



MULTIPLE USES OF VISUAL MOTION. THE CASE FOR STABILITY IN SENSORY CORTEX

R. J. A. VAN WEZEL¹ and K. H. BRITTEN*

Center for Neuroscience and Section of Neurobiology, University of California at Davis, 1544 Newton Court, Davis, CA 95616, USA

Abstract—The problem of ‘readout’ from sensory maps has received considerable attention recently. Specifically, many experiments in different systems have suggested that the routing of sensory signals from cortical maps can be impressively flexible. In this review, we discuss many of the experiments addressing readout of motion signals from the middle temporal area (also known as V5) in the macaque monkey. We focus on two different types of output: perceptual reports (categorical decisions, usually) and motion-guided eye movements. We specifically consider situations in which multiple-motion vectors present in the stimulus are combined, as well as those in which one or more of the vectors in the stimulus is selected for output. The results of these studies suggest that in some situations multiple motions are vector averaged, while in others multiple vectors can be maintained. Interestingly, in most of the experiments producing a single (often average) vector, the output is a movement. However, many perceptual experiments involve the simultaneous processing of multiple-stimulus motions. One prosaic explanation for this pattern of apparently discrepant results is that different downstream structures impose different rules, in parallel, on the output from sensory maps such as the one in the middle temporal area. We also specifically discuss the case of motion opponency, a specific readout rule that has been posited to explain perceptual phenomena such as the waterfall illusion (motion aftereffect). We present evidence from a recent experiment showing that an opponent step must occur downstream from the middle temporal area itself. This observation is consistent with our proposal that significant processing need occur downstream from sensory structures. If a single output is to be used for multiple purposes, often at once, this necessitates a degree of task invariance on the sensory information present even at a relatively high level of cortical processing. © 2002 IBRO. Published by Elsevier Science Ltd. All rights reserved.

Key words: vision, attention, cortex, MT, motion, eye movements.

It has become popular in recent years to emphasize the dynamic operations of the nervous system. As our knowledge of such complexities as neuromodulators and network dynamics grows, it is tempting to employ such complicated tools as cornerstones of the basic operations of the nervous system. In sensory systems, it is easy to make the representations ever more powerful, so that more and more of the impressive behavioral flexibility of animals directly arises from the dynamic and flexible operations of their sensory apparatus. Receptive fields (RFs) change their dimensions, attention or task demands re-route information, and evanescent cell assemblies carrying essential information form and dissolve in tens or hundreds of milliseconds. In this paper, it will be our contention that sensory representations are relatively stable, even when task demands vary, and to make this case we will discuss a series of experiments which specifically target the question of ‘readout’ from sensory representations. By readout, we refer to the man-

ner in which sensory signals are transformed into perceptual decisions or overtly observed behavior. Obviously, such readout appears on the surface to be impressively flexible, since the same sensory stimuli can be used for diverse purposes. From this observation two related questions arise: first, where do changes occur when we change the behavioral context or task demands? Second, do the basic interactions between sensory signals, at the level of sensory cortex, change when they are read out in different ways?

In this paper, we will discuss the question of readout using visual motion as a model system. It forms a good example because the same visual motion areas in the extrastriate cortex of monkeys are involved in different behavioral outputs. We will mostly focus on one visual motion-related area in dorsal extrastriate cortex, the middle temporal area (MT, also known as V5). This key component of the so-called ‘motion system’ of dorsal extrastriate cortex has been studied extensively, using a variety of approaches. MT probably forms the best case for linkage between known cortical structures and well-defined behavioral tasks (for recent reviews, see Albright and Stoner, 1995; Parker and Newsome, 1998; Snowden, 1994; Zeki, 1990). Area MT (and its neighbor, the medial superior temporal area, MST) is involved in forming categorical decisions about stimuli (Newsome and Paré, 1988; Orban et al., 1995; Pasternak and

¹ Present address: Helmholtz Institute, Utrecht University, Neuroethologie Station, Limalaan 30, 3584 CL, Utrecht, Netherlands.

*Corresponding author. Tel.: +1-530-7545080.

E-mail address: khbritten@ucdavis.edu (K. H. Britten).

Abbreviations: MAE, motion aftereffect; MST, medial superior temporal; MT, middle temporal; RF, receptive field.

Merigan, 1994; Salzman et al., 1992) and also in the guidance of eye movements (Groh et al., 1997; Komatsu and Wurtz, 1988; Lisberger and Movshon, 1999; Newsome et al., 1985). By our focus upon one area, we do not discount other areas, which work together with MT to process visual motion, but we use MT as a 'model system' for the extraction of behaviorally relevant information from well-defined cortical maps.

We will use the tight relationship between the visual motion system and motion-dependent tasks, combined with key results from recent literature, to make a simple case. We propose that cortical motion signals from MT are used in a stable and hard-wired manner, although behavior or perception based on motion information can be highly task-dependent. We will make what we believe to be a testable, simplifying proposal on the readout from sensory maps. Put simply, we believe that parallel outputs downstream of MT use distinct readout mechanisms, each geared to the corresponding purpose. Thus, the motor system might average its sensory inputs,

since only a single motion can be made at time, while circuits for making perceptual decisions might instead use a mutually competitive mechanism to reach a categorical decision (e.g. left or right?). Under this scheme, what changes when an animal changes contexts is simply the engagement of different output circuitry, and the basic flow of information from sensory to premotor structures can remain relatively stable. We will also argue that this parallel architecture constrains the amount of signal combinations that can take place on cortical structures before the branch point. We think MT is such a structure, and will make the case using examples from this well-studied area.

In the following sections, we first introduce the motion system ('The nature of the motion map in area MT'), then move towards experiments that specifically target the question of readout ('Reading out a motion map'), including one recent experiment from our laboratory ('Adaptation points to opponency after MT'), and then flesh out our proposal in more specific terms ('The case for stability in sensory representations'). Readers familiar with the visual motion system should feel free to skip ahead to 'Reading out a motion map'.

THE NATURE OF THE MOTION MAP IN AREA MT

The motion areas of extrastriate cortex have been extensively studied over the last 20 years; it is not our purpose here to provide a detailed review of this body of work. Instead, we will focus on the key issues germane to the question of how visual information is extracted from MT. Specifically, we will discuss why we can think of MT as a purely visual sensory area, and discuss the behavioral contexts in which MT has been studied. We will briefly discuss the evidence that documents a causal role for MT in the processing of visual motion; this sets the stage for our main interest in the mechanisms underlying this role. In our discussion of MT, we will attempt to point out concepts which are potentially confusing. Some operations might occur either 'early' (in the formation of directional responses) or 'late' (in the process by which decisions or movements are made from the directional information).

In Fig. 1 we illustrate the general context of 'readout' as it applies to MT, in very general terms. We separate the processing that leads to the formation of MT directional responses (upper dashed box) from the process by which these might be converted to perceptual decisions or used for the guidance of motor acts like pursuit eye movements (lower dashed box). This schematic is not intended to be anatomically precise, but instead illustrates the concepts we will be emphasizing. Specifically, we suggest a hierarchy of function, where the processes that form even fairly high-level sensory representations are largely separate from the mechanisms by which these signals are put to use. One simplification implicit in this schematic is the absence of functionally important feedback. We recognize this omission, and later we explicitly consider the importance of 'top-down' influences on sensory processing. In our view, however, these influences

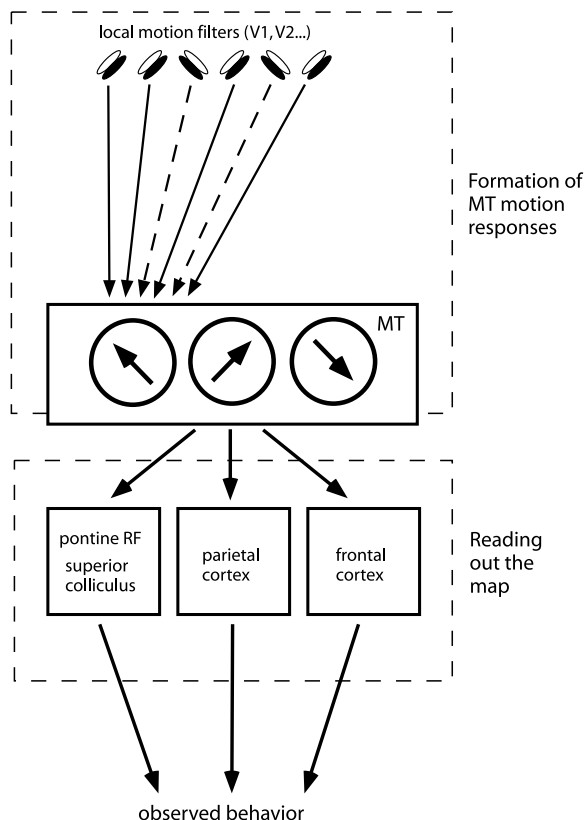


Fig. 1. Schematic of MT inputs and outputs. The dominant set of inputs comes from directionally selective complex cells in V1 (Carandini et al., 1997), and we indicate these connections with the upper set of arrows, coming from the upper icons. The icons indicate the space-time orientation of the filters in V1 (or elsewhere). We also suppose there to be direct or indirect inhibitory connections from local motion filters preferring the opposite direction (oppositely oriented RFs), indicated by the dashed arrows. Below MT, we illustrate the widely divergent outputs of MT, which project to a large number of target structures (only a subset of possible target structures is shown). The dashed-outline box schematically illustrates what we are calling the 'readout mechanism' that ultimately results in observed behavior.

are relatively simple in nature, and do not fundamentally reorganize the motion map in MT.

The generation of MT directionality

MT was originally defined in the owl monkey as a myelin-rich area in temporal cortex receiving strong, direct projections from primary visual cortex (Allman and Kaas, 1971; Dubner and Zeki, 1971). The connections from V1 are of the 'feedforward' type (Felleman and Van Essen, 1991; Maunsell and Van Essen, 1983a). Given the intensity of study of MT, it is remarkable how little is known about the detailed circuitry generating directionality in MT cells. MT cells receive input not only from V1, but also from other prestriate areas (V2, V3A) and from the pulvinar (Stepniewska et al., 1999). Both direct and indirect evidence (Maunsell and Van Essen, 1983a; Movshon and Newsome, 1996; Newsome et al., 1985; Ungerleider and Mishkin, 1982) points to the idea that MT inherits many of its response properties from its afferents.

In many respects, MT cells resemble spatially extended, directionally selective complex cells (like those that are observed in primary visual cortex). They are insensitive to spatial phase and sign of contrast, respond to moving gratings with unmodulated discharge, and often prefer spectrally broadband stimuli to narrow-band ones. Their RF diameters are about equal to their eccentricity, meaning that they are about 10 times the diameter of their V1 afferents, but only about twice that of their V3A afferents (Gattass et al., 1997). Thus, considerable spatial pooling occurs in the formation of MT cell RFs. In addition, they pool across spatial and temporal frequency; this might occur in a way that makes them represent 'true' velocity, irrespective of the combination of spatial and temporal frequency (Movshon et al., 1988). Presumably as a consequence of this extensive summation, MT cells are quite broadly tuned, on average, for stimulus direction, speed and location.

MT cells are also organized into a reasonably regular map of RF location in retinotopic coordinates (Maunsell and Van Essen, 1983b), of stimulus direction (Albright, 1984), and probably of stereoscopic depth as well (DeAngelis and Newsome, 1999). These maps are not nearly as precise and 'crystalline' as their counterparts in earlier areas, but none the less are regular and reliable.

All of this has led to the simplistic working model of MT as an intermediate-level representation of motion: a crude retinotopic map of local velocity, where individual estimates of direction provided by much more local elements in earlier stages are combined into a map of local 'object motion' at a medium level of spatial resolution. This oversimplified view is mostly consistent with available data, fits computational models of motion perception, and is adequate for discussing how such a representation might be employed in supporting tasks which depend on the analysis of local object motion. Before we consider experiments which most directly reveal the manner in which these signals are being used, we need to consider some of the more complex

aspects of the formation of MT responses. These include evidence for interactions between directions of motion, and state- or task-dependent modulation of MT cell responses.

Directional interactions in MT

Motion opponency. As alluded to above, MT apparently combines rather broadly across its inputs. What is the nature of this combination? The first clue can be found in the physiology of MT. MT cells are not only excited by motion in their preferred direction, but they are also inhibited by the opposite ('null') direction of motion. On average, response increments for motion in the preferred direction are about four times the magnitude of response decrements for motion in the null direction (Britten et al., 1993; Snowden et al., 1991). This is seen in the responses of the MT neuron shown in Fig. 2A. In this experiment, the strength of the motion is varied by adjusting the coherence of a moving random dot pattern. Preferred direction strength increases to the right, while null direction strength increases to the left. It can be seen that the slope of the preferred direction side is much higher, indicating greater sensitivity. This phenomenon is also of interest for the concept of motion opponency: opponent models of motion analysis predict this sort of behavior (Adelson and Bergen, 1985; Heeger, 1987; Simoncelli and Heeger, 1998; van Santen and Sperling, 1985).

These antagonistic responses to opposite motion directions also reveal themselves when two oppositely moving stimuli are simultaneously presented within the RF. For instance, two spatially overlapping random dot patterns moving in opposite directions are perceived as two transparent surfaces (hence the common term 'transparent motion'). [For a more extended discussion on transparent motion, and its relationship to MT, we refer to a recent review by Braddick (1997).] Andersen et al. have studied intensively transparent motion and the conditions under which preferred and null directions interact (Bradley et al., 1998; Qian and Andersen, 1994; Qian et al., 1994a, b; Snowden et al., 1991). The results are all consistent with the idea that the preferred and null directions of motion are brought together in some antagonistic manner at the level of MT. We discuss this issue further in 'Three readout combination rules'.

Responses to plaids. In addition to transparently moving random dot patterns, physiologists and psychophysicists interested in the interaction between different motion directions have frequently studied moving 'plaid'. These stimuli are particularly interesting because simple parametric manipulations (of spatial frequency or contrast, for example) often lead to qualitatively distinct percepts. Two superimposed gratings that drift in different directions can either appear to slide freely past each other or, under other conditions, cohere and give the impression of one plaid moving in one direction (Adelson and Movshon, 1982; Movshon et al., 1985; Stoner and Albright, 1993). Single-cell recordings in MT show that about a quarter of the cells respond to

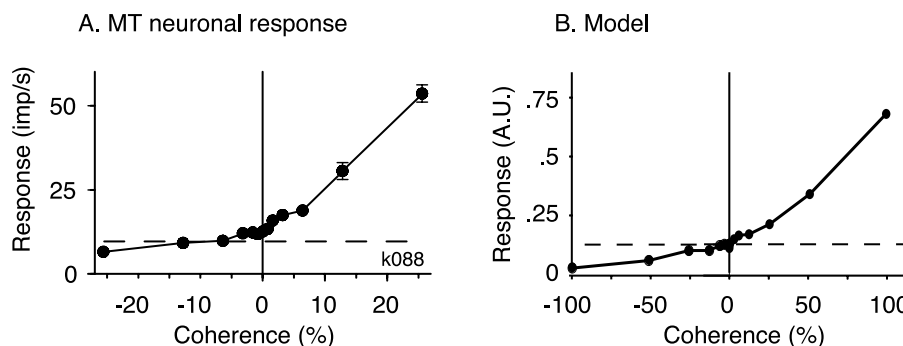


Fig. 2. Neuronal and model responses showing motion opponency to variable strength motion stimuli. The x -axis shows the strength and direction of motion in a dynamic random dot display. The strength is given by a parameter we call the coherence of the display; it is the proportion of dots carrying a specified direction signal. Motion in the preferred direction is indicated by positive coherence values, while motion in the null direction is indicated by negative values. Neuronal response is indicated in units of spikes/s, averaged over the entire 2-s stimulus duration. The slope of these response functions indicates the sensitivity to the corresponding direction of motion. The negative slopes in the null direction show that the neuron's response is declining with increasing stimulus strength. This is probably but not certainly active inhibition, since the response in the null direction drops below the maintained activity (indicated by the horizontal dashed line). (A) Typical neuron, which shows a much higher slope in the preferred direction than in the null direction, which can be called unbalanced opponency. [Adapted from Britten et al. (1993).] (B) Normalization model predictions (Simoncelli and Heeger, 1998) of the results of the same experiment. This model contained explicit opponency between opposed directions of motion, derived from the weights of inputs from earlier motion selective areas. In addition, a normalization step at MT rescaled the responses. [Adapted from Simoncelli and Heeger (1998) with permission from Elsevier Science.]

the direction of the coherently moving plaid pattern, rather than to the directions of the two gratings that create the pattern (Movshon et al., 1985; Rodman and Albright, 1989). However, in V1, all directionally selective neurons responded to the motion of the components, rather than the plaid direction (see Fig. 3A, B; compare to Fig. 3E, F). Furthermore, MT neurons that respond to plaids, respond to its components under exactly the same conditions that are known to cause humans to perceive two separate gratings (Stoner and Albright, 1992).

This family of observations supports two general conclusions. First, something about the way directions are combined at the level of MT allows these cells to recognize the pattern direction, instead of the components (as do prior stages). Second, the similarity of the conditions under which coherence of plaids occurs for MT cells and for human perception suggests that the combination that occurs at the level of MT is sufficient to 'explain' perceptual coherence. Perhaps the processing at or before the level of MT is all that is needed. However, the fact that subsequent processing is not necessary is not proof of its non-existence. At present, these arguments have been left at the qualitative-similarity level, and detailed, quantitative measurement combined with precise definition of the decision rule underlying appearance judgments about plaid coherence would be necessary to reinforce this conclusion.

Averaging and normalization. Several recent experiments have addressed the manner in which MT responds to multiple stimuli, and the common observation is that individual cells produce responses that most strongly resemble an average of the responses to the individual stimuli (Recanzone et al., 1997; Snowden et al., 1991). This appears to be the case for both small moving objects and extended dot fields. Where tested, the averaging does not appear to strongly depend on the individ-

ual directions of the moving stimuli, or on their spatial arrangement within the RF of the cell being tested. However, recent measurements suggest that at least under some conditions, the averaging is not linear, but contains a non-linearity favoring the more effective of two stimuli (Britten and Heuer, 1999). This operation is similar to what Nowlan and Sejnowski call a 'soft-maximum' operation, which can be considered a weak version of a winner-take-all operation (Nowlan and Sejnowski, 1995).

The observations of response averaging, as well as seemingly paradoxical observations like transparent motion perception, in fact are consistent with a single mechanism: response normalization. A recently published model based on a relatively simple architecture by Simoncelli and Heeger (1998) can largely account for the different response properties of MT neurons. Their MT model consists of two stages corresponding to visual areas V1 (local motion elements) and MT (a velocity map). The key feature of the model, that allows it to fit the above-mentioned experimental results (and others as well) is normalized, weighted summation. The first step is a weighted summation, where local motions consistent with a particular velocity are added together, and those consistent with the opposite direction are subtracted. This sum is then normalized by a value corresponding to the total activity. This last step, which is useful in keeping the representation of motion stable in the face of different image contrasts, turns out to be sufficient to account for some of the apparent peculiarities in the responses of MT cells to complex stimuli.

We show the predicted responses of the Simoncelli model to both simple and complex motion stimuli. In Fig. 2B, we show how it responds to random dot fields of varying coherence, and compare this against the response of a typical MT cell. The main features of the model responses are entirely within the range of normal

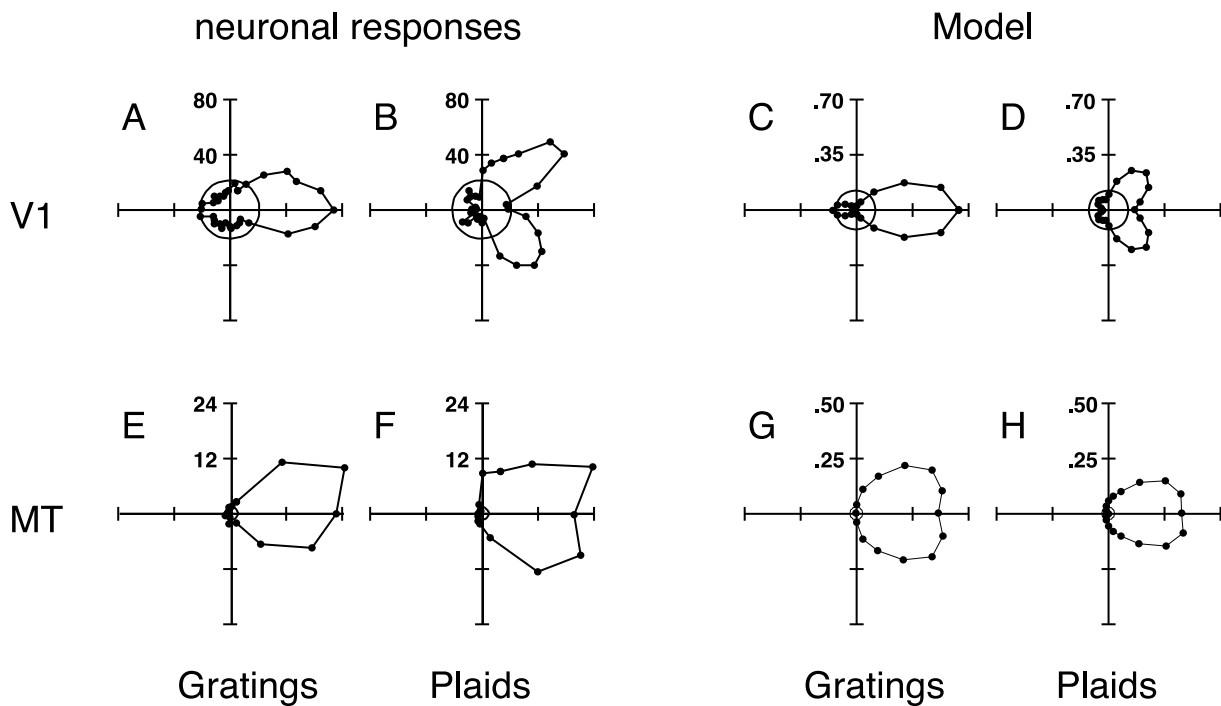


Fig. 3. Data and model predictions for responses to grating and 'plaid' stimuli, or compound stimuli consisting of two superimposed gratings oriented 90° from each other. Data are shown in A, B, E and F (Movshon et al., 1986); model simulations are shown in C, D, G and H (Simoncelli and Heeger, 1998). Each column labeled 'gratings' shows the responses to an individual grating moving in different directions, and the columns labeled 'plaids' depict the responses when the pattern direction was the same; each component was thus moving in a direction 45° away. The V1 cells show bi-lobed responses to the plaids, indicating that they 'see' each of the components in isolation. However, the MT cells respond to the pattern direction, indicating that the two directions of motion have been combined in some way. The model shows the same kind of response to the plaid (H), with a relatively simple summation algorithm to extract the pattern motion. [Adapted from Simoncelli and Heeger (1998) with permission from Elsevier Science.]

MT cell behavior to this class of stimuli, showing that it accounts for these data quite well. While we do not show it, the model also does a good job predicting the increasing responses of MT cells as the stimulus strength is changed in other ways as well, for example by raising the contrast of a sine-wave grating (Sclar et al., 1990), or the number of dots in a dot display (Snowden et al., 1991).

The model also accurately predicts the responses to more complex stimuli containing multiple-stimulus motions, such as the plaid stimuli discussed in 'Responses to plaids'. In Fig. 3, we show the model's responses compared against neuronal responses in V1 and in MT. Neuronal responses are shown in the left four panels, while the model predictions are shown in the right four panels. Both the example MT cell (Fig. 3F) and the model (Fig. 3H) respond with a peak when the plaid pattern moved in the cell's preferred direction, rather than when the individual components did (contrast with the V1 cell). Thus, weighted summation allows the model to 'see' these stimuli like MT cells, which in turn is more in accordance with human perception of these stimuli.

We have emphasized this model because of the relatively simple feature of contrast normalization that appears to predict, counterintuitively, some of the apparent peculiarities of MT cell responses. Moreover, it

shows that profound differences in motion percept need not be caused by task-dependent changes in either motion processing or readout mechanisms, but can more easily be explained by fixed (if non-linear) operations at the level of the sensory representation. We will return to this idea in 'Perceptual experiments addressing readout'.

Extra-retinal modulation of MT activity. Neural activity in MT, and other extrastriate areas, is not only affected by visual sensory input, but also depends on the animal's behavioral state and other extra-retinal factors (for recent review, see Maunsell and Ferrera, 1993). We must consider this carefully, because of the important distinction between task-dependent changes in sensory processing (activity in MT, in our case) and task-dependent changes in the way that information is subsequently dealt with. Can these two be distinguished? We believe so, and most of the rest of this paper considers various forms of evidence that address this question.

Numerous human psychophysical studies show attention effects on motion processing (Cavanagh, 1992; Chaudhuri, 1990; Gogel and Sharkey, 1989; Lankheet and Verstraten, 1995; Shulman, 1993; Watanabe et al., 1998). Indeed, directed attention can produce a percept of motion (Ashida and Verstraten, 1997; Cavanagh, 1992; Culham and Cavanagh, 1994; Shimojo et al.,

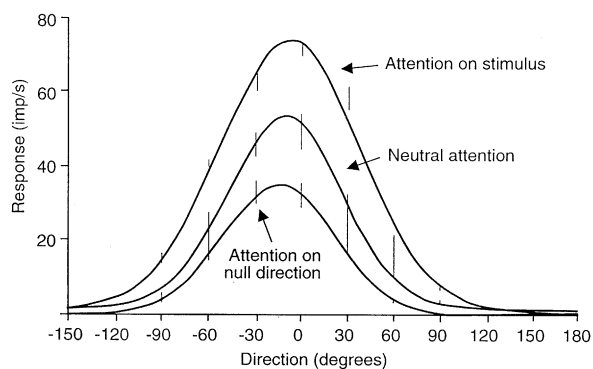


Fig. 4. Modulation of MT cell responses by directed attention. The monkey was presented two stimuli within the RF of a single MT cell, and one was always moving in the cell's null direction. If attention was directed to null direction (lower curve), responses were attenuated. However, if responses were directed towards the stimulus that moved in a variety of directions (upper curve), responses were increased. It should be stressed that all three curves were generated from visually identical stimuli, the only thing that changed the responses was which stimulus was behaviorally significant. [Adapted from Treue and Martínez Trujillo (1999) with permission from Nature.]

1997). Such phenomena are usually considered to be distinct from the local, 'real' motion signals used in the experiments described here, although they might both converge at the level of MT (Albright, 1992; Wilson and Ferrera, 1992; but also see Baloch et al., 1999; O'Keefe and Movshon, 1998).

Recent experiments show that under the right conditions, attentional modulation can be found in MT and MST, either using recording methods in alert monkeys (Seidemann and Newsome, 1999; Treue and Martínez Trujillo, 1999; Treue and Maunsell, 1996) or functional magnetic resonance imaging techniques in humans (Beauchamp et al., 1997; O'Craven et al., 1997; Watanabe et al., 1998). Recording results from Treue and Martínez Trujillo (1999) are shown in Fig. 4. In this experiment, two moving targets were presented in the RF of an MT cell, and monkeys were trained to attend to one or the other and detect a small speed or direction change (ignoring similar changes in the other stimulus). One stimulus was always moving in the cell's null direction, while the other was moving in any of 12 different directions. Under these conditions, attention had a profound effect on the response of this cell, and a similar effect on the population response in both MT and MST. The effect is evident as an overall increase or decrease in response magnitude, without much of a change in the width of the cell's tuning curve. Thus, the effect seems well described as an overall change in the sensitivity or 'gain' of the cells' responses. Similar observations have been made in other extrastriate areas (McAdams and Maunsell, 1999).

Finally, we note for completeness that eye movement or eye position modulates responses of some neurons in the motion pathway. While these effects can be striking in area MST, they are rare or weak in area MT (Bradley et al., 1996; Bremmer et al., 1997; Newsome et al., 1988; Squatrito and Maioli, 1996). Since we concentrate on MT, we will not further address these effects.

Reading out a motion map

The term readout refers to the entire sequence of steps that leads from a sensory representation to a percept or to observable behavior. Three logical steps are involved in such a process. These are signal selection, signal combination, and converting the signal to a response. We will describe these in necessarily rather general terms, because the specifics vary so much according to the details of the behavioral context. In this discussion, we will adopt the terminology of human psychophysics, loosely following Graham (1989). A channel is a group of neurons that are similarly tuned along the sensory dimension (e.g. motion direction), and a signal is the response of those neurons to a relevant stimulus. In terms of neuronal structures, a channel may be equivalent to a cortical column, where the map is orderly (as in MT), but such an anatomical correlate for a channel is not necessary. So, how is information extracted from a representation, whatever its anatomical layout? First, some number of channels must be selected from the plethora available. Secondly, under most conditions, signals must be combined in some way, either within a channel or across channels. Within-channel combination is probably a necessary feature of vertebrate brain organization, because privileged, one-to-one connections between individual sensory neurons and individual motor neurons have not been observed. Between-channel combination (see below) can take a variety of forms, and requires the specification of weights on individual channels. Lastly, the sensory signals must be converted to particular responses. In motor tasks, this conversion is often imagined to be a continuous function (e.g. Robinson et al., 1986) relating sensory and motor rate codes. In perceptual decision tasks, this mapping usually but not always involves a step discontinuity of some kind (a criterion). In psychophysics, the conversion is often referred to as the decision rule, and the signal to which it is applied is often referred to as a decision variable.

We have described these three steps as being discrete and serial stages, but in practice they need not be so distinct. For example, the selection and combination steps are logically very tightly linked, and a real neuronal circuit would probably select, weight, and combine its inputs all at once. In the simplest reflexes, the conversion to response can occur at the same step as well. A well-studied example of this can be found in Mauthner cell-mediated escape reflexes in fish (e.g. Faber et al., 1991). In this circuit, only two synapses intervene between primary afferents and motor neurons in the spinal cord. The selection of sensory signals is formed by their synaptic weights on the dendrites of the Mauthner cell, the combination of signals occurs via summation on the dendrites, and the categorical 'decision' to initiate behavior is made at the axon hillock of the Mauthner cell.

In more complex brain circuits, the extent to which these three operations are distinct is an open question. We believe it to be likely that the selection and conversion steps for different visual motion-guided behaviors are sufficiently distinct that they are likely to be carried out in different places, using distinct circuitry. Thus MT

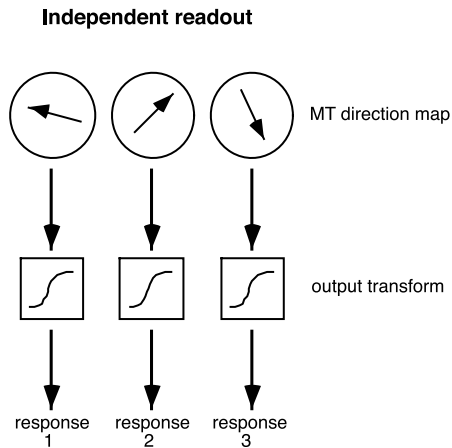


Fig. 5. Schematic for an independent readout mechanism. The circles depict individual neurons, or groups of neurons in MT of similar properties (channels). The square boxes indicate the transform between each channel and the corresponding percept or response. These indicate a 'generic' non-linearity, but the actual function could be very different for different responses. The important feature of this mechanism is that each channel and response is independent of the signal on the other channels, and of the other responses.

could send its outputs in parallel to circuits (say in the oculomotor control circuits in the pons) for the guidance of smooth pursuit movements, to the colliculus for the updating of saccadic movements, and to circuits for formation of categorical decisions, perhaps in the parietal or frontal lobes. These different targets probably instantiate different combination rules, consistent with their distinct functions. The key feature of this hypothesis is that in order to support multiple downstream operations, the amount of signal combination at early stages is necessarily limited. Because the rules by which signals are combined for different purposes might differ, the signals must remain independent up to the branch point. Thus, at the level of MT, the different channels for visual motion must remain relatively independent.

In the sections that follow, we first outline in general terms three distinct readout combination rules (independent, winner-take-all, and summation), and describe how these might manifest themselves in perceptual or motor system experiments. Then, we present a selected subset of experiments that address the readout question. We made no attempt to be encyclopedic, but have instead selected what we believe to be particularly revealing experiments.

Three readout combination rules

Independent combination rule. Independent readout is in fact no combination of signals: each channel provides independent evidence for the existence of stimuli that match the channel's preference. This mechanism is illustrated in Fig. 5. The key feature of this mechanism lies in the independence of the channels, the signals in each channel do not interact at all. Such a model is easy to imagine in the context of a perceptual experiment, where the number of simultaneously present percepts is large (as would be the number of possible subjective reports). However, if the responses are motor behaviors, problems

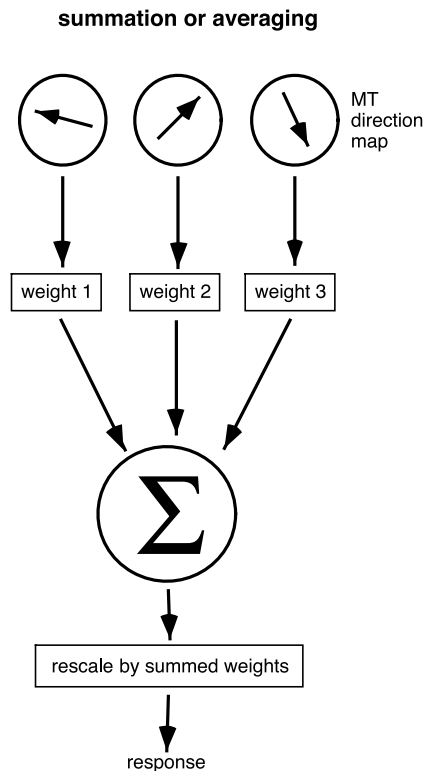


Fig. 6. Schematic of a summation or averaging mechanism. In this figure, each channel contributes to a single summed signal, producing a single response. Note that this response itself is often a vector for motor responses, or it can be a single valued decision variable. This variable can feed other decision rules as well, in more complicated models, but need not do so. Also note that in a distributed 'population vector' model, the output vector is never explicitly represented on any single element: it remains distributed until the response itself is generated.

might arise. For all types of movements (eye or arm movements, for instance), only one movement is possible at a time, and this places constraints upon the degree of independence. The signals leading to the movement might well be separate up to the level of motor planning, but thereupon some form of combination must arise. Because of this constraint, we will principally discuss the independent combination rule in the context of perceptual experiments ('Perceptual experiments addressing readout').

Summation or averaging. Another simple signal combination rule is that of summation or averaging. Both are similar in their general layout, and are illustrated in Fig. 6. The essential operation here is a weighted combination of inputs to produce a single output value. Averaging and summation differ only by a scale factor (the box at the bottom of the figure) which scales the output value to be near the range of its inputs in overall magnitude. This normalization (as described above in 'Averaging and normalization') is useful in keeping neuronally represented values in the range that is usefully captured in neuronal firing rates. Such models and their derivatives are widely used both in perceptual and sensorimotor experiments. The hallmark, and the key prediction for either type of experiment, is revealed when multiple

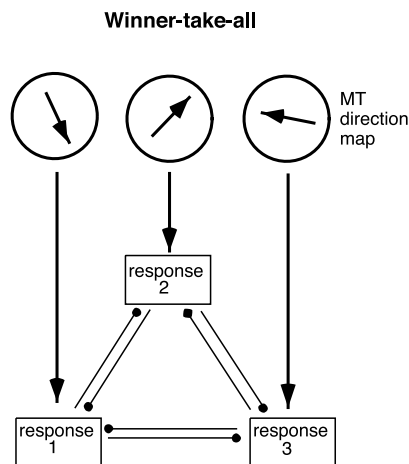


Fig. 7. Schematic of a winner-take-all mechanism. In this algorithm, each channel in MT raises the likelihood of the corresponding response (bold, downwards-directed arrows), as in Fig. 5, but in this case each response inhibits the other responses (light lines with filled circles). In such a network, minor imbalance of the signals in the three channels will tip the balance in favor of the largest response.

stimuli are simultaneously presented. In either form of experiment, the outcome will always reflect the presence of all the stimuli, a prediction that diverges both from the independent channels model above, and the winner-take-all described in the next section. For example, in a perceptual detection experiment, the threshold for some stimulus will be measurable and reliably affected by the presence of another stimulus near or even below threshold, suggesting that the two signals are summed before the conversion step is made (Graham, 1989).

A specific case of the general summation we have been describing is of particular interest in the study of motion, because visual motions are well described by vectors. Vector summation and vector averaging describe specific weighting arrangements (boxes in Fig. 6), and are extremely powerful analytic tools in a variety of sensory and motor processing contexts. We will more specifically discuss the predictions and tests of such mechanisms in 'Eye movement experiments addressing readout'.

Winner-take-all. Winner-take-all describes another combination rule that is closely related to the independent readout described in 'Independent combination rule'. In this case, though, signals from only one of the active channels is used for conversion. This exclusivity between channels is schematically depicted in Fig. 7 by the inhibitory connections (filled circles) between the boxes representing the possible behavioral responses. Again, the predictions from such a mechanism depend on the nature of the measurement being made. In a perceptual context, it usually refers to the fact that the more salient of two simultaneously present stimuli will drive perception; this is most evident near perceptual threshold. In detection experiments, this means that the threshold for detection of a more salient stimulus is not lowered by the addition of another, less salient stimulus (for a good review, see Graham, 1989). In identification or discrimination experiments, such mechanisms predict

that the report of a particular stimulus type will not be affected by the superposition of another stimulus.

In a motor task, the predictions of such a mechanism are also relatively easy to frame, and again require the presence of more than one stimulus. In such cases, the movement (for example an eye movement) will match one or the other of the two stimuli, rather than taking some intermediate value. However, an additional subtlety is frequently important. If the two stimuli are approximately equal in effectiveness, then winner-take-all models might allow either one or the other to win on different trials. If one averages across trials, the outcome might again look intermediate. Thus, the divergence of the predictions of these two models often requires the analysis of behavior on single trials.

Before we proceed to describing experiments testing the predictions of such models, a couple of caveats are required. First, the simple mechanisms we have described are a subset of possible rules, and a large family of other possibilities and intermediate mechanisms exists as well. However, these are the most commonly proposed models, and form good heuristic tools, which can be expanded upon as needed. Secondly, the abstract 'black box' diagrams that we have used to illustrate the mechanisms should not be taken literally. Even if a behavioral experiment provides some support for one or the other of such readout mechanisms (summation versus winner-take-all, for instance), we still need to consider a variety of anatomical possibilities for where the various embodied operations might take place. They need not be as distinct as they appear in schematic form.

In the next two sections we will specifically discuss experimental results which test such models of how sensory signals are used. We separate out experiments with a more perceptual flair from those in which eye movements are being measured.

Perceptual experiments addressing readout. In this and the following section, we have loosely organized the experiments to be consistent with the order of the mechanisms discussed above. Thus, we first describe observations on 'motion transparency', which most clearly support independent representations of multiple stimuli. Then, we move towards experiments that have been interpreted as supporting either summation or winner-take-all models for the perception of motion.

The most compelling evidence for independent readout of different motion directions is the perception of transparent motion. If two dot fields move in different directions, they are clearly both visible at the same place and time, and are perceptually segregated from each other, as long as the directions are sufficiently different (e.g. Andersen, 1989; Clarke, 1977; Gibson et al., 1959; van Doorn et al., 1985). Plainly, this supports the fact that individual channels can simultaneously report the existence of their preferred direction, even at the same spatial location.

However, alternative interpretations of the phenomenon of transparency exist. It is possible that during transparent motion perception only one pattern is seen at a time, and the appearance of transparency is just the

result of attentional switching between the different active channels. Evidence for such an explanation of motion transparency, that doesn't involve an independent readout mechanism, comes from De Bruyn and Orban (1993). They show that under uncertainty and restricted viewing times (85 ms), two different optic flow components (like rotation and divergence) cannot be identified when spatially superimposed. Subjects had no difficulty in identifying the patterns when either the subject was cued, or when the presentation duration was longer (170 ms), or when the patterns were presented in isolation. From these results they concluded that optic flow components can be segregated at a longer stimulus duration, but are not processed simultaneously or independently. A related experiment, in which translational movements were used, was performed by Zohary et al. (1996). They tested observers' perception of the direction of briefly presented stimuli containing either uniform or unbalanced distributions of different component directions. They found that when multiple motions are uniformly distributed, the reported direction closely approximates a vector average prediction. When the distribution of directions became unbalanced, the percept shifted towards the dominant direction, more consistent with a winner-take-all mechanism. However, the exact outcome was highly variable across conditions and subjects, leading the authors to conclude that 'the full distribution of activity' may be available for perceptual judgments.

Other evidence for the failure of complete independence comes from measurements of motion sensitivity under conditions of transparency. Observers show a decrease in motion sensitivity under such conditions (Snowden, 1989; Verstraten et al., 1996), compared to when stimuli are presented alone. Furthermore, when two moving random dot patterns are superimposed, the direction of each pattern is misperceived (Marshak and Sekuler, 1979; Mather, 1980). These observations suggest that the channels that support perception are not fully independent, but do not specify where the interactions take place. Reduced sensitivity under transparent conditions can, at least partly, be explained by interactions in the generation of MT cell responses. Several studies have shown that MT cell responses to a pattern in their preferred direction are reduced when another pattern is transparently moving in the opposite direction (Qian and Andersen, 1994; Qian et al., 1994a, b; Snowden et al., 1991). No attempt has been made yet to make even a qualitative comparison between psychophysically determined reduced motion sensitivity and physiologically determined reduction in activity in MT, preferably measured simultaneously in the same animal. The shift of perceived direction in the presence of another surface also suggests an easy experimental test: combined psychophysical and physiological testing could reveal whether this perceived shift is a result of interactions at the level of MT or later.

One special case of the transition between independent and non-independent perception has received tremendous attention in the last 15 years: the case of moving 'plaid' patterns. We discussed this in the context of MT

physiology in 'Responses to plaids', and here we will focus on the perceptual observations. The conditions under which this coherence occurs have been a subject of intense scrutiny (e.g. Adelson and Movshon, 1982; Alais et al., 1996; Beutter et al., 1996; Stoner and Albright, 1998) and several models have been proposed to account for the phenomenology (e.g. Grzywacz et al., 1995; Koechlin et al., 1999; Nowlan and Sejnowski, 1995). Detailed examination of this body of work is beyond the scope of this paper, but a summary of the consensus findings is instructive. Independent percepts are favored when the speeds, spatial frequencies, temporal frequencies, or temporal phases of the gratings substantially differ. If the component gratings are square waves, and the luminance of the intersections is varied independently of the gratings themselves, independent percepts occur when the relative luminance is consistent with occlusion of one surface by another (Stoner and Albright, 1992). For our present purposes, it is sufficient to know that under some conditions, multiple motion percepts can coexist on, and be read out from a motion map. This documents that under some conditions independent readout is possible, and raises the interesting question that when motion coherence occurs, where does it occur?

Some physiological evidence suggests that MT might be the limiting locus for the perception of motion coherence. Unlike in earlier stages of analysis, MT contains a substantial minority of cells which combine the components and thus signal the pattern direction (as described in 'Responses to plaids'; Movshon et al., 1985). In addition, in experiments where the luminance of the intersections is set to favor independence, the signal for the pattern direction in MT becomes significantly weaker (Stoner and Albright, 1992). These observations suggest that there is some correlation between the degree of segmentation in MT with the degree of segmentation in the percept. In turn, this suggests that a fixed readout from MT, without much additional downstream processing, is sufficient to account for the perceptual observations.

It is also possible that perceptual transparency plays a role in experiments where MT is microstimulated in the context of a motion discrimination task (Salzman et al., 1992; Salzman and Newsome, 1994). In these experiments, monkeys can only report a single direction of motion on a trial, out of either two or eight candidate stimulus motions. In these experiments, weak, high-frequency electrical current pulses are introduced into one column of neurons representing a particular direction in MT, activating this channel. When the microstimulation was added, the monkey tended to respond in favor of either the visual or the microstimulation signal. Results from cases where the artificial and visual signals were 135° apart, and of roughly equal strength, are shown in Fig. 8. There are two clear peaks in the choice distribution, one corresponding to the visual and the other to the microstimulation signal. Were the monkey combining the signals (some kind of average), one would instead see more choices in the intermediate directions.

The authors interpreted this in terms of a winner-take-all readout rule, and the results are certainly consistent

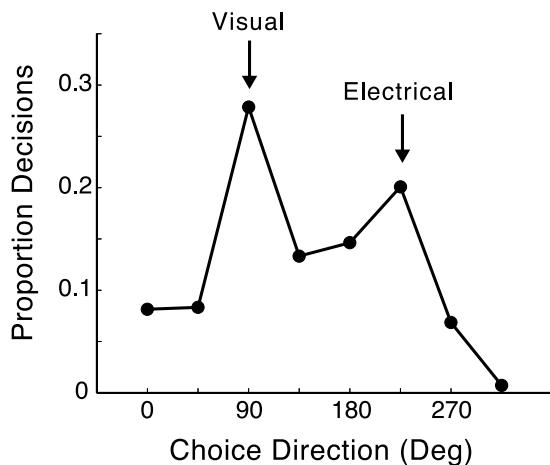


Fig. 8. Results of an electrical microstimulation experiment in an eight-alternative direction discrimination task (Salzman and Newsome, 1994). The stimulus was a variable coherence dot field matched in speed, location, and diameter to the multiunit RF recorded at the microstimulation electrode. After a stimulus, the monkey would indicate its choice by a saccade to a LED aligned with the corresponding direction. The figure presents the average results for 27 individual conditions in 23 experiments. All were rotated so that the preferred direction of the neurons was shown at 225°, and all were cases where the visual stimulus was 90°, and mirror reflected to be clockwise. Furthermore, the conditions were selected so that the visual and electrical influences on the monkeys' choice behavior were approximately equivalent. Under these conditions, monkeys' choices cluster either on the visual or on the electrical direction, but not the intermediate direction. [Adapted from Salzman et al. (1992) with permission of the American Association for the Advancement of Science.]

with this interpretation. However, it remains possible that the monkey perceived both motions simultaneously, akin to transparent motion, but was merely forced to respond in favor of one or the other. This interpretation places the winner-take-all operation downstream from MT, which is consistent with our recent work on adaptation ('Adaptation points to opponency after MT') and also with the results of eye movement experiments ('Eye movement experiments addressing readout').

Recent observations from the same laboratory have emphasized the importance of the task demands in how microstimulation signals interact with visual stimuli. Nichols and Newsome (1998) report that if instead of making a categorical judgment, the monkey is trained to make an eye movement towards the perceived direction, the results are very different. In this experiment, under conditions much like those shown in Fig. 8, the monkey very frequently made intermediate eye movements, suggesting that averaging occurred under these modestly different task conditions.

Another form of physiological measurement that addresses how the signals are employed in perceptual tasks is that of trial-by-trial correlation between neuronal responses and behavioral choices. Two different groups have reported such decisional correlations in MT, though the interpretation each offers is different. The work of Thiele and Hoffman (1996) showed activity correlated with the direction of an arm movement, when this indicated the choice in a direction discrimination task. Their interpretation of this phenomenon was that

it reflected attentional modulation, or the selection of one direction of movement. However, Britten et al. (1996), in describing a very similar observation, interpreted it as a signature of the fact that variable signals in MT produce variable decisions, when the visual stimulus is weak. In other words, they saw it as a direct measurement of readout. The latter group discarded feedback as an explanation by consideration of the time course of the phenomenon: the modulation commenced early in a trial and did not change much during the stimulus period. If the correlation reflected feedback from a decision, then it would be expected to grow during the stimulus period as the decision itself grew more reliable. The time course for the correlation reported by each group differed somewhat, as did the temporal demands of the tasks. In the Thiele and Hoffman (1996) paper, the correlation with decision lasted well after the stimulus was over, into the period of time the animal was making the actual response, whereas in the Britten et al. observations, the correlation disappeared immediately after the stimulus (Britten et al., unpublished observations).

Recent observations of Seidemann and Newsome (1999) perhaps shed some light on this discrepancy. In this microstimulation study, a delay period intervened between the visual stimulus and the monkey's choice. The question under study was to what extent microstimulation during this delay period would influence the monkey's choice behavior. The answer varied greatly across animals: in some, the microstimulation effects were just as large and reliable as they were during the visual stimulus epoch, but in others they completely disappeared. From these results it is clear that monkeys are able to regulate ('gate') the signals from MT in making their decisions, but sometimes they do not.

From these perceptually motivated experiments, we see little evidence for averaging of motion vectors in the formation of a perceptual decision. Either multiple-motion vectors remain simultaneously present, or one is selected and reported. While there is some evidence for modest, quantitative interactions between stimuli, they are clearly not obligatorily averaged by the visual system. This stands in contrast to many, but not all, measurements of the response of motor systems to such multiple-motion stimuli, which we describe in the next section.

Eye movement experiments addressing readout. Experiments that incorporate direct measurements of motor behavior are perhaps the most direct test of readout mechanisms from sensory maps. We know the input to the system (the stimulus), we can measure the representation (the sensory map), and from a measurement of the subsequent behavior we can directly infer the transform that lies between. The appeal of such experiments is perhaps strongest in studies of motion and motion-evoked eye movements, because both inputs and outputs are vectors. An example of the cleanliness with which testable predictions can be framed is shown in Fig. 9, from a recent review article by Groh (1998).

As in perceptual experiments, the most insightful

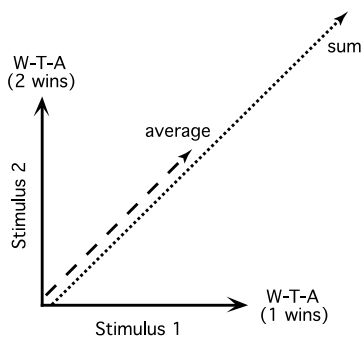


Fig. 9. Schematic of different predictions for pursuit direction when two signals are used [as in the experiment of Groh et al. (1997)]. One signal is from electrical microstimulation, and the other is from the motion of a pursuit target. The two dimensions of the page are horizontal and vertical eye velocities for the resulting eye movement. (Adapted from Groh, 1998.)

experiments incorporate two or more input motions (shown as stimulus 1 and 2 in Fig. 9), which differ in their directions or speeds. Then, since a single motion will be evoked, it can be tested against the predictions. The vector sum and vector average predictions derive from evenly weighted sums of the input vectors; they differ by a scale factor that normalizes the speed of the resultant vector. This framework will guide our discussion of some recent experiments using motion stimuli that address readout from MT.

Several experiments have produced evidence for vector averaging of motion inputs, and one of the most compelling employs electrical microstimulation in MT, as did the Salzman experiments previously described. Groh et al. (1997) microstimulated MT while a monkey pursued a small moving dot. We know from previous experiments (Newsome and Paré, 1988) that lesions to MT impair performance on such tasks, and also that high-current microstimulation reduces the gain of smooth pursuit eye movements, when the stimulus to be pursued is present over the RF of the region of MT being stimulated (Komatsu and Wurtz, 1989). The incisive experiment, however, includes two additional features. First, lower currents are used, which selectively activates a restricted subset of directional signals in MT (approximating a single column or channel (Murasugi et al., 1993)). Secondly, this microstimulation signal is combined with visual stimuli moving in a range of directions and speeds. This is the critical test for the interaction between the visual and electrical signals, and can reliably differentiate between winner-take-all, vector averaging, and vector summation, as illustrated in Fig. 9.

In the Groh et al. experiments, two distinct eye movement measurements were made, both in the same trials. In these 'step-ramp' experiments, the monkeys both begin to pursue the moving target, and make a rapid saccade to its initial position. Both of these eye movements are modified by the direction and speed of the visual stimulus. Stimulus motion directly influences the initial pursuit eye movement and the amplitude of the first saccade; the latter is sometimes termed saccadic velocity compensation. Two things are striking about the results from the Groh experiments. First, that there

was clear evidence for vector averaging for both smooth pursuit and saccadic velocity compensation, although the specific weighting of the averages varied considerably across experiments. Secondly, the vector average for pursuit and saccade velocity compensation was frequently different from each other, in individual experiments. Thus, in the same trials, with the identical pattern of activity on the motion areas produced by a combination of visual and electrical stimulation, different combination rules were employed in the production of saccades and pursuit. This indicates that within one experiment, motion signals in MT are being combined or utilized in different ways, for the production of different classes of eye movements.

Averaging can also describe the response of the pursuit system to two visual motion inputs as well, under the right conditions. A series of studies from Steve Lisberger's laboratory have explored this question. Monkeys were trained to pursue one of two simultaneously present moving targets. Ferrera and Lisberger (1997a) and Lisberger and Ferrera (1997) found clear evidence for averaging in smooth pursuit responses to pairs of moving stimuli. Such averaging occurred when either of two simultaneously presented targets was equally likely to be the target of pursuit, and predicted the initial trajectory of the eye, for approximately the first 150 ms. A short interval later, the monkey would select one target, and its eye velocity would rapidly change to match the selected target.

However, under closely related experimental conditions, the pursuit system appears to operate in a manner more consistent with the winner-take-all prediction. In these experiments (Ferrera and Lisberger, 1995), one of the stimuli was always cued to be the target for the pursuit, and the direction of the initial pursuit eye movement was well matched to the target. In other words, the cued signal 'won' in driving the eye movement. There were modest effects on the latency and acceleration of the pursuit from the presence of the distractor, but in general the behavior was well described by a winner-take-all rule. From these experiments, we can conclude that if the monkey is given prior cueing of the correct pursuit target, then it can behaviorally suppress the irrelevant motion information from the other target. Absent such cueing, averaging occurs. So, the immediate question is where does this selection take place? Two related physiological experiments address this question.

In one revealing experiment exploring physiological correlates of target selection for smooth pursuit eye movements, Ferrera and Lisberger (1997b) measured the responses of MT cells in a pursuit task, where a cue specified which of two targets was to be pursued. Under these conditions, employing a fairly permissive statistical criterion, very few (14%) MT cells showed significant cue effects, with an average modulation amplitude of about 30%. Furthermore, while behavioral latencies were dramatically influenced by the direction of the distractor, the neuronal latencies were barely affected. The authors concluded that these selection signals were inadequate to directly cause the monkeys' behavioral selection, and that further amplification of this

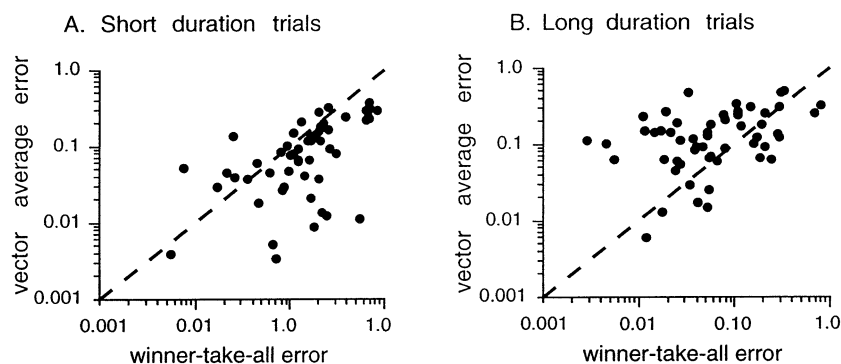


Fig. 10. Transition between winner-take-all and vector averaging in the responses of MT cells. Responses to two moving stimuli passing through the RF of an MT cell were measured under two different time conditions. In the ‘long-duration’ trials, the stimuli were presented 400 ms before they entered the RF of the cell, giving the monkey ample time to discriminate them (they differed in color) and select one for subsequent pursuit, before the response was measured. In the ‘short-duration’ trials, the stimuli were presented and entered the RF almost immediately. These responses were then compared against two different models: winner-take-all and averaging. The error against each model’s prediction was computed using a contrast index: $(R_{\text{obs}} - R_{\text{pred}})/(R_{\text{obs}} + R_{\text{pred}})$. These error metrics are scattered against each other from recordings derived from 27 cells in two monkeys. One can see that the data were much closer to the average prediction (larger winner-take-all error scores) for the short trials, and vice versa for the long-duration trials. [Adapted from Recanzone and Wurtz (1999).]

minor imbalance was needed. The proposed competitive mechanism was very similar to the winner-take-all architecture illustrated in Fig. 7.

Target selection signals for smooth pursuit appear weak under these conditions, but attentional selection can produce profound modulation in MT (recall the results of Treue and Martínez Trujillo (1999) as described in Fig. 4 and ‘Extra-retinal modulation of MT activity’). This seems like a contradiction, and might, on the surface, suggest that subtle changes in task characteristics cause differences in sensory processing. A recent experiment by Recanzone and Wurtz (1999) appears to clarify the apparent discrepancy, however, and points to the importance of time in attentional modulation or target selection. In this experiment, once again two small moving dots moved through the RF of an MT cell, and the monkey was trained to pursue one or the other, which was cued in advance. The key feature of this experiment, though, was that in some cases, ‘long-duration trials’, the two stimuli were offered well in advance (about 400 ms) of the time when the RF was activated. In other, interleaved ‘short-duration trials’, the stimuli appeared just outside the RF, and started to traverse it almost immediately. In both types of trial, the pursuit eye movement was held back until after the stimuli had traversed the RF of the cell. Thus, the trials are identical, except for the additional time given to the monkey in the long-duration trials. The responses of MT neurons were quite different under these two conditions, as shown in Fig. 10. The data were compared against averaging and winner-take-all predictions, by comparing the responses to the paired stimuli against the responses to each component of the pair, presented by itself. The relative errors of these two predictions are compared in Fig. 10. In the short-duration trials, errors to the averaging prediction were much smaller than to the winner-take-all prediction (points above the unity diagonal); in the long-duration trials exactly the opposite relationship was observed. Thus, MT cells appear to

switch from averaging to winner-take-all behavior during the first few hundred milliseconds, between when the stimuli first appear, and when the pursuit movement itself is initiated. It seems very likely that the switch to winner-take-all in these cells is via a similar mechanism as Treue and Maunsell’s attentional modulation. Also, it seems likely that this selection can also explain the change in the initial direction of smooth pursuit (Ferrera and Lisberger, 1995; Ferrera and Lisberger, 1997a). However, this mechanism might not account for the differences in the dynamics of the initiation of pursuit in the presence of multiple targets, which appears to require an additional downstream step. We now describe an experiment from our own laboratory, which also clearly indicates the need for additional downstream competitive mechanisms.

Adaptation points to opponency after MT

In this chapter we will discuss an example from our own laboratory which addresses the readout from MT. The method is based on a visual illusion, called the motion aftereffect (MAE, or ‘waterfall illusion’). After prolonged exposure to one direction of motion, stationary or flickering targets are seen to move in the opposite direction. Ever since early psychophysical descriptions (Wohlgemuth, 1911; for a recent review, see Mather et al., 1998), the MAE has been used as evidence for interactions between different motion directions. Since most adaptation effects are negative, and since motion adaptation produces a positive percept, it strongly suggests mutual inhibition between opposite directions of motion. A variety of specific models have been proposed, and most have some form of mutual inhibition amongst different motion directions. This mutual inhibition has been posited to either just coupled opposed directions (Adelson and Bergen, 1985; Barlow and Hill, 1963; Sutherland, 1961) or directionally non-selective (Grunewald and Lankheet, 1996; Mather, 1980;

Verstraten et al., 1994). The latter rather strongly resembles the normalization we discussed in ‘Averaging and normalization’.

Recently, this line of work has been extended in a way more directly testing motion interactions, by the use of adapting stimuli that contain multiple motions. After adaptation to multiple-moving stimuli human observers perceive a unidirectional MAE, opposite to the vector average of the adaptation directions (Mather et al., 1998; van Doorn et al., 1985; Verstraten et al., 1994). Based on this observation, Mather et al. (1998) suggested a distribution-shift model for motion perception, in which the MAE depends on a vector average of the activity in a map of directions such as that in MT.

Although these psychophysical studies on the MAE have been cited to support a vector averaging mechanism along the visual motion processing pathway, it is not clear that this must be the case. According to recent modeling work (Grunewald and Lankheet, 1996), broadly tuned inhibitory interactions (normalization), at the level of MT is sufficient to produce the unidirectional MAE without additional integration after the map. Grunewald and Lankheet’s model suggests that independent readout (consistent with the observation of transparent motion; see ‘Three readout combination rules’), combined with normalization and broadly tuned directional responses, could account for most known observations on the MAE.

We know from physiological evidence that in MT interactions between neurons tuned to different directions occur (‘Motion opponency’). Most studies have focused on interactions of neurons tuned to opposite directions (opponency). But is this the same opponency that underlies the MAE? Even the terminology is confusing, because ‘opponency’ can either mean the physiological opponency described at the level of directionally selective responses themselves (‘Motion opponency’), or the opponency implicit in a winner-take-all decision rule. For this reason, we will use the term winner-take-all for the decision step, and reserve the term ‘opponency’ for the physiological observations of below-baseline responses to motion opposite the preferred direction. So, the question we are asking is whether there exists an additional winner-take-all step after the level of MT. In order to resolve this question, one needs to record from MT cells under conditions that produce the MAE, and quantitatively relate the adaptation effects in MT to the behavioral results of the adaptation. In other words, is the opponency in MT sufficient to explain the perceptual changes?

This work relates to a long series of studies exploring the linkage between MT activity and behavior (Britten et al., 1992; Celebrini and Newsome, 1994; Croner and Albright, 1999; Shadlen et al., 1996). The general question of all these studies is to what extent one can identify perceptual phenomena with observed physiological phenomena in the motion system. In these studies single unit activity and perceptual choices were measured simultaneously in monkeys performing a two alternative forced choice motion task. The information conveyed by MT neurons to later stages was expressed in so-called ‘neuro-

metric’ functions. Such functions describe the direction discrimination capabilities of an individual neuron, in terms identical to those used to describe observer performance. The neurometric functions derived from the electrophysiological recordings were very similar to the psychometric functions based on the monkey’s behavior. The derivation of these functions assumes a comparison between opposed directions downstream from MT, although related functions can be derived under a criterion model (Shadlen et al., 1996, Appendix 4). In previous work, it was argued that the entire body of data relating activity in MT with perception could be best accounted for under the assumption of a comparison of MT neurons tuned to opposite directions. However, this conclusion relied on indirect arguments, and some of the data used for the comparison between neuronal and psychophysical performance were difficult to measure accurately (e.g. the slopes of neurometric and psychometric functions).

We performed an experiment to explicitly test the readout mechanism, using motion adaptation as a probe. The effect of motion adaptation on neuronal responses can reveal opponent mechanisms that are present in generating the neuronal responses, and comparison with perceptual measurements can test whether any observed physiological effects are sufficient in and of themselves. Only one previous study investigating the effect of motion adaptation in MT showed, on average, a decrease in response after adaptation in the neuron’s preferred direction and an increase after adaptation in the opposite (‘null’) direction (Petersen et al., 1985). The increase in response after adaptation in the null direction suggests that adaptation interacts with opponent mechanisms in MT. However, the stimulus conditions that they used were quite different from those normally used in psychophysical experiments. Therefore, in our experiments, we measured the effect of adaptation both on MT cell responses and human behavioral judgments of motion, under matched stimulus conditions. Ideally, of course, one would directly measure the after-effect in the same monkeys. However, this is technically difficult, because the illusion is inherently subjective. As a result animals can shift their strategy dependent on the adaptation stimulus, simply attempting to maximize their rewards. Thus, one cannot count on the readout mechanism itself being constant under the different adaptation conditions. Human observers, however, can be more counted upon for their honesty under conditions of illusion.

We quantified the effect of motion adaptation on both neuronal and behavioral motion sensitivity, using methods that were recently introduced in human psychophysics (Blake and Hiris, 1993; Raymond, 1993). In the electrophysiological experiments we compared MT cell sensitivity to moving random dot patterns under three different adaptation conditions. Fig. 11 shows responses of a single MT cell to a range of stimulus strengths (as in Fig. 2), after adaptation to preferred, null, or zero motion. After adaptation in the neuron’s preferred direction, the response of this MT cell is decreased, while the response after adaptation in the null direction is indis-

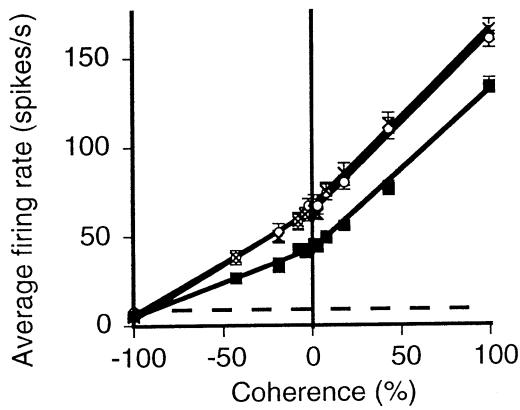


Fig. 11. Effect of motion adaptation on extracellular responses in an example single MT cell. The cell was adapted for 3 s with a 100% coherent random dot pattern. Three different adaptation conditions were compared: adaptation in preferred direction (■), adaptation in null direction (×), adaptation with a static pattern (○). The graph represents the average responses during the subsequent 1-s test-stimulus presentation at different coherence levels. Positive coherence levels denote preferred direction stimuli, negative numbers show null direction. Compared to the static condition, adaptation in the preferred direction reduces the responses for both directions and coherence levels. However, adaptation in the null direction does not affect the cell's response. The error bars represent ± 1 S.E.M., and the horizontal dashed line indicates spontaneous activity. The lines through the data points show the corresponding linear fits.

tinguishable from the response after adaptation with a stationary pattern.

We have repeated this measurement on 74 cells in three monkeys, and the cell in Fig. 11 is quite typical of the observed results. We quantified these response functions, and their changes following adaptation, using linear regression. Changes in intercept show response changes that are approximately constant; changes in slope reveal adaptation effects that vary with stimulus coherence. In our sample of neurons, following preferred direction adaptation, both intercepts and slopes were significantly changed (paired *t*-test, $P < 0.05$), while these were not systematically affected following null direction adaptation. Thus, MT neurons representing the adaptation direction have attenuated responses, while those representing the opposite direction are not affected. We can conclude that the effects of adaptation, at the level of MT, are very selective: they only attenuate the adaptation direction, and leave the other direction unaffected. This result appears inconsistent with models where the perceptual effects of adaptation are a direct consequence of changes in opponent motion processing at the level of sensory processing, but to be sure, we need to measure human psychophysics under the same conditions and compare the physiology to these results.

In Fig. 12 we have plotted the results for human judgments of motion direction with (solid line) and without (dashed line) motion adaptation. In this figure, we plot the proportion of decisions in the adaptation direction, as a function of the strength and direction of the immediately following test stimulus. Neutral stimuli (pure noise) as always are in the middle, and increasing stim-

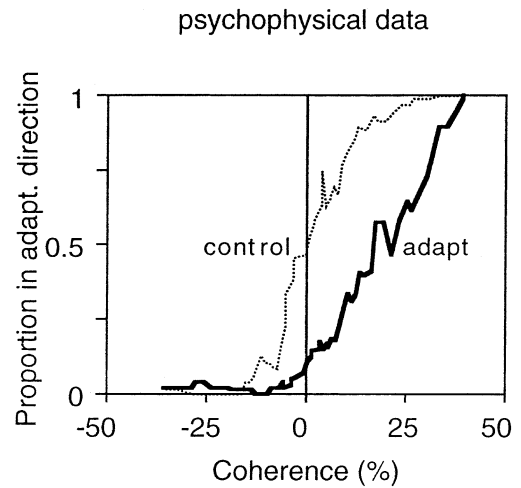


Fig. 12. Psychophysical data from a directional adaptation experiment. Average results of decisions on motion direction during the test phase made by three human subjects. The vertical axis corresponds not to proportion correct, but to the proportion of choices made in favor of the adaptation direction. The lighter line shows the results of trials where no adapting stimulus was presented, while the bold line indicates the results following the presentation of a 3-s adapting stimulus. Each subject was run on slightly differing coherence levels across the shown range, and for presentation, an *N*-point running average was taken.

ulus strength in the adaptation direction are indicated with positive stimulus coherence values (see legend for details). The figure shows average performance for three subjects, two of them were naive with respect to the goals of the experiment. Adaptation clearly has an effect on perception of both test directions: in the adaptation direction and opposite to it. This is in contrast to the physiological results just described, where adaptation only affected neurons representing the adaptation, and had no systematic effect on neurons representing the opposite direction. Since we used exactly the same stimulus conditions in these experiments to measure motion sensitivity of MT cell responses and human motion detection, we must conclude from our data that an additional step after MT exists. We have explicitly modeled two candidate decision rules: winner-take-all and independent, criterion-based readout.

Fig. 13A shows the results of modeling a winner-take-all decision rule and Fig. 13B shows the results of the independent model with a high criterion level. To derive each of these, we took the observed distribution of cell responses to each direction and stimulus strength employed. Because we did not simultaneously record neurons responding to both directions of stimulus motion, we assumed symmetry, and modeled the responses to neurons preferring the direction opposite adaptation with the null direction responses of those preferring the adaptation direction. Then, the observed distributions of responses were passed through the chosen decision rule. In the case of the winner-take-all model, this meant that the two distributions of responses (neurons preferring the two choice directions) were explicitly compared, and the choice was made in favor of the group with the larger total response. This analysis is exactly the same as the 'neurometric' functions

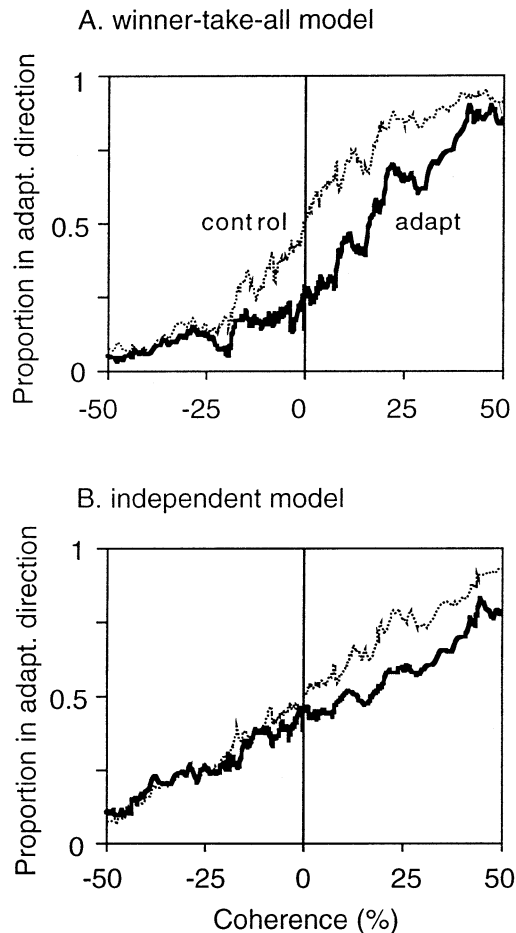


Fig. 13. Model predictions for the psychophysical experiment shown in Fig. 12. (A) Predictions based on a winner-take-all ('opponent') decision rule. To derive these predictions, the distributions of observed responses before and after adaptation of two neurons were compared, and the proportion of times the response was higher in the neurons preferring the adaptation direction is plotted. The broken line shows reference performance after adaptation to a static pattern (control), and the solid line performance following adaptation. (B) Predictions resulting from an independent, criterion-based decision rule. The same responses were evaluated independently and if the responses of either group of neurons exceeded a fixed criterion (in this case, 1 S.D. above the mean response to 0% coherence motion), then the channel detected it's preferred direction. In the event of a 'tie vote', random responses were assigned.

described in earlier studies (Britten et al., 1992). In the independent model, each group was compared against its own criterion, which was defined relative to the response to pure noise (neutral stimulus). If one or the other group alone gave an above-criterion response, then that became the decision, but if a tie occurred a random decision was assigned.

The results in Fig. 13 clearly indicate that the winner-take-all model provides a much better match to the psychophysical results shown in Fig. 12. Specifically, this decision rule predicts adaptation effects to both directions of stimulus motion, but the criterion model predicts highly asymmetrical effects. The rationale for these diverging predictions is easily understood, since the physiology itself shows very asymmetric effects from adapta-

tion. If each channel is treated independently, then the effects of adaptation also become asymmetric. However, if they are compared, like in the winner-take-all model, each loses its own identity, and changes in either channel produce comparable changes in behavior. Hence the adaptation effects become symmetric with stimulus direction.

From these results, we can conclude that an additional interaction must exist, somewhere between the neuronal representation in MT and the formation of the behavioral response. This can take the form of an explicit winner-take-all step (Shadlen et al., 1996) (see Fig. 7, or alternately involve some high-level categorical transformation of the value of a vector average). We do not know from this experiment where this interaction occurs, but we can be sure it lies downstream of MT. The avenue of investigation, however, is open: the same experiment could be done throughout the series of areas to which MT projects, and in this way we could narrow down the candidate structures where the decision rule is applied to sensory representations.

The case for stability in sensory representations

The work highlighted in the foregoing section has called attention to the multiplicity of ways in which multiple motions can be employed in the formation of perception or motor behavior. Now, we wish to return to the main questions we posed earlier. Does the multiplicity of output rules necessitate changes in sensory processing? We believe not, and will suggest how apparently flexible readout can occur in the presence of effectively invariant connections and stable sensory processing. The idea is simple: different downstream loci employ different readout rules, and when the animal 'switches gears' from one set of task demands to another, it is merely engaging an alternate, relatively stable readout mechanism. Abundant evidence exists for the maintenance of multiple motions on the map itself and in perceptual experiments where multiple percepts can be reported. However, most motor tasks and some perceptual judgements allow only one outcome. The difference between these rules is fundamental, and it is most parsimonious to place the differences downstream from the sensory representations themselves. The parallel, divergent outputs from sensory cortical structures such as MT provide the necessary substrate for the parallel operation of multiple-readout rules.

Proposal: the map remains the same. We propose an idea that contains three linked parts:

1. Sensory maps in cortex retain their basic structure in the face of changing task demands; this covers time periods ranging from seconds to minutes in duration. Different signals present in these representations do not change their interactions in different contexts.
2. The basic rules by which different signals are combined and used are also relatively invariant, the multiple-readout rules remain stable, individually.
3. Multiple-readout mechanisms can exist for different

tasks, by virtue of parallel efferent connections from sensory areas. For a specific example, the pursuit system can average multiple-motion directions while parietal or frontal decision circuits are deciding which of two simultaneously present motions is stronger.

This idea has the advantage of simplicity: the basic wiring diagram, both up to an intermediate-level representation such as MT, and leading from that representation out towards behavior, need not change with differing demands. In changing from task to task, one is simply tapping into different, pre-existing circuits. Thus, in development, or in the process of learning a new task, the readout mechanism is established, but on a more rapid time scale, information is always flowing from sensory to diverse associative and premotor structures, which are always poised and ready to operate. We will now discuss how this idea applies to the experiments we have reviewed.

Microstimulation in decision tasks versus pursuit experiments. We have described three very similar experiments from one laboratory, which appear to suggest distinct readout rules. In the experiments of Groh et al. (1997) (see 'Eye movement experiments addressing readout') and the experiments of Nichols and Newsome (1998), microstimulation-induced signals appeared to average with the visual signals in the guidance of eye movements. However, in the experiments of Salzman et al. (1992) and Salzman and Newsome (1994) ('Eye movement experiments addressing readout'; Fig. 6) the same signals appeared to pass through a winner-take-all algorithm in guiding a categorical, eight-alternative choice. Importantly, the 'final common path' for both tasks was very similar: both used saccadic eye movements. So, how does our framework encompass and make sense of this apparent dichotomy? In our view, the signals from microstimulation and the visual stimulus remain intact and relatively distinct (up to the limits imposed by directional bandwidths in MT), but feed different output circuitry. In the categorical task, competitive decision networks force the exclusiveness of the different responses. The monkeys cannot choose more than one target, this is a requirement of the task, so mutual inhibition between the alternatives (at some level) is necessary. In such a task, where no reward will come for an intermediate movement (even if this is what is perceived), it makes little sense to make 'average' movements. Instead, an optimal strategy, given multiple signals present on a sensory map, would be to pick the best response and exclude the alternatives. The real question, then, is whether the exclusion is happening at the level of the sensory map, or later, in the formation of the movements.

In guidance of a matching smooth pursuit eye movement in the experiments of Groh et al., where no choice is involved, no such competition need exist. The demands of the task were intentionally made rather lax (in terms of the accuracy of pursuit required), so as to detect the effects of microstimulation, and minimize the chances for

monkey strategies to produce artifacts. In situations like this, where the motor system is attempting not to execute a choice, but is instead attempting to minimize position and velocity errors between eye and target, it makes sense to integrate broadly, thus averaging out noise and deriving the most accurate estimate of target velocity. Feedback systems in general are very sensitive to internal noise, and averaging is one of the best means of minimizing internal noise. Indeed, this averaging might be hard-wired into the machinery feeding the pursuit system from the motion system (see 'Do motor systems average?').

Thus the two distinct operations, which must share some components (especially near the motor end) probably do not use all the same circuits. Efferent connections of MT (and MST) include a litany of different structures known to be involved in both smooth pursuit and saccadic eye movements. To explore the readout mechanism, and especially its dependence on the demands of the trained task, both traditional single-cell recording and more complex paired recordings could be useful. In animals that are trained for both types of tasks (categorical and pursuit), based upon the same signals, one could explore candidate structures (parietal cortex, the pontine nuclei, frontal eye fields) as the monkey switched from one task to the other. Based on our idea, we would predict two things. First, in structures that both carry directional signals and appear to be selectively involved in a particular behavior (e.g. the flocculus for pursuit), the directional signals would still be present when the animals switched gears to a categorical decision task. This would provide direct evidence that the signals from the motion system were not 'switched off' when the animal stopped pursuit. In a more difficult experiment, one might directly measure effective connectivity between neurons in sensory areas and premotor structures using two electrodes and cross-correlation analysis. We would predict that the correlations would be the same as the animal switched from one task to another, despite large changes in behavioral output. This would even more directly demonstrate invariance of sensorimotor throughput in the face of changing task demands.

Do motor systems average? Neuroscientists studying motor systems have a long history of elegant experiments addressing the kinds of questions we are setting in a sensory context. Revealing experiments have explored how the saccadic system deals with multiple targets (Sparks and Mays, 1983), the population code for intended movement primary motor cortex (Georgopoulos, 1990; Georgopoulos et al., 1986), and numerous others. While intense debate still rages about the exact nature of such premotor representations, and exactly what entities are encoded, it is safe to say that there appears to be broad consensus that averaging is a fundamental property of motor systems. In some sense, this is a biophysical constraint: most body parts can only do one thing at a time. However, these motor systems must be fed with sensory data, which often contain, especially in the types of experiments that we have covered, multiple simultaneous signals. This provokes several questions. First, is it

necessary that these signals be discarded prior to premotor structures? Secondly, if so, where on the sensorimotor pathway does the selection take place? Lastly, how does the selection happen? These are not new questions, and have been discussed extensively, but we wish to specifically bring them up in the context of the motion system and motion-guided behaviors.

Several experiments have used multiple-moving stimuli driving pursuit to address the question of how these are converted into a single motion. These sometimes suggest vector averaging and sometimes suggest winner-take-all, even with very similar experimental configurations. The key feature that causes the apparently different readouts is expectation. If, at the time when pursuit is enabled, there is no prior information as to which of two targets is to be followed, then averaging occurs, and the eye movement will be intermediate between the two input vectors. If, however, the subject is given information about which of two targets is to be pursued, the influence of the distracting stimulus on the direction of pursuit is much reduced. Thus, the pursuit system appears to switch from an averaging operation to a winner-take-all operation, given sufficient information. Where is this switch implemented, and does its operation require any fundamental reorganization of the way information flows from sensory to motor structures?

Lisberger and Ferrera (1997) argue that averaging is the 'default' operation of the smooth pursuit system. Because averaging occurs across hemifields, it also seems likely that averaging must occur after MT, in addition to that which occurs within MT ('Averaging and normalization'). This additional averaging could easily occur very close to the motor 'plant' itself. But, if this is the case, how can the system switch to apparent winner-take-all behavior, given enough prior information? One possibility is that the irrelevant motion signal is attenuated relative to the other by an attentional or target selection mechanism. There is clear evidence that such mechanisms can have profound effects at the level of sensory cortex, and the specific question of how early in sensory structures such modulation occurs is under intense scrutiny at the present. We will specifically address this question, and attempt to link it to our proposal of sensory stability, in the next section.

For the moment, though, let us move on to the question of how one addresses the questions posed at the start of this section: whether, where, and how selection of sensory signals occurs. First, is it necessary to have selection taking place to account for the present data? We believe so: both Recanzone et al. and Lisberger's work suggests that the initial pursuit can be accurate given enough prior information, and not incorporate other simultaneously present motion signals. This demonstrates that target selection is occurring. Regarding where it is occurring, there is fairly compelling evidence that the level of motion modulation at the level of MT is inadequate to account for the behavioral selectivity. Therefore, it seems parsimonious to place much of the selection operations downstream from MT. How would we actually test where such selection is occurring? Obviously, imaging experiments are ideal for questions of this

sort, but until they reach spatial resolution at the single cortical column level (which they are poised to do: Menon and Goodyear, 1999), it is not clear that questions of the interaction of multiple stimuli will be fully resolved. However, it is possible that some answers might not require this level of resolution. For example, Heeger et al. (1999) used functional imaging to test global interactions between different directions of motion, and such evidence might reveal selection effects. However, more direct evidence might come from physiological experiments. Recent work exploring mechanisms of target selection in the saccadic system has emphasized the role of structures in the parietal and frontal lobes (Platt and Glimcher, 1997, 1998); similar experiments quantitatively exploring motion interactions in the context of smooth pursuit are clearly needed.

We have specifically proposed that there is no need for the basic readout rules to change to explain any of the data we have been discussing. To test this idea, more difficult experiments are probably required, testing connection weights between sensory and motor structures. Groh et al. cleverly used microstimulation of MT to reveal vector averaging in pursuit; a similar approach can test whether the weighting of the average is under dynamic control, or is relatively fixed. In a paradigm such as that used by Recanzone and Wurtz, where the monkey's behavior can be rapidly switched from one to the other readout mechanism by the addition of a cue, one could measure the weights of one vector, by activating the same neurons under each context. We would predict that the influence of microstimulating unrelated channels would not depend on whether the animal had been allowed to previously select the target. In fact, the results of Groh et al. already suggest this: the animals certainly averaged in the case of a single visual stimulus being present and selected.

Attention, target selection, and normalization. One of the questions we have been posing in this review is whether modification of signals on sensory maps is sufficient to account for apparent changes in the sensitivity to different stimuli when more than one is present. One candidate mechanism at the level of sensory cortex is that of selective attention. This subject has been reviewed recently (Desimone and Duncan, 1995; Mangun, 1995; Posner and Petersen, 1990), so we will focus on newer work, especially that which relates to the motion system. We will also (in line with the model of Desimone and Duncan), attempt to draw a link between normalization mechanisms and target selection mechanisms in motion tasks. We know modulatory signals are present in MT; the tools are in place to ask whether these signals are of the right kind to explain behavioral target selectivity.

Desimone and Duncan (1995) compellingly suggest that selective attention is the result of a process of 'biased competition' between stimuli, especially when multiple stimuli are within the dimensions of a RF. This is consistent with recent measurements (McAdams and Maunsell, 1999; Reynolds et al., 1999) suggesting that the modulation of sensitivity by attention largely operates through a multiplicative factor, equivalent to

'gain' or sensitivity. These are consistent because simple accounts of competitive processes in sensory cortex (Carandini et al., 1997; Simoncelli and Heeger, 1998) contain mutual inhibitory interactions of a divisive nature. Modulating the effectiveness of such divisive inhibitory circuits would thus be expected to produce a change in gain of the cells being affected. This suggestion has been strongly supported by recent experimental findings from area V4 (Reynolds et al., 1999). Measurements in the motion system to date have not determined whether attentional modulation can also be described as gain changes, but this is the most parsimonious hypothesis at present.

What is lacking from these accounts, and which makes a very ripe target for experimental work, is a quantitative account of the sufficiency of these changes to support behaviorally measured sensitivity changes. Lisberger and Ferrera report very modest signals, but consistent in sign with target selection, in the context of pursuit experiments. However, Treue and Maunsell report much larger effects in the context of a speed discrimination task. What each observation lacks is a quantitative model predicting exactly how the signals in MT contribute to the observed behavior. With such a model, one would be in a position to test our idea rigorously. Does one really need additional modulation downstream from MT? Models can well address such questions of sufficiency. Obviously, theoretical approaches also need to be bolstered by additional measurements, especially from downstream areas. Refreshingly, this work is well under way, at least in the context of perceptual discrimination tasks (Horwitz and Newsome, 1999; Kim and Shadlen, 1999).

Concluding remarks

A neuroscientific equivalent to the Holy Grail is to seek a complete account of behavior from the stimulus through central processing and out to observable behavior. In simple reflexes, good quantitative accounts have been found. However, when it comes to complex volitional behavior, simple hierarchical accounts tend to flounder in the complexity of the intervening brain anatomy. Thus, simple 'bottleneck' accounts of perceptual performance (see Parker and Newsome, 1998) or motor processing have been quite difficult to test. Largely, this is because of the apparent complexity of central readout circuits. In this paper, we have attempted to suggest that simple hierarchical ideas, far from being dead, are still alive and capable of accounting for most modern physiological and behavioral data. This suggestion is based on a naive desire for simplicity: why propose a complex model when a simple one will work? Our proposal contains fairly simple rules, which can be laid down in development and learning, and need not involve complex changes in information processing as animals change behavioral contexts. Whether such simple ideas can stand against the demands of increasingly sophisticated behavioral testing is an open question, one to carry us into the next century of behavioral neuroscience.

Acknowledgements—We wish to thank Wim van de Grind, Kelly Huffman, Martin Lankheet, and Bruno Olshausen for their thoughtful comments on earlier versions of this manuscript.

REFERENCES

- Adelson, E.H., Bergen, J.R., 1985. Spatio-temporal energy models for the perception of motion. *J. Opt. Soc. Am. A* 2, 284–299.
- Adelson, E.H., Movshon, J.A., 1982. Phenomenal coherence of moving gratings. *Nature* 300, 523–525.
- Alais, D., van der Smagt, M.J., Verstraaten, F.A.J., van de Grind, W.A., 1996. Monocular mechanisms determine plaid motion coherence. *Vis. Neurosci.* 13, 615–626.
- Albright, T.D., 1992. Form-cue invariant motion processing in primate visual cortex. *Science* 255, 1141–1143.
- Albright, T.D., 1984. Columnar organization of directionally selective cells in visual area MT of macaques. *J. Neurophysiol.* 51, 16–31.
- Albright, T.D., Stoner, G.R., 1995. Visual motion perception. *Proc. Natl. Acad. Sci. USA* 92, 2433–2440.
- Allman, J.M., Kaas, J.H., 1971. A representation of the visual field in the caudal third of the middle temporal gyrus of the owl monkey (*Aotus trivirgatus*). *Brain Res.* 31, 85–105.
- Andersen, G.J., 1989. Perception of three-dimensional structure from optic flow without locally smooth velocity. *J. Exp. Psychol.* 15, 363–371.
- Ashida, H., Verstraten, F.A.J., 1997. Attentive tracking of motion interaction between low-level and high-level motion. In: *Selection and Integration of Visual Information: Proceedings of the International Workshop on Advances in Research on Visual Cognition*. National Institute of Bioscience and Human-Technology, Tsukuba, Japan, pp. 229–3233.
- Baloch, A.A., Grossberg, S., Mingolla, E., Nogueira, C.A., 1999. Neural model of first-order and second-order motion perception and magnocellular dynamics. *J. Opt. Soc. Am. A* 16, 953–978.
- Barlow, H.B., Hill, R.M., 1963. Evidence for a physiological explanation of the waterfall phenomenon and figural after-effects. *Nature* 200, 1345–1347.
- Beauchamp, M.S., Cox, R.W., DeYoe, E.A., 1997. Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *J. Neurophysiol.* 78, 516–520.
- Beutter, B.R., Mulligan, J.B., Stone, L.S., 1996. The barberplaid illusion: plaid motion is biased by elongated apertures. *Vis. Res.* 36, 3061–3075.
- Blake, R., Hiris, E., 1993. Another means for measuring the motion aftereffect. *Vis. Res.* 33, 1589–1592.
- Braddick, O., 1997. Local and global representations of velocity: transparency, opponency, and global direction perception. *Perception* 26, 995–1010.
- Bradley, D.C., Chang, G.C., Andersen, R.A., 1998. Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature* 392, 714–717.
- Bradley, D.C., Maxwell, M., Andersen, R.A., Banks, M.S., Shenoy, K.V., 1996. Mechanisms of heading perception in primate visual cortex. *Science* 273, 1544–1547.
- Bremmer, S., Ilg, U.J., Thiele, A., Distler, C., Hoffman, K.P., 1997. Eye position effects in monkey cortex. I. Visual and pursuit-related activity in extrastriate areas MT and MST. *J. Neurophysiol.* 77, 944–961.

- Britten, K.H., Heuer, H.W., 1999. Spatial summation in the receptive fields of MT neurons. *J. Neurosci.* 19, 5074–5084.
- Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., Movshon, J.A., 1996. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* 13, 87–100.
- Britten, K.H., Shadlen, M.N., Newsome, W.T., Movshon, J.A., 1992. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745–4765.
- Britten, K.H., Shadlen, M.N., Newsome, W.T., Movshon, J.A., 1993. Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* 10, 1157–1169.
- Carandini, M., Heeger, D.J., Movshon, J.A., 1997. Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.* 17, 8621–8644.
- Cavanagh, P., 1992. Attention-based motion perception. *Science* 257, 1563–1565.
- Celebrini, S., Newsome, W.T., 1994. Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J. Neurosci.* 14, 4109–4124.
- Chaudhuri, A., 1990. Modulation of the motion aftereffect by selective attention. *Nature* 334, 60–62.
- Clarke, P.G.H., 1977. Subjective standstill caused by the interaction of moving patterns. *Vis. Res.* 17, 1243.
- Croner, L.J., Albright, T.D., 1999. Segmentation by color influences responses of motion-sensitive neurons in the cortical middle temporal visual area. *J. Neurosci.* 19, 3935–3951.
- Culham, J.C., Cavanagh, P., 1994. Motion capture of luminance stimuli by equiluminous color gratings and by attentive tracking (see comments). *Vis. Res.* 34, 2701–2706.
- De Bruyn, B., Orban, G.A., 1993. Segregation of spatially superimposed optic flow components. *J. Exp. Psychol.* 5, 1014–1027.
- DeAngelis, G.C., Newsome, W.T., 1999. Organization of disparity-selective neurons in macaque area MT. *J. Neurosci.* 19, 1398–1415.
- Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. *Ann. Rev. Neurosci.* 18, 193–222.
- Dubner, R., Zeki, S.M., 1971. Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus. *Brain Res.* 35, 528–532.
- Faber, D.S., Korn, H., Lin, J.W., 1991. Role of medullary networks and postsynaptic membrane properties in regulating Mauthner cell responsiveness to sensory excitation. *Brain Behav. Evol.* 37, 286–297.
- Felleman, D., Van Essen, D., 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47.
- Ferrera, V.P., Lisberger, S.G., 1995. Attention and target selection for smooth pursuit eye movements. *J. Neurosci.* 15, 7472–7484.
- Ferrera, V.P., Lisberger, S.G., 1997a. The effect of a moving distractor on the initiation of smooth-pursuit eye movements. *Vis. Neurosci.* 14, 323–338.
- Ferrera, V.P., Lisberger, S.G., 1997b. Neuronal responses in visual areas MT and MST during smooth pursuit target selection. *J. Neurophysiol.* 78, 1433–1446.
- Gattass, R., Sousa, A.P., Mishkin, M., Ungerleider, L.G., 1997. Cortical projections of area V2 in the macaque. *Cereb. Cortex* 7, 110–129.
- Georgopoulos, A., 1990. Neural coding of the direction of reaching and a comparison with saccadic eye movements. *Cold Spring Harbor Symp. Quant. Biol.* 55, 849–859.
- Georgopoulos, A.P., Schwartz, A.B., Kettner, R.E., 1986. Neuronal population coding of movement direction. *Science* 233, 1416–1419.
- Gibson, E.J., Gibson, J.J., Smith, O.W., Flock, H.R., 1959. Motion parallax as a determinant of perceived depth. *J. Exp. Psychol.* 58, 40–51.
- Gogel, W.C., Sharkey, T.J., 1989. Measuring attention using induced motion. *Perception* 18, 300–320.
- Graham, N.V.S., 1989. *Visual Pattern Analyzers*. Oxford University Press, Oxford.
- Groh, J.M., 1998. Reading neural representations. *Neuron* 21, 661–664.
- Groh, J.M., Born, R.T., Newsome, W.T., 1997. How is a sensory map readout? Effects of microstimulation in visual area MT on saccades and smooth pursuit eye movements. *J. Neurosci.* 17, 4312–4330.
- Grunewald, A., Lankheet, M.J.M., 1996. Orthogonal motion after-effect illusion predicted by a model of cortical motion processing. *Nature* 384, 358–360.
- Grzywacz, N.M., Watamaniuk, S.N., McKee, S.P., 1995. Temporal coherence theory for the detection and measurement of visual motion. *Vis. Res.* 35, 3183–3203.
- Heeger, D.J., 1987. Model for the extraction of image flow. *J. Opt. Soc. Am. A* 4, 1455–1471.
- Heeger, D.J., Boynton, G.M., Demb, J.B., Seidemann, E., Newsome, W.T., 1999. Motion opponency in visual cortex. *J. Neurosci.* 19, 7162–7174.
- Horwitz, G.D., Newsome, W.T., 1999. Separate signals for target selection and movement specification in the superior colliculus. *Science* 284, 1158–1161.
- Kim, J.N., Shadlen, M.N., 1999. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque (see comments). *Nat. Neurosci.* 2, 176–185.
- Kochlin, E., Anton, J.L., Burnod, Y., 1999. Bayesian inference in populations of cortical neurons: a model of motion integration and segmentation in area MT. *Biol. Cybern.* 80, 25–44.
- Komatsu, H., Wurtz, R., 1988. Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *J. Neurophysiol.* 60, 580–603.
- Komatsu, H., Wurtz, R.H., 1989. Modulation of pursuit eye movements by stimulation of cortical areas MT and MST. *J. Neurophysiol.* 62, 31–47.
- Lankheet, M.J.M., Verstraten, F.A.J., 1995. Attentional modulation of adaptation to two-component transparent motion. *Vis. Res.* 10, 1401–1412.
- Lisberger, S.G., Ferrera, V.P., 1997. Vector averaging for smooth pursuit eye movements initiated by two moving targets in monkeys. *J. Neurosci.* 17, 7490–7502.
- Lisberger, S.G., Movshon, J.A., 1999. Visual motion analysis for pursuit eye movements in area MT of macaque monkeys. *J. Neurosci.* 19, 2224–2246.
- Mangun, G.R., 1995. Neural mechanisms of visual selective attention. *Psychophysiology* 32, 4–18.
- Marshak, W., Sekuler, R., 1979. Mutual repulsion between moving visual targets. *Science* 205, 1399–1401.
- Mather, G., 1980. The movement aftereffect and a distribution-shift model for coding the direction of visual movement. *Perception* 2, 379–392.
- Mather, G., Verstraten, F.A.J., Anstis, S., 1998. *The Motion Aftereffect: a Modern Perspective*. MIT Press, Cambridge, MA.
- Maunsell, J.H.R., Ferrera, V.P., 1993. Extraretinal representations in visual areas of macaque cerebral cortex. In: Ono, T. et al. (Eds.), *Brain Mechanisms of Perception and Memory: from Neuron to Behavior*. Oxford University Press, Oxford, pp. 104–118.
- Maunsell, J.H.R., Van Essen, D.C., 1983a. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* 3, 2563–2586.
- Maunsell, J.H.R., Van Essen, D.C., 1983b. Functional properties of neurons in the middle temporal visual area (MT) of the macaque monkey: I. Selectivity for stimulus direction, speed and orientation. *J. Neurophysiol.* 49, 1127–1147.

- McAdams, C.J., Maunsell, J.H.R., 1999. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.* 19, 431–441.
- Menon, R.S., Goodyear, B.G., 1999. Submillimeter functional localization in human striate cortex using BOLD contrast at 4 Tesla: implications for the vascular point-spread function. *Magn. Reson. Med.* 41, 230–235.
- Movshon, J.A., Adelson, E.H., Gizzi, M.S., Newsome, W.T., 1985. The analysis of moving visual patterns. In: Chagas, C., Gattass, R., Gross, C. (Eds.), *Pattern Recognition Mechanisms*, (Pontificiae Academiae Scientiarum Scripta Varia 54, 117–151). Vatican Press, Rome. (Reprinted in *Experimental Brain Research, Supplementum* 11, 117–151, 1986).
- Movshon, J.A. et al., 1986. The analysis of moving visual patterns. In: Chagas, C. et al. (Eds.), *Experimental Brain Research, Suppl. II: Pattern Recognition Mechanisms*. Springer, New York, pp. 117–151.
- Movshon, J.A., Newsome, W.T., 1996. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J. Neurosci.* 16, 7733–7741.
- Movshon, J.A., Newsome, W.T., Gizzi, M.S., Levitt, J.B., 1988. Spatio-temporal tuning and speed sensitivity in macaque visual cortical neurons. *Invest. Ophthalmol. Vis. Sci.* 29 (Suppl.), 327.
- Murasugi, C.M., Salzman, C.D., Newsome, W.T., 1993. Microstimulation in visual area MT: effects of varying pulse amplitude and frequency. *J. Neurosci.* 13, 1719–1729.
- Newsome, W.T., Paré, E.B., 1988. A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* 8, 2201–2211.
- Newsome, W.T., Wurtz, R.H., Dursteler, M.R., Mikami, A., 1985. Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *J. Neurosci.* 5, 825–840.
- Newsome, W.T., Wurtz, R.H., Komatsu, H., 1988. Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *J. Neurophysiol.* 60, 604–620.
- Nichols, M.J., Newsome, W.T., 1998. Influence of electrical microstimulation in area MT on direction of motion estimates in a direction identification task. *Soc. Neurosci. Abstr.* 24, 1746.
- Nowlan, S.J., Sejnowski, T.J., 1995. A selection model for motion processing in area MT of primates. *J. Neurosci.* 15, 1195–1214.
- O'Craven, K.M., Rosen, B.R., Kwong, K.K., Treisman, A., Savoy, R.L., 1997. Voluntary attention modulates fMRI activity in human MT-MST. *Neuron* 18, 591–598.
- O'Keefe, L.P., Movshon, J.A., 1998. Processing of first- and second-order motion signals by neurons in area MT of the macaque monkey. *Vis. Neurosci.* 15, 305–317.
- Orban, G.A., Saunders, R.C., Vandenbussche, E., 1995. Lesions of the superior temporal cortical motion areas impair speed discrimination in the macaque monkey. *Eur. J. Neurosci.* 7, 2261–2276.
- Parker, A.J., Newsome, W.T., 1998. Sense and the single neuron: probing the physiology of perception. *Annu. Rev. Neurosci.* 21, 227–277.
- Pasternak, T., Merigan, W.H., 1994. Motion perception following lesions of the superior temporal sulcus in the monkey. *Cereb. Cortex* 4, 247–259.
- Petersen, S.E., Baker, J.F., Allman, J.M., 1985. Direction-specific adaptation in area MT of the owl monkey. *Brain Res.* 346, 146–150.
- Platt, M.L., Glimcher, P.W., 1997. Responses of intraparietal neurons to saccadic targets and visual distractors. *J. Neurophysiol.* 78, 1574–1589.
- Platt, M.L., Glimcher, P.W., 1998. Response fields of intraparietal neurons quantified with multiple saccadic targets. *Exp. Brain Res.* 121, 65–75.
- Posner, M., Petersen, S., 1990. The attention system of the human brain. In: Cowan, W. et al. (Eds.), *Annual Review of Neuroscience. Annual Reviews, Palo Alto, CA*, pp. 25–42.
- Qian, N., Andersen, R.A., 1994. Transparent motion perception as detection of unbalanced motion signals. II. Physiology. *J. Neurosci.* 14, 7367–7380.
- Qian, N., Andersen, R.A., Adelson, E.H., 1994a. Transparent motion perception as detection of unbalanced motion signals. III. Modeling. *J. Neurosci.* 14, 7381–7392.
- Qian, N., Andersen, R.A., Adelson, E.H., 1994b. Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. *J. Neurosci.* 14, 7357–7366.
- Raymond, J.E., 1993. Movement direction analysers: independence and bandwidth. *Vis. Res.* 33, 767–775.
- Recanzone, G.H., Wurtz, R.H., 1999. Shift in smooth pursuit initiation and MT and MST neuronal activity under different stimulus conditions. *J. Neurophys.*, in press.
- Recanzone, G.H., Wurtz, R.H., Schwarz, U., 1997. Responses of MT and MST neurons to one and two moving objects in the receptive field. *J. Neurophysiol.* 78, 2904–2915.
- Reynolds, J.H., Chelazzi, L., Desimone, R., 1999. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci.* 19, 1736–1753.
- Robinson, D.A., Gordon, J.L., Gordon, S.E., 1986. A model of the smooth pursuit eye movement system. *Biol. Cybern.* 55, 43–57.
- Rodman, H.R., Albright, T.D., 1989. Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Exp. Brain Res.* 75, 53–64.
- Salzman, C.D., Murasugi, C.M., Britten, K.H., Newsome, W.T., 1992. Microstimulation in visual area MT: effects on direction discrimination performance. *J. Neurosci.* 12, 2331–2355.
- Salzman, C.D., Newsome, W.T., 1994. Neural mechanisms for forming a perceptual decision. *Science* 264, 231–237.
- Sclar, G., Maunsell, J.H.R., Lennie, P., 1990. Coding of image contrast in central visual pathways of the macaque monkey. *Vis. Res.* 30, 1–10.
- Seidemann, E., Newsome, W.T., 1999. Effect of spatial attention on the responses of area MT neurons. *J. Neurophysiol.* 81, 1783–1794.
- Shadlen, M.N., Britten, K.H., Newsome, W.T., Movshon, J.A., 1996. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* 16, 1486–1510.
- Shimojo, S., Miyauchi, S., Hikosaka, O., 1997. Visual motion sensation yielded by non-visually driven attention. *Vis. Res.* 37, 1575–1580.
- Shulman, G.L., 1993. Attentional effects on adaptation of rotary motion in the plane. *Perception* 22, 947–961.
- Simoncelli, E.P., Heeger, D.J., 1998. A model of neuronal responses in visual area MT. *Vis. Res.* 38, 743–761.
- Snowden, R.J., 1989. Motions in orthogonal directions are mutually suppressive. *J. Opt. Soc. Am. A* 6, 1096–1101.
- Snowden, R.J., 1994. Motion processing in the primate cerebral cortex. In: Smith, A.T., Snowden, R.J. (Eds.), *Visual Detection of Motion*. Academic Press, London, pp. 51–83.
- Snowden, R.J., Treue, S., Erickson, R.G., Andersen, R.A., 1991. The response of area MT and V1 neurons to transparent motion. *J. Neurosci.* 11, 2768–2785.
- Sparks, D.L., Mays, L.E., 1983. Spatial localization of saccade targets. I. Compensation for stimulation-induced perturbations in eye position. *J. Neurophysiol.* 49, 45–63.
- Squatrito, S., Maioli, M.G., 1996. Gaze field properties of eye position neurones in areas MST and 7a of the macaque monkey. *Vis. Neurosci.* 13, 385–398.
- Stepniewska, I., Qi, H.X., Kaas, J.H., 1999. Do superior colliculus projection zones in the inferior pulvinar project to MT in primates? *Eur. J. Neurosci.* 11, 469–480.

- Stoner, G.R., Albright, T.D., 1992. Neural correlates of perceptual motion coherence. *Nature* 358, 412–414.
- Stoner, G.R., Albright, T.D., 1993. Image segmentation cues in motion processing: implications for modularity of vision. *J. Cogn. Neurosci.* 5, 129–149.
- Stoner, G.R., Albright, T.D., 1998. Luminance contrast affects motion coherency in plaid patterns by acting as a depth-from-occlusion cue. *Vis. Res.* 38, 387–401.
- Sutherland, N., 1961. Figural after-effects and apparent size. *Q. Exp. Psychol.* 13, 222–228.
- Thiele, A., Hoffman, K.-P., 1996. Neuronal activity in MST and STPp, but not MT changes systematically with stimulus-independent decisions. *NeuroReport* 7, 971–976.
- Treue, S., Martínez Trujillo, J.C., 1999. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 200, 575–577.
- Treue, S., Maunsell, J.H.R., 1996. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 382, 539–541.
- Ungerleider, L.G., Mishkin, M., 1982. Two cortical visual systems. In: Ingle, D.J. et al. (Eds.), *Analysis of Visual Behavior*. MIT Press, Cambridge, MA, pp. 549–580.
- van Doorn, A.J., Koenderink, J.J., van de Grind, W.A., 1985. Perception of movement and correlation in stroboscopically presented noise patterns. *Perception* 14, 209–224.
- van Santen, J.P.H., Sperling, G., 1985. Elaborated Reichardt detectors. *J. Opt. Soc. Am. A* 2, 300–321.
- Verstraten, F.A.J., Fredericksen, R.E., van de Grind, W.A., 1994. Movement aftereffect of bi-vectorial transparent motion. *Vis. Res.* 34, 349–358.
- Verstraten, F.A.J., Fredericksen, R.E., van Wezel, R.J.A., van de Grind, W.A., 1996. Directional motion sensitivity under transparent motion conditions. *Vis. Res.* 36, 2333–2336.
- Watanabe, T., Harner, A.M., Miyauchi, S., Sasaki, Y., Nielsen, M., Palomo, D., Mukai, I., 1998. Task-dependent influences of attention on the activation of human primary visual cortex. *Proc. Natl. Acad. Sci. USA* 95, 11489–11492.
- Wilson, H.R., Ferrera, V.P., 1992. A psychophysically motivated model for two-dimensional motion perception. *Vis. Neurosci.* 9, 79–97.
- Wohlgemuth, A., 1911. On the aftereffect of seen movement. *Br. J. Psych. Monogr. (Suppl. 1)*, 1–117.
- Zeki, S.M., 1990. The motion pathways of the visual cortex. In: Blakemore, C. (Ed.), *Vision: Coding and Efficiency*. Cambridge University Press, Cambridge, pp. 321–345.
- Zohary, E., Scase, M.O., Braddick, O.J., 1996. Integration across directions in dynamic random dot displays: vector summation or winner take all? *Vis. Res.* 36, 2321–2331.

(Accepted 3 December 2001)