Journal of Mathematical Psychology 82 (2018) 26–37

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

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

Classification images as descriptive statistics

Peter Neri

Laboratoire des Systèmes Perceptifs (CNRS UMR 8248), Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris, France

HIGHLIGHTS

- Classification images are psychophysical estimates of perceptual mechanisms that resemble 'filters'.
- They are almost invariably connected with human discrimination via a template-matching operation, however the connection is far more opaque than envisaged by this operation.
- Extension to higher-order statistical properties of the classified noise is necessary for adequately constraining potentially underlying circuit models.
- Classification images are best thought of as rich descriptors of data structure, rather than intuitively interpretable snapshots of system components.

ARTICLE INFO

Article history:
Received 20 February 2017
Received in revised form 22 October 2017

Keywords:
System identification
Reverse correlation
Perceptual weight
Psychophysics
Sensory processing

ABSTRACT

Classification images have become popular tools in psychophysics, yet difficulties associated with their interpretation have often hindered their application. Alternative methods for characterizing perceptual filters have been proposed, and the discussion has often focussed on the degree to which classification images are optimal statistical estimators of system components (e.g. kernels). This technical note argues that those difficulties become irrelevant once the tool is situated within a data-driven interpretational framework. Within this framework, classification images and their nonlinear derivatives are understood *not* as transparent estimates of system components, but instead as transparent descriptors of data structure. The many pitfalls associated with the former approach, and the power of the latter, are demonstrated via combination of counter-intuitive computer simulations with empirical examples from published literature. A change in perspective over the manner in which this tool is understood and utilized may lead to a more productive engagement with this methodology.

1. Introduction

1.1. What is a classification image, and why is it useful?

We constantly 'filter' the world around us. Our sensors (eyes, ears, nose, tongue, skin) are bombarded by signals of various kinds, and our ability to discriminate between two such signals (e.g. blue versus red colours) relies on perceptual filters that retain one signal and throw out the other. Our brain then exploits the activity of many such filters to perform specific actions for the purpose of successfully interacting with the environment. In its simplest account, this filtering process can be summarized by a trace (bell-shaped curve in Fig. 1A) that records the response of the perceptual filter (plotted on the y axis) to different values of an environmental characteristic, such as the position of an object along the horizon (plotted on the x axis). From Fig. 1A we infer that this specific filter is selective for objects sufficiently close to the middle position along the x axis (stimulus in Fig. 1C), but stops responding when the object is moved further away from the midpoint (Fig. 1D).

The filtering stage outlined above returns a continuous value. Our behavioural decisions, however, do not come in this format: we either decide to run away from a predator, or stay put; we either eat a potentially poisonous food item, or we drop it. In other words, most decisions we take on how we use our sensory representation to interact with the world are discrete (typically binary): we either choose to take an action, or we choose not to. How do we go from our perceptual representation, which comes in the form 'it is 2× more likely that a predator is hiding behind that bush than not', to the decision 'run away!'?

The simplest model of how this conversion may happen involves a threshold (Green & Swets, 1966): if the ratio between the likelihood of 'predator' versus 'non-predator' is greater than some value, e.g. 1, we run away; if it is smaller than that value, we stay put. Because we tend to produce this kind of response somewhat erratically, i.e. our estimate of the likelihood is not always identical under the same environmental conditions due to noise in our sensors and our decisional process (Green, 1964;

E-mail addresses: neri.peter@gmail.com, peter.neri@ens.fr.
URL: https://sites.google.com/site/neripeter/.

https://doi.org/10.1016/j.jmp.2017.10.004

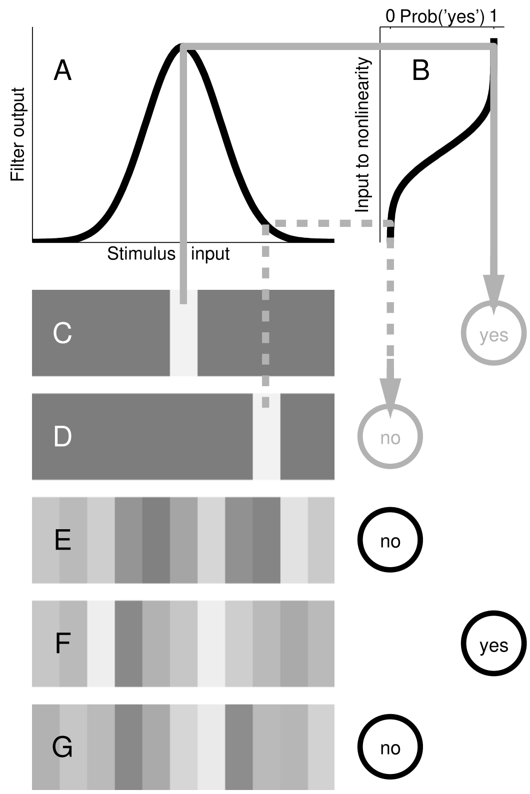

Fig. 1. The linear–nonlinear model consists of a linear filtering stage (A) followed by a nonlinear static nonlinearity (B). The linear filter specifies weights associated with different portions of the incoming stimulus (e.g. different bar positions along the *x* axis) and converts each stimulus into a single decision variable. The static nonlinearity takes the decision variable as its own input, and converts it into a psychophysical choice ('yes' versus 'no' in this example) according to a specified probabilistic rule. Solid/dashed grey lines indicate processing paths through this model for well-matched and mismatched example stimuli (C/D). When stimuli contain multiple bars (E–G), the filter in A acts as a weighting function that sums across bars.

Neri, 2010a), any model of what decision we take must be itself probabilistic: it can only predict that we will run with probability x. To this end, the model in Fig. 1A converts the output from the filter (on the y axis) onto the probability that it will lead to one of two binary choices (e.g. 'yes' versus 'no'). The 'link' function is called a static nonlinearity (an example is shown in Fig. 1B). This function is necessary if one is to re-format the output of the filtering stage into the currency of real-world actions.

The combination of the filtering stage and the static nonlinearity in Fig. 1A–B is termed a 'linear–nonlinear' (LN) model (Murray, 2011; Ostojic & Brunel, 2011): the 'linear' part is the filtering stage (A), the 'nonlinear' part is the static nonlinearity (B). This is a minimal model: anything simpler will not provide a description of sensory processing that is even passable (Neri, 2015). It is therefore understandable that this model is used as a reference point in computational accounts of sensory processing by neurons (Ostojic & Brunel, 2011) as well as observers (Murray, 2011): it is a sensible building block to start with; more complex models can be constructed using blocks of this kind (Carandini et al., 2005) if called for by the phenomenon under study (e.g. Fournier, Monier, Pananceau, & Fregnac, 2011). In particular with relation to the topic discussed in this article, the LN model is often regarded as the theoretical foundation for computing classification images in sensory psychophysics (Ahumada, 2002; Murray, 2011), which brings us to the question: what are classification images?

If we accept the LN model as an adequate representation of the sensory process at hand, say human vision, the classification image is an 'image' of the filtering stage in Fig. 1A: when the underlying sensory filter takes on the shape in Fig. 1A, so will the classification image (Ahumada, 2002). In other words, *if our viewpoint is informed by the LN model*, the classification image technique is a tool for deriving a picture of the filtering stage (Murray, 2011). Why should we want to obtain such a picture?

Classical approaches to sensory processing in animals, e.g. Fechnerian psychophysics, have traditionally emphasized performance: the experimenter focuses on measuring *how well* the animal can detect/discriminate among different signals (Green & Swets, 1966). From these measurements, inferences are sometimes made about the possible shape of the filtering stage, but this is typically achieved via indirect routes (e.g. poorly constrained models with several free parameters) or not at all: the transduction from stimulus to filter output is modelled as a static nonlinear function, effectively incorporating it into the N portion of the LN model and shifting the focus of the investigation onto this component alone (Solomon, 2009). With classification images, the opposite approach is taken: the focus is shifted onto the filtering stage, while the decisional nonlinearity that maps filter output onto choice is bypassed (Neri, 2010b). In this sense, the two approaches are complementary and should be used synergistically whenever possible (Neri, 2011b, 2014b).

There are two critical ingredients that enable this technique to take a snapshot of the filtering stage in a way that is not accessible to e.g. Fechnerian psychophysics. First, the injection of a controlled small perturbation into the stimulus: external noise. To provide a simple example, if observers are asked to discriminate between a bar in the middle (Fig. 1C) and a bar to the side (Fig. 1D), the luminance of each image bar along the x axis may be independently jittered by a random source (see toy examples in Fig. 1E–G). In this way, the output of the filtering stage in the observer's brain will not always be the same in response to the central bar, due to small fluctuations introduced by the added pixel noise; further, the decision taken by the observer on the basis of the filter output will also vary from trial to trial (see 'yes'/'no' responses corresponding to Fig. 1E–G), and those variations will depend on the fluctuations introduced by the noise. Sometimes, the added pixel noise will make a bar-in-the-middle stimulus look very much like a bar-to-the-side stimulus; on those trials, the observer will likely classify the bar-in-the-middle stimulus as bar-to-the-side. On other trials, the added noise will emphasize those features of the bar-in-the-middle stimulus that set it apart from bar-to-the-side; on those trials, the observer will likely classify the bar-in-the-middle stimulus as containing a bar in the middle. The term 'classification' in 'classification images' comes from the classification carried out by the observer as just described.

The addition of noise *per se*, however, is not in itself new: there is a long tradition of using stimulus noise to study its impact on performance (Ahumada, 1987; Pelli & Farell, 1999). The additional ingredient is that, when adding noise, the experimenter keeps track of the *specific* noise sample that was added on every separate perturbation that led to a classification by the observer (Ahumada, 1967; Ahumada & Marken, 1971). This is different than classical approaches where, say, 1000 trials are run at some noise intensity x_1 to measure observer performance p_1 (whatever metric is used to assess it), this process is repeated for different noise intensity values x_2, x_3 and so on, and finally the relationship between x and p becomes the main subject of investigation. In those approaches, the specific noise samples presented during the 1000 trials at intensity x_1 are all lumped into one class without regard for the fact that, on some of those 1000 trials, the specific noise sample that was added to the target signal may have made it easier to detect, while the opposite may have been true for other noise samples in the 1000 trial sequence. In the classification image technique, different noise samples (even if generated by a noise source of