A Continuous Information Theoretic Approach to the Analysis of Cutaneous Receptor Neurons*

Marc S. Fuller[1] and William J. Williams[2]

1 Dept. of Electrical Engineering, Worcester Polytechnic Institute, Worcester, USA
2 Dept. of Electrical Engineering, University of Michigan, Ann Arbor, USA

Abstract. Information theoretic and statistical techniques for determining the number of discernible levels in cutaneous receptor neurons are reviewed. Reasons for the large variance in these results are discussed. A new continuous information theoretic analysis technique is presented that overcomes many of the problems in the other methods of analysis discussed. Comparison of this new method of analysis with a statistical technique developed by Schreiner et al. (1978) clearly shows some of the misconceptions that are associated with statistical analysis techniques, and why these problems cannot arise in the new information theoretic technique discussed here.

1. Introduction

The introduction of information theory by Shannon (1948) has led to many advances in communication theory. Despite the success of information theory in the technological realm, its early perceived promise in evaluating the information processing capabilities of biological systems has not been realized. Some attempts to apply information theory to biological systems have proved interesting, but conflicting in their conclusions. For example, estimates on the information carrying capacity of a single neuron vary from a few bits/s (Jacobson, 1950) to thousands of bits/s (Stein, 1967). The wide variation in results is, in part, a result of the biological variable chosen as the carrier of the information. Although there is some agreement when the input and output of the system is settled upon, however, confusion still exists concerning the results and interpretations of many experiments performed to shed light on the information carrying

capacity of a neural system. In particular the application of discrete information theory to continuous systems, must be suspect. Yet this is the method used by most investigators (Werner and Mountcastle, 1965; Kenton and Kruger, 1971; Walloe, 1970).

More recently investigators have tried using methods other than information theory to predict information transfer in neurons. Schreiner et al. (1978) have recently examined the number of levels resolvable by means of a mean rate code for somatosensory cortical neurons. They used statistical, rather than an information theoretic approach in their study. As we shall see this may lead to even more confusion. Although the information theoretic approach has not been as useful as had been hoped in evaluating the information carrying capacity of neural systems, it does, if properly used, place upper limits on the amount of information transfer in certain well defined situations such as outlined by Schreiner et al. (1978).

We present in this paper a method of determining the upper limit on information transfer for the system outlined by Schreiner et al. (1978) and contrast the results obtained with the results obtained using the heuristic approach used by Schreiner et al. (1978). Secondly we will show why this new technique does not have the pitfalls of those techniques used in other information theoretic investigations of the neural system. This will at least give one a bench mark for the validity of other analyses.

2. Statistical Analysis

Several investigators (Werner and Mountcastle, 1969; Walloe, 1970; Schreiner et al., 1978) have noted that the distribution of a tonic receptor neuron's interval histogram is gaussian with mean rate, μ, related to standard deviation, σ, as
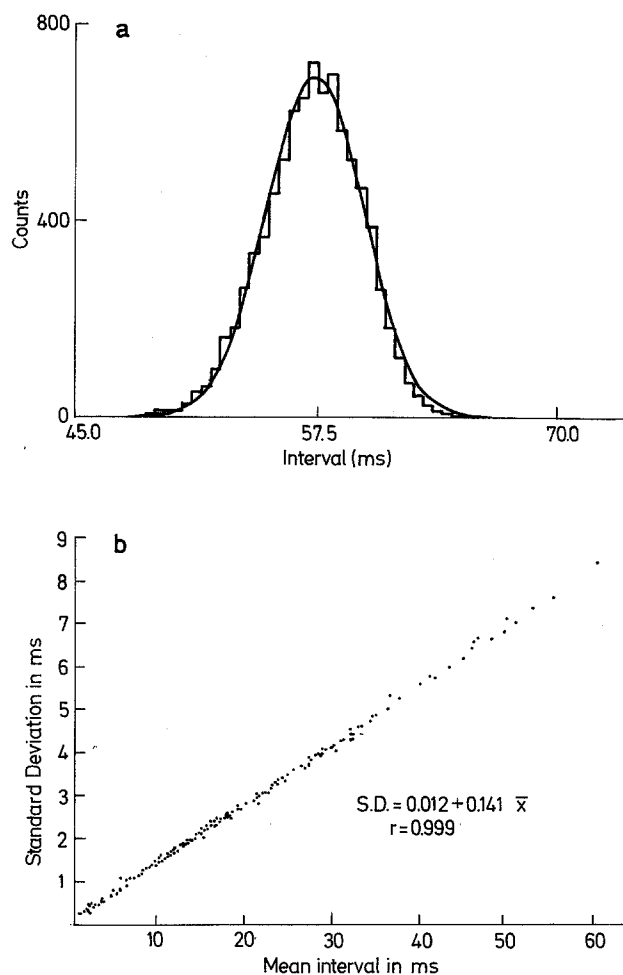
$$\sigma = A\mu + B, \tag{1}$$

Fig. 1. a Interval histogram of a tonic joint receptor fit to a gaussian curve. The gaussian curve's mean and variance are the same that of the histogram. These data are from a cat muscle spindle. (Unpublished data from McCall.) **b** The linear relation between mean and standard deviation of a tonic unit. This plot is based on data from a Type II unit (Werner and Mountcastle, 1965)

Table 1. Comparison of Schreiner et al.'s statistical results and the information theoretic results described here. Note that the number of levels predicted by Schreiner et al. are generally higher than those predicted by the information theoretic technique. The division of stimulus levels is arbitrary and done only to make comparison with the statistical technique possible

Unit slope (A), intercept (B)	High stimulus levels		Low stimulus levels	
	Schreiner	Information theoretic	Schreiner	Information theoretic
86-5-A A = 1.57 B = −8.74	2	2.13	1	1.27
74-2-D A = 0.65 B = −3.98	7	3.34	5	1.69
40-4-A A = 0.71 B = −0.25	3	1.78	3	1.39
40-2-A A = 0.55 B = 0.21	4	2.01	4	1.56
74-1-C A = 0.51 B = 2.60	3	1.67	4	1.47
90-1-A A = 0.34 B = 0.34	4	1.97	5	1.77

where A and B are experimentally determined for each neuron. Figure 1 shows an example of this response for a particular neuron. Given this model one may easily use a variety of techniques to analyze the information content of the neural pattern.

Schreiner et al. (1978) have performed a statistical analysis on neurons that follow the model described in the previous section. Their technique involves using Cochran and Cox's modified t-test to determine when two firing rates can be distinguished. Table 1 shows data compiled by Schreiner et al. (1978) from 6 neurons and the predictions they made on the number of distinguishable intensity levels. Also in Table 1 are some information theoretic predictions that will be

discussed later. Note that the interval ranges were arbitrarily divided in half by Schreiner et al. (1978) and that we will continue to use this arbitrary division for comparison purposes only. It appears that the total number of intervals calculated using the statistical technique gives one a measure of the neurons sensitivity. However, we shall show that this type of statistical analysis may be deceiving, and in this particular case overly generous in estimating the number of distinct transmission levels.

3. A Continuous Information Theoretic Model

Information theory is an alternate method that can be used to predict the number of distinguishable levels in a stochastic system as mentioned in the introduction. By applying information theory to a given stimulus-response pairs, joint probability distributions can be used to predict the maximum number of distinguishable levels if one decodes or interprets the data correctly. The application of information theory to neural systems is certainly not new. We have mentioned a few of the many investigators that have

applied it to similar experimental data. These investigators all attempted to use discrete information theory applied to stimulus-response distributions that are clearly continuous. These techniques arbitrarily divide both stimuli and responses into arbitrary bins, and then calculate information transfer. These bins are then subdivided and information transfer recalculated. If the transinformation appears to remain constant as the subdivisions further divide the data, then the calculations are assumed to be valid. However, the results of this technique may be incorrect for two reasons. First, information transfer would, at best, increase logarithmically with the number of subdivisions, so it becomes difficult to determine if the changes in transinformation are real or simply caused by statistical variations in the data. Secondly, there is no guarantee that there is not one small area within a single bin that is "rich" in information while the transinformation appears to be constant as subdivisions occur. The obvious solution to these problems is to use a continuous information theoretic analysis. The strong point of Schreiner et al.'s (1978) analysis is that it does use a continuous model for the neural system. In this section we shall use the same model and develop a continuous information theoretic analysis to predict the number of levels. For a given stimulus-response pair the transinformation is defined as

$$I(S, R) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(s, r) \log_2 [p(s, r)/(p(s)p(r))] dr ds. \quad (2)$$

Using the neural model given above, we note that if a given stimulus cause neural activity with mean rate s, the actual distribution r, given s is

$$p(r|s) = (2\pi\sigma_r^2)^{-1/2} \cdot \exp[-(r-s)/(2\sigma_r)], \quad (3)$$

where $\sigma_r = As + B$ as assumed. Now the transinformation between S and R is

$$I(S, R) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(r|s) \cdot p(s) \log[p(r|s)/p(r)] dr ds, \quad (4)$$

where $p(r|s)$ is defined in (3) and the distribution of $p(s)$ is determined by the stimulus probability distribution and the relation between stimulus intensity and mean rate. This relation may vary depending upon the particular sensory receptor under observation. For the sake of example, let us assume that $p(s)$ is uniformly distributed from some s_{min} to s_{max}. This will allow us to be consistent with the model used by Schreiner et al. (1978), and in any case the stimulus distribution may be arbitrarily altered to make this assumption true. In

this case (4) may be restated as:

$$\int_{s_{min}}^{s_{max}} \int_{-\infty}^{\infty} (p(r|s)/s_{max} - s_{min}) \cdot \log[p(r|s)/p(r)] dr ds, \quad (5)$$

where

$$p(r) = \int_{s_{min}}^{s_{max}} p(r|s)/(s_{max} - s_{min}) ds. \quad (6)$$

Using the Simpson method of integration on a Hewlett Packard 21 MX computer we have calculated the transinformation, in bits, for the same 6 units that were analyzed by Schreiner et al. (1978). The number of distinguishable levels was then calculated by raising 2 to the number of bits calculated. Table 1 compares the results of this method with the statistical technique just described. Note that the statistical predictions are generally higher than ours. This result is disturbing because the information theoretic calculation are predictions of the best a communication system can do with ideal decoding. This discrepancy can be resolved if one recalls that error in transmission is allowed in the statistical analysis. First note that Schreiner et al. (1978) chose an alpha of 0.2, so there is a 20% probability of a Type I statistical error. Secondly, they have ignored the contribution of Type II statistical errors. In this case, because of the symmetry of the problem, the probability of Type I and Type II errors is equal. Therefore, the total probability of error for any received interval is 40%. Schreiner et al. (1978) do not make it apparent that each level has only a 60% chance of being correctly received. Once this is clear, it is certainly not surprising that the number of levels they predicted is much larger than those predicted by the information theoretic technique presented here.

4. Conclusions

One can now see why an information theoretic technique, if properly applied, is attractive. It provides an absolute upper limit on the amount of information that can be transmitted by a system with a given set of statistical distributions. In contrast, statistical techniques leave one with an uncertainty about the true meaning of the results. The choice of confidence intervals is arbitrary and every choice leads to different results.

References

Jacobson, H.: The information capacity of the human ear. Science 112, 243 (1950)

Kenton, B., Kruger, L.: Information transmission in slowly adapting mechanoreceptive fibers. Exp. Neurol. **31**, 114–139 (1971)

Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. **63**, 81–97 (1956)

Schreiner, R.C., Eggick, G.K., Whitsel, B.L.: Variability in somatosensory cortical neuron discharge: effects on capacity to signal different stimulus conditions using a mean rate code. J. Neurophysiol. **41**, 338–349 (1978)

Shannon, C.E.: A mathematical theory of communication. Bell Syst. Techn. J. **27**, 379–423 (1948)

Stein, R.B.: The information capacity of nerve cells using a frequency code. Biophys. J. **7**, 797–826 (1967)

Walloe, L.: On the transmission of information through sensory neurons. Biophys. J. **8**, 745–763 (1970)

Werner, G., Mountcastle, V.B.: Neural activity in mechanoreceptive cutaneous afferents: stimulus response relations, Weber functions and information transmission. J. Neurophysiol. **28**, 358–397 (1965)

Prof. M. S. Fuller
Dept. of Electrical Engineering
Worcester Polytechnic Inst.
Worcester, MA 01609
USA