Theoretical Considerations for the Analysis of Population Coding in Motor Cortex

Terence D. Sanger

MIT, E25-534, Cambridge, MA 02139 (tds@ai.mit.edu)

Key words: Motor cortex; Population Coding; Population Vector; Distributed Coding; Locally Tuned Cells; Cosine Tuning; Reaching Movements

Recent evidence of population coding in motor cortex has led some researchers to claim that certain variables such as hand direction or force may be coded within a Cartesian coordinate system with respect to extrapersonal space. These claims are based on the ability to predict the rectangular coordinates of hand movement direction using a "population vector" computed from multiple cells' firing rates. I show here that such a population vector can always be found given a very general set of assumptions. Therefore the existence of a population vector constitutes only weak support for the explicit use of a particular coordinate representation by motor cortex.

1 Introduction

Recent results suggest that the representation of arm movement in motor cortex involves the simultaneous activity of many cells, and that the pattern of activation over the group of cells specifies the motion which occurs (Caminiti *et al.* 1990, Kalaska and Crammond 1992, for review). These results have led many researchers to ask whether movement variables are coded internally in terms of a particular coordinate system such as a Cartesian or polar representation of extrapersonal space, a representation of muscle lengths around relevant joints, or some other set of coordinates. In the following, I distinguish between a coded variable (such as hand position) and the coordinates used to represent that variable (such as Cartesian coordinates). I will summarize certain experiments which demonstrate that hand movement direction is represented within motor cortex, but I claim that these experiments cannot be used to determine the coordinate system in which movements are coded.

I discuss a set of experiments that investigated the relationship between cell firing rates during free arm movements in awake monkeys, and the direction in which the hand was moved to a target in space (Georgopoulos *et al.* 1988, Kettner *et al.* 1988, Schwartz *et al.* 1988). These experiments led to the following results:

R1: The firing rate of 89.1% (486/568) of the cells tested in motor cortex varied consistently with the direction of hand motion within a limited region of space, and many cells would be simultaneously active for any given direction.

R2: A statistically significant component of the variance of the firing rate of 83.6% of the cells could be accounted for by a broadly-tuned function of the form

$$d_i(M) \approx b_i + k_i \cos(\theta_i - \theta_M) \tag{1}$$

where $d_i(M)$ is the firing rate of cell i for hand motion in the direction of a unit vector M, θ_i is the direction of motion in which the cell has maximal response, $\theta_i - \theta_M$ is the angle[1] between the direction of hand motion θ_M and the cell's preferred direction, and b_i and k_i determine the average firing rate and modulation depth, respectively. (Here and in the following, capital letters indicate vector quantities.)

R3: The preferred directions θ_i are approximately uniformly distributed with respect to directions in the workspace.

R4: The hand direction vector M can be approximated in Cartesian coordinates by a population vector P computed from a linear combination of the cell firing rates.

R5: The coefficients of this linear combination are given by unit vectors C_i along the

[1] A difference of 3D angles is defined by $\theta_i - \theta_M = \cos^{-1}(C_i \cdot M)$, where C_i and M are both unit vectors.

preferred direction θ_i for each cell, so that

$$M \approx P = \sum_{i=1}^{N} C_i d_i \tag{2}$$

where the d_i have been normalized to account for resting firing rate and response amplitude, and both the movement direction M and the preferred direction vectors C_i are given in Cartesian coordinates with respect to the external workspace.

Together, these results might suggest that a Cartesian representation of the direction of hand motion is coded in motor cortex (Schwartz *et al.* 1988).

I will show that results R2, R4, and R5 are direct consequences of results R1, R3, and the experimental design. This in no way reduces the importance of these experiments, but rather emphasizes the fact that results R1 and R3 contain the most significant information. Although their importance was recognized in (Georgopoulos *et al.* 1988), the fact that they imply the other results was not. Previous investigations have studied the conditions under which results R4 and R5 hold, and it has been shown that the population vector predicts the direction of hand motion if the tuning curve is symmetric and the distribution of preferred directions is uniform (Georgopoulos *et al.* 1988) or has no second harmonic components (Mussa-Ivaldi 1988). I derive a necessary and sufficient condition which is even broader, since it requires only that the three components of the preferred directions be uncorrelated with each other over the population. Before doing this, I will first show that the cosine tuning curves found in (Schwartz *et al.* 1988) may be an artifact of the analytic techniques used.

2 Single Unit Tuning Curves

In (Schwartz *et al.* 1988), the firing rate of each tuned cell $d_i(M)$ is approximated by a linear combination of the normalized Cartesian coordinates of the target toward which the monkey is reaching. These coordinates are relative to the initial hand position X_0 and are given by a unit vector in the direction of motion $M = (m_x, m_y, m_z)$. The linear approximation is

$$d_i(M) \approx b_i + b_{ix} m_x + b_{iy} m_y + b_{iz} m_z \tag{3}$$

and an F test showed that the variance of 83.6% of all cells was at least partly accounted for by this linear regression. The preferred direction vector C_i is calculated from

$$k_i = (b_{ix}^2 + b_{iy}^2 + b_{iz}^2)^{1/2} \tag{4}$$

$$C_i = < b_{ix}/k_i, b_{iy}/k_i, b_{iz}/k_i > \tag{5}$$

and we can now write

$$d_i(M) \approx b_i + k_i C_i \cdot M \tag{6}$$

which is equivalent to equation 1. Note that cells with $k_i = 0$ are not sensitive to the direction of movement and were not analyzed further, so $k_i \neq 0$.

To understand results R1 and R2, I perform a simplified analysis of movement in 2 dimensions (the extension to 3 dimensions is straightforward but complicates the notation significantly). For a fixed initial hand position and with all other variables held constant, consider any arbitrary firing rate function $d(\theta_M)$ which depends on the direction of hand movement θ_M. θ_M is a periodic variable, so the output of $d(\theta_M)$ will be periodic, and if 8 uniformly spaced directions are tested then the complete behavior can be described by a discrete Fourier series with periods up to $\pi/2$

$$d(\theta_M) = \sum_{k=0}^{4} \alpha_k \cos(k\theta_M + \phi_k) \tag{7}$$

where ϕ_k is the phase for each angular frequency component k. Note that for $k > 1$ the terms have no directional component, since they consist of either 2, 3, or 4 "lobes" symmetrically placed around the circle. Thus a linear regression on the Cartesian coordinates $x = cos(\theta_M), y = sin(\theta_M)$ will be unaffected by the values of α_2, α_3, and α_4 and will depend only on α_0 and α_1. To see this, note that linear regression computes the three expected values:

$$\begin{aligned} E[d(\theta_M)] &= E\left[\sum_{k=0}^{4} \alpha_k \cos(k\theta_M + \phi_k)\right] \\ &= \alpha_0 \end{aligned}$$

$$E[d(\theta_M)x] = E\left[\sum_{k=0}^{4} \alpha_k \cos(k\theta_M + \phi_k)\cos(\theta_M)\right]$$

$$
\begin{aligned}
&= E[\alpha_1 \cos(\theta_M + \phi_1)\cos(\theta_M)] \\
&= \left(\frac{\alpha_1}{2}\right)\cos(\phi_1)
\end{aligned}
$$

$$
\begin{aligned}
E[d(\theta_M)y] &= E\left[\sum_{k=0}^{4} \alpha_k \cos(k\theta_M + \phi_k)\sin(\theta_M)\right] \\
&= E[\alpha_1 \cos(\theta_M + \phi_1)\sin(\theta_M)] \\
&= \left(\frac{\alpha_1}{2}\right)\sin(\phi_1)
\end{aligned}
$$

where the expectation operator $E[]$ is taken over all tested directions θ_M. The preferred direction is therefore equal to ϕ_1 and is independent of ϕ_2, ϕ_3, or ϕ_4. Even if more than 8 directions are tested, the linear regression will respond only to the ϕ_1 component.

The "goodness of fit" to the linear regression is the extent to which the $k = 0$ and $k = 1$ terms capture the behavior of $d(\theta_M)$. However, it is important to realize that a statistically significant F test does not indicate a good fit to a linear model in the sense of having small prediction error variance. Fit is determined by mean squared error which distributes according to a χ^2 statistic. The F test only estimates the probability that the linear model accounts for some portion of the total variance. This is equivalent to testing if α_1 is significantly different from 0. A significant F test does not imply that α_1 describes the dominant response behavior, and α_2, α_3, or α_4 might well be larger. If a set of tuning curves were generated randomly by selecting the coefficients α_k independently from a normal distribution, then one would expect 95% of the tuning curves to have statistically significant values of α_1. Thus the observed value of 83.6% (93% in (Caminiti *et al.* 1990)) does not support statistical arguments that the population has been "engineered" to have directional tuning.

Since this method of analysis ignores terms for $k > 1$, it in effect low-pass filters the tuning curves. So the cosine tuning results from the method of analysis and may not be justified by the original data. These considerations show that result R2 does not provide any information beyond result R1, since R2 would be true for a randomly chosen set of tuning curves satisfying R1 which were analyzed in this way.

True cosine tuning could be verified by fitting equation 1 to data samples from many different directions and measuring the average mean-squared approximation error over the population using a χ^2 statistic. A similar test was done in the 2-dimensional case, where it was found that 75% of 241 cells had a normalized mean-squared approximation error less than 30% of the total variance(Georgopoulos *et al.* 1982). Although this is not a statistically good fit to the population, there may have been individual cells whose response was well predicted by cosine tuning.

What is the significance of the cells which were well fit by a cosine tuning curve? As shown in equation 6, these cells have a response d which is approximately linearly related to the hand movement vector M. We can thus claim either that these cells are in fact linear in the movement direction, or else that they are linear in the testing region but may be nonlinear if tested in other regions of space. So if we write the response as $d(X_0, X)$ where X_0 is the initial hand position and X is the target, then we know that $d(X_0, X)$ must be sufficiently smooth that it appears locally linear for the positions X which were tested.

Over larger distances, d may not be well approximated linearly, but it can still be written as

$$d(M) \approx b + kC(X_0) \cdot M \tag{8}$$

where $C(X_0)$ emphasizes that the preferred direction may become dependent on the initial position, as was indeed found in (Caminiti *et al.* 1990). But equation 8 is a general representation for arbitrary smooth functions, so even an accurate fit to a locally linear function does not allow one to claim much beyond the fact that the preferred direction remains approximately constant over the tested region.

3 Population Vectors

Result R4 that there exists a linear combination of the firing rates which can predict the Cartesian coordinates of hand motion follows as a direct consequence of well-known results on coarse coding and the theory of radial basis functions (Poggio and Girosi 1990, for example), since a raised cosine function of angle can be thought of as a local basis function centered on the preferred direction. An alternate way to prove this fact follows. Define an $N \times 3$ matrix Q whose rows are the preferred direction vectors C_i. Let D be an N-dimensional column vector formed from the firing rates of all the cells by the formula

$[D]_i = (d_i - b_i)/k_i$ as in (Georgopoulos *et al.* 1988). Then from equation 6 we can write

$$D \approx QM \tag{9}$$

We seek a 3xN weighting matrix H such that a population vector of the form HD predicts hand direction M according to

$$M \approx HD \approx HQM \tag{10}$$

There are many matrices H which will satisfy this equation. One possibility is to use linear least-squares regression, giving

$$H = (Q^T Q)^{-1} Q^T \tag{11}$$

where the inverse $(Q^T Q)^{-1}$ will always exist so long as there are three linearly independent preferred direction vectors. We now have

$$M \approx HD \approx HQM = (Q^T Q)^{-1} Q^T QM = M \tag{12}$$

as desired. This equation means that so long as there exist three linearly independent direction vectors, the hand direction will be approximately linearly related to the cell firing rates d_i in any coordinate system for M which satisfies equation 9. So far I have shown that result R1 implies both results R2 and R4, given the method of analysis.

In (Georgopoulos *et al.* 1988) the columns of H were not found by performing a regression of the cell firing rates against the hand direction according to equation 11, but instead were assumed *a priori* to be equal to the preferred direction C_i for each cell, so that $H = Q^T$. I now discuss under what conditions result R5 holds, so that this particular linear combination will give the right answer. The population vector is given by equation 2 which we can rewrite in vector notation as

$$M \approx P = Q^T D \approx Q^T QM \tag{13}$$

and if this holds for all directions M then we must have $Q^T Q = I$. This is a necessary condition for the existence of a population vector. In (Georgopoulos *et al.* 1988) a more restrictive sufficient condition satisfying equation 2 is that the distribution of preferred directions is uniform over the sphere. Another necessary and sufficient condition based on Fourier analysis of the distribution of preferred directions for the planar case is given in (Mussa-Ivaldi 1988).

To understand the meaning of equation 13, we can write each component of Q^TQ as

$$[Q^TQ]_{jk} = \sum_{i=1}^{N}[C_i]_j[C_i]_k \tag{14}$$

and $I = Q^TQ$ implies that $\sum_{i=1}^{N}[C_i]_j[C_i]_k = 0$ whenever $j \neq k$. This expression is the correlation of the j^{th} and k^{th} components of the preferred direction vectors C_i, so a necessary and sufficient condition for equation 2 to work is that the x, y, and z components of these vectors are uncorrelated and have equal variance. The result that equation 2 is satisfied is thus implied by the approximately uniform distribution of cell preferred directions in result R3. Note that for other coordinate systems, even if the components of the C_i's are correlated there will still exist a linear combination $H \neq Q^T$ of the firing rates which will predict the desired values, although the matrix H may need to be found by regression using equation 11. But if both results R1 and R3 hold, then result R5 must hold.

Suppose that rather than using the predicted value QM we use the true measured value D and this includes significant non-cosine (nonlinear) terms. Then we have

$$D = QM + E$$

where E is a vector with components

$$e_i(\theta_M) = \sum_{k=2}^{4} \alpha_{ik}\cos(k\theta_M + \phi_{ik})$$

If the terms α_{ik} and ϕ_{ik} are distributed independently of the components of C_i, then $Q^TE = 0$ and these terms will not affect the value of the population vector. So even if the individual cells do not have cosine tuning, the population vector will correctly predict hand direction if the terms for $k > 1$ do not correlate with the terms for $k = 1$ in the expansion given in equation 7.

If the experiments are repeated with differing initial positions as in (Caminiti *et al.* 1990), then the preferred directions C_i may change. This will lead to a new matrix Q' so that $D' = Q'M$. Population vector analysis under the new conditions will give $P' = Q'^TQ'M$, so again the requirement for success is that the components of the new preferred directions are uncorrelated. The fact that population vectors "proved to be good predictors of movement direction regardless of where in space the movements were performed" (Caminiti *et al.* 1990, page 2039) provides no information beyond the knowledge that the components of the preferred directions remain uncorrelated as the initial hand position changes.

4 Coordinate-free Representations

One might ask if the experiments described above could be modified to determine the "true" coordinate system used by motor cortex to describe hand movement direction. However, I claim that for certain classes of distributed representation this is not a well-defined question. Distributed representations of measured variables can be coordinate-free in the sense that they do not imply any particular coordinate system. To see this, let X be any variable represented in cortex (such as hand movement direction), and let $D(X)$ be a vector-valued function representing the outputs of a large set of basis functions $d_i(X)$ which describe the behavior of (motor) cortical cells. $D(X)$ is then a distributed representation of the variable X. Now, consider a vector function $T(X)$ which measures X in a particular coordinate system ($T(X)$ might give the 3 Cartesian components of hand movement direction, for example). If there exists a matrix H such that $HD(X) \approx T(X)$, then one can say that the distributed representation D codes the coordinate system T. Yet this will hold for any $T(X)$ which is close to the linear span of the basis functions $d_i(X)$, so we cannot claim that D encodes any single coordinate system for X within this span better than another.

5 Conclusion

In this letter, I have extended the generality of previous results (Georgopoulos *et al.* 1988, Mussa-Ivaldi 1988) to show that cosine tuning curves will be found for large classes of arbitrary response functions if they are analyzed according to the statistical techniques in (Schwartz *et al.* 1988), and that the existence of a population vector as found in (Georgopoulos *et al.* 1988) is determined by very general necessary and sufficient conditions which depend only on the distribution of preferred directions rather than on any intrinsically coded coordinate system. The concept that a distributed representation codes a particular coordinate system may not be well-defined, since certain types of representations can be considered "coordinate-free". These considerations imply that experiments of the type described may yield population vectors which predict many different 3-dimensional coordinates (such as Cartesian, polar, muscle lengths, or joint angles).

It is important to understand that the considerations presented here in no way reduce the importance of the results reported in (Schwartz *et al.* 1988, Georgopoulos *et al.* 1988, Kettner *et al.* 1988, Caminiti *et al.* 1990) and elsewhere. The fact that results R2, R4, and R5 are direct consequences of R1 and R3 only serves to underscore the significance of these two results. They show that large populations of motor cortical cells respond to hand motion in a predictable way, and that the preferred directions are approximately uniformly distributed with respect to a Cartesian representation of extrapersonal space. No additional conclusions can be drawn from the population vector, since its existence is a mathematical consequence of these two facts. However, if the distribution of preferred directions is nonuniform with respect to other coordinate systems or if the distribution can be modified through experience, then this would provide significant information about cortical representations. In addition, if cosine tuning can be verified by explicitly fitting cell tuning curves to a linear regression model, then further studies may discover constraints which explain why more than 486 linear cells are needed to code for only 3 linearly independent components of hand direction.

Acknowledgements

I would like to thank Sandro Mussa-Ivaldi, Ted Milner, Emilio Bizzi, Richard Lippmann, Marc Raibert, and the reviewers for their comments and criticism. This report describes research done within the laboratory of Dr. Emilio Bizzi in the department of Brain and Cognitive Sciences at MIT. The author was supported during this work by the division of Health Sciences and Technology, and by NIH grants 5R37AR26710 and 5R01NS09343 to Dr. Bizzi.

References

Caminiti R., Johnson P. B., Urbano A., 1990, Making arm movements within different parts of space: Dynamic aspects in the primate motor cortex, *J. Neuroscience*, 10(7):2039–2058.

Georgopoulos A. P., Kalaska J. F., Caminiti R., Massey J. T., 1982, On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex, *J. Neurosci.*, 2(11):1527–1537.

Georgopoulos A. P., Kettner R. E., Schwartz A. B., 1988, Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population, *J. Neurosci*, 8(8):2928–2937.

Kalaska J. F., Crammond D. J., 1992, Cerebral cortical mechanisms of reaching movements, *Science*, 255:1517–1523.

Kettner R. E., Schwartz A. B., Georgopoulos A. P., 1988, Primate motor cortex and free arm movements to visual targets in three-dimensional space. III. Positional gradients and population coding of movement direction from various movement origins, *J. Neurosci.*, 8(8):2938–2947.

Mussa-Ivaldi F. A., 1988, Do neurons in the motor cortex encode movement direction? An alternative hypothesis, *Neurosci. Lett.*, 91:106–111.

Poggio T., Girosi F., 1990, Regularization algorithms for learning that are equivalent to multilayer networks, *Science*, 247:978–982.

Schwartz A. B., Kettner R. E., Georgopoulos A. P., 1988, Primate motor cortex and free arm movements to visual targets in three-dimensional space. I. Relations between single cell discharge and direction of movement, *J. Neurosci*, 8(8):2913–2927.