Central Auditory Processing and Neural Modeling,
edited by P. Poon and J. Brugge (Plenum Press, 1998)
pp. 161-174

MODELS OF DIRECTION ESTIMATION WITH SPHERICAL-FUNCTION APPROXIMATED CORTICAL RECEPTIVE FIELDS

Rick L. Jenison

Department of Psychology
University of Wisconsin-Madison
Madison, WI 53706

INTRODUCTION

Spatial sensitivity of auditory cortex to free-field sound sources has been recognized for many years. Azimuthal sensitivity of auditory cortex to free-field tones, noise, and clicks has been studied by several investigators (see Eisenman (1974) and Middlebrooks and Pettigrew (1981) for early work). Imig, Irons and Samson (1990) and Rajan, Aitkin, and Irvine (1990ab) have shown that about 80% of the neurons can be classified as azimuthal sensitive. Neurons could be generally classified as having contralateral (most prevalent), ipsilateral, central, multipeaked, or omnidirectional response properties. With the development of advanced signal processing techniques, virtual-space stimuli, rather than free-field stimuli, have been used to characterize, in detail, spatial receptive fields in the cat (Brugge, Reale, Hind, Chan, Musicant, and Poon, 1994). A benefit of using virtual stimuli, aside from the efficiency of stimulus generation, is that the stimuli can be manipulated in systematic ways to dissociate specific factors shaping a cortical spatial receptive field.

Virtual-space receptive fields (VSRFs) of cortical neurons are often broad and typically lateralized to one hemisphere. Throughout the VSRF, responses at effective directions vary in strength and latency. For a sizable proportion of recorded neurons gradient structure has been observed in the spatial distribution of first-spike latency. Furthermore, it was hypothesized that this structure may carry directional information across an ensemble of such neurons (Brugge, Reale, and Hind, 1996).

A spherical basis function approximation technique (Jenison and Fissell, 1996) was applied as a first step toward testing hypotheses concerning the information bearing latency structure of VSRFs. Populations of modeled neurons were then analyzed with the goal of obtaining an optimal estimation, in a probabilistic sense, of sound direction. For a population of cortical neurons with statistically independent latency, maximum-likelihood estimation of direction is asymptotically optimal. Dependence of maximum-likelihood estimation performance on the spatial response properties of cortical neurons and population size were

investigated using the Cramer-Rao lower bound (CRLB). The CRLB represents the minimum possible variance of unbiased estimates of an unknown parameter (or parameters). In this study, the parameters are dimensions of azimuth and elevation. Latency variability was assumed to be independent and follow a Gaussian distribution whose variance is a function of first-spike-latency. The utility of the maximum-likelihood approach is that modeled estimation error can be compared directly to that obtained psychophysically.

VIRTUAL-SPACE RECEPTIVE FIELDS

Receptive fields with graded and overlapping response profiles are ubiquitous in biological systems (McIlwain, 1976). The question as to how a population of cells, each of which exhibit broad spatial tuning, can account for psychophysical discrimination of a few degrees of angle is the focus of this chapter. Under general anesthesia, AI neurons generally respond to transient virtual acoustic stimuli with a single or on occasion several spike discharges that are tightly time-locked to the stimulus (Brugge, Reale, Hind, Chan, Musicant, and Poon, 1994). Any directional information contained within the VSRF is hypothesized to be found in the graded structure of the receptive field (Brugge, Reale, and Hind, 1996). It has been observed that at low intensities the maximum response of the spatial receptive field tends to be centered near the acoustic axis. The acoustic axis is the direction at which the transformation of sound pressure by the pinna achieves maximum amplification at a given frequency. Three general characteristics of the VSRF were observed. First, there is a systematic intensity-dependent expansion of the VSRF as intensity increases. Second, latency tends to be at a minimum near the contralateral acoustic axis. Third, the absolute value of the latency shortens with increasing intensity of the stimulus. Thus, response latency may conjointly convey information specifying direction of the sound source and its intensity. The interaction between the acoustic axis induced by pinna filtering and the characteristic frequency of an AI neuron has the effect of modulating the spatial location of the VSRF centroid. Although VSRFs are often quite broad, the gradient structure within the field affords systematic variability that could be exploited by some form of population decoding (Heiligenberg, 1987; Zohary, 1992). Using a virtual acoustic stimulus procedure to map the gradient structure of the receptive field with high spatial resolution, Brugge et al. (1996) found a sizable population of recorded AI neurons with discharge timing following an orderly, but at times complex, pattern. Factors including interaural-time-delay, intensity, and spectrum could play major roles in determining this structure. An approach to functionally characterizing gradient structure of the VSRF is addressed next using techniques common to mathematical approximation theory .

SPHERICAL-FUNCTION APPROXIMATION

Ever since the pioneering work of Hubel and Wiesel (1962, 1974) on cortical functional architecture and receptive field organization, it has been known that critical tuning variables characterizing response selectivities of visual cortical cells include the dimensions of 2-D location, size, and orientation. These spatial variables have been combined into a framework that models cortical simple cell receptive fields as Gabor functions (Marcelja, 1980; Daugman, 1980) or difference-of-Gaussians (Rodieck, 1965; Marr, 1982). A similar approach is proposed for characterizing response latencies of auditory cortical neurons using a form of spherical approximation based on the so-called von Mises basis functions (Jenison and Fissell, 1995; 1996). In the present case, only first-spike latency is approximated, but the methodology extends easily to spike-train approximation.

Artificial neural networks and approximation techniques typically have been applied to problems conforming to a multidimensional Cartesian input space. However, approximation problems that exist on a unit sphere, such as directional response measures, are most accurately framed within a spherical or polar coordinate system, rather than projected implicitly onto a two-dimensional Cartesian system. Spherical data arise in many areas of science such as geophysics and meteorology, and projection geometry has been the focus of cartographic interest for problems involving a transformation of the spherical earth into a planar map. A well-known problem to cartographers is that area, direction and shape cannot be preserved simultaneously in these projections, hence some degree of spatial distortion must be incurred, unless an alternative basis is used.

Radial Basis Function (RBF) approximation can be viewed as expanding an input-output mapping (or surface) into a linear combination of weighted nonlinear basis functions (Powell, 1987). We have extended radially-symmetric functions on the plane to that of a rotationally-symmetric function on the sphere. This basis function is based on a spherical probability density function that has been used to model line directions distributed unimodally with rotational symmetry. The function is well-known in the spherical inferential statistics literature and is commonly referred to as either the von Mises-Arnold-Fisher or Fisher distribution (Mardia, 1972; Fisher, Lewis, and Embleton, 1987). The kernel form of the probability density function was first introduced by Arnold (1941) in his unpublished doctoral dissertation. The expression for the von Mises basis function (VMBF), dropping the constant of proportionality and elevational weighting factor from the probability density function, is

$$VM(\theta, \phi, \alpha, \beta, \kappa) = e^{\kappa[\sin\phi \sin\beta \cos(\theta-\alpha) + \cos\phi \cos\beta]} \tag{1}$$

where the input parameters correspond to a sample direction in azimuth and elevation (θ, ϕ), a centroid in azimuth and elevation (α, β), and the concentration parameter, κ. Application of the von Mises function requires an azimuthal range in radians from 0 to 2π and elevational range from 0 to π. Any sample direction (θ, ϕ) on the sphere will induce an output from each VMBF proportional to the angle between the sample direction and the centroid of the VMBF (α, β). κ is a shape parameter called the concentration parameter, where the larger the value the narrower the function width after transformation by the expansive *exponential*. While other spherical functions have been proposed for approximation on the sphere (e.g. thin-plate pseudo-spline (Wahba, 1981)), the VMBF serves as a convenient spherical analogue of the well-known multidimensional Gaussian on a plane. It resembles a bump on a sphere and behaves in a similar fashion to the planar Gaussian with the centroid corresponding to the mean and $1/\kappa$ corresponding to the standard deviation. It differs from the thin-plate spline in that it has a parameter for controlling the width or concentration of the basis function, which allows the VMBF to focus resolution where needed.

The approximation function, $m_\ell(\theta,\phi)$, that yields in the present case spike latency in response to the stimulus, is given by

$$m_\ell(\theta,\phi) = \sum_{j=1}^{J} w_{\ell j}\, VM(\theta,\phi,\alpha_j,\beta_j,\kappa_j) \tag{2}$$

where $VM(\theta,\phi,\alpha_j,\beta_j,\kappa_j)$ is the output of the jth von Mises basis function and $w_{\ell j}$ is the

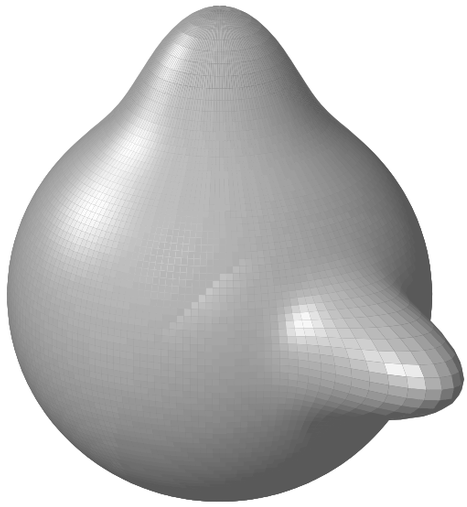

Figure 1. 3-D rendering of two von Mises basis functions positioned on the unit sphere. The basis functions are free to move about the sphere, changing adaptively in position, width, and magnitude.

weight on the jth basis function. For a multivariate output, $m_\ell(\theta,\phi)$ is the ℓth element in the output vector. For the case of first-spike latency, ℓ is restricted to 1, however for the case of multiple spikes, ℓ would index each observed spike latency (or interval). Two arbitrary von Mises basis functions (J = 2) are shown positioned on the unit sphere in Figure 1 to illustrate Equation 2. The approximation parameters (α, β, κ, and w) for a fixed number of basis functions can be adaptively learned by applying a gradient-descent method[1] to the desired cost-function for the training set of P input-output pairs, where the pth observed patterns t_p is modeled as

$$t_p = m(\theta_p,\phi_p) + \varepsilon_p \tag{3}$$

such that the error or noise term ε_p is independent with zero mean and unit variance. For this case we require the approximation sum-of-squared-error to be minimized. The expression of this cost-function for the pth L-dimensional known or measured sample (or in adaptive learning parlance, the teaching pattern) is

$$E_p = \frac{1}{2}\sum_{\ell}^{L}\left(t_{\ell p} - m_\ell(\theta_p,\phi_p)\right)^2 \tag{4}$$

where $t_{\ell p}$ is the ℓth element of the pth teaching pattern and $m_\ell(\theta_p,\phi_p)$ is the pth approximated pattern. In the present case, the teaching pattern is the observed first-spike latency, for a known sound direction.

Parameter values are learned through successive presentation of P input-output teaching pairs and application of the specific update rules for each parameter of the von Mises basis function. The amount of change made to each parameter following successive presentation is based on the error gradient (i.e., the negative derivative of the error with respect to that parameter,) which was derived analytically using the chain rule for partial differential equations (see Jenison, 1995; Jenison and Fissell, 1995; Jenison and Fissell, 1996). Over the course of iterative training, the centroid of each basis function will move on the surface of the sphere,

and the concentration (width) of the basis function will change, progressively minimizing the cost-function (Equation 4).

The update rules for learning the von Mises basis parameters have the following form:

$$\Delta w_{\ell j} = -\frac{\partial E}{\partial w_{\ell j}} = \left(t_\ell - m_\ell(\theta,\phi)\right) VM\left(\theta,\phi,\alpha_j, \beta_j, \kappa_j\right) \tag{5}$$

$$\Delta\alpha_j = -\frac{\partial E}{\partial \alpha_j} = \sum_{\ell=1}^{L} \left\{\left[t_\ell - m_\ell(\theta,\phi)\right] w_{\ell j}\right\} VM\left(\theta,\phi,\alpha_j, \beta_j, \kappa_j\right) \times \kappa_j \left[\sin\phi \sin\beta_j \sin\left(\theta - \alpha_j\right)\right] \tag{6}$$

$$\Delta\beta_j = -\frac{\partial E}{\partial \beta_j} = \sum_{\ell=1}^{L} \left\{\left[t_\ell - m_\ell(\theta,\phi)\right] w_{\ell j}\right\} VM\left(\theta,\phi,\alpha_j, \beta_j, \kappa_j\right) \times \kappa_j \left[\sin\phi \cos\beta_j \cos\left(\theta - \alpha_j\right) - \cos\phi \sin\beta_j\right] \tag{7}$$

$$\Delta\kappa_j = -\frac{\partial E}{\partial \kappa_j} = \sum_{\ell=1}^{L} \left\{\left[t_\ell - m_\ell(\theta,\phi)\right] w_{\ell j}\right\} VM\left(\theta,\phi,\alpha_j, \beta_j, \kappa_j\right) \times \left[\sin\phi \sin\beta_j \cos\left(\theta - \alpha_j\right) + \cos\phi \cos\beta_j\right]. \tag{8}$$

VMBF Approximation Of Virtual-Space Receptive Fields

First-spike latency has been shown to be a measure of the direction dependent response of single neurons in the auditory cortex, and because of the spherical nature of the VMBF approach, VMBF approximation is likely to be the best “tool” for functionally modeling the directional response properties. In this approach, a sufficient number of basis functions are chosen, such that the systematic internal fine structure of the spatial receptive field is characterized by the approximation, and the unsystematic noise is isolated as residuals. This is commonly accomplished by keeping the number of basis functions used in the approximation to a small number, in most cases in the neighborhood of nine to 35. Following asymptotic learning of the spatial receptive field as a function of direction, a dense test set is used to interrogate the approximation for how well it has generalized to novel directions. Figure 2 shows the rendering of the measured first-spike latency as a function of direction (left) and the functional approximation using the VMBF approach with 35 basis functions (right). In this rendering, first-spike latency is reflected about a reference latency (the neuron’s longest latency), such that longer radii correspond to shorter latencies.

Functional Approximation of Latency Variance

Following VMBF approximation to the VSRF, residuals about the modeled prediction are examined — a common practice in regression analysis. Figure 3 shows the scatter of latencies about the modeled approximation. The variance about the estimate is clearly not homoscedastic (constant variance) as assumed by the VMBF model, but tends to increase as a function of the first-spike latency approximation, which is a trend supported in the literature

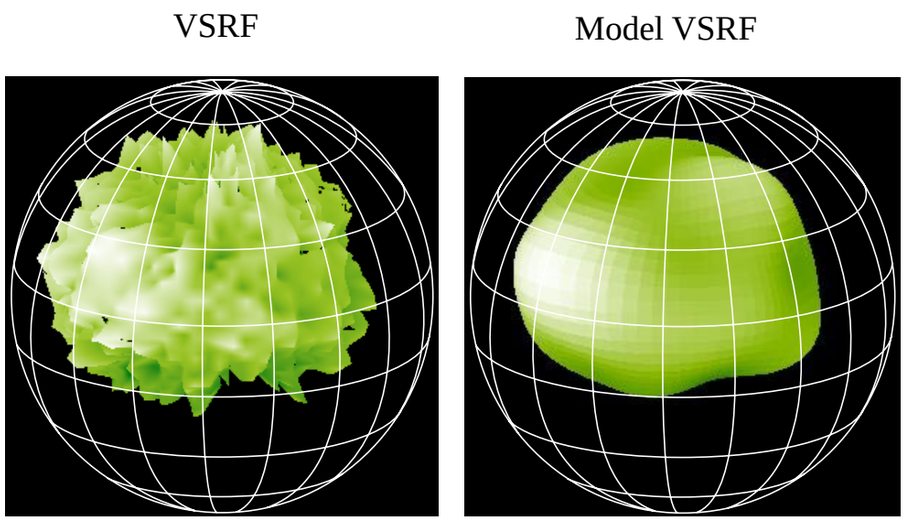

Figure. 2. VMBF approximation to first-spike latency, as a function of direction. Approximation surface is shown projected through the data on spherical coordinates.

for both latency and discharge-rate (e.g., Phillips and Hall, 1990; Tolhurst, Movshon, and Dean, 1983). Although heteroscedasticity is often recognized simply by examining the residuals by eye, more quantitative techniques have been mined from the field of econometrics for the present application (see Muller and Stadtmuller, 1987; Hall and Carroll, 1989). In the present case, variance of the residuals is considered to be a linear function of the VMBF approximation $m(\theta,\phi)$, where

$$\text{var}\,(\varepsilon_p) = g\left(m\left(\theta_p,\phi_p\right)\right) \tag{9}$$

and

$$g\left(m\left(\theta_p,\phi_p\right)\right) = a\, m\left(\theta_p,\phi_p\right) + b. \tag{10}$$

The estimates of *a* and *b* of the function $g(m(\theta,\phi))$ are obtained via minimization of the squared-deviates, which utilize measurements from all the sample directions to obtain the estimates. The distribution of first-spike latency residuals (or deviates) will be conveniently modeled by a Gaussian probability density using Equation 10 as the variance term and Equation 2 as the mean term, as shown in Figure 3. The curved dotted line shows the square-root of the variance function (the standard deviation), which is a nonlinear transformation of the linear variance function. It should be noted that this approach to fitting response variability differs from the more common approach in the literature where the variance (or standard deviation) is calculated with respect to a mean latency (or rate) and then fit as a function of mean latency (or rate). In the present case, the squared-deviates are fit directly as a function of the modeled latency.

Residual analysis is important for two reasons. First, violating the assumption of homoscedasticity (Equation 3) requires a weighting strategy to lessen the contribution of large variability to the least-squares approximation. Second, the variance function is an integral part of an optimal observer model, which is described in the next section.

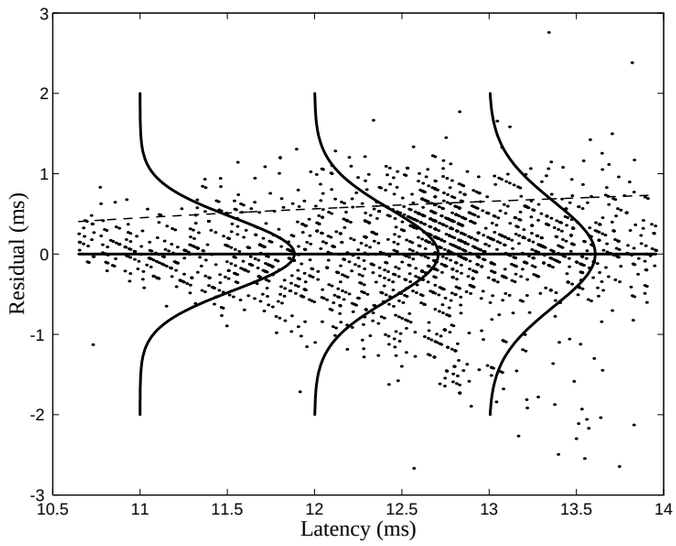

Figure. 3. Residuals/deviates (first-spike latency) about VMBF approximation (shown in Fig, 2). Deviates are assumed to be Gaussian about the approximation and heteroscedastic. Heteroscedastic variance is modeled as a linear function of approximated latency, however the standard deviation will appear as curved (dotted line).

OPTIMAL OBSERVER

A very useful application of spherical approximation to first-spike latency and the linear approximation to latency variance about the spherical model is the derivation of theoretically optimal observer to estimate sound direction. Much has been written on the topic of estimation theory, the roots of which extend back to Sir Ronald Fisher (1925). Estimation theory is concerned with the problem of finding a best value for an unknown parameter from a measured sample. The accuracy of the estimate is limited by the quality of the measurements and the gradients of the function of the unknown parameters. For a large population of neurons whose responses are independent, maximum-likelihood estimation procedures are optimal in a probabilistic sense. Although maximum-likelihood estimation provides theoretical bounds on performance, the exact mathematical procedures are unlikely to have a neural basis. Still, maximum-likelihood estimation provides a bench mark for evaluating psychophysical performance, as well as evaluating future neural network models. This approach has been employed in interaural discrimination models (Colburn, 1973), and more recently in function-based models of binaural processing (Johnson, Dabak, and Tsuchitani, 1990; Dabak and Johnson, 1992). Gradient-descent trained artificial neural networks also have been used to evaluate the available directional information in the spike pattern of a single cortical cell (Middlebrooks, Clock, Xu, and Green, 1994). The present approach differs from these other approaches in that it addresses spherical direction, rather than direction restricted to the horizontal plane, and considers ensemble rather than single cell responses. Methods of maximum-likelihood estimation also have proved useful in predicting psychophysical results of orientation discrimination based on models of visual cortex (Paradiso, 1988).

Maximum-likelihood Estimation

Given a model of the spatial receptive field, $m_i(\theta, \phi)$, for the ith neuron and a model of variance growth, $g\left(m_i(\theta, \phi)\right)$, we consider a population vector of latency observations, denoted

$\mathbf{x} = [x_1, x_2, \cdots, x_N]$ conditioned on two real parameters θ and ϕ, azimuth and elevation. We assume that each neuron in the population of N neurons has a probability density function of the following Gaussian form

$$f(x_i|\theta,\phi) = \frac{1}{\sqrt{2\pi\, g(m_i(\theta,\phi))}}\, e^{\frac{-(x_i - m_i(\theta,\phi))^2}{2\, g(m_i(\theta,\phi))}}. \tag{11}$$

The likelihood function is defined as a function of the parameters θ and ϕ, where the function is the joint density or product density of the vector of measurements **x** under the assumption of independent noise between cells[2]

$$L(\theta,\phi) = \prod_{i=1}^{N} f(x_i|\,\theta,\phi) \tag{12}$$

It should be emphasized that the likelihood function is a function of the parameters, given the vector of measurements. The method of maximum-likelihood entails choosing estimates of θ and ϕ, denoted $\hat{\theta}$ and $\hat{\phi}$ that maximize the likelihood function $L(\theta,\phi)$. To find maximum-likelihood estimates of θ and ϕ continuous in some interval, we differentiate $L(\theta,\phi)$ with respect to each parameter, set each result equal to zero, and solve for the desired unknowns θ and ϕ. Note that any value that maximizes the more tractable log-likelihood, $\ln L(\theta,\phi)$, will also maximize $L(\theta,\phi)$. Some simple algebra yields the following expressions

$$\frac{\partial}{\partial\theta}\ln L(\theta,\phi) = \frac{\partial}{\partial\theta}\left[\ln\prod_{i=1}^{N} f(x_i|\,\theta,\phi)\right] = \sum_{i=1}^{N}\frac{\frac{\partial}{\partial\theta}f(x_i|\,\theta,\phi)}{f(x_i|\,\theta,\phi)} = 0 \tag{13}$$

and

$$\frac{\partial}{\partial\phi}\ln L(\theta,\phi) = \frac{\partial}{\partial\phi}\left[\ln\prod_{i=1}^{N} f(x_i|\,\theta,\phi)\right] = \sum_{j=1}^{N}\frac{\frac{\partial}{\partial\phi}f(x_i|\,\theta,\phi)}{f(x_i|\,\theta,\phi)} = 0 \tag{14}$$

The theoretical picture looks like this — an ensemble of neurons fire, with independent first-spike latency, and the task of the observer is to choose estimates of azimuth and elevation that have the maximum-likelihood of being correct, given the observed ensemble of responses. Each neuron has been modeled as a mixture of von Mises basis functions that yields the modeled function $m_i(\theta,\phi)$ for the ith cortical neuron. As noted in the earlier residual analysis, the variability about the VMBF model was approximated by a variance that is linearly dependent on the modeled first-spike response. As a first approximation, all modeled neurons will make the simplifying assumption of having the same estimates of a and b. Differentiating

the log-likelihood functions with respect to azimuth and elevation, yield the following gradients

$$\frac{\partial}{\partial\theta}\ln L(\theta,\phi) = -\frac{1}{2}\sum_{i=1}^{N}\left\{\frac{\left(\frac{\partial}{\partial\theta}m_i(\theta,\phi)\right)\left(am_i^2(\theta,\phi)+a^2m_i(\theta,\phi)+2\,b\,m_i(\theta,\phi)-2\,b\,x_i-a\,x_i^2+a^2+ab\right)}{\left(am_i(\theta,\phi)+b\right)^2}\right\} \quad (15)$$

$$\frac{\partial}{\partial\phi}\ln L(\theta,\phi) = -\frac{1}{2}\sum_{i=1}^{N}\left\{\frac{\left(\frac{\partial}{\partial\phi}m_i(\theta,\phi)\right)\left(am_i^2(\theta,\phi)+a^2m_i(\theta,\phi)+2\,b\,m_i(\theta,\phi)-2\,b\,x_i-a\,x_i^2+a^2+ab\right)}{\left(am_i(\theta,\phi)+b\right)^2}\right\} \quad (16)$$

where

$$m_i(\theta,\phi) = \sum_{j=1}^{J} w_{ij}\ VM(\theta,\phi,\alpha_{ij},\beta_{ij},\kappa_{ij}) \quad (17)$$

$$\frac{\partial}{\partial\theta}m_i(\theta,\phi) = \sum_{j=1}^{J} -\kappa_{ij}\sin\phi\,\sin\beta_{ij}\sin(\theta-\alpha_{ij})\,w_{ij}VM(\theta,\phi,\alpha_{ij},\beta_{ij},\kappa_{ij}) \quad (18)$$

$$\frac{\partial}{\partial\phi}m_i(\theta,\phi) = \sum_{j=1}^{J} \kappa_{ij}\left[\cos\phi\ \sin\beta_{ij}\cos(\theta-\alpha_{ij}) - \sin\phi\cos\beta_{ij}\right] w_{ij}VM(\theta,\phi,\alpha_{ij},\beta_{ij},\kappa_{ij}). \quad (20)$$

Complete analytical solutions for $\hat{\theta}$ and $\hat{\phi}$ are difficult to derive in this case, therefore a constrained nonlinear optimization procedure (sequential quadratic programming) was used to find the estimates, numerically subject to the constraints that $0 \le \hat{\theta} < 2\pi$ and $0 \le \hat{\phi} \le \pi$. Consequently, the analytical gradients (Equations 15 and 16) serve to guide the numerical search for the maximum of the log-likelihood function $\ln L(\theta,\phi)$.

Cramer-Rao Lower Bound (CRLB)

Ideally, one would choose estimators $\hat{\theta}$ and $\hat{\phi}$ such that the expected values satisfy $E[\hat{\theta}] = \theta$ and $E[\hat{\phi}] = \phi$, so-called unbiased estimators. It is also desirable that the variance about the expected value of the estimators, $Var[\hat{\theta}]$ and $Var[\hat{\phi}]$, be as small as possible. The Cramer-Rao lower bound (CRLB) is a lower bound on the variance of any unbiased estimator

(Blahut, 1987), where

$$\mathrm{Var}\left[\hat{\theta}\right] \geq \frac{1}{\mathrm{E}\left[\frac{\partial}{\partial\theta}\ln\prod_{i=1}^{N} f\left(x_i|\,\theta,\phi\right)\right]^2} \tag{21}$$

and

$$\mathrm{Var}\left[\hat{\phi}\right] \geq \frac{1}{\mathrm{E}\left[\frac{\partial}{\partial\phi}\ln\prod_{i=1}^{N} f\left(x_i|\,\theta,\phi\right)\right]^2}. \tag{22}$$

If the CRLB can be derived analytically, it can be used to directly compute the minimum possible variability about any value estimated by a theoretical ideal observer. The power of this analysis is the following: whatever form of biological information processing ensues on the population of cortical responses, the very best that can be done, in terms of estimation error, is the CRLB. Furthermore, it can be shown that when the estimator is maximum-likelihood, the estimator asymptotically approaches the CRLB as the number of measurements becomes large. Suffice it to say that the analytical CRLB does indeed exist for Equations 21 and 22, however the derivation is lengthy and beyond the scope of this chapter.

Maximum-likelihood Estimation of Direction from Ensemble Responses

A single VSRF for first-spike latency as a function of direction was approximated using the VMBF technique. This approximation was modified such that the centroids of different approximations were positioned at the locations of the 65 VSRFs mapped by Brugge et al. (1996). A coarse approximation using a single VMBF was chosen as a first approximation to the directional dependency of first-spike latency. Future approximations will use higher-order VMBF approximations to capture finer detail of the internal structure of the VSRF.

Figure 4 shows the scatter of the Monte-Carlo simulated maximum-likelihood estimates about the true direction 0˚ azimuth, 0˚ elevation, and also at 45˚ azimuth, 45˚ elevation. The Monte-Carlo simulation was performed by randomly selecting a first-spike latency normal deviate from each VMBF modeled VSRF for the given direction of the sound source, and maximum-likelihood estimation of direction was performed. This process was executed 200 times and plotted in spherical coordinates. Surrounding the scatter of estimates is the so-called 2-σ ellipse denoting 2 standard deviations from the true direction, computed independently from the analytical CRLB. This confirmation is significant, as it allows us to compute analytical lower bounds of estimate variability without having to run thousands of Monte-Carlo simulations (Jenison, Reale, and Brugge, 1997). An analytical solution is also advantangeous in that it affords the opportunity to mathematically investigate the factors contributing to estimator variability.

The systematic effect on directional acuity of increasing the number of cells, from three to 65, in an ensemble of responses to a sound source at the origin (0˚ azimuth and 0˚ elevation) is shown in Figure 5. One hundred samples[3] of N cell ensembles selected at random from the total available 65 were selected for each N-cell ensemble, and the average of the computed lower bounds on the estimation error was calculated for each ensemble. The standard error for the 100 random ensembles is shown with respect to its average as a function of N. The

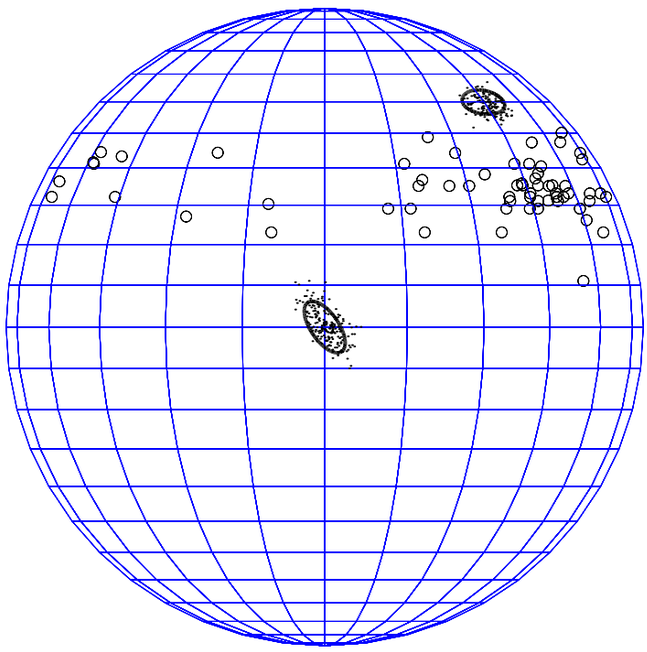

Figure 4. Maximum-Likelihood Estimates of sound location 0˚ azimuth and 0˚ elevation, and 45˚ azimuth and 45˚ elevation (dots). Cramer-Rao Lower Bounds are shown as 2-σ (2 s.d.) ellipses. Estimates and CRLB are based on an ensemble of 65 VMBF approximated neurons (centroids shown as circles).

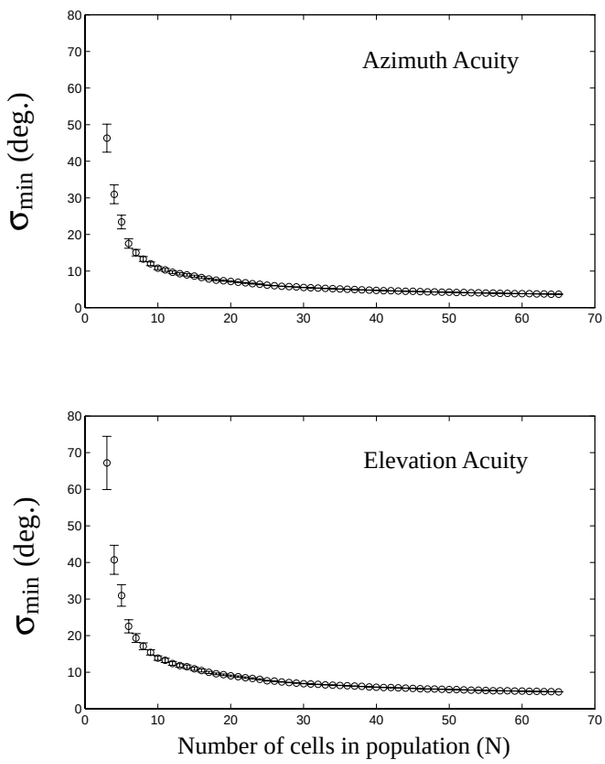

Figure 5. Cramer-Rao lower bound on the standard deviation of the estimates as a function of ensemble size. Each ensemble of size N is selected at random from the total pool of 65 cells. The mean CRLB is computed over 100 samples of size N. The standard error bars are shown for each value of N.

well-known dependence of maximum-likelihood error (standard deviation) on $\frac{1}{\sqrt{N}}$ is evident in both measures of azimuthal and elevational acuity, and illustrates the diminishing return in acuity beyond ensembles of 20 cells. In the cat, May and Huang (1996) observed angle standard deviations on the order of 2.2˚ in azimuth and 3.3˚ in elevation at 0˚ azimuth and at 0˚ elevation. In humans, Makous and Middlebrooks (1990) observed standard deviations on the order of 2.9˚ in azimuth and 3.6˚ in elevation at 0˚ azimuth and at 5˚ elevation.

CONCLUSIONS

A new approach has been presented for mathematically characterizing response selectivities of auditory cortical neurons using a form of spherical approximation based on von Mises basis functions. One important use of the formal von Mises basis function approximation is in the derivation of optimal observers (estimators). Maximum-likelihood estimators of direction from a simulated ensemble of first-spike latencies in auditory cortex generate unbiased estimates with asymptotic performance approaching that of the Cramer-Rao lower bound. The analytical lower bound on variability is approximately what one observes psychophysically. Although making absolute comparisons could be viewed as unwarranted given the number of parametric assumptions, the exercise does demonstrate psychophysically observed acuity with a relatively small ensemble of cells with broad receptive fields characteristic of cortical neurons.

An unsettled issue with respect to the use of first-spike latency as a dependent measure for sound direction is the apparent lack of temporal marking for the onset of the stimulus. One possible resolution is that relative, not absolute, latencies serve as the code, operating with respect to a periodic clock. In the absence of a periodic marker, relative latencies in an ensemble of cells could be decoded with respect to time relative to the initial spike in the ensemble. In this decoding scheme all responses would contribute to a common estimate of a single ensemble delay from stimulus onset. So, in addition to estimating azimuth θ and elevation ϕ, the optimal observer would also need to *estimate* an ensemble delay λ, thus eliminating the need for *knowing* the absolute temporal mark.

END NOTES

[1] The "gain" term $w_{\ell j}$ could be estimated via the pseudo-inverse, however because the basis functions are nonlinear, the other parameters must be estimated by some form of gradient-descent.

[2] Independent noise is unlikely for cells with common input (Dickson and Gerstein, 1974). However, non-independence could be addressed in the model by substituting the variance vector with the complete covariance matrix.

[3] The number of all possible combinations of size N from 65 is $\binom{65}{N}$, which can obviously be quite large. It is unnecessary to consider all possible combinations, so long as a random selection is obtained. The choice of 100 was arbitrary.

ACKNOWLEDGMENTS

This work was supported in part by NIH grants DC02804 and DC00116. I greatly appreciate access to the cortical recordings provided by John F. Brugge, Richard A. Reale, and Joseph E. Hind, as well as their participation in the modeling work.

REFERENCES

Arnold KJ (1941) On spherical probability distributions Unpublished Ph D Thesis Massachusetts Institute of Technology

Blahut RE (1987) Principles and practice of information theory Reading MA: Addison-Wesley.

Brugge JF, Reale RA, Hind JE, Chan JCK, Musicant AD, Poon PWF (1994 Simulation of free-field sound sources and its application to studies of cortical mechanisms of sound localization in the cat Hear Res 73: 67-84.

Brugge JF, Reale RA, Hind JE (1996) The structure of spatial receptive fields of neurons in primary auditory cortex of the cat . J Neurosci 16:4420-4437.

Colburn HS (1973) Theory of binaural interaction based on auditory-nerve data I General strategy and preliminary results on interaural discrimination. J Acoust Soc Am 54: 1458-1470.

Dabak A, Johnson DH (1992) Function-based modeling of binaural processing: Interaural phase. Hear Res 58:200-212

Daugman J G (1980) Two-dimensional spectral analysis of cortical receptive field profiles Vision Res 20:847-856.

Dickson JW, Gerstein GL (1974) Interactions between neurons in auditory cortex of the cat . J Neurophys 37:1239-1261.

Eisenman LM (1974) Neural coding of sound location: an electrophysiological study in auditory cortex using free field stimuli. Brain Res 75: 203-214.

Fisher RA (1925) Theory of statistical estimation. Proc Camb Phil Soc 22:700-725.

Fisher NI, Lewis T, Embleton BJJ (1987) Statistical Analysis of Spherical Data Cambridge:Cambridge University Press.

Hall P, Carroll RJ (1989) Variance function estimation in regression: the effect of estimating the mean. J R Statist Soc 51:3-14.

Heiligenberg W (1987) Central processing of sensory information in electric fish. J Comp Physiol A 161:621-631.

Hubel DH, Wiesel TN (1962) Receptive fields binocular interaction and functional architecture in the cat's visual cortex. J Physiol 160:106-154.

Hubel DH, Wiesel TN (1977) Functional architecture of macaque monkey visual cortex Proc R Soc London Ser B 198:819-847.

Imig TJ, Adrian HO (1977) Binaural columns in the primary field AI of cat auditory cortex. Brain Res 138: 241-257.

Imig TJ , Irons WA , Samson FR (1990) Single-unit selectivity to azimuthal direction and sound pressure of noise bursts in cat high-frequency primary auditory cortex. J Neurophys 63: 1448-1466.

Jenison R L (1995) von Mises basis function neural networks for spherical approximation Proceedings of the World Congress on Neural Networks. Hillsdale, NJ :Lawrence Erlbaum Associates 610-614.

Jenison RL, Fissell K (1995) A Comparison of Gaussian and von Mises Basis Functions for Approximating Spherical Acoustic Scatter. IEEE Trans Neural Networks 65:2123-2127.

Jenison RL (1996) A spherical basis function neural network for approximating acoustic scatter. J Acoust Soc Am 99:2001-2005.

Jenison RL, Fissell K (1996) A spherical basis function neural network for modeling auditory space. Neural Comp 8:115-128.

Jenison RL, Brugge JF, Reale RA (1997) Maximum-Likelihood estimation performance of cortical spatial receptive fields. Assoc Res Otol Abs 20: 198.

Johnson DH, Dabak A, Tsuchitani C (1990) Function-based modeling of binaural processing: Interaural level. Hear Res 49:301-320.

Knudsen EI, Konishi M (1978) A neural map of auditory space in the owl Science Wash DC 200:795-797.

Knudsen EI, Lac SD, Esterly S D (1987) Computational maps in the brain. Ann Rev Neurosci 10:41-65
Mardia KV (1972) Statistics of directional data. London:Academic Press.
Marcelja S (1980) Mathematical description of the responses of simple cortical cells. J Opt Soc Am 70:1297-1300.
Makous JC , Middlebrooks JC (1990) Two-dimensional sound localization by human listeners. J Acoust Soc Am 87:2188-2200.
Marr D (1982) Vision. W H San Francisco, CA:Freeman.
Martin RL, Webster WR (1987) The auditory spatial acuity of the domestic cat in the interaural horizontal and median vertical planes. Hear Res 30:239-252.
May BJ and Huang AY (1996) Sound orientation behavior in cats. I. Localization of broadband noise. J Acoust Soc Am 100:1059-1069.
McIlwain J T (1976) Large receptive fields and spatial transformation in the visual system In: International Review of Physiology Neurophysiology II (Porter R, ed), pp 223-248. Baltimore:University Park Press.
Middlebrooks JC, Clock AE, Xu L, Green DM (1994) A panoramic code for sound location by cortical neurons. Science 264:842-844.
Middlebrooks JC, Dykes RN, Merzenich MM (1980) Binaural response-specific bans in primary auditory cortex AI of the cat: topographical organization orthogonal to isofrequency contours. Brain Res 181: 31-48.
Middlebrooks JC, Pettigrew JD (1981) Functional classes of neurons in primary auditory cortex of the cat distinquished by sensitivity to sound location. J Neurosci 1: 107-113.
Moody J , Darken C J (1989) Fast learning in networks of locally-tuned processing units. Neural Comp 1:281-294.
Muller HG, Stadtmuller U (1987) Estimation of heteroscedasticity in regression analysis. Ann of Statist 15:610-625.
Musicant AD, Chan JCK, Hind JE (1990) Direction-dependent spectral properties of cat external ear: New data and cross-species comparisons. J Acoust Soc Am 87: 757-781.
Paradiso MA (1988) A theory for the use of visual orientation information which exploits the columnar structure of striate cortex. Biol Cybern 58:35-49.
Phillips DP, Hall SE (1990) Response timing constraints on the cortical representation of sound time structure. J Acoust Soc Am 88:1403-1410.
Powell MJD (1987) Radial basis functions for multivariable interpolation: A review In : radial basis functions for multivariable interpolation: A review (Mason J C, Cox MG eds)), pp 143-166. Oxford:Clarendon Press.
Rajan R, Aitkin LM, Irvine DRF (1990) Azimuthal sensitivity of neurons in primary auditory cortex of cats. I. Types of sensitivity and the effects of variations in stimulus parameters. J Neurophys 64: 872-887.
Rajan R, Aitkin LM, Irvine DRF (1990) Azimuthal sensitivity of neurons in primary auditory cortex of cats. II. Organization along frequency-band strips. J Neurophys 64: 888
Rodieck, R W (1965) Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vision Res. 5: 585-601. -902.
Tolhurst DJ, Movshon JA, Dean AF (1983) The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res 23:775-785.
Wahba G (1981) Spline interpolation and smoothing on the sphere. SIAM J Sci Stat Comput 2:5-16.
Zohary E (1992) Population coding of visual stimuli by cortical neurons tuned to more than one dimension. Biol Cybern 66:265-272.