ELSEVIER

Cognitive Science 28 (2004) 141–146

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

A picture is worth thousands of trials: rendering the use of visual information from spiking neurons to recognition

Frédéric Gosselin [a,*], Philippe G. Schyns [b]

[a] *Département de Psychologie, Université de Montréal, 90 Vincent-d'Indy, C.P. 6128 succ., Centre-ville, Montreal, Que., Canada H3C 3J7*

[b] *Department of Psychology, University of Glasgow, Glasgow G12 8QB, UK*

Received 16 January 2004; received in revised form 22 January 2004; accepted 22 January 2004

Keywords: Vision; Attention; Reverse correlation; Bubbles; RISE; Change detection

> Mastery of technique is so important that [. . .] it may be stated that the greatest discoveries are in the hands of the finest and most knowledgeable experts on one or more of the analytical methods.
>
> –(Santiago Ramón y Cajal, 1897)

If what Ramón y Cajal (1999) suggested at the end of the 19th century still applies today—and we certainly believe that it does—researchers in vision science are now facing tremendously exciting times. Techniques to measure the activity of the brain—especially the visual brain—are now more numerous, accessible and affordable than ever. Brain imaging techniques such as functional magnetic resonance imagery (fMRI), positron emission tomography (PET), electroencephalography (EEG), event related potential (ERP), optical imagery, magnetoencephalography (MEG), tools to induce localized and transient disruptions in the normal functioning of the brain like transcranial magnetic stimulation (TMS), and accurate video eye-trackers have now joined methods of single-cell recording, psychophysics, visual cognition and neuropsychology in the toolbox of the vision scientist.

In contrast to unquestionable progress in measurement techniques, up until recently little progress had been made to relate brain and behavioral measurements to the properties of the distal stimuli eliciting these responses. But times are changing: Vision science is undergoing a small-scale revolution that is the topic of this special issue of *Cognitive Science*. New tools[1] are being developed to relate stimulus properties with behavioral and brain events, and ascribe a functional role to the latter.

* *E-mail address:* frederic.gosselin@umontreal.ca (F. Gosselin).

0364-0213/$ – see front matter
doi:10.1016/j.cogsci.2004.01.003

1. The problem

The articles of this special issue attempt to answer to the following question: What aspect of a distal visual stimulus is responsible for a measurable response of the observing mechanisms, from brain cells to recognition, in biological systems and machines? While it is obvious that visual stimuli elicit numerous measurable brain and behavioral responses in observers, it remains a challenge for cognitive science to determine the stimulus properties that elicit and modulate the amplitude of these responses, and to provide a common language of information with which to relate behavioral and brain events.

To illustrate this generic problem, consider the example of single-cell studies concerned with the mechanisms of visual categorization. Various researchers have established that IT neurons respond specifically to complex objects such as faces and hands (Desimone, Albright, Gross, & Bruce, 1984; Gross, Rocha-Miranda, & Bender, 1972; Perrett, Rolls, & Cann, 1982). However, further investigations revealed that the effective stimulus was represented in a much lower-dimensional space of abstract features (Desimone et al., 1984; Kobatabe & Tanaka, 1994; Tanaka, Saito, Fukada, & Moriya, 1991; Tsunoda, Yamame, Nishizaki, & Tanifuji, 2001) or parameters (Op de Beeck, Wagemans, & Vogels, 2001; Sigala & Logothetis, 2002; see also Pasupathy & Connor, 2002 for V4 neurons). In a related vein, EEG/MEG and neuroimaging studies have established sensitivity of brain signals to faces (i.e., the N170 in ERP, Bentin, Allison, Puce, Perez, & McCarthy, 1996; Carmel & Bentin, 2002), or activity in the middle fusiform gyrus in neuroimaging (Gauthier, Tarr, Moylan, Skudlarski, Gore & Anderson, 2000; Kanwisher, McDermott, & Chun, 1997), but further studies also revealed sensitivity to other expert object categories (Gauthier, Skudlarski, Gore, & Anderson, 2000) and novel objects (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999), leaving unresolved the question of the specific stimulus determinants (but see Schyns, Jentzsch, Johnson, Schweinberger, & Gosselin, 2003; Smith, Gosselin, & Schyns, in press).

2. The contributions

The articles of this special issue describe techniques adapted to the complexity of this generic problem at different levels of visual integration. They share the characteristic of estimating the use of visual information with little bias (at the cost of many trials). We and others find useful to classify these information estimates according to three main types: represented (R), available (A) and potent (P), articulated as follows (see Gosselin & Schyns, 2002, 2004; Murray & Gold, 2004, for further discussions):[2]

$$\mathrm{R} \otimes \mathrm{A} \approx \mathrm{P}$$

Reverse correlation provides estimates of information internally represented (R). It has been successfully applied to estimate represented information at different levels of visual organization, ranging from the receptive fields of single cortical cells in neuroscience (Ringach & Shapley, this issue), textures and faces discrimination in photometric space (Gold, Bennett, & Sekuler, this issue), photometric features of gender, expression, and identity recognition

Table 1
The six basic decisions (i.e., what is the stimulus set, the search space, the noise/sample, the observer, the task, and the response?) for each one of the four techniques measuring represented (top) and the three techniques measuring potent (bottom) visual information

Authors	Stimulus set	Search space	Noise/sample	Observer	Task	Response
Gold, Bennet, and Sekuler	Two faces and two textures	Photometric space	White Gaussian	Individual results of two humans	2AFC[a]	Key-press: 2 choices
Mangini and Biederman	One ambiguous face per task	Photometric space	Colored “Sinusoidal”	Average of 36 humans per task	Discrimination of identity, gender and expression	Key-press: 2 choices and 2 confidence levels
Olman and Kersten	Stick-figure animals	3D stick-figure animal space	Positional Gaussian	One human	Typicality judgments	Key-press: 2 choices
Ringach and Shapley	None	Space–time	“Subspaces”: grating space and others	Monkey cells	Passive viewing	Spiking frequency
Sadr and Sinha	Five objects	Photometric space	Phase alignment	Average of four to eight humans	Object naming	Key-press: 2 choices
Tse	Array of two-color elements	Space–time	One element changing color at different times	Average and individual results of three humans	Color change detection	Key-press: 2 choices
Vinette, Gosselin, and Schyns	30 faces	Space–time	Space–time bubbles	Average and individual results of 10 humans	Identification	Key-press: 10 choices

[a] In a 2-alternative-force-choice (2AFC) trial, an observer must discriminate between two two-stimuli intervals: stimulus A–stimulus B and stimulus B–stimulus A; the order of the stimuli represents either an arrangement in space or in time. Researchers use 2AFC tasks to minimize response bias.

(Mangini & Biederman, this issue), and other important dimensions of object recognition (Olman & Kersten, this issue).

Other techniques discussed here have instead focused on clarifying the potent components (P) of a distal stimulus that modulate measurable responses. These techniques attempt to reduce real-world stimuli to their response-triggering information. *Bubbles* (Vinette, Gosselin, & Schyns, this issue) and RISE (Sadr & Sinha, this issue) are techniques that prune the stimulus to reveal its effective information for behavior in face and object recognition tasks. Similarly, Tse (this issue) samples the visual field with a change detection task to estimate the dynamics of the deployment of attention.

Finally, several contributors (Gold, Bennett, & Sekuler, this issue; Olman & Kersten, this issue; Vinette, Gosselin, & Schyns, this issue) also discuss the information that is available (A) to resolve a task. In relation to P (or R), estimates of A provide the formal benchmarks of available information against which to rank usage by organisms (see Kersten, Mamassian, & Yuille, in press, for a more thorough discussion of the ideal observer approach; see Geisler & Diehl, 2003, for an interpretation of available information as adaptive information).

When applying the techniques just described to new experiments researchers typically address six basic questions (Gosselin & Schyns, in press): (1) what is the external visual information, (2) in which space will stimuli be searched, (3) what is the noise that will probe the search space, (4) who are the observers, (5) what is the task and (6) what are the responses. Table 1 summarizes how the contributions of this special issue related to R (the first four entries) and those related to P (the last three entries) address these questions.

3. Conclusion

Reverse correlation, *Bubbles*, RISE, and change detection techniques offer the potential to visualize and measure the information used at different levels of visual integration. No book, special issue, or even review article describe all the techniques covered in this special issue for a broad readership. Instead, their presentation has been so far confined to specialist journals (e.g., Eckstein & Ahumada's, 2002, *Journal of Vision* special issue on reverse correlation), limiting their impact. We hope that the specialist reader will benefit from the multi-disciplinary aspect of the collection of articles, and that the honour undergraduate or graduate student will learn about possible applications of these techniques to his/her own research and that all will live true to Ramón y Cajal's words.

Notes

1. Even though these tools have been mostly developed in vision, their application is not restricted to this field (e.g., Eckstein & Ahumada, 2002; Gosselin & Schyns, 2002; Simoncelli, 2002). Historically, Wiener (1958) initiated the approach in engineering (with reverse correlation) and suggested to extend it to study the brain. de Boer and Kuyper (1968) carried out the initial application in neuroscience, and Ahumada and

Lovell (1971) later examined the psychophysics of auditory processes. In principle, it should be possible to adapt this approach to haptic, taste and olfactory signals.

2. Murray and Gold (2004) have demonstrated that, for a Linear Amplifier Model (LAM) observer, a pixel by pixel product replaces the '$\otimes$' operator in this generic equation. RAP, however, extends beyond the LAM (e.g., to nonlinear information).

Acknowledgments

This research was supported by an NSERC (R0010085) and an NATEQ (R0010287) grant awarded to Frédéric Gosselin; and by an ESRC grant (R000237901) awarded to Philippe G. Schyns.

References

Ahumada, A. J., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, *49*, 1751–1756.

Bentin, S., Allison, T., Puce, A., Perez, A., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition*, *83*, 1–29.

de Boer, R., & Kuyper, P. (1968). Triggered correlation. *IEEE Transaction of Biomedical Engineering*, *15*, 169–179.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the Macaque. *Journal of Neuroscience*, *4*, 2051–2062.

Eckstein, M. P., Ahumada, A. J. (Eds.). (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision*, *2*, i-i, http://journalofvision.org/2/1/i/, doi:10.1167/2.1.i.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruit brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191–197.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*, 568–573.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*, 495–504.

Geisler, W. S., & Diehl, R. L. (2003). A Bayesian approach to the evolution of perceptual and cognitive systems. *Cognitive Science*, *27*, 379–402.

Gosselin, F., & Schyns, P. G. (2002). RAP: A new framework for visual categorization. *Trends in Cognitive Sciences*, *6*, 70–77.

Gosselin, F., & Schyns, P. G. (2004). No troubles with Bubbles: A reply to Murray and Gold. *Vision Research*, *44*, 471–477.

Gosselin, F., & Schyns, P. G. (in press). A user's guide to *Bubbles*.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *Journal of Neurophysiology*, *35*, 96–111.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.

Kersten, D., Mamassian, P., & Yuille, A. (in press). Object perception as Bayesian inference. *Annual Review of Psychology*.

Kobatabe, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, *71*, 856–867.

Murray, R. F., & Gold, J. M. (2004). Troubles with bubbles. *Vision Research*, *44*, 461–470.

Op de Beeck, H., Wagemans, J., & Vogels, R. (2001). Inferotemporal neurons represent low-dimensional configurations of parametrized shapes. *Nature Neuroscience*, *4*, 1244–1252.

Pasupathy, A., & Connor, C. (2002). Population coding of shape in area V4. *Nature Neuroscience*, *5*, 1332–1338.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329–342.

Ramón y Cajal, S. (1999). *Advice for a young investigator.* Translated by N. Swanson & L. W. Swanson from *Reglas y Consejos sobre Investigación Cientifica: Los tónicos de la voluntad* (1897). Massachusetts: MIT Press.

Schyns, P. G., Jentzsch, I., Johnson, M., Schweinberger, S. R., & Gosselin, F. (2003). A principled method for determining the functionality of ERP components. *Neuroreport*, *14*, 1665–1669.

Sigala, N., & Logothetis, N. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature*, *415*, 318–320.

Simoncelli, E. P. (2002). Seeing patterns in noise. *Trends in Cognitive Sciences*, *7*, 51–53.

Smith, M., Gosselin, F., & Schyns, P. G. (in press). Receptive fields for flexible face categorizations. *Psychological Science.*

Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, *66*, 170–189.

Tsunoda, K., Yamame, Y., Nishizaki, M., & Tanifuji, M. (2001). Complex objects are represented in macaque inferotemporal cortex by combinations of feature columns. *Nature Neuroscience, 4,* 832–838.

Wiener, N. (1958). *Nonlinear problems in random theory.* New York: Wiley.