive field of the neuron [25]. Neurons at adjacent points in the cortex are activated by stimuli presented at adjacent regions of the visual space. Consequently, there is a continuous but deformed map of the coordinates of the visual space to the coordinates of the cortex. Extracellular recordings from the visual cortex of freshwater turtles, on the other hand, produce a different and interesting result [86], described as follows.

"*Neurons at each cortical locus are activated by visual stimuli presented at every point in the binocular visual space, although the latency and shape of the response waveforms vary as the stimulus is presented at different loci in the visual space.*"

This suggests that position in the visual space is represented not by a retinotopic map, but by a wave of neural activity. Experiments conducted by Senseman *et al.* [81], [86], [87] and Prechtl *et al.* [76]–[78] support this viewpoint. These experiments used voltage sensitive dyes and multielectrode arrays to show that presenting a visual stimulus to the retina of an *in vitro* preparation of the turtle eye and brain produces a wave of depolarization that propagates anisotropically across the cortex.

More recent studies [30] have indicated a connection of sensory stimuli and traveling electrical waves with their possible role in computation. Such waves have also been described in the primary visual [23], [88] and somatosensory cortices [27], [36], [72] of mammals, including humans [98]. Kopell and coworkers [53] have studied *rhythms of the nervous system* using dynamically coupled oscillator network synchronization [2], [13], [54], [65] and have linked these networks to important behavioral and cognitive states including attention, working memory, associative memory, object recognition, sensory motor integration, and perception, among others. It has long been known [93] that the firing of hippocampal cells during rat movement is strongly correlated with the theta rhythm (i.e., a field potential oscillation with a frequency range of 4–12 Hz). More recently [70], an interesting relationship between the theta rhythm and the firing of place cells that selectively respond to specific portions of the environment has been observed. Place cells have been subsequently used by Brown and coworkers [17], [18] in a statistical paradigm to decode position from ensemble firing patterns of the rat hippocampus.

Our goal in this section is to explore the connection between visual sensory inputs and associated waveform generated using a recently introduced [63], [64] large-scale model of the turtle visual cortex. The emphasis is primarily on the role of cortical waves (see Fig. 7) for the purpose of encoding and decoding.

A. A Large Scale Model of the Turtle Visual Cortex

The turtle visual cortex has three layers (an outer layer 1, an intermediate layer 2, and an inner layer 3) and contains at least 11 morphologically distinct types of neurons, only some of which are well characterized. These are the pyramidal cells, the subpial cells, the stellate cells, and the horizontal cells. Pyramidal cells have their somata located in intermediate layer 2 of the cortex and are predominantly excitatory. The other three types of cells are inhibitory; subpial cells have their somata and dendrites situated in the outer half of layer 1, stellate cells have their somata situated in the inner half of layer 1, and horizontal cells have their somata situated in layer 3. Visual input from the retina is routed through an intermediate structure called the *lateral geniculate complex* (LGN). Pyramidal cells, subpial cells, and stellate cells receive direct [feedforward, Fig. 6(b)] projections from the

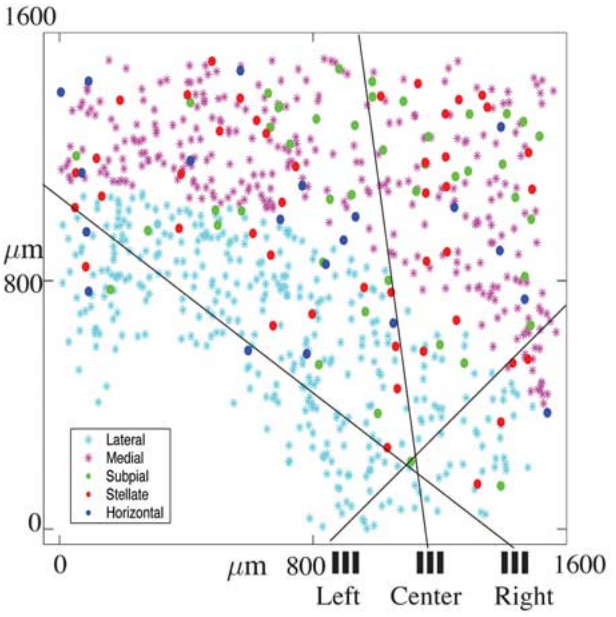

Fig. 5. *Distribution of cells in each of the three layers of the turtle cortex projected on a plane. The lateral geniculate (LGN) cells are distributed linearly (shown at the right side of the bottom edge of the cortex) and the solid line shows how they interact with cells in the cortex. The distribution of subpial cells is not shown.*

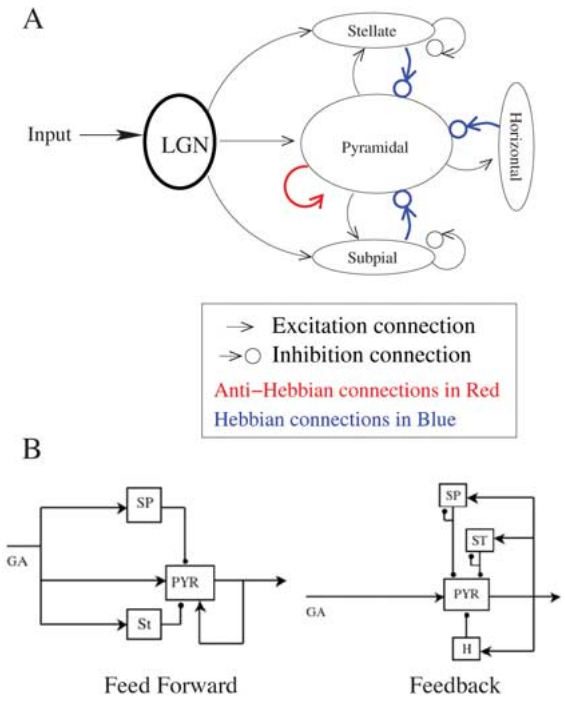

Fig. 6. ***Cortical circuit of freshwater turtles. Interconnection between neurons in various layers of the visual cortex is shown.***

LGN. Pyramidal cells make projections to subpial, stellate, horizontal, and other pyramidal cells. Subpial, stellate, and horizontal cells make feedback projections to pyramidal cells [Fig. 6(b)]. See Fig. 5 for a distribution of cells in each of the three layers projected on a plane and Fig. 6(a) for a schematic diagram of the interconnection between the cells (for details, see [94]).

The large-scale model is constructed by representing each neuron through a multicompartmental model based on the anatomy of the neurons. Each compartment is modeled by a standard membrane equation and implemented in GENESIS [14]. The model that we have constructed and shown in Fig. 5, consists of 200 LGN cells, 680 pyramidal cells, 44 subpial cells, 50 stellate cells, and 20 horizontal cells. For a detailed description of the constructed model, see [64] and [94].

B. Simulation of Cortical Waves With Stationary Stimuli

A stationary point target stimulus has been simulated by presenting a 50-ms square current pulse to a set of adjacent LGN neurons. For the purposes of our simulation, we first considered 20 equidistant positions of stimuli across the LGN (three of these positions—left, right, and center—are shown in Fig. 5). The simulation time was set to 800 ms. Membrane potentials of individual pyramidal neurons were recorded and saved in a data file and visualized as movies as shown in Fig. 7. The cell coordinates, indicating the position of the neuron in the cortex, follow a statistical distribution that can be estimated from experimental data (see [64] for details). In this paper, we assume that these distributions are known.

C. Detection From the Cortical Waves Using Principal Component Analysis

Let $I(x, y, t)$ denote the spatiotemporal signal of response of the model to different stationary stimuli as detailed in the last subsection. It follows that $I(x, y, t)$ can be viewed as a collection of movie frames (snapshots). Given that every frame is $n \times m$ pixels and every movie has p frames, it follows that the dimension of $I(x, y, t)$ could be as high as $n \times m \times p$. Our main interest is to compare two different responses (movies) in order to quantify the differences and similarities between them. We now describe the principal components-based technique for such a comparison.

Principal components have been widely used in various disciplines, including image and signal processing, data compression, fluid dynamics, partial differential equations [26], weather prediction, etc. [43]. In image processing, the method is used for removing a redundancy (decorrelated pixels) from images [80]. The transformation itself is linear, representing a rotation of a coordinate system so that neighboring pixels in the new coordinate system are less correlated. Moreover, the rotation proposed by the method is optimal as it leads to a complete removal of the correlation from neighboring pixels, which is equivalent to diagonalizing the image correlation matrix. Consequently, the image can be approximated in a low-dimensional subspace, using only selected basis vectors, also called principal eigenvectors. Depending on the context, this method goes by various names: Karhunen–Loeve (KL) decomposition, proper orthogonal decomposition, Hotelling

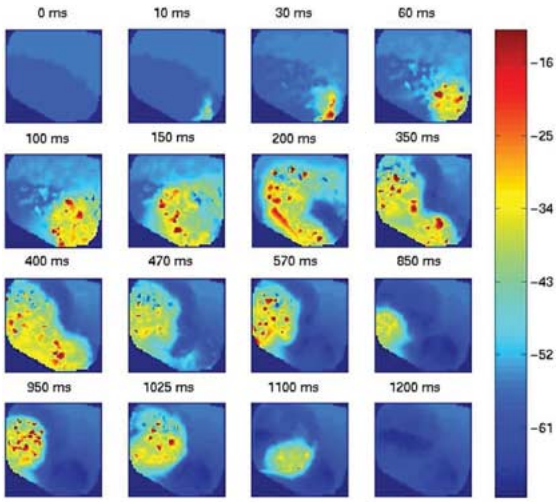

Fig. 7. ***A traveling wave of cortical activity from the model cortex.***

decomposition, and singular value decomposition. We shall refer to it as KL decomposition.

In the KL decomposition, an $n \times p$ pixel frame is written as a vector of size $n \times p$. Therefore, the kth spatiotemporal signal I^k can be viewed as a collection of m frames

$$I^k = \{u_1^k, u_2^k, \ldots, u_m^k\}$$

where each vector u_j^k is an element of $\mathbb{R}^{np}$. The qth order successive reconstruction of the spatiotemporal signal $I^k(x, y, t)$ is given by

$$\hat{I}^k(x, y, t) = \sum_{i=1}^{q} \alpha_i^k(t) M_i(x, y) \tag{19}$$

where $M_i(x, y) \in \mathbb{R}^{np}$ are the principal modes.

The analysis presented in [63] observed that the first few principal components capture most of the energy content of a movie. In particular, a third-order approximation utilizing three principal modes $M_1(x, y)$, $M_2(x, y)$, $M_3(x, y)$, carries over 99% of the energy content. These three modes have been sketched in the left-hand column of Fig. 8. The kth movie is closely represented by the vector time trajectory

$$\boldsymbol{\alpha}^{\mathbf{k}} = \left(\alpha_1^k(t), \alpha_2^k(t), \alpha_3^k(t), t = 1, 2, \ldots, m\right)^T$$

in $\mathbb{R}^{3m}$, where t varies between each of the m time frames. In the right-hand column of Fig. 8, the time coefficients $\alpha_1^k(t)$, $\alpha_2^k(t)$, $\alpha_3^k(t)$ have been sketched.

Because the positions of neurons in the model cortex are randomized with each new simulation, the vector function $\boldsymbol{\alpha}^{\mathbf{k}}$ can be viewed as a random process. Statistical analysis of this process can be facilitated if the process is further parameterized using a second KL decomposition. The rth order successive approximation of the kth random vector $\boldsymbol{\alpha}^k$ is given by

$$\begin{bmatrix} \alpha_1^k(t) \\ \alpha_2^k(t) \\ \alpha_3^k(t) \end{bmatrix} = \sum_{j=1}^{r} \beta_j^k \Phi_j(t).$$

It follows that the two KL decompositions lead to a convenient representation of the kth movie $I_k(x, y, t)$ by a vector in R^r given by

$$\boldsymbol{\beta}^{\mathbf{k}} = \left(\beta_1^k, \beta_2^k, \ldots, \beta_r^k\right)^T.$$

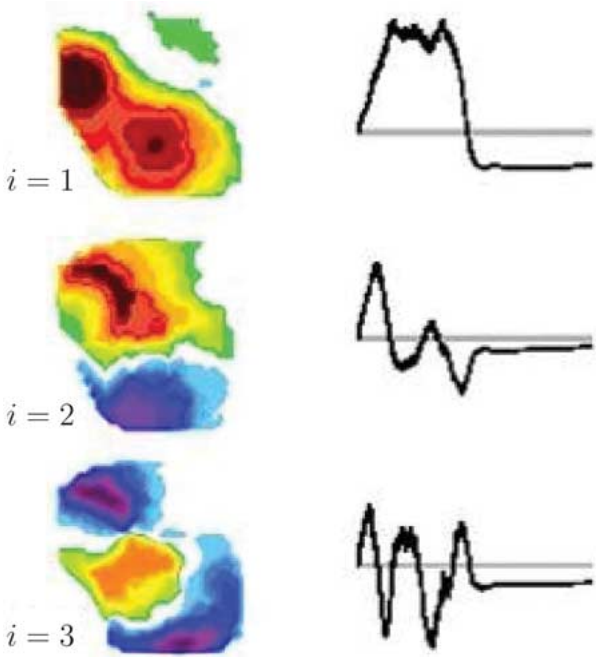

Fig. 8. *The left-hand column shows the three principal spatial modes. The right-hand column shows the corresponding time coefficients.*

In our analysis, we use a third-order approximation ($r = 3$) and each data set is represented by a point in $\mathbb{R}^3$, conveniently called the β space.

Because of the randomness of the model, presentation of the same stimulus does not produce the same response in general. Correspondingly, these responses do not produce the same point in the β space, but appear clustered. These clusters move in the β space as the position of the stimulus changes. This is illustrated in Fig. 9 wherein only the second and the third coordinates of the cluster have been shown and 20 different positions of the stimulus have been chosen uniformly from left to right along the LGN array.

The cluster of points in the β space can be used to compute the conditional density functions, and these functions can subsequently be used in statistical detection of stimulus location using hypothesis testing algorithms [91] (see Fig. 10 for the associated decision space with three stimulus points located at the left, center, and right positions of the LGN array. The decision spaces are sketched for three different time windows from top to bottom. The overlapping clusters in the bottom-most picture show that stimulus location is not detected accurately.).

D. Hebbian and Anti-Hebbian Control of Patterns

We have already seen in this section that a group of neurons in the turtle visual cortex has the ability to sustain a traveling wave. Typically this wave results as an interaction between a feedforward and a feedback circuit

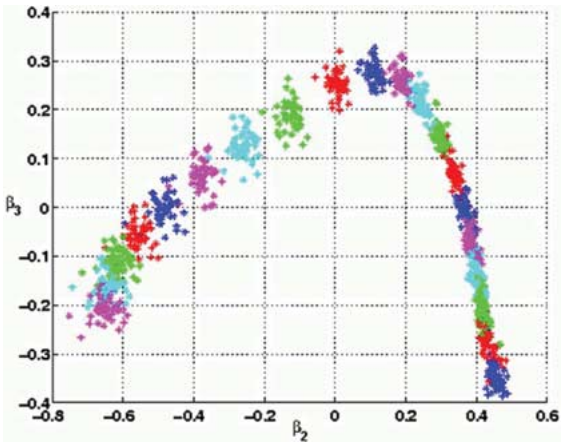

Fig. 9. ***Cluster of points in β space corresponding to different stationary stimuli. Each cluster contains 50 points as a result of simulating the cortex model 50 times with each of 20 stimuli. Two-dimensional versions of the plots are shown for clarity using the second and third components of the β vector.***

(see Fig. 6), the details of which have been explained in [94]. Roughly speaking, the feedforward circuit controls the origination and propagating speed of the traveling wave and the feedback circuit controls the propagation duration. Waves are typically generated in the pyramidal cells as a result of an external input current that results in an increase in membrane potential. Pyramidal cells locally excite each other, resulting in a region of neural activity which tends to propagate in all directions. Left unabated, these pyramidal cells would excite the entire cortex. Fortunately, the feedforward circuit incorporates inhibitory actions from the stellate and subpial cells. Although the precise roles of the two inhibitory cells are different and somewhat unclear, they control the timing of wave generation. There are inhibitory actions that inhibit the wave using a feedback circuit due to three different cells: subpial, stellate, and horizontal. The feedback inhibition reduces and eventually kills the neuronal activity at the spot where the activity is greatest. The combined effect of the two circuits gives the appearance of a traveling wave. Eventually these waves are killed by a strong gaba (a type of synaptic input) initiated inhibition that originates after a long delay.

Using the large-scale model of the visual cortex that consists of excitatory and inhibitory cells described above, we observed that the neuron population remained hyperpolarized (i.e., maintained a very low membrane potential) long after the initial wave had been killed. The cortex remained unresponsive to future visual inputs, an undesirable property. One way to remedy this problem is to detect this period of hyperpolarization and increase the synaptic interaction between the excitatory pyramidal cells. This would amplify the tiny input into the pyramidal cells, forcing these cell populations to get out of hyperpolarization. This was achieved successfully, using Hebbian and anti-Hebbian adaptation.

In *Hebbian adaptation*, the synaptic strength between two cells increases in proportion to the product of the pre- and post-synaptic activities. Likewise, in *anti-Hebbian adaptation*, the synaptic strength between two cells decreases in proportion to the product of the pre- and post-synaptic activities. In our model, the excitatory interconnection between pyramidal cells is chosen to be anti-Hebbian. This produces increasingly larger synaptic weights between pyramidal cells once the wave has *died out*. The inhibitory interactions between the stellate/subpial/horizontal and the pyramidal cells are chosen to be Hebbian. These produce increasingly stronger inhibition to active pyramidal cells. In Fig. 11, we show anti-Hebbian

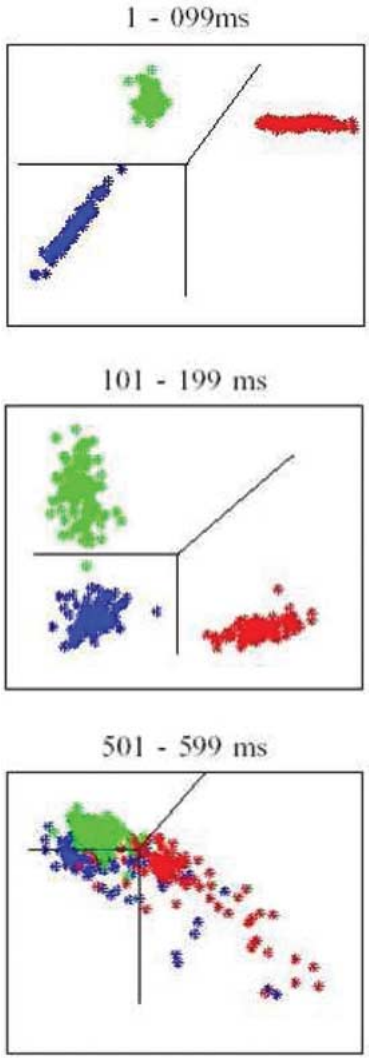

Fig. 10. ***Decision space for three hypothesis detection. The coordinates are log likelihood ratios computed for three different time windows.***

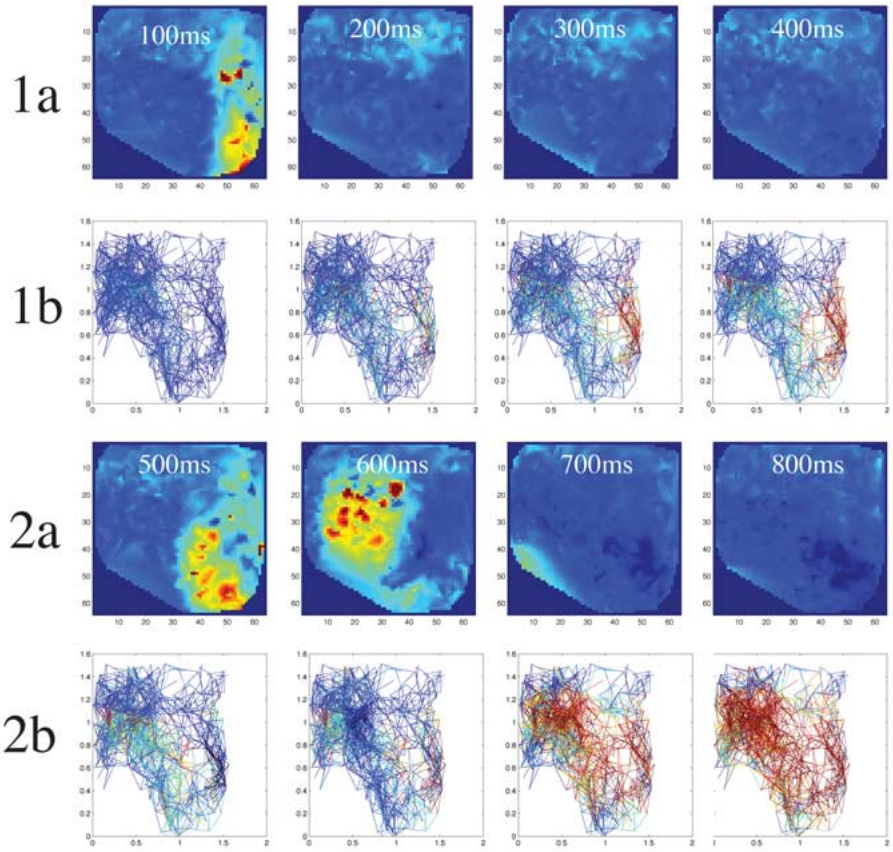

Fig. 11. *Pyramidal to pyramidal anti-Hebbian synaptic response to changes in the pyramidal activity. (1a): Frames of pyramidal cell activity due to pulse input to the LGN at 0 ms lasting for 150 ms. (1b): Frames of weight responses corresponding to the activities in 1a. (2a): Frames of pyramidal cell activity due to pulse input to the LGN at 400 ms following the first pulse lasting for 150 ms. (2b): Frames of synaptic weight responses corresponding to activities in 2a.*

action on the pyramidal cells. Rows 1a and 2a show wave activity as a function of time. After about 700 ms, the first round of waves has been inhibited and the pyramidal cells are hyperpolarized. The weights between the cells are very large, as indicated by the red lines in rows 1b and 2b of Fig. 11. A subsequent input causes a second round of waves (not shown in the figure).

VI. MEMORY AND RECALL OF EVENT SEQUENCE WITH A NETWORK OF HIPPOCAMPAL CELLS

In this section, we show the cortex in relation to other functional units of the animal's brain viz. the hippocampus, as shown in Fig. 12. A salient stimulus from the cortex captures the animal's attention and drives the hippocampus to a theta state [57]. In this state, the hippocampus produces a 4–8 Hz wave in the local field potential (LFP). Note that LFP is the integration of extracellular field potential over a large population of cells. It is an electrical potential of up to several millivolts produced by a cell population. In the theta state, the LFP of the hippocampus shows a rhythm that is called the theta rhythm. In this state, the hippocampus is believed to be actively engaged in memory formation [20].

In addition to the theta rhythms, the hippocampus also has a family of place cells [85], [99], which fire selectively in correlation with the rat's position while navigating a known environment. As the rat passes through a place field, over a matter of seconds, many successive theta

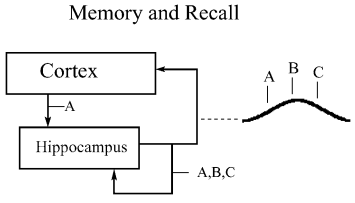

Fig. 12. *Sequence memory in the hippocampus. The associations from A to B to C are stored in the hippocampal synaptic connections which leads to the recall of the A,B,C sequence in one theta cycle when the place cells for A are activated by the cortex.*

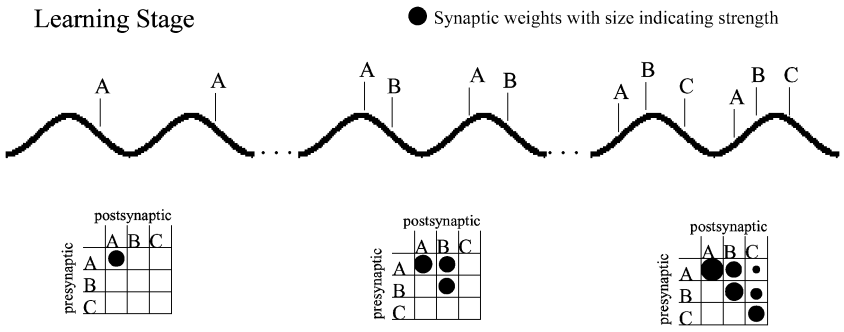

Fig. 13. ***Formation of the A to B to C asymmetric synaptic sequence associations by slowly progressing cortex stimulation and theta phase precession.***

cycles pass. It has been observed that the phase of the theta cycle at which the place cells fire progresses gradually, a phenomenon called the "theta phase precession" [84] (see Fig. 13).

It has been proposed [57] that information in the hippocampus is carried by the phase at which the place cell fires with respect to the theta rhythm, and not just by its firing rate. The phase timing of the place cells encoding the place field location information helps us to solve a well-known problem of sequence encoding. A sequence of locations visited by the animal, is encoded by the firing of place cells separated by phase over a single cycle of the theta rhythm. It is shown in Fig. 13 that the events *A*, *B*, and *C* may have taken place in intervals of 1 s to 1 min. The associated place cells fire with a difference of up to about 100 ms, shown in Fig. 13 at the third cycle.

Synaptic plasticity was first observed in the brain by Bliss and Lomo [11] in the rabbit hippocampus, which led to the view of the hippocampus as a key center for learning in the brain. It was later established by Bi and Poo [10] that the form of learning in the hippocampus, in particular learning a sequence of events, is the spike time dependent synaptic plasticity (STDP). STDP states that for two events whose associated place cell spiking activity falls within the range of about 40 ms, the synaptic association between the corresponding place cells is strengthened. In Fig. 13, we show the synaptic association between events *A*, *B*, and *C*. When only event *A* occurs, the synaptic strength between place cells that represent *A* is strengthened. When event *B* follows event *A*, cells associated with events *A* and *B* both fire (*A* slightly leading *B* in phase) within one cycle of the theta rhythm. This strengthens the synaptic association from *B* to *A*. The situation is similar when all three events *A*, *B*, and *C* occur in succession.

As a result of the association between place cells corresponding to the three locations, these locations are memorized in the order they were visited. At the time of recall, the current sensory input could provide a recall cue. For example, starting from position *A*, the animal could recall that it had previously visited positions *B* and *C* after *A*, and the recall takes place within one cycle of the theta rhythm.

VII. DETECTION OF PATTERNS USING KURAMOTO TYPE OSCILLATOR NETWORKS

The purpose of this section is to introduce yet another computing paradigm, emerging from a network of oscillators, for the purpose of decoding from cortical waves. Elements of the oscillator network interact with each other via phases rather than amplitudes; memorized patterns correspond to synchronized states. Each unit of the oscillator network oscillates with the same frequency and a prescribed phase relationship. For pattern recognition with a network of oscillators, phase differences, instead of phases, play a crucial role. The mechanism of recognition is related to phase locking. To illustrate the main idea, we would like to review a model proposed by Kuramoto [56].

A. Phase Locking With a Network of Kuramoto Models

Consider a dynamical system of the form

$$\dot{\phi}_i = \omega + \sum_{j=1}^{N} s_{ij} \sin(\phi_j - \phi_i + \psi_{ij}) \tag{20}$$

where ϕ_i, $i = 1, \cdots, N$ (assume $N = 2$ for illustration), are phase variables taking values in the interval $[-\pi, \pi)$. The parameters s_{ij} and ψ_{ij} are assumed to satisfy $s_{ij} = s_{ji}$, $\psi_{ij} = -\psi_{ji}$. The index i refers to the ith unit and these units are coupled. In order to understand the dynamics of (20), we define a new variable $\phi = \phi_1 - \phi_2$ and rewrite (20) as follows:

$$\dot{\phi} = -2s_{12} \sin(\phi - \psi_{12}). \tag{21}$$

The stationary points of (21) are given by $\phi - \psi_{12} = k\pi$, out of which the stable points are given precisely by

$$\phi - \psi_{12} = 2k\pi, \quad k = 0, \pm 1, \pm 2, \cdots. \tag{22}$$

For ϕ_1, ϕ_2 in the interval $[-\pi, \pi)$, $\phi = \psi_{12}$ and $\phi = \psi_{12} + 2\pi$ are the two stable points if $\psi_{12} < 0$, and $\phi = \psi_{12}$ and $\phi = \psi_{12} - 2\pi$ are the two stable points if $\psi_{12} > 0$. Up to mod 2π, the two stable points of ϕ are actually the same indicating that (21) converges globally to an unique equilibrium point.

B. Memory With Two Elements

Let us discuss the problem of detecting n patterns with a Kuramoto model using two units (i.e., $N = 2$). In order to use (21) for the purpose of memorizing n patterns, we would require that it have (at least) n equilibria. This can be achieved by rescaling the phase variables as

$$\bar{\phi}_1 = \frac{1}{n}\phi_1, \quad \bar{\phi}_2 = \frac{1}{n}\phi_2.$$

Rewriting (20) with respect to the new variables, we obtain

$$\dot{\bar{\phi}}_1 = \frac{1}{n}\omega + \frac{1}{n}s_{12}\sin(n\bar{\phi}_2 - n\bar{\phi}_1 + \psi_{12})$$
$$\dot{\bar{\phi}}_2 = \frac{1}{n}\omega + \frac{1}{n}s_{21}\sin(n\bar{\phi}_1 - n\bar{\phi}_2 + \psi_{21}).$$

By defining $\bar{\phi} = \bar{\phi}_1 - \bar{\phi}_2$, we obtain analogously the following equation:

$$\dot{\bar{\phi}} = -\frac{2}{n}s_{12}\sin(n\bar{\phi} - \psi_{12}). \tag{23}$$

Up to mod 2π, the n stable stationary points of (23) are given by $\bar{\phi}_k^e = (\psi_{12}/n) + (2(k-1)\pi/n)$ if $\psi_{12} < 0$. Additionally, it can be verified that if

$$\bar{\phi}_k^e - \frac{\pi}{n} < \bar{\phi}(0) < \bar{\phi}_k^e + \frac{\pi}{n} \tag{24}$$

then $\bar{\phi}(t)$ converges to the kth stable stationary point $\bar{\phi}_k^e$. The phase difference variable $\bar{\phi}(t)$ can be plotted as a unit complex number $e^{i\bar{\phi}(t)}$. In Fig. 14, such a plot is shown when the rescaling parameter is 3. This gives rise to three stable stationary points at $\bar{\phi}_k^e = (\psi_{12}/3) + (2(k-1)\pi/3)$, $k = 1, 2$, and 3.

The main idea is to utilize the convergence properties of (23) to distinguish among n complex patterns. Let us define the following n vectors in $\mathbb{C}^2$ as

$$p_1 = \begin{pmatrix}\pi_1 \\ \pi_2\end{pmatrix} \text{ and } p_k = \begin{pmatrix}e^{+i\frac{(k-1)\pi}{n}}\pi_1 \\ e^{-i\frac{(k-1)\pi}{n}}\pi_2\end{pmatrix} \tag{25}$$

for $k = 2, 3, \cdots, n$, where π_1 and π_2 are any two complex numbers such that

$$|\pi_1| = |\pi_2| = 1$$

and

$$arg(\pi_1\bar{\pi}_2) = \frac{\psi_{12}}{n}.$$

The complex vectors p_k, $k = 1, 2, \ldots, n$, are n memorized complex patterns associated with n stable equilibria $\bar{\phi}_k^e$, $k = 1, 2, \ldots, n$. Let us define a mapping

$$\xi : \mathbb{C}^2 \longrightarrow \mathbb{R} \tag{26}$$

as follows:

$$\begin{pmatrix}w_1 \\ w_2\end{pmatrix} \longmapsto arg(w_1\bar{w}_2).$$

It would follow that $\xi(p_k) = \bar{\phi}_k^e = (\psi_{12}/n) + (2(k-1)\pi/n)$. Thus the n patterns p_k, $k = 1, 2, \cdots, n$, are mapped to the n stable equilibria of (23) under the map ξ. Patterns which are close to any p_k would be attracted towards the corresponding kth equilibrium. This principle can therefore be used as memory.

However, we are not interested in a set of complex patterns. Rather, we would like to memorize patterns of real vectors. Assume that we have n vectors v_k, $k = 1, 2, \cdots, n$, in $\mathbb{R}^Q$ which we would like to memorize. We consider a map

$$T : \mathbb{R}^Q \longrightarrow \mathbb{C}^2 \tag{27}$$

such that

$$v_k \longmapsto p_k, \quad k = 1, 2, \cdots, n$$

where p_k-s are defined as above. So the memorized patterns are associated with phase difference equilibria via the map ξT, where

$$\xi T(v_k) = \bar{\phi}_k^e.$$

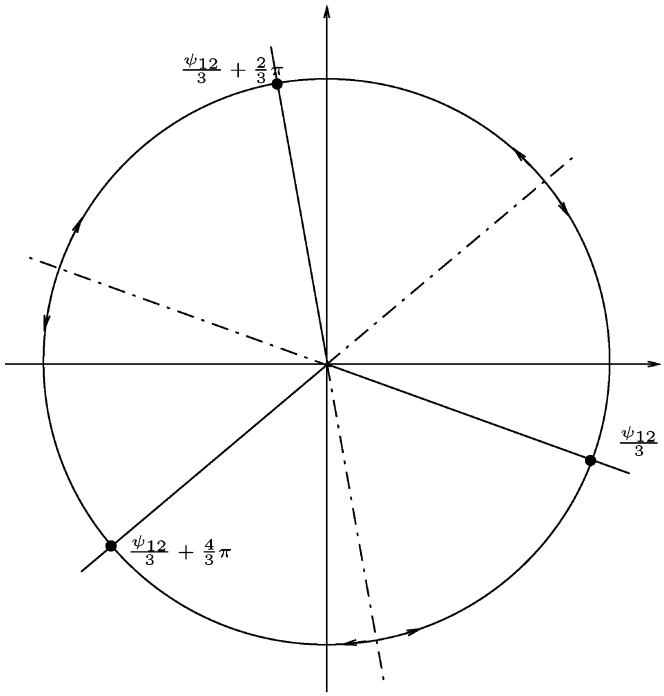

Fig. 14. *Phase variable $\bar{\phi}(t)$ is plotted as a unit complex number $e^{i\bar{\phi}(t)}$ with the rescaling parameter 3, showing three stable equilibria which result in three regions of convergence for the dynamical system (23) under initial conditions constrained by (24).*

Thus, the dynamics of (23) can be used to memorize the n patterns. To recognize (recall) a pattern v in $\mathbb{R}^Q$, the phase variables ϕ_i-s of the two oscillatory units can be initialized with $T(v)$ and $\bar{\phi}(t)$ converges to one of the equilibria.

C. Memorizing Patterns With Higher Order Dynamics

As discussed in the previous subsection, n patterns can be memorized with nonlinear dynamics of a second-order Kuramoto model using a phase difference that is rescaled by n. In general, the nonlinear dynamics of a higher order Kuramoto model can be used to memorize more patterns using the same rescaling, or the same number of patterns using lower rescaling, since more equilibria can be achieved in higher order Kuramoto models. We now consider a locally connected Kuramoto model with N units described as follows:

$$\begin{aligned}
\dot{\phi}_1 &= \omega + s_{12}\sin(\phi_2 - \phi_1 + \psi_{12}) \\
&\quad + s_{1N}\sin(\phi_N - \phi_1 + \psi_{1N}) \\
\dot{\phi}_i &= \omega + s_{i(i-1)}\sin\big(\phi_{(i-1)} - \phi_i + \psi_{i(i-1)}\big) \\
&\quad + s_{i(i+1)}\sin\big(\phi_{(i+1)} - \phi_i + \psi_{i(i+1)}\big) \\
&\quad \text{for } i = 2, \cdots, (N-1) \\
\dot{\phi}_N &= \omega + s_{N(N-1)}\sin\big(\phi_{(N-1)} - \phi_N + \psi_{N(N-1)}\big) \\
&\quad + s_{N1}\sin(\phi_1 - \phi_N + \psi_{N1})
\end{aligned} \tag{28}$$

where $s_{ij} = s_{ji}$, $\psi_{ij} = \psi_{ji}$, and ϕ_i, $i = 1, \cdots, N$ are phase variables taking values in the interval $[-\pi, \pi)$. The index i refers to the ith unit and these units are coupled. Among the phase difference variables between oscillatory units $\phi_{ij} = \phi_i - \phi_j$, $i, j = 1, 2, \cdots, N$, there are only $N-1$ independent variables and they are denoted by $\Phi_d = [\phi_{12}\ \phi_{23} \cdots \phi_{(N-1)N}]^T$. Stability analysis would give us that, up to mod 2π, the unique stable equilibrium is given by $[\psi_{12}\ \psi_{23} \cdots \psi_{(N-1)N}]^T$. Analogously, if the phase variables are rescaled by M with $\bar{\phi}_i = \phi_i/M$, $i = 1, 2, \cdots, N$, then the new phase difference vector $\bar{\Phi}_d = \Phi_d/M$. It would follow that there are $M^{(N-1)}$ equilibria up to mod 2π for the new phase difference variables in the following form:

$$\begin{pmatrix} \frac{\psi_{12}}{M} + \frac{2(k_1-1)\pi}{M} \\ \frac{\psi_{23}}{M} + \frac{2(k_2-1)\pi}{M} \\ \vdots \\ \frac{\psi_{(N-1)N}}{M} + \frac{2(k_{(N-1)}-1)\pi}{M} \end{pmatrix}$$

where $k_i \in \{0, 1, 2, \cdots, M-1\}$ with $i = 1, 2, \cdots, N-1$. So, to memorize n patterns, we may choose a Kuramoto model described by (28) with the least number of oscillatory units for a given rescaling, or choose the lowest rescaling to phase variables for a given Kuramoto model only if $M^{(N-1)} \geq n$.

To illustrate the pattern recognition mechanism, we use the third-order Kuramoto model with rescaling parameter 2 such that up to four patterns can be memorized or recognized. We define the phase difference variables as follows:

$$\phi_{ij} = \phi_i - \phi_j$$

and we have

$$\begin{bmatrix} \dot{\phi}_{12} \\ \dot{\phi}_{23} \\ \dot{\phi}_{31} \end{bmatrix} = \begin{bmatrix} -2s_{12} & s_{23} & s_{31} \\ s_{12} & -2s_{23} & s_{31} \\ s_{12} & s_{23} & -2s_{31} \end{bmatrix} \begin{bmatrix} \sin(\phi_{12} - \psi_{12}) \\ \sin(\phi_{23} - \psi_{23}) \\ \sin(\phi_{31} - \psi_{31}) \end{bmatrix}.$$

Up to mod 2π, the above equation has a unique stable stationary point at ψ_{12}, ψ_{23}, ψ_{31}. Scaling the phase variables as

$$\bar{\phi}_1 = \frac{1}{2}\phi_1, \quad \bar{\phi}_2 = \frac{1}{2}\phi_2, \quad \bar{\phi}_3 = \frac{1}{2}\phi_3$$

we can write down the dynamics of $\bar{\phi}_{12}$, $\bar{\phi}_{23}$, $\bar{\phi}_{31}$, where $\bar{\phi}_{ij} = \bar{\phi}_i - \bar{\phi}_j$.

Note that all the three-phase difference variables are not independent. So we may only consider $\bar{\phi}_{12}$ and $\bar{\phi}_{23}$ without loss of generality. The four stable stationary points are given by

$$\begin{pmatrix} \frac{\psi_{12}}{2} \\ \frac{\psi_{23}}{2} \end{pmatrix}, \begin{pmatrix} \frac{\psi_{12}}{2} + \pi \\ \frac{\psi_{23}}{2} \end{pmatrix}, \begin{pmatrix} \frac{\psi_{12}}{2} \\ \frac{\psi_{23}}{2} + \pi \end{pmatrix}, \begin{pmatrix} \frac{\psi_{12}}{2} + \pi \\ \frac{\psi_{23}}{2} + \pi \end{pmatrix}. \tag{29}$$

Four patterns of vectors in $\mathbb{C}^3$ are defined as follows:

$$P_1 = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{pmatrix}, \quad P_2 = \begin{pmatrix} e^{\frac{i\pi}{2}}\pi_1 \\ e^{\frac{-i\pi}{2}}\pi_2 \\ e^{\frac{-i\pi}{2}}\pi_3 \end{pmatrix}$$
$$P_3 = \begin{pmatrix} e^{\frac{i\pi}{2}}\pi_1 \\ e^{\frac{i\pi}{2}}\pi_2 \\ e^{\frac{-i\pi}{2}}\pi_3 \end{pmatrix}, \quad P_4 = \begin{pmatrix} e^{\frac{i\pi}{2}}\pi_1 \\ e^{\frac{-i\pi}{2}}\pi_2 \\ e^{\frac{i\pi}{2}}\pi_3 \end{pmatrix}$$

where we assume that

$$\psi_{12} = 2(arg\pi_1\bar{\pi}_2), \quad \psi_{23} = 2(arg\pi_2\bar{\pi}_3).$$

Analogously, the four complex patterns are associated to the four stable stationary equilibria. From here onwards, it is easy to define a transformation that would map patterns of real vectors in $\mathbb{R}^Q$ to complex patterns in $\mathbb{C}^3$.

VIII. AN ILLUSTRATIVE EXAMPLE

In this section, we illustrate through an example many of the important encoding and decoding techniques presented earlier in this paper. We use a binocular system to show how a visual scene can be projected onto the retinal planes of two visual sensors (which are cameras), how the projected images can be fused using sparse representation and principal components to generate cortical signals (used as an input to the model cortex), and how the cortical activities can be used to decode the depths of objects in the visual scene using Kuramoto models. This section does not illustrate memory and recalling using place cells in the hippocampus.

We consider a specific visual scene that the cameras detect with a fish as the target (shown in Fig. 15). The fish is kept at various depths in the visual field and is observed by a pair of cameras. Our eventual goal is to be able to detect the depth of the fish in the visual scene.

The first step is the problem of image acquisition by the two cameras. We achieve this by projecting the scene perspectively onto the retinal planes of the cameras. Knowing the positions of the two cameras r_i, $i = 1, 2$, the coordinates of the target with respect to the individual cameras, $[x_i', y_i', z_i']^T$, can be computed via a translation and a rotation operator as follows:

$$\begin{bmatrix} x_i' \\ y_i' \\ z_i' \end{bmatrix} = R_i^T \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - r_i \right)$$

where $[x, y, z]^T$ are the coordinates of the target with respect to a world coordinate frame. The images captured by the cameras can be computed utilizing the following projection rule:

$$\begin{bmatrix} x_i' \\ y_i' \\ z_i' \end{bmatrix} \longmapsto \begin{bmatrix} -x_i' \frac{f}{z_i'} \\ -y_i' \frac{f}{z_i'} \end{bmatrix}$$

where f is the focal length, assumed to be 1 for simplicity.

Note that the rotation matrix R_i^T has to satisfy Listing's constraint and can be easily calculated from the map defined in (4) between S^3 and $SO(3)$.

In this example, we consider the simple case in which two cameras are located on the X-axis symmetrically and the target is located on the Z-axis at three different nominal depths d with variation δ. Three nominal depths are chosen at $50f$, $75f$, and $100f$. Fig. 16 shows the projections of the third image frame through the two

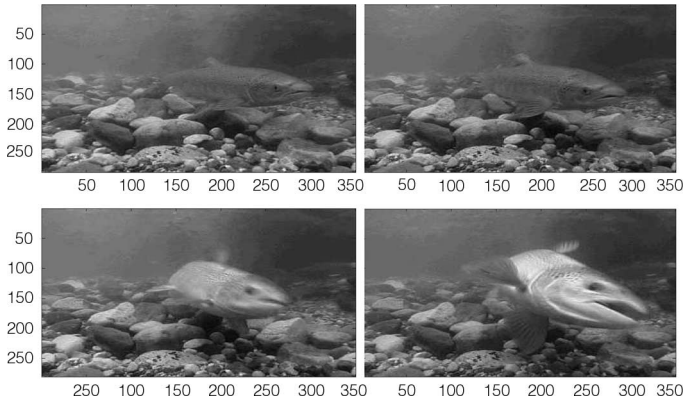

Fig. 15. *A sequence of four images over time in this illustration. Image 1 and 2 are on the top from left to right. Image 3 and 4 are on the bottom from left to right.*

cameras from the depth at $d = 50f$, $75f$, and $100f$, respectively.

As discussed in the introduction and Section IV, the spatiotemporal sequences of images on the image planes of the two cameras need to be fused and encoded. To do so, the images are sparsely represented with a set of overcomplete basis functions and the sparse codes from the two channels are fused and compressed using principal components, as already described in Section IV. The fusion process, illustrated in Fig. 17, is now described. The sparse codes from each of the horizontal blocks in the images are stacked together as follows:

$$\left[A^1(t), \cdots A^J(t)\right]$$

where

$$A^j(t) = \begin{pmatrix} A_1^j(t) \\ A_2^j(t) \end{pmatrix}$$

where the superscript j indicates the blocks along the horizontal direction. The vectors $A^j(t)$, obtained by the process of fusion, are represented using principal components and the associated β-coefficients are the codes that represent the left and right camera views of the image. These codes are used as an input to the model cortex for wave generation.

Using the visual cortex model with an embedded noise, 50 cortical responses are simulated for each depth d together with five different variations, for a total of 150 cortical waves generated.

We hypothesize that the depth information is encoded in the responses of the model cortex. Our next problem is to decode the depth parameter using the nonlinear dynamics of the Kuramoto models. The cortical responses

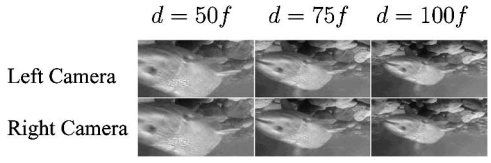

Fig. 16. *Projections of the third image frame in the scene (Fig. 15) on the retinal planes through the two cameras in the binocular visual system from different depths at $d = 50f$, $75f$, and $100f$, respectively.*

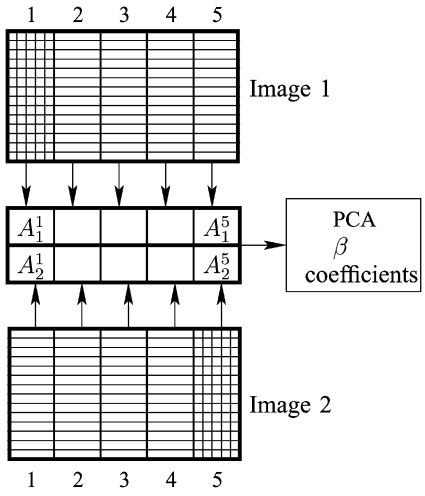

Fig. 17. *Each image is divided into several blocks and each block into several patches. The sparse codes of the corresponding blocks from the two images (one from the left camera and the other from the right) are stacked and compressed using principal components (PCA) to generate the β coefficients.*

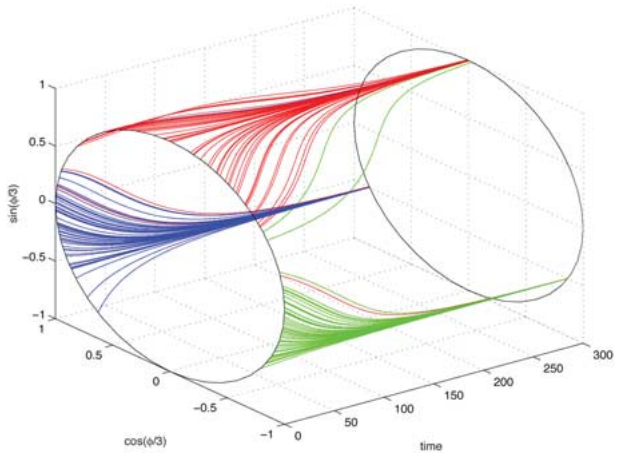

Fig. 18. *Depth detection of objects using a second-order Kuramoto model upon the points on β-space. Blue curves correspond to depths around 50f, red curves correspond to depths around 75f, and green curves correspond to depths around 100f. There are 50 curves for each color.*

are first smoothed with a low-pass filter and represented by points in β-space using a double KL-decomposition within the time window described in Section V. The points are used to initialize a Kuramoto model of two units using the map (27) to detect the depth of the target. The map T optimally transforms the points on β-space to the complex vectors p_k, $k = 1, 2$, and 3 defined in (25).

It is evident from Fig. 18 that the rescaled phase difference $\bar{\phi} = \phi/3$ converges to one of the three equilibria associated with the three depths. Plots of phase difference variable $\bar{\phi}(t)$ are sketched in terms of sin and cos functions over time. It can be seen that most of the detection results are correct.

IX. DISCUSSION

Pattern generating and oscillatory circuits are a source of considerable research interest (see [45] and [83]) in engineering and biology and we would like to make certain connections between these lines of research and the main content of this paper. In biology, central pattern generators provide the source of timing with respect to which an animal is able to perform various physiological tasks such as locomotion. The basic idea is that patterns can generate actuating control signals (see Brockett [15], [16]) and that control signals can be switched by suitably switching from one pattern to another (see Dayawansa [59], [71]). It is therefore of interest in engineering to study pattern generating circuits (see Krishnaprasad [50]) and control these circuits to match a prescribed spatiotemporal pattern (see Dayawansa's [59] work on solitons). The pattern generating circuits may be constructed using an array of microelectromechanical (MEMS) devices. Pattern generation is also of independent interest in biomathematics, particularly for the purpose of modeling dynamical systems that would produce a suitable pattern in the steady state. There are many excellent reviews on this subject (see [62]). Such circuits are now being implemented using optically interconnected parallel computing systems [49], [61].

In this paper, our interest in patterns is motivated from a need to *encode* visual spatiotemporal signals (see Knight [52] for a review article on this topic). We have discussed in sufficient detail the role of the visual cortex in freshwater turtles as an example of a pattern generating circuit capable of generating traveling waves. Of course the visual cortex is not the only place where patterns are generated. Another good example is the hippocampus (see [42], [85], and [100]) where the spatial location of an animal is believed to be encoded as localized activity in place cells.

The precise benefit of encoding via traveling wave (as opposed to localized activity such as in the hippocampus) is perhaps less clear. Performing a careful analysis of the cortical traveling wave, we have observed [29] that a typical wave lasts for about 800 ms, out of which during the first 200 ms the visual images on the retina are being actively encoded. For the next 400 ms the encoded images are retained as is evidenced by our ability to decode target locations in these images. Finally, during the last 200 ms the traveling wave does not carry any information. In summary, we can say the following:

The cortical wave samples the visual field over time, encodes events over a short time window, and retains it for a longer time. This process is serially repeated over and over.

No systematic study has ever been conducted to explain the precise reason for the cortical sampling described above. We speculate that during the active propagation of the wave, lasting for the first 600 ms, the traveling waves are used to predict future positions of a moving target. This prediction occurs in an open loop based on encoding prior information during the first 200 ms of the wave generation. Of course, such a prediction is prone to error and requires correction in real-time. In the turtle visual circuit, this prediction step is believed to be carried out in the optic tectum which receives inputs from the cortex. There is also a direct input from the retina to the tectum, which is believed to provide the correction based on more recent data. The detail mechanism behind how the target locations are predicted and corrected is currently unknown and is under investigation.

In order to classify patterns using dynamic models, in this paper we have introduced a class of models called the Kuramoto model. We have also argued that these models are particularly interesting because one can show asymptotic convergence of the phase difference variables to multiple points of equilibria. An important question that one might ask is how these models can be implemented.

We would like to remark that Kuramoto type oscillators can be implemented using a network of coupled laser

oscillators (see [46] for details). For the purpose of this discussion, we note that such a system of coupled oscillators provides a platform for optically interconnected parallel computing (see Ishikawa [49] for details). A major advantage in the optical domain is that the driving power is independent of interconnection distance with practically no attenuation losses.

The network of coupled lasers (see [46] for details) can be written with the dimensionless rate equations as follows:

$$\dot{E}_i = (1 + i\alpha)N_i E_i + i\omega E_i + \sum_{j=1}^{n} c_{ij} E_j \tag{30}$$

$$\dot{N}_i = \mu\left[P - N_i - (1 + 2N_i)|E_i|^2\right] \tag{31}$$

where E_i and N_i are the complex electric field and the excess carrier number of the ith laser. The derivatives are with respect to s, where $s = t\tau_p^{-1}$ is the time measured in units of the photon lifetime τ_p, $\mu = \tau_p/\tau_s$ is the ratio of photon to carrier time scales, and where τ_s is the carrier lifetime. Other notations are defined as follows: P is the pumping above threshold, α is the linewidth enhancement factor, ω is normalized optical frequency, and c_{ij} are complex connection coefficients.

It is convenient to use polar coordinates $E_i = r_i e^{i\phi_i}$ and $c_{ij} = s_{ij}e^{i\psi_{ij}}$ to rewrite the model (30) and (31) in the form

$$\dot{\phi}_i = \alpha N_i + \omega + \sum_{j=1}^{n} s_{ij}\frac{r_j}{r_i}\sin(\phi_j + \psi_{ij} - \phi_i) \tag{32}$$

$$\dot{r}_i = N_i r_i + \sum_{j=1}^{n} s_{ij} r_j \cos(\phi_j + \psi_{ij} - \phi_i) \tag{33}$$

$$\dot{N}_i = \mu\left[P - N_i - (1 + 2N_i)|r_i|^2\right]. \tag{34}$$

In the case of weak connection, the dynamical analysis of the model (32), (33), and (34) shows that

$$(r_i(t),\ N_i(t)) \longrightarrow (\sqrt{P},\ 0)$$

and the phase $\phi_i(t) \to \omega t + \phi_i^0$, where ϕ_i^0 is determined by the initial conditions.

If all $r_i(t) \to r_0$, then the phase model (32) for $\alpha = 0$ has the form that satisfies the model proposed by Kuramoto [56] and can therefore be used to memorize a pattern. It would be of interest to implement the laser network, both in simulation and experimentally, to ascertain the extent to which these networks can indeed be used in pattern recognition. This is also a subject of future research.

X. CONCLUSION

The main features of this paper are described as follows. First, we consider a set of visual sensors and describe the process of "formation sensing" wherein a group of sensors visually attend to a moving target in a formation. We study the problem of how the sensor group can optimally orient in order to acquire a visual target. The second problem we study is the problem of sparse coding and sensor fusion. The main idea is to choose an over-complete set of basis functions for sparse coding. Sparse codes from individual sensors are fused using a minimal set of bases from principal components. In the third part, we consider the problem of activity wave generation. Activated by a visual point target or a sequence of natural scenes on the retina, an activity pattern is produced on the visual cortex. We study detailed micro circuits of this pattern formation process. We consider the problem of recognition using a family of synchronized coupled oscillators and discuss their implementation using a network of coupled lasers. Finally, we also consider the role of hippocampal place cells in the problem of recalling a sequence of places previously visited by the animal. ■

Acknowledgment

The first author would like to thank P. Ulinski from the University of Chicago and graduate students Z. Nenadic, X. Du, J. Joseph, and Z. Freudenburg for their contributions to our understanding of the turtle visual cortex. The authors thank S. Ghosh for his help in proofreading, and the anonymous reviewers and editors of this special issue for straightening many of the scattered ideas presented in earlier versions of this paper.

REFERENCES

[1] H. Abelson, D. Allen, D. Coore, C. Hanson, E. Rauch, G. J. Sussman, and R. Weiss, "Amorphous," *Commun. ACM*, vol. 43, no. 5, pp. 74–82, May 2000.

[2] C. Acker, N. Kopell, and J. White, "Synchronization of strongly coupled excitatory neurons: Relating network behavior to biophysics," *J. Comput. Neurosci.*, vol. 15, pp. 71–90, 2003.

[3] C. H. Anderson and D. C. Van Essen, "Shifter circuits: A computational strategy for dynamic aspects of visual processing," *Proc. Nat. Acad. Sci., Neurobiol.*, vol. 84, pp. 6297–6301, 1987.

[4] V. I. Arnol'd, *Mathematical Methods of Classical Mechanics, no. 60 in Graduate Texts in Mathematics*, 2nd ed. New York: Springer-Verlag, 1989.

[5] J. J. Atick and A. N. Redlich, "Towards a theory of early visual processing," *Neural Comput.*, vol. 2, pp. 308–320, 1990.

[6] J. J. Atick, "Could information theory provide an ecological theory of sensory processing?" *Network*, vol. 3, pp. 213–251, 1992.

[7] J. J. Atick and A. N. Redlich, "What does the retina know about natural scenes?" *J. Neural Comput.*, vol. 4, pp. 196–210, 1992.

[8] H. B. Barlow, "Possible principles underlying the transformations of sensory messages," in *Sens. Commun.*, W. A. Rosenblith, Ed. Cambridge, MA: MIT Press, 1961, pp. 217–234.

[9] ——, "Unsupervised learning," *J. Neural Comput.*, pp. 295–311, 1989.

[10] G. Bi and M. Poo, "Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type," *J. NeuroSci.*, vol. 18, pp. 10 464–10 472, 1998.

[11] T. V. Bliss and T. Lomo, "Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path," *J. Physiol.*, vol. 232, no. 2, pp. 331–356, 1973.

[12] W. M. Boothby, *An Introduction to Differential Manifolds and Riemannian Geometry*, 2nd ed. New York: Academic, 1986.

[13] C. Borgers and N. Kopell, "Synchronization in network of excitatory and inhibitory neurons with sparse, random connectivity," *Neural Comput.*, vol. 15, pp. 509–538, 2003.

[14] J. M. Bower and D. Beeman, *The Book of Genesis*. Santa Clara, CA: TELOS, 1998.

[15] R. W. Brockett, "On the rectification of vibratory motion," *Sens. Actuators*, vol. 20, pp. 91–96, 1989.

[16] ——, "Pattern generation and the control of nonlinear systems," *IEEE Trans. Autom. Contr.*, vol. 48, no. 10, pp. 1699–1711, Oct. 2003.

[17] E. N. Brown, L. M. Frank, D. Tang, M. C. Quirk, and M. A. Wilson, "A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells," *J. Neurosci.*, vol. 18, pp. 7411–7425, 1998.

[18] E. N. Brown, D. P. Nguyen, L. M. Frank, M. A. Wilson, and V. Solo, "An analysis of neural receptive field dynamics by point process adaptive filtering," *Proc. Nat. Acad. Sci.*, vol. 98, pp. 12 261–12 266, 2001.

[19] N. Bulusu, D. Estrin, and J. Heidemann, "GPS-less low cost outdoor localization for very small devices," *IEEE Pers. Commun.*, vol. 7, no. 5, pp. 28–34, Oct. 2000.

[20] G. Buzaki, "Theta oscillations in the hippocampus," *Neuron.*, vol. 33, pp. 325–340, 2002.

[21] B. W. Cook, S. Lanzisera, and K. S. J. Pister, "SoC issues for RF smart dust," *Proc. IEEE*, vol. 94, no. 6, pp. 1177–1196, Jun. 2006.

[22] D. Culler, D. Estrin, and M. Srivastava, "Overview of sensor networks," *Computer*, vol. 37, no. 8, pp. 41–49, 2004.

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

[24] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, 1st ed. Cambridge, MA: MIT Press, 2001.

[25] F. Delcomyn, *Foundations of Neurobiology*. New York: W. H. Freeman, 1998.

[26] M. Dellnitz, M. Golubitsky, and M. Nicol, "Symmetry of attractors and the Karhunen-Loeve decomposition," in *Trends and Perspectives in Applied Mathematics*, L. Sirovich, Ed. New York: Springer-Verlag, 1994, ch. 4, pp. 73–108.

[27] 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.

[28] J. P. Desai, J. P. Ostrowski, and V. Kumar, "Modeling and control of formations of nonholonomic mobile robots," *IEEE Trans. Robot. Autom.*, vol. 17, no. 6, pp. 905–908, Dec. 2001.

[29] X. Du, B. K. Ghosh, and P. S. Ulinski, "Encoding and decoding target locations with waves in the turtle visual cortex," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 566–577, Apr. 2005.

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

[31] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next century challenges: Scalable coordination in sensor networks," in *Proc. ACM Conference on Mobile and Computer Networking (Mobicom)*. New York: ACM Press, 1999, pp. 263–270.

[32] D. Estrin, D. Culler, K. Pister, and G. Sukhatme, "Connecting the physical world with pervasive networks," *Pervasive Comput.*, vol. 1, no. 1, pp. 59–69, 2002.

[33] O. Faugeras and Q. T. Luong, *The Geometry of Multiple Images*. Cambridge, MA: MIT Press, 2001.

[34] D. J. Field, "Scale-invariant and self-similar "wavelet" transforms: An analysis of natural scenes and mammalian visual systems," in *Wavelets, Fractals and Fourier Transforms*, M. Farge, J. Hunt, and C. Vascillicos, Eds. Oxford, U.K.: Oxford Univ. Press, 1993, pp. 151–193.

[35] S. Funiak, C. Guestrin, M. Paskin, and R. Sukthankar, "Distributed localization of networked cameras," in *ISPN'06*, Apr. 19–21, 2006.

[36] 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.

[37] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*. Englewood Cliffs, NJ: Prentice Hall, 1984.

[38] A. A. Handzel and T. Flash, "The geometry of eye rotations and Listing's law," in *Advances in Neural Information Processing Systems*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds. Cambridge, MA: MIT Press, 1996, vol. 8, pp. 117–123.

[39] D. O. Hebb, *The Organization of Behavior: A Neuropsychological Theory*. Mahwah, NJ: Lawrence Erlbaum Associates, 2002.

[40] J. Heikkila and O. Silven, "A four step camera calibration procedure with implicit image correction," in *Proc. Computer Vision Pattern Recogn.*, San Juan, Puerto Rico, 1997, pp. 1106–1112.

[41] J. L. van Hemmen and T. J. Sejnowski, *23 Problems in Systems Neuroscience*. Oxford, U.K.: Oxford Univ. Press, 2006.

[42] Y. Hirakura, Y. Yamaguchi, H. Shimizu, and S. Nagai, "Dynamic linking among neural oscillators leads to flexible pattern recognition with figure-ground separation," *Neural Netw.*, vol. 9, no. 2, pp. 189–209, 1996.

[43] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structure, Dynamical Systems and Symmetry*. Cambridge, U.K.: Cambridge Univ. Press, 1996.

[44] F. C. Hoppensteadt and E. M. Izhikevich, *Weakly Connected Neural Networks*. New York: Springer-Verlag, 1997.

[45] ——, "Pattern generation via synchronization in phase-locked loop neural networks," *IEEE Trans. Neural Netw.*, vol. 11, no. 3, pp. 734–738, May 2000.

[46] ——, "Synchronization of laser oscillators, associative memory, and optical neurocomputing," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 62, pp. 4010–4013, 2000.

[47] I. I. Hussein and A. M. Bloch, "Dynamic coverage optimal control for interferometric imaging spacecraft formations," in *Proc. 43rd IEEE Conf. Decision and Control*, 2004, pp. 1982–1987.

[48] A. Ihler, J. Fisher, R. Moses, and A. Willsky, "Nonparametric belief propagation for self calibration in sensor network," in *IPSN'04*, Apr. 26–27, 2004.

[49] M. Ishikawa and N. McArdle, "Optically interconnected parallel computing systems," *Computer*, vol. 31, no. 2, pp. 61–68, 1998.

[50] E. Justh and P. S. Krishnaprasad, "Pattern-forming systems for control of large arrays of actuators," *J. Nonlinear Sci.*, vol. 11, pp. 239–277, 2001.

[51] A. Kansal, E. Yuen, W. J. Kaiser, G. J. Pottie, and M. B. Srivastava, "Sensing uncertainty reduction using low complexity actuation," in *ACM Symp. Information Processing in Sensor Networks*. New York: ACM Press, 2004, pp. 388–395.

[52] B. W. Knight, "Dynamics of encoding in neuron populations: Some general mathematical features," *Neural Comput.*, vol. 12, pp. 473–518, 2000.

[53] N. Kopell, "We got rhythm: Dynamical systems of the nervous system," *Notices Amer. Math. Soc.*, vol. 47, pp. 6–16, 2001.

[54] N. Kopell and G. B. Ermentrout, "Chemical and electrical synapses perform complementary roles in synchronization of neuronal networks," *Proc. Nat. Acad. Sci.*, vol. 101, pp. 1542–1548, 2004.

[55] J. Kuipers, *Quaternions and Rotation Sequences*. Princeton, NJ: Princeton Univ. Press, 1998.

[56] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*. New York: Springer-Verlag, 1984.

[57] J. E. Lisman and N. A. Otmakhova, "Theta oscillations in the Hippocampus," *Hippocampus*, vol. 11, pp. 551–568, 2001.

[58] Q. T. Luong and O. D. Faugeras, "Camera calibration, scene motion and structure recovery from point correspondences and fundamental matrices," *Int. J. Comput. Vis.*, vol. 22, no. 3, pp. 261–289, 1997.

[59] D. H. S. Maithripala, J. M. Berg, and W. P. Dayawansa, "Nonlinear dynamic output feedback stabilization of electrostatically actuated MEMS," in *Proc. 42nd IEEE Conf. Decision and Control*, 2003, pp. 61–66.

[60] C. Martin and L. Schovanec, "Muscle mechanics and dynamics of ocular motion," *J. Math. Syst., Estim. Contr.*, vol. 8, pp. 1–15, 1998.

[61] N. McArdle, M. Naruse, H. Toyoda, Y. Kobayashi, and M. Ishikawa, "Reconfigurable optical interconnections for parallel computing," *Proc. IEEE*, vol. 88, no. 6, pp. 829–837, Jun. 2000.

[62] J. D. Murray, *Mathematical Biology I: An Introduction*, 3rd ed. New York: Springer-Verlag, 2004.

[63] 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.

[64] ——, "Propagating waves in visual cortex: A large scale model of turtle visual cortex," *J. Comput. Neurosci.*, vol. 14, pp. 161–184, 2003.

[65] T. Netoff, M. Banks, A. Dorval, C. Acker, J. Haas, N. Kopell, and J. White, "Synchronization and phase locking in hybrid neuronal networks of the

hippocampal formation," *J. Neurophysiol.*, vol. 93, pp. 1197–1208, 2005.

[66] B. A. Olshausen, C. H. Anderson, and D. C. Van Essen, "A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information," *J. Neurosci.*, vol. 13, no. 11, pp. 4700–4719, 1993.

[67] ——, "A multiscale dynamic routing circuit for forming size and position invariant object representations," *J. Comput. Neurosci.*, vol. 2, no. 1, pp. 45–62, 1995.

[68] B. A. Olshausen and D. J. Field, "Natural image statistics and efficient coding," *Network*, vol. 7, pp. 333–339, 1996.

[69] ——, "Sparse coding with an overcomplete basis set: A strategy employed by V1?" *Vision Res.*, vol. 37, no. 23, pp. 3311–3325, 1997.

[70] J. O'Keefe and M. L. Recce, "Phase relationship between hippocampal place units and the EEG theta rhythm," *Hippocampus*, vol. 3, no. 3, pp. 317–330, 1993.

[71] R. Palamakumbura, S. Maithripala, W. P. Dayawansa, and H. Inaba, "Control of travelling pulses in mems arrays: Numerical evidence of practical asymptotic stabilization," in *Proc. Amer. Control Conf.*, 2005, pp. 2481–2486.

[72] 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.

[73] A. D. Polpitiya, "Geometry and control of human eye movements," Sc.D. thesis, Washington Univ., St. Louis, MO, 2004.

[74] A. D. Polpitiya, W. P. Dayawansa, C. F. Martin, and B. K. Ghosh, "Geometry and control of human eye movements," *IEEE Trans. Autom. Contr.*, to be published.

[75] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Commun. ACM*, vol. 43, no. 5, pp. 51–58, 2000.

[76] J. C. Prechtl, "Visual motion induces synchronous oscillations in turtle visual cortex," *Proc. Nat. Acad. Sci., USA*, vol. 91, pp. 12 467–12 471, 1994.

[77] J. C. Prechtl, T. H. Bullock, and D. Kleinfeld, "Direct evidence for local oscillatory current sources and intracortical phase gradients in turtle visual cortex," *Proc. Nat. Acad. Sci.*, vol. 97, pp. 877–882, 2000.

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

[79] M. Rahimi, R. Baer, I. Obimdinachi, J. Garcia, J. Warrior, D. Estrin, and M. Srivastava, "Cyclops: *in situ* image sensing and interpretation in wireless sensor networks," in *Proc. 3rd ACM Conf. Embedded Networked Sensor Systems*, Nov. 2–4, 2005, pp. 192–204.

[80] K. R. Rao and P. C. Yip, *The Transform and Data Compression Handbook*. Boca Raton, FL: CRC, 2001.

[81] K. A. Robbins and D. M. Senseman, "Visualizing differences in movies of cortical activity," in *Proc. Visualization '98*, 1998, pp. 217–224.

[82] D. Robinson, "The mechanics of human saccadic eye movement," *J. Physiol.*, vol. 174, pp. 245–264, 1964.

[83] F. Rogister, K. S. Thornburg Jr., L. Fabiny, M. Moller, and R. Roy, "Power-law spatial correlations in arrays of locally coupled lasers," *Phys. Rev. Lett.*, vol. 92, no. 9, 2004.

[84] N. Sato and Y. Yamaguchi, "Memory encoding by theta phase precession in the hippocampal network," *Neural Comput.*, vol. 15, pp. 2379–2397, 2003.

[85] ——, "Online formation of a hierarchical cognitive map for object-place association by theta phase coding," *Hippocampus*, vol. 15, no. 7, pp. 963–978, 2005.

[86] 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.

[87] D. M. Senseman and K. A. Robbins, "Modal behavior of cortical neural networks during visual processing," *J. Neurosci.*, vol. 19, pp. 1–7, 1999, RC3.

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

[89] C. Sultan, S. Seereeram, and R. K. Mehra, "Energy optimal multi-spacecraft relative reconfiguration of deep space formation flying," in *Proc. 43rd IEEE Conf. Decision and Control*, 2004, pp. 284–289.

[90] D. Tweed and T. Villis, "Geometric relations of eye position and velocity vectors during saccades," Vision Res., vol. 30, pp. 111–127, 1990.

[91] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968.

[92] R. Y. Tsai, "A versatile camera calibration technique for high accuracy 3-D machine vision metrology using off the shelf TV cameras and lenses," *IEEE J. Robot. Autom.*, vol. RA-3, no. 4, pp. 323–344, Aug. 1987.

[93] C. H. Vanderwolf, "Hippocampal electrical activity and voluntary movement in the rat," *Electroencephalogr. Clin. Neurophysiol.*, vol. 26, pp. 407–418, 1969.

[94] W. Wang, B. K. Ghosh, and P. S. Ulinski, "Two cortical circuits control propagating waves in visual cortex," *J. Comput. Neurosci.*, vol. 19, pp. 263–289, 2005.

[95] W. Wang and B. K. Ghosh, "Motion reconstruction in natural scenes from cortical activity waves," *16th IFAC World Conf.*, 2005.

[96] X. Wang and S. N. Balakrishnan, "Optimal and hierarchical formation control for UAV," in *Proc. 2005 Amer. Control Conf.*, 2005, pp. 4685–4689.

[97] B. Warneke *et al.* "Smart dust: Communicating with a cubic millimeter computer," *Computer*, vol. 34, no. 1, pp. 44–51, 2001.

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

[99] Y. Yamaguchi, "A theory of hippocampal memory based on theta phase precession," *Biol. Cybern.*, vol. 89, pp. 1–9, 2003.

[100] Y. Yamaguchi and H. Shimizu, "Pattern recognition with figure-ground seperation by generation of coherent oscillations," *Neural Netw.*, vol. 7, no. 1, pp. 49–63, 1994.

[101] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.

ABOUT THE AUTHORS

Bijoy K. Ghosh (Fellow, IEEE) received the B.Tech. degree from Birla Institute of Technology and Science, Pilani, India, in 1977 and the M.Tech. degree from the Indian Institute of Technology, Kanpur, India, in 1979, both in in electrical and electronics engineering, and the Ph.D. degree in engineering from the Decision and Control Group of the Division of Applied Sciences, Harvard University, Cambridge, MA, in 1983.

From 1983 to 2006, he was a faculty member in the Department of Electrical and Systems Engineering, Washington University, St. Louis, MO, as a Professor, and directed the center for BioCybernetics and Intelligent Systems. Presently he is a Professor in the Department of Mathematics and Statistics, Texas Tech University, Lubbock. His research interests are in multivariable control theory, machine vision, robotic manufacturing, and bioinformatics and control.

Ashoka D. Polpitiya (Member, IEEE) received the B.Sc. degree in electrical and electronic engineering from the University of Peradeniya, Sri Lanka, in 1996 and the M.S. and D.Sc. degrees in systems science and mathematics from Washington University, St. Louis, MO, in 2000 and 2004, respectively.

From 2004 to 2006, he was a Postdoctoral Research Associate with the Cellular Injury and Adaptation Laboratory, Washington University, School of Medicine, St. Louis, MO. He is currently a Senior Research Scientist at the Pacific Northwest National Laboratory, Richland, WA. His research interests are in the areas of systems biology and bioinformatics, application of control theory in biological systems, and geometric control.

Wenxue Wang (Member, IEEE) received the B.Sc. degree in automatical control from Beijing Institute of Technology, Beijing. China, in 1996, the M.Sc. degree in control theory from the Institute of Systems Science, Chinese Academy of Sciences, Beijing, in 1999, and the M.Sc. and D.Sc. degrees in systems science and mathematics from Washington University, Saint Louis, MO, in 2002 and 2006, respectively.

He is currently a Postdoctoral Research Associate with the Center for BioCybernetics and Intelligent Systems in the Department of Mathematics and Statistics, Texas Tech University, Lubbock. His research interests are in the areas of computational neuroscience, neural networks, systems biology, and application of systems science in biological systems.