Vision Research 123 (2016) 26–32

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Classification images in a very general decision model

Richard F. Murray

Department of Psychology and Centre for Vision Research, York University, 4700 Keele Street, LAS 0009, Toronto, Ontario M3J 1P3, Canada

ARTICLE INFO

Article history:
Received 30 January 2015
Received in revised form 24 March 2016
Accepted 21 April 2016
Available online 13 May 2016

Keywords:
Classification images
Decision making
Modelling

ABSTRACT

Most of the theory supporting our understanding of classification images relies on standard signal detection models and the use of normally distributed stimulus noise. Here I show that the most common methods of calculating classification images by averaging stimulus noise samples within stimulus-response classes of trials are much more general than has previously been demonstrated, and that they give unbiased estimates of an observer's template for a wide range of decision rules and non-Gaussian stimulus noise distributions. These results are similar to findings on reverse correlation and related methods in the neurophysiology literature, but here I formulate them in terms that are tailored to signal detection analyses of visual tasks, in order to make them more accessible and useful to visual psychophysicists. I examine 2AFC and yes-no designs. These findings make it possible to use and interpret classification images in tasks where observers' decision strategies may not conform to classic signal detection models such as the difference rule, and in tasks where the stimulus noise is non-Gaussian.

1. Classification images in a very general decision model

Classification images have proven to be a useful tool for investigating visual processing in a wide range of tasks (Ahumada, 1996, 2002; Murray, 2011). In a classification image experiment we introduce many small fluctuations into a stimulus, and measure the influence of these fluctuations on observers' responses. One appealing feature of this approach is that it probes observers' strategies in a very open-ended way. Instead of using, say, proportion correct or reaction time measurements to choose between two or three candidate models, a classification image experiment gives a highly flexible description of how observers make visual judgements, and can reveal features of visual processing that may not have been anticipated by the experimenter (e.g., Ahumada, 1996; Gold, Murray, Bennett, & Sekuler, 2000; Neri & Heeger, 2002).

However, most of the theory for understanding classification images is based on a few standard models from signal detection theory (e.g., Abbey & Eckstein, 2002; Murray, Bennett, & Sekuler, 2002). As a result, this highly flexible method actually seems to depend on rigid assumptions about visual processing, such as the assumption that observers make 2AFC decisions by calculating a decision variable from each stimulus interval, and choosing the interval with the higher decision variable. Furthermore, there have long been doubts about whether these assumptions are always correct (e.g., Treisman & Leshowitz, 1969; Yeshurun, Carrasco, & Maloney, 2008), and this raises the question of what classification images tell us about observers' strategies when these assumptions fail.

In addition, some interesting results have come from studies where standard methods of calculating classification images are applied to new tasks that are not described well by the models that were originally used to justify the standard methods. For example, classification images have been measured in visual search tasks (Rajashekar, Bovik, & Cormack, 2006; Saiki, 2008), which are not instances of the yes-no or 2AFC tasks that underlie the justifications for standard classification image methods, and for which there is no broad agreement about the correct psychophysical model. Here classification images are used outside the domain where they are well understood theoretically, and so again there is room for questions about exactly what they tell us about observers' strategies.

Similarly, Pritchett and Murray (2015) used classification images to estimate observers' decision variables on individual trials, and then they used these estimates to study observers' decision rules in 2AFC tasks. The previous literature suggests that this approach is problematic, because the classification image methods that Pritchett and Murray used have been justified using a specific model of 2AFC decision making (the difference rule), whereas it is precisely the decision rule in 2AFC tasks that Pritchett and Murray are attempting to investigate.

The most widely used methods of calculating classification images are based on averages of Gaussian stimulus noise within stimulus-response classes of trials (Abbey & Eckstein, 2002;

E-mail address: rfm@yorku.ca

http://dx.doi.org/10.1016/j.visres.2016.04.003
0042-6989/

Ahumada, 2002; Murray et al., 2002). I will call these conditional average methods. (An example of a method outside this category is estimating classification images using the generalized linear model, e.g., Knoblauch and Maloney (2008).) Here I show that conditional average methods of calculating classification images are far more general than has been previously demonstrated, and that they give unbiased estimates of observers' templates for a wide range of decision rules and non-Gaussian stimulus noise distributions. The main assumption behind these results is simply that the observer's responses are mediated by the dot product of a template with the stimulus. To show that non-Gaussian noise can be used in classification image experiments, I also assume that each noise element has only a small influence on the observer's responses.

First I discuss classification images measured using Gaussian noise in a 2AFC task, and then using Gaussian noise in a yes-no task. Finally I discuss classification images measured using non-Gaussian noise in a yes-no task.

1.1. Classification images in a 2AFC task

I use upper case letters for matrices and random variables, and lower case letters for scalar constants. I use bold font for images and templates, which I represent as column vectors.

1.1.1. The task

In a 2AFC task there are two stimulus intervals, which I will label as $k = 1, 2$. In each interval the observer views a stimulus $\mathbf{s}_k + \mathbf{N}_k$, where $\mathbf{s}_k$ is a signal and $\mathbf{N}_k$ is noise. I assume that $\mathbf{N}_k$ is a linear transformation of independent and identically distributed (i.i.d.) Gaussian noise: $\mathbf{N}_k = A\mathbf{M}_k$, where $\mathbf{N}_k$ is an $n \times 1$ vector, A is an $n \times m$ matrix and $\mathbf{M}_k$ is an $m \times 1$ vector of Gaussian noise with each element an i.i.d. sample from $N(0, \sigma_M^2)$. In practice, A is usually a convolution, which can produce i.i.d. noise (if A is the identity matrix) or correlated noise. The two signals $\mathbf{g}_1$ and $\mathbf{g}_2$ appear in random order, and I represent the signal order with a random variable S that takes value 1 or 2 to indicate which stimulus interval $\mathbf{g}_1$ appeared in. Thus the signal in interval 1 is $\mathbf{s}_1 = \mathbf{g}_S$ and the signal in interval 2 is $\mathbf{s}_2 = \mathbf{g}_{3-S}$. The observer's task is to judge which stimulus order was shown. I represent the observer's responses with a random variable R that takes value 1 or 2 to indicate which stimulus interval the observer judged signal $\mathbf{g}_1$ to be in.

1.1.2. The observer model

I assume that the observer's responses are based on decision variables D_1 and D_2 that are calculated from the two stimulus intervals. I assume that the stimulus affects the decision variable for stimulus interval k via a dot product of the stimulus with a template $\mathbf{t}_k$.

$$E_k = (\mathbf{s}_k + \mathbf{N}_k)^{\mathrm{T}}\mathbf{t}_k \tag{1}$$

Here T is matrix transposition. I call E_k the 'external component of the decision variable'. I allow different templates $\mathbf{t}_k$ for the two stimulus intervals. The decision variable D_k is some function of the random variable E_k and a multivariate random variable $\mathbf{V}_k$ that represents trial-to-trial fluctuations that are independent of the stimulus noise, such as internal noise that the observer adds to the external component of each decision variable.

$$D_k = f(E_k, \mathbf{V}_k) \tag{2}$$

To describe the observer's decision rule, I define the decision space as

$$H(x_1, x_2) = P(R = 2 | D_1 = x_1, D_2 = x_2) \tag{3}$$

This is the probability of the observer choosing response 2 given the values of the decision variables D_1 and D_2. I do not rely on the usual assumption that 2AFC decisions are based on the difference rule (Tanner & Swets, 1954), which says:

$$H(x_1, x_2) = \begin{cases} 1 & \text{if } x_2 \leqslant x_1 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Here the criterion $x_2 \leqslant x_1$ assumes that the templates $\mathbf{t}_k$ have a higher response to signal $\mathbf{g}_2$ than to $\mathbf{g}_1$; if they have a higher response to $\mathbf{g}_1$, then the criterion is $x_1 \leqslant x_2$. Instead of assuming Eq. (4), as previous studies have done (e.g., Abbey & Eckstein, 2002), I allow the decision space $H(x_1, x_2)$ to be an arbitrary function from $\mathbb{R}^2$ to $[0, 1]$.

In this observer model there is a redundancy between the internal variability $\mathbf{V}_k$ and the decision space H, because any randomness in the observer's responses caused by $\mathbf{V}_k$ could be absorbed into the decision space. However, I will keep this redundancy because it allows us to separately describe channel noise using $\mathbf{V}_k$ (e.g., additive Gaussian noise) and decision noise using H (e.g., randomness due to probability matching (Murray, Patel, & Yee, 2015)), and so it makes the observer model more easily relatable to common signal detection models.

1.1.3. The classification image

Conditional average methods of calculating a classification image in a 2AFC task are based on the average of stimulus noise samples within stimulus-response classes of trials. In Appendix A I derive the conditional expected value $E[\mathbf{M}_1 | S = 1, R = 2]$, where $\mathbf{M}_1$ is the Gaussian i.i.d. noise used to generate the stimulus noise $\mathbf{N}_1 = A\mathbf{M}_1$ in the first stimulus interval. I show that:

$$E[\mathbf{M}_1 | S = 1, R = 2] \propto A^T\mathbf{t}_1 \tag{5}$$

That is, the expected value of the noise vector $\mathbf{M}_1$ is either zero or proportional to the observer's template, transformed by the transpose of the matrix A used to generate the stimulus noise. If A^T is invertible, then the conditional expected value of $(A^T)^{-1}\mathbf{M}_1 = A^{-T}\mathbf{M}_1$ is either zero or proportional to the template $\mathbf{t}_1$. This means that the average of the samples of $A^{-T}\mathbf{M}_1$ on all trials where $S = 1$ and $R = 2$ gives an unbiased estimate of the template $\mathbf{t}_1$ that the observer uses in the first stimulus interval. Alternatively, if we wish to use the stimulus noise $\mathbf{N}_1 = A\mathbf{M}_1$ to estimate the template, then we can take the average of the samples of $A^{-\mathrm{T}}A^{-1}\mathbf{N}_1$ on all trials where $S = 1$ and $R = 2$.

The derivation in Appendix A shows that Eq. (5) is true regardless of the observer's decision space (i.e., the function H in Eq. (3)). Thus conditional average methods do not rely on specific assumptions about observers' decision rules, such as the difference rule in Eq. (4), but instead can be used whenever the stimulus affects the observer's responses via a dot product, as in Eq. (1).

This result can be explained simply and informally as follows. Starting with Eq. (1), the external component of the decision variable is

$$E_k = (\mathbf{s}_k + \mathbf{N}_k)^{\mathrm{T}}\mathbf{t}_k \tag{6}$$

$$= \mathbf{s}_k^{\mathrm{T}}\mathbf{t}_k + (A\mathbf{M}_k)^{\mathrm{T}}\mathbf{t}_k \tag{7}$$

$$= \mathbf{s}_k^{\mathrm{T}}\mathbf{t}_k + \mathbf{M}_k^{\mathrm{T}}(A^{\mathrm{T}}\mathbf{t}_k) \tag{8}$$

The shift in parentheses from Eqs. (7) to (8) shows that applying a template $\mathbf{t}_k$ to filtered noise $A\mathbf{M}_k$ gives the same result as applying a transformed template $A^{\mathrm{T}}\mathbf{t}_k$ to i.i.d. noise $\mathbf{M}_k$. If the stimulus affects the observer's responses only via a template, then i.i.d. noise that is orthogonal to the template can have no effect on the observer's responses, and the expected value of the i.i.d. noise in a stimulus-response class of trials can only be zero or proportional to the template. Thus the expected value of $\mathbf{M}_1$ in a stimulus-response