POSITION AND VELOCITY ESTIMATION IN THE VISUAL CORTEX

Bijoy K. Ghosh[1] **Zoran Nenadic**

Department of Systems Science and Mathematics
Campus Box 1040
Washington University
Saint Louis, MO 63130
U.S.A.

Abstract: Early extracellular recordings from turtle visual cortex have shown the existence of a wave of depolarizing activity as a response to a localized light flashes presented in the turtle visual field. Experiments using voltage sensitive dye techniques have supported this observation. The dynamics of the cortical wave might be responsible for encoding the parameters of an unknown input stimulus. We develop a data compression algorithm for representing the cortical response elicited by a family of stimuli. To process image data we use the principal components analysis method and we provide a consistent way of choosing the number of principal modes necessary for effective data representation.

Keywords: turtle visual cortex, KL decomposition, detection and estimation

1. INTRODUCTION

Extracellular recordings from turtle visual cortex showed that the presentation of a stimulus (usually localized or diffused retinal light flashes) in the visual space of a turtle produces a wave of activity that propagates across the turtle visual cortex (Mazurskaya, 1974). This is contrary to the concept of receptive field in mammals, where a stimulus localized in the visual field produces a localized activity in the brain (Delcomyn, 1998). The existence of a wave of depolarizing activity in the turtle visual cortex has been verified using voltage sensitive dye techniques (Senseman, 1996). The voltage sensitive dye signals could be visualized as a series of snapshots, and the movies corresponding to different stimuli could be compared using the principal components analysis techniques (Senseman and Robbins, 1998). It has been observed that the movies could be successfully approximated using only a couple of principal eigenvectors (modes) and that the differences could be visualized in a low dimensional subspace of time coefficients. This raised the possibility that the position of a visual stimulus is encoded in the spatiotemporal dynamics of the wave, and that the estimates could be made from observations.

Visual cortex of freshwater turtle contains three topologically distinct layers. Intermediate layer (layer 2) contains mainly excitatory, pyramidal cells. Pyramidal cells are organized in two subtypes, lateral pyramidal and medial pyramidal cells. Outer layer (layer 1) contains predominantly inhibitory, stellate cells. Inner layer (layer 3) contains primarily inhibitory, horizontal cells. Layer 2 and layer 1 receive direct synaptic projections (visual input) from lateral geniculate complex LGN. Pyramidal cells make excitatory projections to other pyramidal cells as well as stellate cells and horizontal cells. Stellate cells make inhibitory connections to other stellate cells and pyramidal cells. Horizontal cells make inhibitory connections to pyramidal cells.

[1] corresponding author: ghosh@zach.wustl.edu

2. COMPUTATIONAL MODEL OF TURTLE VISUAL CORTEX

Based on anatomical and physiological data, we have developed a computational model of turtle visual cortex (Nenadic, 2001). The model is synthesized from the models of individual cortical neurons, information on their spatial distributions and connectivity patterns. The model is assembled in GENESIS, a specialized software for modeling medium to large scale networks of biological neurons. The response of the network to different stimuli has been observed and recorded. In particular we have tested the behavior of the network to both localized stationary stimuli and stimuli moving with constant velocity across the visual field.

Our model contains 945 neurons, among which, 679 pyramidal cells, 45 stellate cells, 20 horizontal cells and 201 LGN neurons. Individual neurons are modeled using compartmental approach, and different types of cells have been assigned different number of compartments, ranging from one compartment for LGN neurons to 16 compartments for lateral pyramidal cells. The spatial densities of different types of neurons have been obtained from biological measurements, and the coordinates of individual cells are drawn at random from the distributions corresponding to the spatial densities. All three layers have been projected onto a single plane and the cells are interconnected. Generally, a cell is connected to its neighboring cells within so-called sphere of influence. Synaptic strengths across the sphere of influence are graded by a Gaussian function. The LGN is modeled as a linear structure, with axons extending into the cortex, and the axons are modeled as a time delay. For an extensive description of the model the reader is referred to (Nenadic, 2001).

Localized stationary stimuli are simulated by a simultaneous presentation of a square current pulses of certain amplitude and duration to a set of adjacent LGN neurons. Moving stimuli are simulated by delaying localized stimuli sequentially from left to right, where the delay parameter is kept constant and is inversely proportional to the velocity of a moving stimulus. To make problem tractable, the inputs are labeled; the most left stationary input is *input 1* and the most right input is *input N*. Likewise, the input moving from left to right with the slowest velocity is labeled as *input 1*, and the input moving with the fastest velocity is *input N*. Other inputs within the two families are linearly arranged between *input 1* and *input N*. The responses of the network to different stimuli have been simulated and recorded.

3. VISUALIZING THE CORTICAL RESPONSE

As discussed above, we record the response of the network to various stimuli. In particular we observe the time responses of cortical neurons i.e. pyramidal, stellate and horizontal cells. However, for data analysis we use only the time responses of the pyramidal cells (lateral and medial), which are also the most numerous cells in the turtle visual cortex. Furthermore, experimental data are available for pyramidal cells, therefore it is possible to compare our data against experimental results. The raw spike data has been low-pass filtered prior to any analysis. The reason for this is that the spikes introduce a high frequency component in the time signal, which may give rise to non-smooth signals in further analysis. Also, low-pass filtering removes the dc component of the signal, which again is relevant for further signal processing.

Let $S\colon [0,\, T] \to \mathcal{R}^n$ be a function of the spike train sequences corresponding to the network of n pyramidal cells, where T is the total time of simulation. Then the function $R\colon [0,\, T] \to \mathcal{R}^n$ is the time response function of n pyramidal cells, and is defined via convolution operator

$$\mathbf{R}(t) = \int_{-\infty}^{t} \mathbf{S}(\tau)\, h(t-\tau)\, d\tau,$$

where $h(t)$ is the impulse response of a *bona fide* low-pass filter. Given the coordinates $(x_i,\, y_i)$ of n pyramidal cells, the response function $\mathbf{R}(t)$ can be written as a spatiotemporal signal

$$I(x, y, t) = \sum_{i=1}^{n} r_i(t)\, \delta(x - x_i,\, y - y_i), \qquad (1)$$

where $\mathbf{R}(t) = \begin{bmatrix} r_1(t) \\ r_2(t) \\ \vdots \\ r_n(t) \end{bmatrix}$, and $\delta(x, y)$ is a Dirac distribution on $\mathcal{R}^2$. The signal in (1) can be resampled to a uniform grid of coordinates, and can be visualized as a sequence of snapshots using MATLAB imaging tools. Among different resampling methods, we use a triangle-based linear interpolation. The time responses of individual pixels are color coded, and one such sample has been shown in figure (1).

4. MODAL DECOMPOSITION OF CORTICAL RESPONSE

Our model contains 679 pyramidal cells, and the simulation time is set to $1,500\, ms$ with the sampling time $dt = 1\, ms$. Therefore, the response function $\mathbf{R}(t)$ can be discretized in time and represented as a matrix in $\mathcal{R}^{679\times 1500}$. By playing movies corresponding to different visual stimuli we

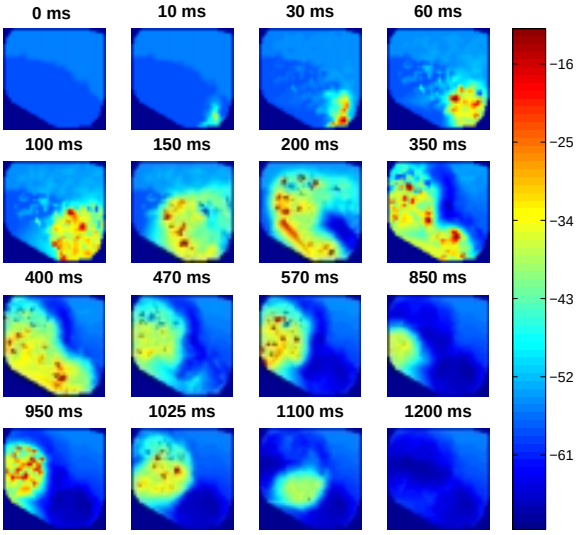

Fig. 1. Selected snapshots from a particular movie showing the wave propagation. The time is measured with respect to the onset of a stimulus. The stimulus is switched off after $150\,ms$, and the signal is expressed in mV and color coded.

notice that they are different, and the differences between them have to be quantified. Due to the large dimension of the data sets, a convenient compression method has to be found. The entries of the response matrix $\mathbf{R}$ are highly correlated; the correlation in space is due to the connectivity of individual neurons (response of a particular cell is affected by the responses of the cells in its sphere of influence), and the correlation in time is because each cell represents a time invariant dynamical system whose response depends on the response in the previous instance. Because of this redundancy, the representation of the response matrix $\mathbf{R}$ is by no means optimal, and can be rewritten in a subspace of lower dimension (less degrees of freedom). One such method that gives an optimal approximating subspace is the principal components method which will be introduced next.

The principal components analysis was introduced independently by many authors at different times. The method is widely used in various disciplines such as image and signal processing, data compression (Rao and Yip, 2001), fluid dynamics (Holmes *et al.*, 1996), partial differential equations (Dellnitz *et al.*, 1994), weather prediction, etc. The transformation itself is linear, and represents the rotation of a coordinate system, so that the data 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 a diagonalization of the image correlation matrix. Consequently, the image can be approximated in a lower dimensional subspace, using only selected basis vectors, also called principal eigenvectors. Depending on the context, the method goes by the names: Karhunen-Loeve (KL) decomposition, proper orthogonal decomposition, Hotelling decomposition and singular value decomposition. In the further discussion, we shall refer to it as KL decomposition.

After discretization in time, the response matrix corresponding to k^{th} data set $\mathbf{R}^k(t)$ can be found as $\mathbf{R}^k \in \mathcal{R}^{n\times m}$ where n is the number of cells and m is the number of time samples. The average correlation matrix for a set of N different data sets (possibly resulting out of N different stimuli) can be found as

$$\mathbf{C_1} = \frac{1}{N\,m}\sum_{i=1}^{N}\sum_{j=1}^{m}(\mathbf{v}_j^i)(\mathbf{v}_j^i)^T \qquad (2)$$

where $\mathbf{R}^k = [\mathbf{v}_1^k|\mathbf{v}_2^k|\cdots|\mathbf{v}_m^k]$ and $\mathbf{v}_j^k \in \mathcal{R}^{n\times 1}$. The matrix in (2) is symmetric and positive semidefinite, therefore its eigenvalues are all real and non-negative and can be ordered in a descending sequence. The corresponding eigenvectors are orthonormal and form a basis in $\mathcal{R}^n$. This basis will be referred to as *global basis*. The eigenvectors corresponding to the largest p eigenvalues of $\mathbf{C}_1$ are called the principal eigenvectors (modes) and the p^{th} order successive reconstruction of the response function $\mathbf{R}^k(t)$ is given by

$$\hat{\mathbf{R}^k}(t) = \sum_{i=1}^{p}\alpha_i^k(t)\,\mathbf{M}_i, \qquad (3)$$

where $\mathbf{M}_i \in \mathcal{R}^{n\times 1}$ are the principal modes and the time coefficients $\alpha_i^k(t)$ are given by $\alpha_i^k(t) = \langle \mathbf{v}^k(t), \mathbf{M}_i\rangle$. Thus, the k^{th} data set $\mathbf{R}^k(t)$ is characterized by a set of p time coefficients $\alpha_i^k(t)$, and the compression can be written as $\mathbf{R}^k(t) \rightarrow \left[\alpha_1^k(t), \alpha_2^k(t), \cdots, \alpha_p^k(t),\right]$. In the existing literature on modal decomposition of turtle cortical waves of both experimental (Senseman and Robbins, 1999) and computational preparations (Nenadic and Ghosh, 2001), it has been reported that a 'reasonably' good reconstruction of the cortical response can be obtained with exceptionally few principal modes, e.g. 3 or 4. Our analysis indicates a different result. In the experimental preparation of Senseman, the input space was not as diverse as in the case of computational model. That means that the movies used for generating the global basis were highly correlated, so a large degree of data compression was possible. The previous version of computational model (Nenadic, 2001) was lacking sufficient connectivity, therefore the waves observed in this network were primarily the consequences of input stimulus, but not so much of interactions among the cortical cells. The current state of the model, as introduced earlier in this article, has a very prolonged transient response. Namely the waves are dying out approximately $800 - 1000\,ms$ after

the offset of a stimulus, indicating that there is a strong interaction among cortical neurons, and that the wave phenomenon is considerably shaped by the cortico-cortical connections. To determine a minimal number of spatial modes necessary for a 'faithful' representation of a response function $\mathbf{R}^k(t)$, we use the concept of distance.

We define a distance between time response functions $\mathbf{R}^k(t)$ and $\mathbf{R}^j(t)$ as

$$d(\mathbf{R}^k(t),\, \mathbf{R}^j(t)) = \int_0^T (\mathbf{R}^k(t) - \mathbf{R}^j(t))^T (\mathbf{R}^k(t) - \mathbf{R}^j(t))\, dt \tag{4}$$

The idea is to choose the order of reconstruction which will preserve the distance between two data sets. This guarantees that two data sets which were neighbors (in $\mathcal{L}^2$ sense) in the space of response functions will remain neighbors in the space of time coefficients $\alpha(t)$, and this space will conveniently be denoted by A-space. Let us define two compressions

$$\mathbf{R}^k(t) \to \left[\alpha_1^k(t), \alpha_2^k(t), \cdots, \alpha_p^k(t),\right]$$

and

$$\mathbf{R}^j(t) \to \left[\alpha_1^j(t), \alpha_2^j(t), \cdots, \alpha_p^j(t),\right]$$

as described by (3). Then the distance between two approximations $\hat{\mathbf{R}}^k(t)$ and $\hat{\mathbf{R}}^j(t)$ can be written as

$$d(\hat{\mathbf{R}}^k(t),\, \hat{\mathbf{R}}^j(t)) = \sum_{i=1}^{p} \int_0^T (\alpha_i^k(t) - \alpha_j^k(t))^2\, dt \tag{5}$$

The result given by (5) is easy to prove and follows from the fact that principal modes are orthonormal.

We have seen how a time response function $R\colon [0,\, T] \to \mathcal{R}^n$ can be compressed to a function $\hat{R}\colon [0,\, T] \to \mathcal{R}^p$ $(p < n)$ by virtue of the KL decomposition. The time function thus obtained still does not have a convenient form, and can be further decomposed for proper analysis. The idea is to use the KL decomposition once again, this time on data sets represented by $\hat{\mathbf{R}}(t)$. After the dimension p of approximating subspace has been found, we can construct a global basis for the second KL decomposition by calculating the correlation matrix

$$\mathbf{C}_2 = \frac{1}{N} \sum_{i=1}^{N} \mathbf{A}^i\, \mathbf{A}^{iT}$$

where $\mathbf{A}^i = \left[\mathbf{a}_1^i,\, \mathbf{a}_2^i,\, \cdots,\, \mathbf{a}_p^i\right]^T$ and where $\mathbf{a}_j^i$ represents the time function $\alpha_j^i(t)$ discretized in time, i.e. $\mathbf{a}_j^i = \left[\alpha_j^i(1),\, \alpha_j^i(2),\, \cdots,\, \alpha_j^i(m)\right]$. Despite its simplicity, this method has a serious drawback in practical implementation. Namely, the dimension of $\mathbf{A}^i$ is very large, $\mathbf{A}^i \in \mathcal{R}^{p \times m}$, which implies that the matrix $\mathbf{C}_2$ has $p \times m$ eigenvalues/eigenvectors. Such a computation can be easily performed for a low order function approximation (p small). However, we have already indicated that such approximation is not useful, therefore solving for the eigenvalues and eigenvectors of $\mathbf{C}_2$ may be computationally expensive if not impossible on commercial digital computers. We propose to approximate the eigenvectors of $\mathbf{C}_2$ (basis vectors for second KL decomposition) by the method of averaging. The method consists of a recursive algorithm where the first principal mode of $\mathbf{C}_2$ is approximated by a vector that represents the average of the collection of vectors $\{\mathbf{A}^1,\, \mathbf{A}^2\, \cdots,\, \mathbf{A}^N\}$, i.e.

$$\hat{\Phi}_1 = \frac{\sum_{i=1}^{N} \mathbf{A}^i}{||\sum_{i=1}^{N} \mathbf{A}^i||},$$

and the corresponding residual vectors are given by $\mathbf{e}^i = \mathbf{A}^i - \langle \mathbf{A}^i, \hat{\Phi}_1 \rangle$. Because the residual vectors satisfy the following property $\sum_{i=1}^{N} \mathbf{e}^i = \mathbf{0}$, their average cannot be taken as a basis vector. Equivalently said, the residual vectors point in many different directions and consequently their average gives rise to a zero vector. Thus, we need to change the directions of certain number of residual vectors, by flipping their signs. Let $\mathbf{e}^k$ be any nonzero residual vector from the collection $\{\mathbf{e}^1,\, \mathbf{e}^2\, \cdots,\, \mathbf{e}^N\}$. We calculate the cosine of the angle between $\mathbf{e}^k$ and $\mathbf{e}^i$ $(\forall i = 1, 2, \cdots, N)$, and change the sign of those vectors with corresponding negative cosine. If

$$S = \{j\colon \langle \mathbf{e}^j,\, \mathbf{e}^k \rangle < 0\},$$

then $\mathbf{e}^j = -\mathbf{e}^j$. Then the approximation to the second principal mode can be found as

$$\hat{\Phi}_2 = \frac{\sum_{i=1}^{N} \mathbf{e}^i}{||\sum_{i=1}^{N} \mathbf{e}^i||}.$$

This procedure can be repeated as many times as necessary to obtain a desired number of principal modes. Note that the vectors thus obtained are orthonormal and form a basis in a Euclidean space of proper dimension. Once the approximation to the global basis is known, the q^{th} order successive approximation of $\mathbf{A}^k(t)$ is given by

$$\hat{\mathbf{A}}^k(t) = \sum_{i=1}^{q} \beta_i^k\, \hat{\Phi}_i(t),$$

where $\mathbf{A}^k(t) = \left[\alpha_1^k(t),\, \alpha_2^k(t)\, \cdots, \alpha_N^k(t)\right]$, and $\beta_i^k = \langle \mathbf{A}^k(t),\, \hat{\Phi}_i(t) \rangle$. After the double data compression each time response function is represented by a set of q real numbers $\mathbf{R}^k(t) \to \left[\beta_1^k,\, \beta_2^k, \cdots,\, \beta_q^k\right]$. The coefficients β belong to a q dimensional Euclidean space, called B-space. As before, the distance can be written as

$$d(\hat{\mathbf{A}}^k(t), \hat{\mathbf{A}}^j(t)) = \sum_{i=1}^{q} (\beta_i^k - \beta_i^j)^2, \tag{6}$$

and this result readily follows from the fact that the basis vectors are orthonormal.

5. SIMULATION RESULTS

We performed simulation for both stationary and moving stimuli. Stationary stimuli are localized from left to right, and are labeled from 1 to 19. They are parameterized by their center point, width and amplitude of the square current pulse. The width is kept constant to 20 LGN cells and the amplitude is set to $0.2\,nA$. The moving stimuli are labeled from 1 to 18 and represent a localized flash of light moving with a constant velocity from left to right. The stimuli are parameterized by the velocity, the slowest stimulus being *input 1* and the fastest stimulus being *number 18*. The distances are calculated based on formulas (4),(5) and (6). The number of modes p and q are found such that the distance is preserved. For the stationary input case $p = 170$ is found to be the compression that preserves the distance. In the second KL decomposition, $q = 37$ provides sufficiently close distances. Global basis is calculated based on $N = 37$ data sets, and the time responses have been truncated at $T = 1,000\,ms$, with the time sample $dt = 1\,ms$. We now form 3 matrices $\mathbf{D}_1$, $\mathbf{D}_2$ and $\mathbf{D}_3$ of distances in three different subspaces, namely the space of time response functions, A-space and B-space. Matrix $\mathbf{D}_1$ is a matrix of distances in the space of time response functions, so that $\mathbf{D}_1(i,j) = d(\mathbf{R}^i(t), \mathbf{R}^j(t))$ where the distance is calculated by (4). Likewise, $\mathbf{D}_2(i,j) = d(\hat{\mathbf{R}}^i(t), \hat{\mathbf{R}}^j(t))$ and $\mathbf{D}_3(i,j) = d(\hat{\mathbf{A}}^i(t), \hat{\mathbf{A}}^j(t))$, where the two distances have been evaluated using (5) and (6), respectively. The three distance matrices have been visualized and displayed in figure 2, where darker shades represent smaller distance and vice versa. The three images from figure (2) are quite similar, as are the distance matrices $\mathbf{D}_1$, $\mathbf{D}_2$ and $\mathbf{D}_3$. Similar analysis has been applied to the response to moving stimuli, and the two numbers are found to be $p = 170$ and $q = 18$. The corresponding distance matrices are shown in figure (3).

Fig. 2. Three distance matrices are displayed as images. The top figure refers to the distance matrix $\mathbf{D}_1$. The distance matrix $\mathbf{D}_2$ is given in the middle and $\mathbf{D}_3$ is shown in the bottom.

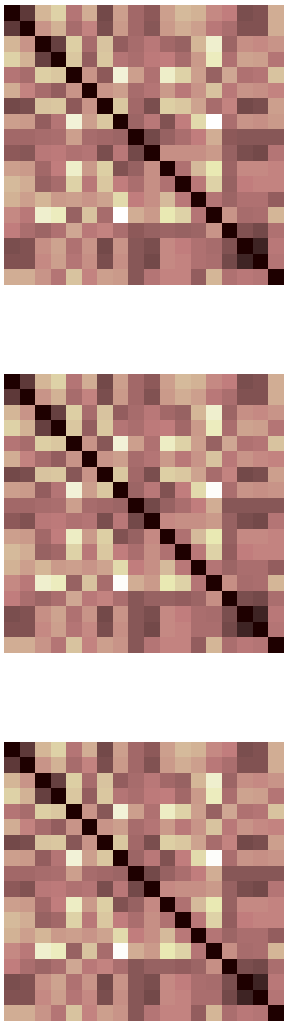

Fig. 3. Three distance matrices are displayed as images. The top figure refers to the distance matrix $\mathbf{D}_1$. The distance matrix $\mathbf{D}_2$ is given in the middle and $\mathbf{D}_3$ is shown in the bottom.

6. CONCLUSIONS

In this article we introduced a large scale computational model of turtle visual cortex. Such a model fairly accurately captures some of the important features of a real visual cortex (Nenadic, 2001). In the second part of the article we introduce the KL decomposition, which proved to be quite useful in the analysis of generated spatiotemporal wave forms. Unlike the analysis of experimental data and existing computational model, which indicated that the waves could be analyzed using only a few principal modes, our analysis has provided a different result. We use the concept of distance in $\mathcal{L}^2$ (Euclidean distance for discrete algorithms), to find the minimum number of principal modes that will preserve the distance from original space to a subspace of reduced dimension. This might be important if statistical inferences on the parameters of an unknown input are to be made (Nenadic, 2001) with a certain level of confidence. We also introduce the approximate algorithm for finding the best approximating subspace in the sense of KL decomposition. This algorithm, even though yielding a suboptimal solution, proves more efficient over the classical eigenvalue/eigenvector problem, as it calculates only necessary (minimum) number of modes, starting from the first one (very rarely do we need all principal modes). The work outlined above can be extended to a detection problem, where one could build a statistics on the coefficients β by imposing randomness in the model. Once the conditional densities of observations given stimuli are known, Bayesian and maximum likelihood methods can be used for detection.

7. REFERENCES

Delcomyn, F. (1998). *Foundations of Neurobiology*. W. H. Freeman & Co. New York.

Dellnitz, M., M. Golubitsky and M. Nicol (1994). In: *Trends and Perspectives in Applied Mathematics* (L. Sirovich, Ed.). Chap. 4, pp. 73–108. Springer-Verlag. New York.

Holmes, P., J.L. Lumley and G. Berkooz (1996). *Turbulence, Coherent Structure, Dynamical Systems and Symmetry*. Cambridge University Press. Cambridge.

Mazurskaya, P.Z. (1974). Organization of receptive fields in the forebrain of *emys orbicularis*. *Neurosci. Behav. Physiol.* **7**, 311–318.

Nenadic, Z. (2001). *Signal Processing and Computation in Biological Neural Networks*. Washington University. Doctoral Thesis.

Nenadic, Z. and B.K. Ghosh (2001). Signal processing and control problems in the brain. *IEEE Control Systems Magazine* **21**, 28–41.

Rao, K.R. and P.C. Yip (2001). *The Transform and Data Compression Handbook*. CRC Press. Boca Raton.

Senseman, D.M. (1996). Correspondence between visually evoked voltage sensitive dye signals and activity recorded in cortical pyramidal cells with intracellular microelectrodes. *Vis. Neurosci.* **13**, 963–977.

Senseman, D.M. and K.A. Robbins (1998). Visualizing differences in movies of cortical activity. *IEEE Trans. Visualization. Compt. Graphics* **4**, 217–224.

Senseman, D.M. and K.A. Robbins (1999). Modal behavior of cortical neural networks during visual processing. *J. Neuroscience* **4**, 1–7.