Image segmentation: **A tug-of-war for the eyeball**
S. Treue* and U.J. Ilg†

Separating objects from their background is one of the central abilities of the visual system. Recent evidence has revealed how populations of neurons, some of which have receptive fields with an antagonistic center–surround structure, and some of which do not, might contribute to this ability.

Addresses: *Cognitive Neuroscience Laboratory, *†Department of Cognitive Neurology, University of Tübingen, Auf der Morgenstelle 15, 72076 Tübingen, Germany.
E-mail: treue@uni-tuebingen.de, uwe.ilg@uni-tuebingen.de

The image of our environment projected onto our retinas is encoded by a mosaic of photoreceptors. While this might seem to be not very different from the raster image created by a scanner, our perception is not one of an unstructured bitmap of pixels of different color and luminance, but rather of a world of objects. The ubiquity and apparent ease with which our visual system segments its input into objects and background belies the difficulty of the task, which is in fact one of vision's greatest challenges.

This 'figure–ground' segmentation is based on differences between the object and its background. In real life situations, these differences can occur across a plethora of visual dimensions. An object and its background might differ in their brightness, texture, distance from the observer, color, motion and/or many other features. In all cases though, the visual system has to perform a comparison to find such object boundaries. This requires detecting edges — lines of sudden changes in one or several particular features across space. Because such edges can occur along any visual dimension, and because figure–ground segmentation is of such central importance for vision, mechanisms for figure–ground segmentation are likely to be widespread in the visual system.

Motion is a particularly strong cue for figure–ground segmentation. We use the relative motion between an object and its background, not only for segmenting the object, but also for estimating its velocity so that we are able to make saccades (fast foveating eye movements) and for determining the speed of the pursuit eye movement needed to stabilize the object on the retina. It has been suggested that the receptive field structure of direction-selective neurons in area MT of primate visual cortex provides the neural basis for motion-based figure–ground segmentation. MT is a much-studied area that is specialized for the processing of visual motion information by its high density of direction-selective neurons. Many of these neurons have receptive fields with an antagonistic center–surround structure: a stimulus that extends beyond the classical receptive field will impinge on a surround that is generally inhibitory, reducing the response of the neuron. This means that large field motion stimuli evoke little or no response at all from such neurons.

The most potent stimulus for these local motion neurons is, in fact, often movement in the preferred direction inside the classical receptive field and the opposite (anti-preferred) direction outside the classical receptive field. Such neurons seem well suited to signal the presence of small moving objects, but the motion of the background, which often covers large portions of the visual field, cannot be signaled by these neurons. This ability might rest with the MT cells that have another type of receptive field: these 'wide field' neurons respond best to large fields of motion with no apparent drop in response for stimuli extending well beyond their classical receptive field. Besides encoding background motion, these neurons might also contribute to the perception of 'induced motion', a phenomenon where a stationary spot embedded in a moving background is often perceived as moving in the direction opposite to the background [1].

The hypothesis that object and background motion are encoded and signaled by different subpopulations of neurons has now received direct behavioral support. In a recent study Born *et al.* [2] have exploited a feature of these two neuronal subpopulations that was described several years ago by Born and Tootell [3]. Using the 2-deoxyglucose labeling technique, they were able to demonstrate that global and local motion neurons form distinct stripes in area MT. This means that any given site within MT is dominated by neurons with one or the other receptive field type. This allows the selective activation of one subpopulation of neurons by inserting a stimulating microelectrode into either a global or local motion site. A similar microstimulation approach has been used very successfully in a number of studies exploiting the local patchiness of neuronal subpopulations preferring different directions of motion.

To investigate how the signals of local and wide field neurons might contribute to figure–ground segmentation based on motion, Born *et al.* [2] exploited the important role that motion information plays in the accurate planning of saccadic and pursuit eye movements. When planning a saccade to a moving peripheral target, the visual system

Figure 1
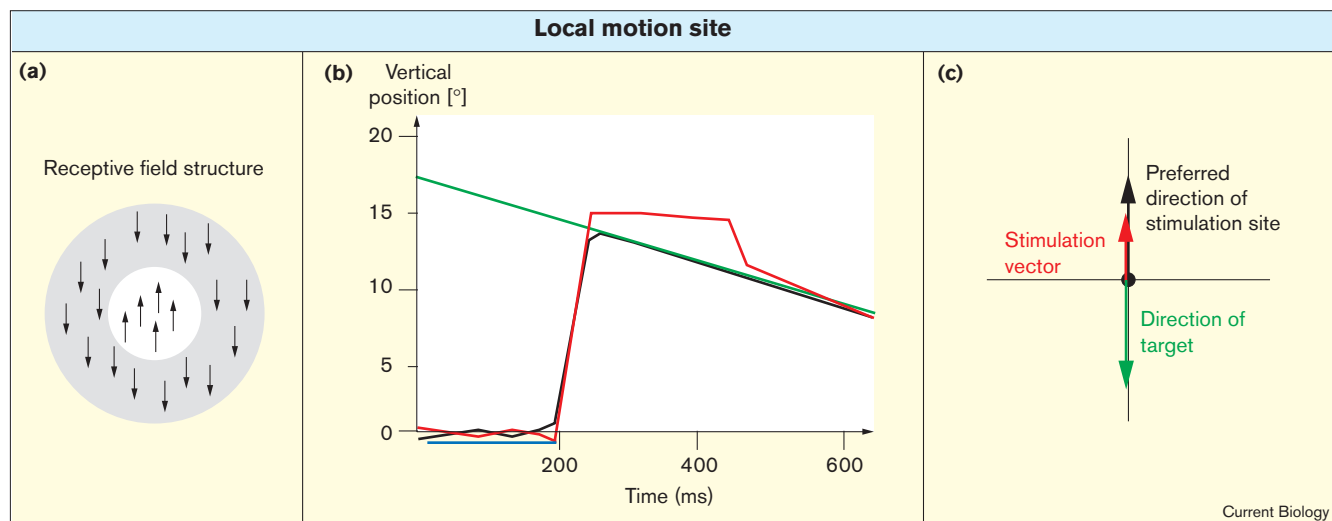
Effects of microstimulation at a local-motion site. **(a)** The receptive field structure of a local motion neuron that prefers upward motion in its classical receptive field (inner circle). The gray annulus represents a surround that is generally inhibitory, but sometimes enhances responses to the central stimulation in response to motion in the center's anti-preferred direction. **(b)** Idealized plot of target (green) and eye position in a control (black) and microstimulated (red) step-ramp trial. The target appeared above the initial fixation point and moved towards it, the anti-preferred direction Under these conditions, stimulation at local motion sites leads to an incomplete compensation for the target's displacement during the delay between its appearance and the saccade, and a pursuit velocity that was too slow (visible in the shallower slope of the red eye position trace) and required a corrective saccade. (In about half of all sites, microstimulation resulted in a non-directional slowing of pursuit velocity; these sites were excluded from further analysis.) **(c)** Vectors representing target movement (green), preferred direction of a site (black), and hypothesized effect of stimulation (red), respectively. At local motion sites, microstimulation introduced a 'stimulation vector' in the sites' preferred direction. Averaging this vector with the target-induced vector led to an underestimation of target velocity that was strongest when the target moved in a site's anti-preferred direction (as shown in this example). The initial saccade therefore underestimated the target's *anti-preferred* movement and the post-saccadic pursuit was too slow.

takes the direction and velocity of motion of that target into account. This allows for an appropriate pursuit velocity immediately after the saccade, and for an accurately sized eye movement that is adjusted for the target's displacement during the unavoidable delay between the signal acquisition and the saccade execution (see the black eye position traces in Figures 1b and 2b).

By determining the adjustment an observer makes in the initial pursuit velocity and in the saccade as a function of the direction and velocity of a moving target, researchers can thus behaviorally assess motion processing in a manner untainted by the execution of a motor program. In their study, Born *et al.* [2] microstimulated area MT in monkeys performing a classic step-ramp paradigm. In this set up, the animal is trained to fixate a small dot. A target spot will then appear eccentrically and move in one of several directions and speeds. The animal's task is to foveate this moving target and maintain pursuit. This design allows the presentation of a pursuit target inside non-foveal receptive fields. The neurons in the receptive fields of which the target appears are the ones supplying the information about the target's and background's direction and speed.

In their experiments, Born *et al.* [2] placed a microelectrode into MT and established the local receptive field type and preferred direction. They then applied electrical stimulation on half of the trials from the time of target onset to the time the animal made the initial saccade to the target — that is, in the period when the eccentric target's velocity is estimated. They could now establish the influence of the stimulation on the eye movements. At about half of the sites in MT, the effect was a non-directional general slowing of pursuit. For the remaining sites, just as predicted by the hypothesis outlined above, when stimulating local motion cells the fastest pursuit velocities were observed when the target moved in the preferred direction of the neurons. Stimulating wide field sites had the opposite effect: in these cases, the fastest pursuit was observed in the anti-preferred direction of the stimulated neurons. This likely reflects the influence of the induced motion effect that creates a motion signal in the direction opposite to that of the background.

The effects Born *et al.* [2] observed suggest that the microstimulation created a velocity vector that was combined with the physical target's velocity vector. Rather than following one or the other of these two vectors — a winner-take-all

Figure 2
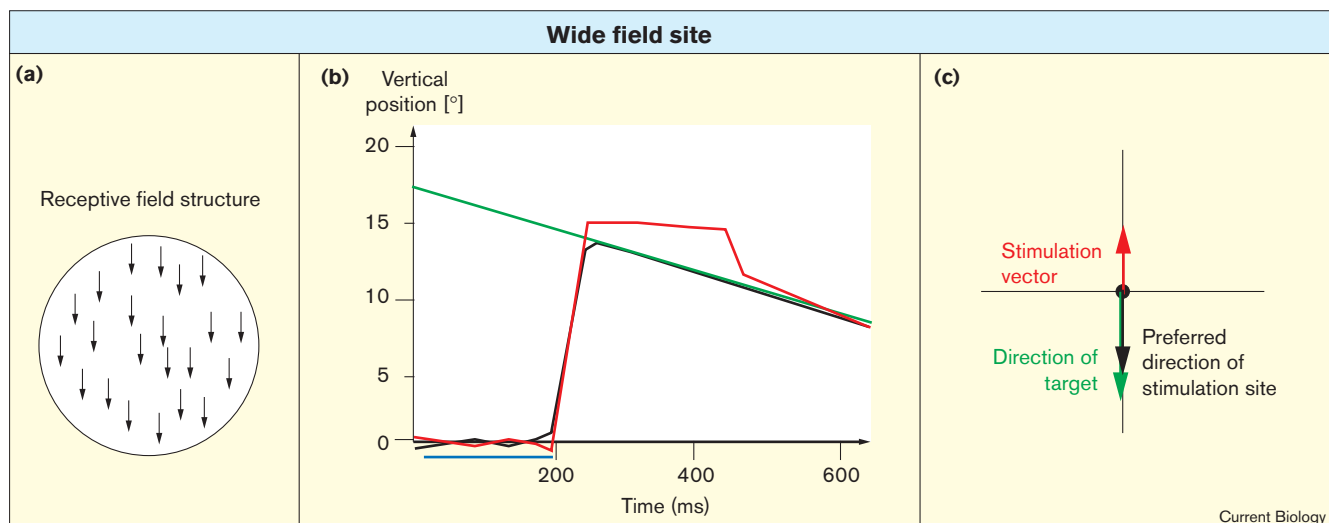
Effects of microstimulation at a wide-field site. **(a)** The receptive field structure of a wide-field motion neuron preferring downward motion. The classical receptive field is indicated by the circle; extending a stimulus beyond these boundaries had no effect on responses, indicating the absence of an inhibitory surround. **(b)** Idealized plot of target (green) and eye position in a control (black) and microstimulated (red) step-ramp trial. Trial conditions and results were the same as those in Figure 1, but note that here the target motion is in the sites' preferred direction. **(c)** Vectors representing target movement (green), preferred direction of the site (black), and hypothesized effect of stimulation (red), respectively. At wide-field sites, microstimulation introduced a 'stimulation vector' in the sites' anti-preferred direction, presumably because these sites signal background motion in the preferred direction. Averaging this vector with the target-induced vector led to an underestimation of target velocity that was strongest when the target moved in a site's preferred direction. Therefore the initial saccade underestimated the target's *preferred* movement and the post-saccadic pursuit was too slow.

mechanism — the oculomotor system seems to average them together. This is in agreement with previous evidence for vector averaging in the initiation of pursuit eye movements [4–6], even when stimuli combining first-order and second-order motion components are pursued [7].

The interaction between background motion and pursuit is complex and various studies have reported a range of effects, most likely reflecting differences between the paradigms used. Born *et al.* [2] therefore conducted a control experiment in which they simply replaced the microstimulation with a period of actual large-field background motion, and compared the effect with that of applying stimulation at the wide-field sites. In both cases, pursuit was shifted in the direction opposite to that of the background motion, either real or stimulated. These results are consistent with the hypothesis that wide-field neurons code background motion and that their activation results in a target motion signal in the opposite direction from the real or simulated background motion.

The visual signal encoded in the retinas contains many dimensions that are processed in a large set of visual areas in primate cortex. In the past, systems neuroscience has concentrated on accounting for visual abilities by single types of neurons that could be the basis of a given task. But often signals need to be combined, either across

dimensions or across space, to derive new information. Therefore, as our understanding of information processing in primate visual cortex has grown, the focus of interest has increasingly shifted towards more complex visual tasks that cannot be attributed to single neurons but require populations of neurons [8].

To go further, and show how neurons with *different* properties are combined to provide a population account for such abilities is a major step in system neuroscience. Born *et al.*'s [2] elegant approach of microstimulating defined neuronal subpopulations and behaviorally assessing the effects on sensory signal processing is such a major step. They have provided strong evidence of how two subpopulations of neurons that differ in their receptive field structure contribute two complementary sets of information and interact in the segmentation of moving objects from the background. Applying such an approach across visual cortex should spawn further progress in our understanding of complex sensory information processing.

References
1. Duncker K: **Über induzierte Bewegung.** *Psychologische Forschung* 1929, **12**:180-259.
2. Born RT, Groh JM, Zhao R, Lukasewycz SJ: **Segregation of object and background motion in visual area MT: effects of microstimulation on eye movements.** *Neuron* 2000, **26**:725-734.
3. Born RT, Tootell RBH: **Segregation of global and local motion processing in primate middle temporal visual area.** *Nature* 1992, **357**:497-499.

4. Lisberger SG, Ferrera VP: **Vector averaging for smooth pursuit eye movements initiated by two moving targets in monkeys.** *J Neurosci* 1997, **17:**7490-7502.

5. Recanzone GH, Wurtz RH: **Shift in smooth pursuit initiation and MT and MST neuronal activity under different stimulus conditions.** *J Neurophysiol* 1999, **82:**1710-1727.

6. Groh JM, Born RT, Newsome WT: **How is the sensory map read out? Effects of microstimulation in visual area MT on saccades and smooth pursuit eye movements**. *J Neurosci* 1997, **17:**4312-4330.

7. Lindner A, Ilg UJ: **Initiation of smooth pursuit eye movements elicited by first and second-order motion stimuli.** *Exp Brain Res* 2000, in press.

8. Treue S, Hol K, Rauber HJ: **Seeing multiple directions of motion – physiology and psychophysics.** *Nat Neurosci* 2000, **3:**270-276.