

2006 Special Issue

Mathematics bridging cognition and single-cell physiology

Claus Bundesen

Center for Visual Cognition, University of Copenhagen, Linnésgade 22, DK-1361 Copenhagen K, Denmark

We need well-founded theories linking attentional phenomena at the cognitive, psychological level to neurophysiological phenomena at the level of individual neurons, and we want the theories to be as precise as possible. Is it currently feasible to formulate such theories in mathematical terms? Here I argue, by example, for a positive answer. My example is the neural theory of visual attention called NTVA (Bundesen, Habekost, & Kyllingsbæk, 2005).

1. TVA

NTVA is a further development of the theory of visual attention called TVA (Bundesen, 1990). TVA provides quantitative accounts of human performance (reaction times and error rates) in a broad range of experimental paradigms of single-stimulus recognition and attentional selection from multi-object displays.

In TVA, both visual identification and attentional selection of objects in the visual field consist in making visual categorizations—that is, encoding visual categorizations into short-term memory. A visual categorization has the form “object x has feature i ” or, equivalently, “object x belongs to category i ”. When a visual categorization of an object completes processing, the categorization enters visual short-term memory (VSTM) if memory space for the categorization is available in VSTM. The capacity of VSTM is limited to about 4 different objects (see, e.g., Luck and Vogel (1997), Shibuya and Bundesen (1988) and Sperling (1960)).

Clearing VSTM effectively starts a race among objects in the visual field to become encoded into VSTM. Each object x is represented in the encoding race by all possible categorizations of the object. By the central equation of TVA,

$$v(x, i) = \eta(x, i) \beta_i \frac{w_x}{\sum_{z \in S} w_z}, \quad (1)$$

the rate, $v(x, i)$, at which a particular visual categorization, “ x belongs to i ”, is encoded into VSTM, is given by a product of three terms. The first term, $\eta(x, i)$, is the strength of the sensory evidence that x belongs to category i ; the second term, β_i , is a perceptual decision bias associated with category i ($0 \leq \beta_i \leq 1$); and the third term is the relative attentional weight of object x (i.e., the weight of object x , w_x , divided by the sum of weights across all objects in the visual field, S).

2. NTVA

NTVA gives Eq. (1) an interpretation at the level of individual neurons. A typical neuron in the visual system is assumed, first, to be specialized to represent a single feature and, second, to represent the properties of only one object at any given time. The feature for which the neuron is specialized can be a more or less simple “physical feature” or a “microfeature” in a distributed representation. Similarly, depending on the level of processing, the object represented by the neuron can vary widely in complexity (e.g., an edge, a bar, a group of bars, a letter, a word, a face). The object selection of the neuron occurs by dynamic remapping of the cell’s receptive field so that the functional receptive field contracts around the selected object (cf. Moran and Desimone (1985), Reynolds, Chelazzi, and Desimone (1999)). The remapping is done so that the probability that the neuron comes to represent a particular object equals the attentional weight of the object divided by the sum of the attentional weights of all objects in the classical receptive field.

The *activation* of a neuron by the appearance of an object in its receptive field is defined as the increase in firing rate above a baseline rate representing the undriven activity of the neuron. If the baseline rate is zero, the activation is just the firing rate. Independently of the distribution of neurons among objects (*filtering*), the activation in neurons representing a particular feature (say, the activation in all feature- i neurons—all neurons representing feature i) is scaled up or down (*pigeonholing*; cf.

E-mail address: claus.bundesen@psy.ku.dk.

Martinez-Trujillo and Treue (2004), McAdams and Maunsell (1999) and Treue and Trujillo (1999)).

Eq. (1) describes the combined effects of filtering and pigeonholing on the total activation of the population of neurons representing the categorization “object x has feature i ”. The v value on the left-hand side of the equation is the total activation of the neurons that represent the categorization at the level of processing where objects compete for entrance into VSTM. At this level of processing, the classical receptive fields of neurons are so large that each one covers the entire visual field. On the right-hand side of the equation, both β_i and $w_x / \sum_{z \in S} w_z$ range between 0 and 1, so $\eta(x, i)$ equals the highest possible value of $v(x, i)$. Thus, $\eta(x, i)$ equals the total activation of the set of all feature- i neurons when every feature- i neuron represents object x (say, x is the only object in the visual field) and the featural bias in favor of i is maximal (i.e., $\beta_i = 1$).

When the proportion of feature- i neurons representing object x , $w_x / \sum_{z \in S} w_z$, is less than 1, then the total activation representing the categorization “object x has feature i ” is scaled down by multiplication with the factor $w_x / \sum_{z \in S} w_z$ on the right-hand side of the equation. The total activation representing the categorization depends not only on the number of neurons representing the categorization but also on the level of activation of the individual neurons representing the categorization. The bias parameter β_i is a scale factor that multiplies activations of all individual feature- i neurons, so the total activation representing the categorization “object x has feature i ” is also directly proportional to β_i .

Thus, in the neural interpretation proposed for Eq. (1), the total activation representing the categorization “object x has feature i ” is directly proportional to both the number of neurons representing the categorization and the level of activation of the individual neurons representing the categorization. The number of neurons is controlled by the relative attentional weight of x , $w_x / \sum_{z \in S} w_z$ (filtering), but the activation of the individual neurons is controlled by β_i (pigeonholing).

NTVA explains a wide range of findings from single-cell studies of attentional effects on visual representations in primates: attention effects with single versus multiple stimuli in the receptive field, spatial and nonspatial attention effects, interactions with luminance contrast, multiplicative modulations of firing rates, and baseline shifts. Being a neural interpretation of TVA, NTVA also provides quantitative accounts of human performance (reaction times and error rates) in a broad range of experimental paradigms of single-stimulus recognition and attentional selection from multiobject displays. Thus, NTVA provides a mathematical framework to unify the two fields of research—mathematical formulas bridging cognition and single-cell physiology.

References

- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*, 523–547.
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, *112*, 291–328.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.
- Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, *14*, 744–751.
- McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, *19*, 431–441.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*, 782–784.
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, *19*, 1736–1753.
- Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception & Performance*, *14*, 591–600.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, *74* (11, Whole No. 498).
- Treue, S., & Trujillo, J. C. M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*, 575–579.