

Climbing the cortical ladder from sensation to perception

Stefan Treue

German Primate Center, Kellnerweg 4, 37077 Goettingen, Germany

A recent study using displays that are ambiguous for motion direction demonstrates that the current perceptual interpretation of such a stimulus is encoded in the highest areas of visual cortex whereas earlier areas encode only its sensory properties. This finding implies that cortical processing pathways perform a transition from a sensory representation to a representation that emphasizes the input's perceptual interpretation and ultimately the organism's behavioral state.

The visual system of primates is highly structured, containing several dozen distinct areas. These areas are organized into a hierarchical system for the analysis of sensory information in which processing pathways, that is, chains of serially connected areas, can be identified. One of the central questions of systems neuroscience is how the task of analyzing the visual input is divided amongst the members of such cortical pathways [1]. Recent findings by Williams *et al.* [2] support the hypothesis that a visual pathway is more than a series of sensory processing steps, and in fact represents a gradient from a sensory-centered representation in the early cortical areas to an internal representation of the visual world in higher cortical areas that reflects the organism's current behavioral state and its perceptual interpretation of the sensory input.

Hierarchical visual pathways

The areas of visual cortex can be placed in a hierarchical network by analyzing the laminar distribution of their respective inputs and outputs [3,4]. Such an anatomical analysis, together with data about the response properties of neurons and the effects of localized lesions in these areas have provided the basis for the current view that the visual cortex is organized around two main processing pathways emanating from the primary visual cortex and reaching to the temporal and parietal lobe respectively [5,6]. A large number of studies have identified the sensory preferences, that is to say, the tuning properties of neurons in the various areas. The view that has emerged is that the different visual areas have distinct preferences providing different facets of analysis of the sensory input. The two general trends that can be identified as one moves up through the pathways are an increase in (i) the receptive-field size (the area of visual space to which the neuron responds) and (ii) the complexity of the neuronal preferences.

The growth in receptive-field size represents an increasing spatial integration of information by the neurons and a corresponding loss of retinotopic

organization. In higher areas of visual cortex such a topography is virtually absent, with receptive fields routinely covering a whole quadrant or even a hemifield. The growth in the preference for complex stimulus aspects is equally dramatic. Whereas neurons in primary visual cortex show tuning for basic image properties (e.g. the orientation of lines and edges, the direction of motion or other simple stimulus features), neurons in higher areas respond selectively to complex motion patterns such as expansion or rotation, and to complex forms that can be as specific as particular objects or individual faces.

From stimulus to perception

More recently there has been a shift from such investigations of the sensory aspects to signals in visual cortex reflecting perceptual aspects or cognitive influences on sensory processing [7]. Given the much more detailed understanding of sensory information processing in the dorsal visual pathway, especially the analysis of visual motion information in the middle temporal area (MT) in monkeys, it is not surprising that several studies have concentrated on or included this area and investigated issues of visual motion perception. A particularly appealing approach that can reveal perceptual influences on the responses of sensory neurons has been to look at correlations between neuronal activity and an animal's perceptual state when confronted with ambiguous stimuli.

A recent study by Williams and colleagues [2] used this approach with an elegantly simple design. They compared the responses of neurons in three areas of the dorsal visual pathway – MT, the medial superior temporal area (MST), and the lateral intraparietal area (LIP) – to a motion display whose directionality is entirely perceptual because the stimulus contains equal motion strength for opposite directions. This was achieved by creating apparent motion stimuli (quick successions of stationary images), in which a regular and repeating arrangement of dots was displaced by various fractions of the inter-dot spacing (Fig. 1). Because of the regular array of the dots the visual system is faced with a correspondence problem when trying to link individual dots through successive iterations of the displacement. Generally, the perceptual interpretation of such ambiguous displays is determined by those matches that mean the dots move the shortest distance. Thus, a stable and reliable interpretation of the direction of motion can be achieved in the displays used by Williams *et al.* so long as the dot displacement is not exactly half of the inter-dot spacing. In the particular case when the displacement is exactly half the spacing, either direction of motion is equally likely but, interestingly, human observers usually report a strong sense of unidirectional coherent motion

Corresponding author: Stefan Treue (treue@gwdg.de).

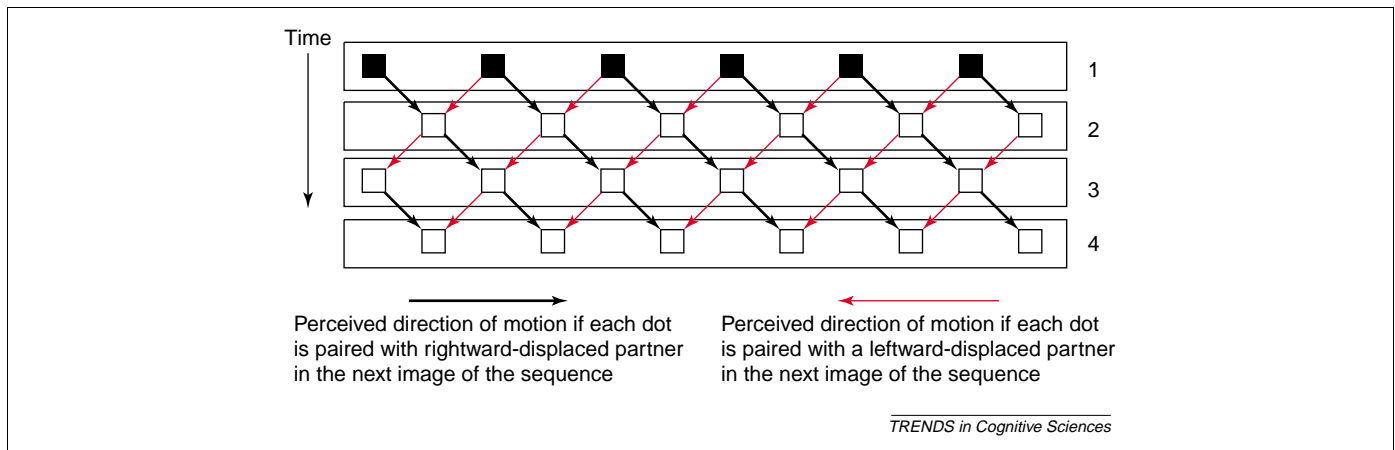


Figure 1. Representation of the ambiguous motion stimulus. The stimulus used by William *et al.* [2] consists of an array of regularly spaced dots within a circular aperture on a computer monitor. Over time these dots are displaced by a variable proportion of the inter-dot distance. If the displacement puts a dot exactly equidistant to two previous dots (as shown here) the display can be interpreted as having moved in one or the other direction with equal certainty. This ‘correspondence problem’ is illustrated by showing a segment of the stimulus through three displacements (strips numbered 1 to 4). The red and black arrows indicate the two possible perceptual interpretations of the stimulus motion.

across the stimulus. Presumably this reflects the visual system’s bias to pair every dot with a partner after the displacement. As this unidirectional motion percept is not stimulus-based but rather reflects a perceptual process it can be used to establish perceptual influences on neuronal firing rates.

In their study William *et al.* found a gradient of such influences. In MT, neuronal responses did not differ between trials in which the animal reported opposite perceived directions of motion. But to some degree in MST, and particularly in LIP, many neurons responded more strongly when the animal reported the perception of the neuron’s preferred direction. In LIP, the directional preference of a given neuron was equal when determined using unambiguous stimuli and when using ambiguous stimuli perceived to be moving in the preferred or anti-preferred directions. Interestingly, the higher responses in LIP neurons in trials in which the monkey’s perceived direction matched the neuron’s preferred direction of motion, were already present before the stimulus appeared on the screen. As the direction of the next stimulus was unpredictable this would appear to reflect the monkey’s expectation, or a bias, concerning the direction of motion in the upcoming trial.

Ambiguous stimuli

William *et al.*’s findings show that, in higher areas in the cortical pathways, neuronal representations shift from being dominated by the properties of the sensory input to those reflecting the animal’s perceptual state in interpreting patterns that allow orthogonal perceptual interpretations. There have been two previous recording studies using similarly ambiguous patterns [8–10] but data were collected only from MT neurons. In these experiments monkeys were trained to report the perceived direction of rotation of a random-dot pattern that represented a transparent rotating cylinder with dots on the front and back surface. Because of the 2-D projection of the dots’ 3-D trajectory onto the computer screen, the speed gradient across such displays creates a vivid impression of depth with a clear separation between a ‘front’ and a ‘back’

surface, which move in opposite directions. Which of the two motion directions is assigned to the front and which to the back is a perceptual choice that varies over time so that on successive trials either perceptual interpretation can occur. Both of the studies reported a correlation between the recorded neurons’ activation and the perceived direction of rotation.

Focusing attention onto a particular stimulus feature such as its motion direction enhances responses in MT neurons preferring that feature [11]. Similar effects have been observed in the temporal pathway [12]. If such a feature-based attentional effect also includes specificity for a particular depth plane the correlation between the monkey’s perceived sense of rotation and the neuronal firing rate might reflect the workings of attentional modulation [10]. The lack of a significant correlation for MT neurons in the William *et al.* study might therefore reflect the low attentional load in their experiment, because their stimulus contained neither stereoscopic depth nor two transparently moving surfaces. Given the increased attentional modulation known to exist in higher areas [13,14] the attentional modulation in the present study might have only reached significance in MST and LIP.

Response fluctuations or attention?

Such an account would provide a link to a series of studies using another class of ambiguous stimuli. Newsome and colleagues used noisy, moving random-dot patterns with various amounts of net motion in one of two opposite directions to demonstrate a close link between an animal’s choice of the dominant direction and the activity of MT [15] and MST [16] neurons. They found very similar effects in the two cortical areas. It should be noted that, unlike the random-dot rotating cylinder stimulus and the stimulus used by Williams *et al.*, for Newsome’s high-noise displays the difference between a stimulus causing one directional decision and a stimulus causing the opposite decision can be very small, as these two stimuli only differ in the exact proportion of dots moving in the two directions. Newsome and colleagues have therefore argued that the trial-to-trial

variability in firing rates in areas such as MT and MST (even when the exact same stimulus is presented) is the basis for the animal's decision and accounts for the observed correlation between neuronal and psychophysical behavior. This interpretation is strengthened by the finding that microstimulating directional cells in area MT creates a bias in the monkey's perceptual report towards the stimulated neurons' preferred direction [17]. However, the influence of feature-based attention could provide an alternative interpretation of the link between neuronal and behavioral performance. On this view, the animal's shift of attention towards one of the two possible directions (or depth planes in the cylinder experiments) would enhance the response of neurons preferring this direction (or depth).

In summary, there is converging evidence that as one ascends through the hierarchy of areas in the visual cortex there is a transition from a neural representation dominated by the sensory input to a representation emphasizing the input's perceptual interpretation and the animal's behavioral state. This gradient is apparent using several experimental paradigms but it remains an open issue whether independent mechanisms (for the resolution of depth ambiguities, of directional ambiguities and for decisions between opposite choices in high-noise conditions) contribute to it, or if feature-based attention is the major and unifying cause of the modulating mechanism.

References

- 1 Kaas, J.H. (1989) Why does the brain have so many visual areas? *J. Cogn. Neurosci.* 1, 121–135
- 2 Williams, Z.M. *et al.* (2003) Parietal activity and the perceived direction of ambiguous apparent motion. *Nat. Neurosci.* 6, 616–623
- 3 Van Essen, D.A. and Maunsell, J.H.R. (1983) Hierarchical organization and functional streams in the visual cortex. *Trends Neurosci.* 6, 370–375
- 4 Felleman, D.J. and Van Essen, D.C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47
- 5 Ungerleider, L.G. and Mishkin, M. (1982) Two cortical visual systems. In *Analysis of Visual Behavior* (Ingle, D. *et al.*, eds), pp. 549–586, MIT Press
- 6 DeYoe, E.A. and Van Essen, D.C. (1988) Concurrent processing streams in monkey visual cortex. *Trends Neurosci.* 11, 219–226
- 7 Parker, A.J. and Newsome, W.T. (1998) Sense and the single neuron: Probing the physiology of perception. *Annu. Rev. Neurosci.* 21, 227–277
- 8 Bradley, D.C. *et al.* (1998) Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature* 392, 714–717
- 9 Dodd, J.V. *et al.* (2001) Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J. Neurosci.* 21, 4809–4821
- 10 Parker, A.J. *et al.* (2002) Neuronal activity and its links with the perception of multi-stable figures. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357, 1053–1062
- 11 Treue, S. and Martinez Trujillo, J.C. (1999) Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399, 575–579
- 12 McAdams, C.J. and Maunsell, J.H.R. (2000) Attention to both space and feature modulates neuronal responses in macaque area V4. *J. Neurophysiol.* 83, 1751–1755
- 13 Treue, S. and Maunsell, J.H.R. (1999) Effects of attention on the processing of motion in macaque visual cortical areas MT and MST. *J. Neurosci.* 19, 7603–7616
- 14 Cook, E.P. and Maunsell, J.H.R. (2002) Attentional modulation of behavioral performance and neuronal responses in middle temporal and ventral intraparietal areas of macaque monkey. *J. Neurosci.* 22, 1994–2004
- 15 Newsome, W.T. *et al.* (1989) Neuronal correlates of a perceptual decision. *Nature* 341, 52–54
- 16 Celebrini, S. and Newsome, W.T. (1994) Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J. Neurosci.* 14, 4109–4124
- 17 Salzman, C.D. *et al.* (1990) Cortical microstimulation influences perceptual judgements of motion direction. *Nature* 346, 174–177

1364-6613/\$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tics.2003.09.003

Wholes, holes, and basic features in vision

James R. Pomerantz

Department of Psychology, Rice University, Houston, Texas 77005, USA

A key issue for theories of perception is specifying the primitives used by the visual system to isolate and identify the objects in an image. Although local features are typically suggested, there is good reason to look for global, configural features as primitives too. Chen *et al.*'s specific proposal of topological features is both explicit and capable of capturing important global information. It may seem surprising that topology can be detected by honeybees, but Chen's results are in keeping with other findings from humans that global properties are sometimes perceived better than local ones and thus might be basic.

The visual system informs organisms about their surroundings, but how does it sort out the vast quantity of information contained in an image? When facing a new scene, where does the visual system start? Most theories of vision begin with local features, such as oriented line segments, and build up from there, using attention to integrate features [1,2]. We have known, at least since the Gestalt psychologists however [3], that global properties often dominate perception. A recent paper by Lin Chen and his colleagues [4] suggests an intriguing method by which global properties might be detected early in the sequence.

The centrality of grouping

It seems sensible that the visual system must segment a full scene into grouped regions before it attempts to

Corresponding author: James R. Pomerantz (pomeran@rice.edu).