Research Focus

Seeing patterns in the noise

Eero P. Simoncelli

Howard Hughes Medical Institute, Center for Neural Science, and Courant Insitute for Mathematical Sciences, New York University, 4 Washington Place, Rm. 809, New York, NY 10003, USA

How do observers detect the presence of objects or features in visual images? Stochastic stimuli (for example, white noise) have become popular choices for providing a linear characterization of early sensory mechanisms. A recent paper by Neri and Heeger takes this type of methodology a step further, and succeeds in isolating and characterizing non-linear mechanisms responsible for the detection and identification of a specific visual target.

What is it that allows us to detect the presence of objects or features in visual images? And are the mechanisms responsible for detection also responsible for identification of those features? Traditional attempts to answer such questions are based on psychopysical experiments, in which the experimentor measures the detectability or identifiability of some fixed target as a function of target brightness and perhaps other target attributes. Early on in such research, experimentors realized the importance of randomizing the presentation of stimuli, in order to avoid unwanted history-dependence in their measurements. Over thirty years ago, a number of authors proposed a more extreme form of randomization, in which stimuli are drawn randomly from an ensemble and presented in rapid succession [1,2]. The stimuli are then labelled according to responses (for example, 'yes' or 'no' in a psychophysical detection task), and the properties of these 'response-triggered' stimulus sets are analyzed.

Related approaches have also been developed, in which a stimulus is buried in noise, and one analyzes the influence of each particular sample of noise on the subjective response (see, for example, [3,4]). There has been a resurgence of interest in these techniques (see [5] for a number of examples), partly as a result of the development of computer hardware and software capable of both real-time random stimulus generation and computationally intensive statistical analysis.

An example: detecting and identifying a target

A recent article by Neri and Heeger provides an intriguing example of the use of this type of technique to reveal non-linear mechanisms used for detection and identification of a vertical bar [6]. The stimuli are movies, each nine frames long, containing a set of eleven abutting vertical bars. The intensities of the bars are chosen randomly on each frame from a uniform distribution. An example stimulus is illustrated in Fig. 1. After each movie is presented, the subject reports whether they believe the central location of

the middle frame contained a target bar of known polarity (bright or dark). The purpose of the experiment is to determine those aspects of the 9×11 space−time array of intensities that determine the subject's response.

Consider the intensity of a single bar in the movie. This intensity takes on a random value for each trial, drawn from some probability distribution. On each trial, the observer indicates whether or not they have seen the target. The result is a partition of the bar intensities into two sets: those for which the subject answered 'yes', and those on which they answered 'no'. A comparison of these two conditional distributions of bar values can tell us something about the relationship between that particular pixel and the subject's behavior on this task.

Some hypothetical situations are illustrated with simulated data in Fig. 2. Figure 2a shows the distribution of bar intensities across all trials. If the subject's responses do not depend on this bar intensity, then the distribution of stimuli associated with 'yes' (or 'no') responses should be the same shape as the full stimulus distribution (up to statistical sampling error) (Fig. 2b). Conversely, if these response-triggered distributions do not match that of the full set of stimuli, we can infer something about the relationship between that bar and the response. For example, Fig. 2c shows response-triggered distributions
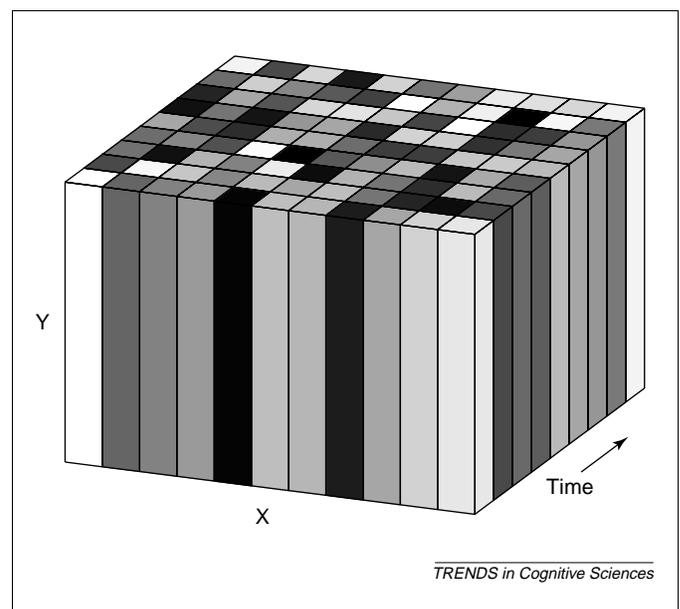

Fig. 1. Depiction of a typical stimulus as used in Neri and Heeger's experiments [6]. Each stimulus was a movie consisting of nine frames of eleven vertical bars, whose intensities were chosen randomly from a uniform distribution on each frame. Subjects were asked whether they saw a bright bar in the middle of the stimulus block, on the middle frame.

Corresponding author: Eero P. Simoncelli (eero@cns.nyu.edu).
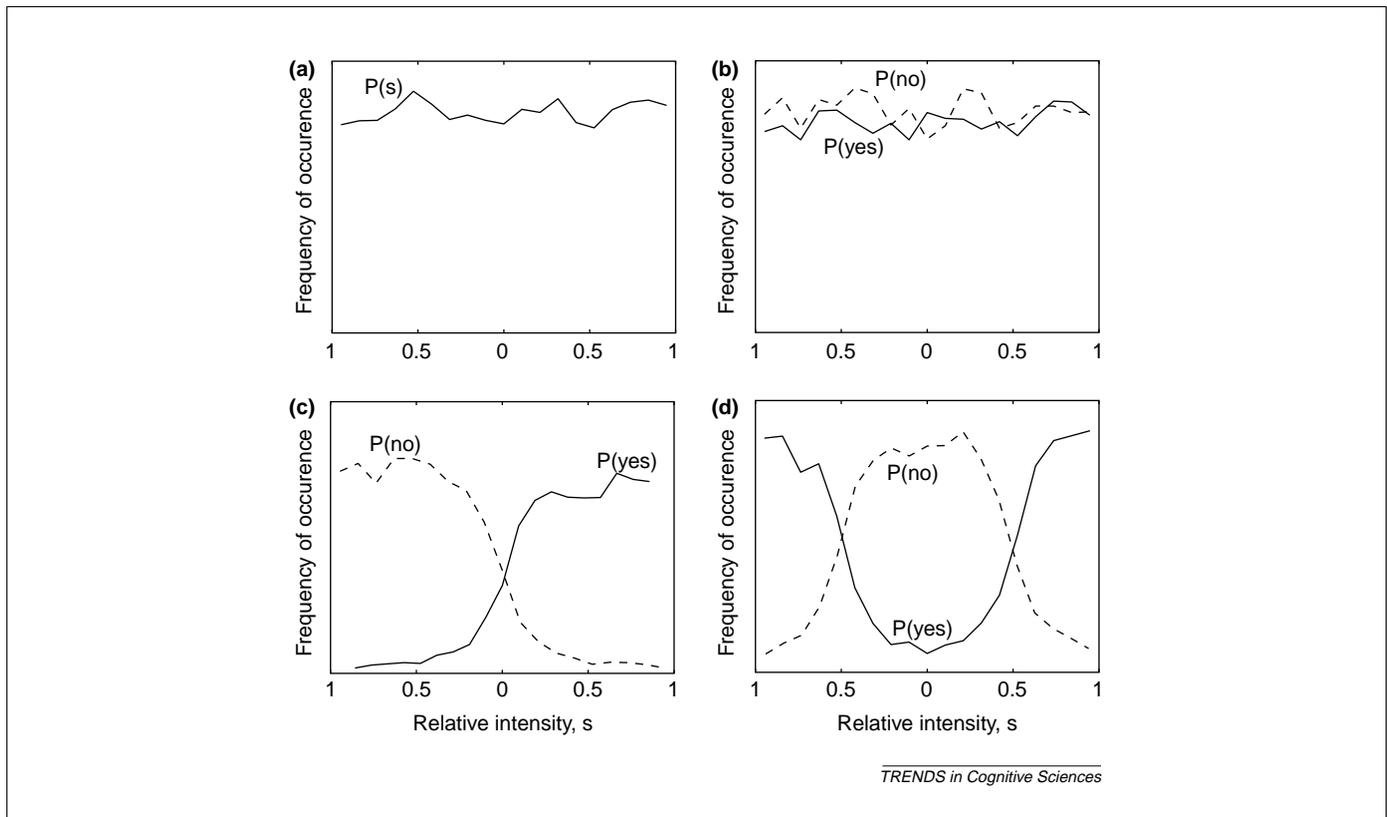
Fig. 2. Hypothetical examples illustrating the analysis used by Neri and Heeger for a single bar within the stimulus (i.e. at a single location in a single frame). (a) Distribution of intensity values (s), relative to mean background intensity, for 10 000 trials. (b,c,d) Three hypothetical distributions of bar intensity values conditioned on subject response (solid = yes, dashed = no). (b) a bar that has no influence on the subjective response. In this case, both response-triggered intensity distributions are the same shape as the raw stimulus distribution, apart from statistical variability; (c) a bar whose intensity affects the subject's response: large intensities are more likely to elicit 'yes' responses; (d) a bar whose contrast (deviation from background intensity) affects subject's response: large constrasts are more likely to elicit 'yes' responses.

for a case where higher bar intensity increases the probability of the subject's responding 'yes'. Similarly, Fig. 2d shows a case in which 'yes' responses are more likely when the bar intensity has a larger contrast (i.e. when it deviates further from the background value). Bar intensities near zero are more likely to elicit a 'no' response, and those far from zero are more likely to elicit a 'yes'.

These three example behaviors can be summarized using the difference between the means and variances of the two response-triggered distributions. Situation (d) produces a large difference in variance, but small difference in mean. Situation (c) produces a large difference in mean but small difference in variance. And situation (b) produces little difference in either mean or variance.

Mechanisms for detection and idenfication

Neri and Heeger performed the mean and variance analysis for every bar in their movie stimuli, and assembled the results to form two summary movies. They refer to these as a 'mean kernel' (the difference in mean intensity of the 'yes'- and 'no'-triggered distributions for each location and frame of the stimulus) and a 'variance kernel' (the same calculation for the variances of the response-triggered distributions). The mean kernel shows a center–surround type of organization: a positive region spread over the middle frames, surrounded by two negative regions that are also spread over the middle frames.

If one assumes that each of the stimulus bars influences the subjective response independently, and that the probability of a 'yes' response varies monotonically with the intensity of each bar, then this mean kernel represents the most potent stimulus for generating a 'yes' response. In physiological applications of reverse-correlation, this kernel can provide a linear characterization of a neuron's receptive field (with a few assumptions) [7]. Psychophysically, the interpretation is similar: the mean kernel provides a linear description of the mechanism that is generating subject responses. Note that an ideal detector for the target would simply measure the intensity at the central location of the middle frame, and thus the mean kernel for such a detector would consist only of a positive central bar. The experimentally measured kernel implies that the visual system is sub-optimal for this task, and implies that subject responses are generated by a mechanism with a center–surround organization, as found in neurons in retina, lateral geniculate nucleus, or primary visual cortex.

The more surprising result comes from the estimation of the variance kernel, which shows a positive region in the center of the stimulus spread over the earliest frames (i.e. preceding the target frame). That is, subjects are more likely to respond 'yes' when the early frames of the stimulus movie contain high contrast bars in the center. This means, for example, that a bar with large positive or negative intensity occurring on the first or second frame makes it more likely the subject will say that they saw a

positive-intensity target in the middle frame. Again, this implies a sub-optimal strategy for detecting the target, as these early bar intensities are independent of the presence or absence of the target. Neri and Heeger interpret this to mean that these early high-contrast signals are engaging a separate 'attentional' mechanism that is used to detect the target. They ran a clever second experiment in which subjects had both to detect and to identify the polarity of the target, and the results demonstrated that the variance kernel accounted for the detection task and the mean kernel accounted for the identification task. The experiment thus lends support to the hypothesis that the visual system answers the questions of 'what?' and 'where?' using separate mechanisms [8].

Conclusion

The results of Neri and Heeger's experiments are intriguing, and provide an elegant demonstration of the power of stochastic stimuli in characterizing visual mechanisms. It is worth considering the drawbacks of this approach, as well as possible generalizations. First, designing and executing this type of experiment is quite difficult, and relies on a number of decisions about how to instruct subjects, how much and what kind of training to allow, how strong a target signal to use, and whether to provide feedback. Second, the summary of their method in this brief review has been simplified to ignore the distinction between cases in which a target was present and those when there was no target, because most of the analysis presented in the paper by Neri and Heeger was done in this fashion (although they do present results for the 'False Alarm' case alone, which seem consistent with the simplified yes/no case). Analysis and interpretation of these sub-cases is more difficult, but can potentially offer further insights into the nature of the underlying visual mechanisms. It would also be interesting to extend the analysis to include interactions between stimulus bars (i.e. estimation of response-triggered *covariance*), as has been done in physiological settings [9–11]. This could provide a richer characterization of the underlying mechanisms, at the expense of requiring more data for reliable estimation.

Finally, it would be interesting to see this technique applied to the detection, discrimination or identification of more complex stimulus features, such as those defined by orientation or motion (e.g. see [12]). Ultimately, refinement of these techniques could allow us to formulate a precise description of the mechanisms underlying all aspects of vision, from detection of complex features, to attentional and recognition processes.

References

1 deBoer, E. and Kuyper, P. (1968) Triggered correlation. *IEEE Transact. Biomed. Eng.* 15, 169–179
2 Marmarelis, P.Z. and Marmarelis, V.Z. (1978) *Analysis of Physiological Systems: The White Noise Approach*, Plenum Press
3 Ahumada, A.J. and Lovell, J. (1971) Stimulus features in signal detection. *J. Acoust. Soc. Am.* 49, 1751–1756
4 Ahumada, A.J. (1996) Perceptual classification images from vernier acuity masked by noise. *Perception* 26, 18
5 Eckstein, M.P. and Ahumada, A.J. (2002) Classification images: a tool to analyze visual strategies. *J. Vision* 2(1), Special Issue: *Classification Images*, DOI 10:1167/2.1.1x
6 Neri, P. and Heeger, D.J. (2002) Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nat. Neurosci.* 5, 812–816
7 Chichilnisky, E.J. (2001) A simple white noise analysis of neuronal light responses. *Netw. Comput. Neural Syst.* 12, 199–213
8 Sagi, D. and Julesz, B. (1985) 'Where' and 'what' in vision. *Science* 228, 1217–1219
9 de Ruyter van Steveninck, R. and Bialek, W. (1988) Coding and information transfer in short spike sequences. *Proc. R. Soc. Lond. B. Biol. Sci.* 234, 379–414
10 de Ruyter van Steveninck, R. and Bialek, W. (1988) Coding and information transfer in short spike sequences. *Proc. R. Soc. Lond. B. Biol. Sci.* 234, 379–414
11 Schwartz, O. *et al.* (2002) Characterizing neural gain control using spike-triggered covariance. In *Adv. Neural Information Processing Systems* (Vol. 14) (Dietterich, T.G. *et al.*, eds), pp. 269–276, MIT Press
12 Ringach, D.L. *et al.* (1997) A subspace reverse-correlation technique for the study of visual neurons. *Vision Res.* 37, 2455–2464

Imagined movements that leak out

Margaret Wilson

Department of Psychology, University of California, Santa Cruz, CA 95064, USA

In a case study that fundamentally alters our understanding of motor imagery, Schwoebel *et al.* report a patient who unintentionally carries out imagined movements. Furthermore, his 'imagery' movements are more accurate than his intended movements, which suggests that the inhibitory signal that normally prevents us from acting out our motor imagery can be selectively blocked. Removing this inhibition allows us to observe motor imagery 'in action', and reveals that motor imagery and motor planning for execution are not identical.

In the last twenty years it has become accepted that 'imagery' of perceptual or motor events involves mental representations that, in some important sense, resemble the 'real thing'. Visual imagery, for example, causes activation in visual processing areas of the brain [1], and motor imagery causes activation in motor areas [2]. But

Corresponding author: Margaret Wilson (mlwilson@cats.ucsc.edu).