

Fisher Information and Optimal Odor Sensors

Manuel A. Sánchez-Montañés ^a, Tim C. Pearce ^b

^a*E.T.S. de Informática, Universidad Autónoma de Madrid, Madrid 28049, Spain*

^b*Department of Engineering, University of Leicester, Leicester LE1 7RH, United Kingdom*

Abstract

We discuss how the Fisher Information Matrix (FIM) may be used as a part of an optimization procedure for selecting odor sensors within a population so as to maximize the accuracy with which the overall sensory system may estimate the stimulus. While the same approach may be equally applied to any sensory system that exploits a population coding of the stimulus in order to optimize its performance, we demonstrate how this technique may be used to optimize the performance of an artificial olfactory system in two simple examples.

Key words: Optimization of sensory arrays, Noise, Fisher Information, Chemical sensors, Odour

1 Introduction

In this work we address the optimization of an array of noisy odor sensors, in order to make as accurate estimate of the real odor exposed to the system as possible. The techniques described may be equally applied to the analysis of biological and artificial olfactory systems. In our system there is an array of odor sensors, each sensor having some parameters that may be tuned; these parameters fix the response characteristics of the receptor to different odors, that is, define its tuning curve to the stimulus. By using an array of such sensors, each with different tuning curves, it is possible to implement a sensory system that can appreciate a wide range of stimuli with relatively few sensors. Depending upon the tuning curves of the individual sensor elements, the accuracy of the overall sensory system in estimating the stimulus varies in addition

Email addresses: Manuel.Sanchez-Montanes@ii.uam.es (Manuel A. Sánchez-Montañés), t.c.pearce@leicester.ac.uk (Tim C. Pearce).

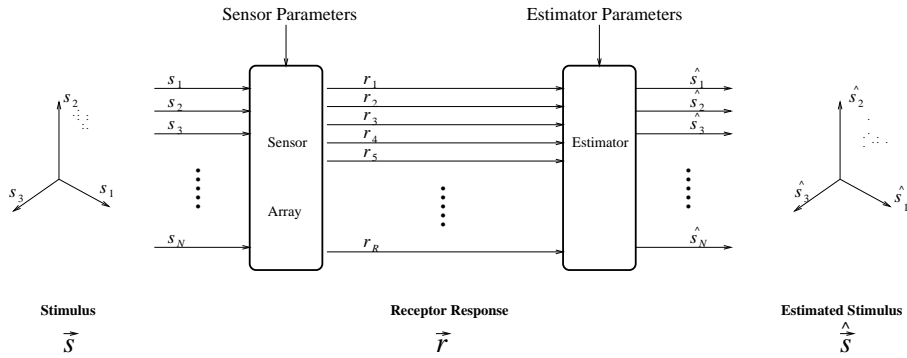


Fig. 1. A hypothetical statistical estimator takes the response vector from a sensor array or receptor population and uses this for estimating the stimulus. The tuning curves for each of the sensors are represented as parameters to the sensor array.

to the range of stimuli that may be appreciated. The usual goal in choosing which sensors to incorporate into an artificial olfactory system is to maximize the accuracy with which the sensory system can estimate the stimulus or optimally discriminate between neighboring stimuli (5). By considering a hypothetical unbiased statistical estimator that uses the receptor responses in order to estimate the individual stimuli within a complex odor mixture, we can test how different tuning curves of the sensor array can effect the accuracy of stimulus estimation. This arrangement is shown in figure 1 where each sensor, k , generates a response, r_k , to the stimulus \vec{s} . In order to optimize the estimation accuracy for a given estimator, we need to select the optimal tuning curve for each sensor, shown in figure 1 by adjusting the sensor parameters. As we will see in the following section, there is a lower bound to the estimation error defined by the Fisher Information Matrix, which only depends on the sensor parameters. The goal is then to find the set of sensor parameters for the array such that this bound is minimized.

2 Fisher Information Matrix and the best estimator

When a multi-component odor stimulus, \vec{s} , is exposed to the input array, the receptor k gives a response r_k following some probability distribution $p(r_k|\vec{s})$. That is, the response of each receptor to the stimulus is noisy.

The entries of the Fisher Information Matrix (FIM), $J_{ij}(\vec{s})$, are defined as (3):

$$J_{ij}(\vec{s}) = \int d\vec{r} p(\vec{r}|\vec{s}) \left(\frac{\partial}{\partial s_i} \ln p(\vec{r}|\vec{s}) \right) \left(\frac{\partial}{\partial s_j} \ln p(\vec{r}|\vec{s}) \right) \quad (1)$$

Then for every unbiased estimator that uses the data \vec{r} for estimating the stimulus \vec{s} :

$$\text{var}(\hat{s}_i|\vec{s}) \geq (J^{-1}(\vec{s}))_{ii} \quad (2)$$

where “var” means variance, and \hat{s}_i is the estimation of the component i of \vec{s} , $i = 1, \dots, N$. This result is called the “Cramér-Rao bound” (3) and limits the performance of the best unbiased estimator we can build.

The expected value of the quadratic error of the estimator is equal to the summation of the expected value of the quadratic error of the estimation of each s_i . Then, using eq. 2, we obtain:

$$\text{var}(\hat{\vec{s}}|\vec{s}) = \sum_{i=1}^N \text{var}(\hat{s}_i|\vec{s}) \geq \sum_{i=1}^N (J^{-1}(\vec{s}))_{ii} \quad (3)$$

so the performance of the best estimator we can build is defined by the entries of the FIM J_{ij} . But we know from (1) that the entries of the FIM depend on the probability density of \vec{r} given \vec{s} . This distribution depends on the values of the parameters of each sensor, including its tuning, so the idea is to select the set of sensor parameters that minimize the right-hand side of (3) so that the optimal unbiased estimator we can build has minimal quadratic error.

3 Fisher Information Matrix for an array of sensors with independent noise

The entries of the FIM for an individual sensor are given by:

$$J_{ij}^k(\vec{s}) = \int dr_k p(r_k|\vec{s}) \left(\frac{\partial}{\partial s_i} \ln p(r_k|\vec{s}) \right) \left(\frac{\partial}{\partial s_j} \ln p(r_k|\vec{s}) \right) \quad (4)$$

It can be easily shown that when the array of sensors has uncorrelated noise the FIM of the array is equal to the summation of the individual matrix for each receptor. This is valid in a general sense, that is, the noise and the tuning curves of the receptors can be different across the array.

We must now calculate the FIM for three important classes of noise by substituting the appropriate probability density function into eq. 4 and rearranging.

Case 1: Analog sensor with Gaussian noise:

$$J_{ij}^k(\vec{s}) = \frac{1}{\sigma_k^2(\vec{s})} \frac{\partial f_k(\vec{s})}{\partial s_i} \frac{\partial f_k(\vec{s})}{\partial s_j} + 2 \frac{1}{\sigma_k^2(\vec{s})} \frac{\partial \sigma_k(\vec{s})}{\partial s_i} \frac{\partial \sigma_k(\vec{s})}{\partial s_j} \quad (5)$$

where $f_k(\vec{s})$ defines the mean value of the sensor output given \vec{s} , and $\sigma_k(\vec{s})$ is the standard deviation of the Gaussian noise for that sensor. Note that, in principle, the noise dispersion can depend on the stimulus. This is an important general case that may be assumed for measurements taken from metal-oxide semiconductor and conducting polymer chemosensors.

Case 2: Analog sensor with Laplacian noise:

$$J_{ij}^k(\vec{s}) = \frac{1}{\alpha_k^2(\vec{s})} \frac{\partial f_k(\vec{s})}{\partial s_i} \frac{\partial f_k(\vec{s})}{\partial s_j} + \frac{1}{\alpha_k^2(\vec{s})} \frac{\partial \alpha_k(\vec{s})}{\partial s_i} \frac{\partial \alpha_k(\vec{s})}{\partial s_j} \quad (6)$$

where $\alpha_k(\vec{s})$ is the dispersion of the Laplacian noise for that sensor. The Laplacian case is most appropriate for describing fluorescence-based optical chemosensors used within artificial olfactory systems (4).

Case 3: Analog sensor with Poisson noise:

$$J_{ij}^k(\vec{s}) = \frac{1}{f_k^2(\vec{s})} \frac{\partial f_k(\vec{s})}{\partial s_i} \frac{\partial f_k(\vec{s})}{\partial s_j} \quad (7)$$

where $f_k(\vec{s})$ is the Poisson frequency. The Poisson case is a useful first approximation to the spiking statistics of Olfactory Receptor Neurons (ORNs).

4 Optimization of the sensor array using the Fisher Information

It can easily be shown, from eq. 3, that the global expected error of any unbiased estimator is

$$\langle \epsilon^2 \rangle \geq \int d\vec{s} p(\vec{s}) \sum_{i=1}^N \left(J^{-1}(\vec{s}) \right)_{ii} \quad (8)$$

where $p(\vec{s})$ is the *a priori* density of probability of the stimuli \vec{s} and ϵ is the error in the estimate. Therefore, if we know this probability distribution, we can minimize the right term as a function of the sensor parameters in order to make the minimum error optimal according to the Cramér-Rao Bound.

5 Example: 2 odor sensors and 2 odor components

To illustrate our ideas while simplifying the calculations we will consider the sensors to generate an analog response that is corrupted by Gaussian noise, of variance σ^2 , that is identical for both sensors and has no dependence upon the stimulus. The analysis is further simplified by considering odor sensors that respond linearly to changes in stimulus concentration (in the range of interest). These assumptions are valid for odor sensors based upon electrochemical cell

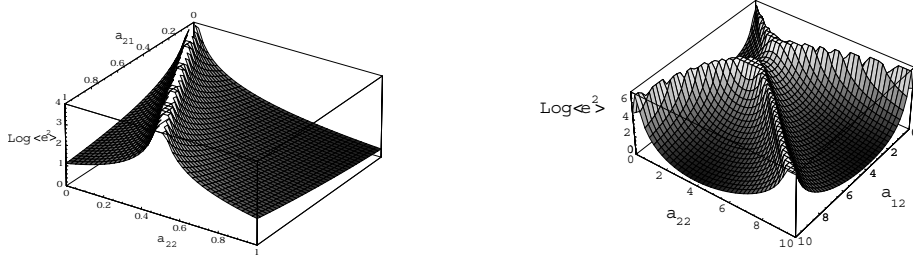


Fig. 2. The effect on the quadratic error of estimation, $\langle \epsilon^2 \rangle$, to variations in the tuning of one sensor within an odor sensing array of two sensors, after fixing the sensitivities for the other sensor. Left: array of linear sensors with Gaussian noise. $a_{11} = 1$, $a_{12} = 0.5$. Right: array of sigmoidal sensors with Poisson noise. $b_{11} = b_{21} = b_{12} = b_{22} = 1$, $a_{11} = 5$, $a_{21} = 2$.

technology (1). Then the sensitivity of sensor k to stimulus component i is a constant $a_{ki} \equiv \frac{\partial f_k(\vec{s})}{\partial s_i}$. Using (5) we can calculate the FIMs for each sensor

$$J^1 = \frac{1}{\sigma^2} \begin{pmatrix} a_{11}^2 & a_{11}a_{12} \\ a_{11}a_{12} & a_{12}^2 \end{pmatrix} \quad J^2 = \frac{1}{\sigma^2} \begin{pmatrix} a_{21}^2 & a_{21}a_{22} \\ a_{21}a_{22} & a_{22}^2 \end{pmatrix}$$

By adding these to form the FIM for the array, then substituting into (8) and rearranging:

$$\langle \epsilon^2 \rangle = \sigma^2 \frac{a_{11}^2 + a_{12}^2 + a_{21}^2 + a_{22}^2}{(a_{12}a_{21} - a_{11}a_{22})^2} \quad (9)$$

Note that this expression does not depend on the probability distribution of the stimuli because the FIMs do not depend on the stimulus in this case.

We wish to choose a_{11} , a_{12} , a_{21} , and a_{22} in order to minimize the estimation error. Clearly a unique solution is not possible, but by choosing the sensitivities of one of the sensors, say a_{1i} , to both stimuli, we can visualize the effect on the estimation error as we vary the tunings for the other sensor, a_{2i} (fig. 2, left). The results are intuitive by considering the situation when one sensor possesses sensitivity terms which are multiples of the other (i.e. the sensors are identical after normalization). In this case, the array is unable to distinguish between the individual stimuli so the estimation error tends asymptotically towards infinity, reflecting the impossibility of building an unbiased estimator in this case. This is represented by the ridge along the center of (fig. 2, left), where the ratio between the sensitivity terms $a_{21} : a_{22}$ is $2 : 1$. If we constrain each of the sensitivity terms to the range $[0, 1]$ (i.e. sensor can only increase in value and its sensitivity is constrained) then the best performance is obtained when $a_{21} = 0$ and $a_{11} = 1$, that is, when the second sensor maximally differentiates itself from the first sensor within the specified constraints.

As a second case we studied an array of two sigmoidal sensors with Poisson noise (e.g. biological sensors). The Poisson frequency is defined by:

$$f_k = \frac{1}{1 + \exp(a_{k1}(b_{k1} - s_1))} + \frac{1}{1 + \exp(a_{k2}(b_{k2} - s_2))} \quad (10)$$

which is based on the observed sigmoidal concentration dependence of ORNs (2). By fixing the offsets (concentration thresholds) for both sensors and the sensitivity tunings to the first compound for both sensors we can visualize the effect on changing the sensitivity tunings to the second compound on the overall estimation error for the array. In fig. 2, right, we can see the optimal error when the stimulus $s_1 = s_2 = .5$ is presented.

6 Discussion

In this work we propose a method for both optimizing and analyzing cross-sensitive sensor arrays that consists in maximizing the trace of the inverse of the FIM. This is equivalent to choosing the array parameters so that the expected error of an optimal unbiased estimator is the lowest possible (Cramér-Rao bound). The method is limited to arrays with a number of sensors of at least the number of dimensions of the sensory space due to the constraints of the form of the FIM and its inverse. An important aspect of Fisher Information not discussed here is its relationship with the Kullback-Leibler distance, which is a measure of the dissimilarity of two probability distributions (3). Future work will explore the optimization of the sensory array using this concept.

References

- [1] N. Barsan, J.R. Stetter, M. Findlay, W. Gopel High performance gas sensing of CO: Comparative tests for semiconducting (SnO₂-based) and for amperometric gas sensors, *Analytical Chemistry* 71 (1999) 2512-7.
- [2] J. Boeckh, K.-D. Ernst, Contribution of single unit analysis in insects to an understanding of olfactory function, *J. Comp. Physiol. A.* 161 (1987) 549-565.
- [3] T.M. Cover, J.A. Thomas, *Information theory* (Wiley, New York, 1991).
- [4] T.C. Pearce, T.A. Dickenson, D.R. Walt, J.S. Kauer, Combining Information Across Input Afferent Streams Gives Rise to Hyperacuity in Biological and Artificial Olfactory Systems, *AChemS XX*, Sarasota, Florida, April 22-26 (1998).
- [5] T.C. Pearce, Odor to sensor space transformations in biological and artificial noses, *Neurocomputing* 32-33 (2000) 941-952.