Neuroscience and Biobehavioral Reviews 36 (2012) 1–25

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review

# The behavioral receptive field underlying motion integration for primate tracking eye movements

Guillaume S. Masson*, Laurent U. Perrinet

*Team Dynamics of Visual Perception and Action (DyVA), Institut de Neurosciences Cognitives de la Méditerranée (INCM), CNRS & Aix-Marseille Université, 13402 Marseille Cedex 20, France*

ARTICLE INFO

*Article history:*
Received 27 December 2010
Received in revised form 11 March 2011
Accepted 13 March 2011

*Keywords:*
Motion integration
Tracking eye movements
Optic flow
Sensorimotor transformation
Gain control
Spatial summation
Isomorphism

ABSTRACT

Short-latency ocular following are reflexive, tracking eye movements that are observed in human and non-human primates in response to a sudden and brief translation of the image. Initial, open-loop part of the eye acceleration reflects many of the properties attributed to low-level motion processing. We review a very large set of behavioral data demonstrating several key properties of motion detection and integration stages and their dynamics. We propose that these properties can be modeled as a behavioral receptive field exhibiting linear and nonlinear mechanisms responsible for context-dependent spatial integration and gain control. Functional models similar to that used for describing neuronal properties of receptive fields can then be applied successfully.



**Contents**



* Corresponding author at: Team Dynamics of Visual Perception and Action (DyVA), Institut de Neurosciences Cognitives de la Méditerranée (INCM), UMR 6193, CNRS, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France. Tel.: +33 491 16 43 15; fax: +33 491 16 44 98.
*E-mail address:* guillaume.masson@incm.cnrs-mrs.fr (G.S. Masson).

0149-7634/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neubiorev.2011.03.009

## 1. Introduction: object motion computation for gaze stabilization

Visual motion is critical for the guidance of slow eye movements that help visual perception by stabilizing the images onto the retina. However, in a crowded and constantly changing visual environment, stabilizing the whole retinal image is largely inappropriate. Our visual system must parse these images into separate objects, select the one of interest and accurately measure its motion in order to smoothly rotate the eyes at the appropriate speed and direction. Primates are equipped with extremely fast and accurate oculomotor behaviours. Since the mid 20th century, the primate oculomotor system is probably one of the best known of the sensorimotor systems (see Leigh and Zee, 2006). Many review articles are available from the literature. For instance, Keller and Heinen (1991) and Ilg (1997) gave exhaustive reviews of the classification of the different types of smooth eye movements and their neural substrates. Krauzlis and Stone (1999) documented the behavioral evidences that smooth and saccadic eye movements are closely linked together. More recently, Krauzlis (2004) has recasted the neural framework of smooth pursuit eye movements and its links with the saccadic systems. However, very few review articles have been dedicated to the more specific topic of visual motion processing in the context of sensorimotor transformation. Miles and colleagues (e.g. Kawano, 1999; Miles, 1993, 1997, 1998; Miles et al., 2004) have published several reviews that summarizes their pioneering work on optic flow processing and reflexive slow eye movements. Still, a review article focusing the consequences of early visual motion computation is missing.

This review takes a different viewpoint about eye movement control. We are interested in the fundamental visual mechanisms underlying this sensori-to-motor transformation. As pointed out by Miles (1993) most models of smooth eye movements collapse this complex processing into a single box that measures retinal target velocity, or higher derivatives needed to feed the motor control mechanisms. Thus, understanding the visual motion processing stage of visuomotor behaviors is a major challenge in systems neuroscience. Lisberger et al. (1987) offered a classical introduction to this fascinating question. In a more recent review, he summarizes the results obtained from his group supporting the idea that voluntary pursuit is initiated based upon a rapid linear read-out of area MT neurons (Lisberger, 2010). Thus, it appears after two decades of intensive work that many of the dynamical aspects of eye movements are in fact constrained by the neural solutions of visual motion processing (see also Masson, 2004; Ilg, 2008; Masson and Ilg, 2010). In this perspective, tracking eye movements offer a miniature model of perception–action coupling.

Herein, our objective is to provide an updated view of the complex front-end processing that computes object motion for visuo-motor control. Our review documents the experimental work conducted by several groups to elucidate the properties of ocular following responses in both human and non-human primates. Thus, we will restrict our scope to the initial, open-loop part of the reflexive responses. In this well-defined framework, it becomes possible to relate neural and behavioral dynamics with a very high level of details (Lisberger et al., 1987; Miles, 1997; Kawano, 1999). Each aspect of the motor responses can be mapped with a specific aspect of visual information processing. These linear and non-linear mechanisms can be summarized within a concept called the *behavioral receptive field* that concisely describes how information is integrated within populations of cortical neurons in order to extract the relevant signals (i.e. speed and direction) about a single object motion (Barthélemy et al., 2006). This idea is reminiscent of the earlier notion of perceptive fields but offer a more precise framework, in particular regarding the temporal dynamics of information processing. Such behavioral receptive field incorporates many of the fundamental aspects of sensory information processing such as spatial and temporal filtering, spatial summation, nonlinear gain control and multiple feature integration. Each of these aspects can be fitted with mathematical tools already used at both population and neuronal levels to describe classical and non-classical receptive fields.

Such behavioral receptive field can be seen as a read-out of low-level cortical computations done through the recurrent connectivity between areas V1 and MT/MST. These first 100 ms of ocular following offers a minimalist, but very efficient window onto these mechanisms. Whenever it is possible, we will compare the tuning properties and their temporal dynamics of the behavior with the neuronal mechanisms observed at each of these cortical stages. Although the scope of this review might sounds a bit too much focused on a particular type of eye movements, we advocate herein the usefulness of such a delicate approach to elucidate the detailed temporal architecture of the brain.

The review is organized as follows. We will first recast ocular following as one example of short-latency, reflexive smooth movements of the eyes that have been identified in both human and non-human primates. From neurophysiological studies conducted in macaque monkeys, we will present the underlying cascade of cortical and sub-cortical steps driving response onset, aiming at extracting the key stages that are essential to understand the behavioral results that will be presented subsequently. We will then define the behavioral receptive field, in the perspective of an older theoretical concept from visual perception, the perceptive field. Both concepts are defined from behavioral (i.e. motor versus perceptual) phenomenology, but we propose that the ocular following approach allows us to better constrain the core idea and therefore will open the door for realistic modeling. Indeed, we will review behavioral properties of ocular following when tackling the classical computational stages of visual motion processing: (i) what are the detection mechanisms, (ii) how local motion in pooled to extract direction and speed information and (iii) how are these mechanisms dependent upon visual context? At each stage we highlight the key results that will be then used to define the mathematical description of its input–output transfer function. We will then close this review by highlighting some open questions.

## 2. Ocular following: reflexive tracking in human and non-human primates

Since the pioneering work of Dodge (1903), several sub-types of visually driven smooth eye movements have been defined. This classical taxonomy distinguishes between reflexive optokinetic eye movements that are driven by large field visual motion to form the slow phases of OKN, and voluntary smooth pursuit eye movements that are elicited by local visual motion. They both have short latency, mostly around or below 100 ms in primates. Eye speed

during tracking phases barely exceeds 40°/s. It is linearly related to target speed below this limit and saturates for higher input velocity (see Lisberger et al., 1987 for a review). Reflexive tracking responses have long been seen as being sub-cortically mediated and driven by the *en masse* motion of the visual scene, as reproduced by optokinetic drums. Investigating slow phases of primates OKN, several groups reported that complex motion processing is actually involved, such as disparity- or attention-based motion selection (e.g. Murasugi et al., 1989; Howard and Simpson, 1989; Mestre and Masson, 1997). The group of Fred Miles at the NIH discovered, first in monkeys (Miles et al., 1986) and later in humans (Gellman et al., 1990), that a rapid shift of a large visual scene can trigger machine-like smooth eye movements at very short-latency. To distinguish them from both voluntary pursuit and OKN, they coined the term of *ocular following responses* (OFRs). Ocular following responses are thought to correspond to the early phase of the optokinetic reflex that backs up translational vestibulo-ocular reflex (Cohen et al., 1977; Lisberger et al., 1981). These results opened the way to the long series of experimental work that will be presented below.

Although some idiosyncratic differences can be found between subjects, in particular regarding the absolute amplitude of the responses, the basic tuning properties of OFRs are highly reproducible. A practical consequence is that input–output transfer functions can be mapped with a remarkable resolution. Most often, human and monkey OFRs exhibit similar visual properties (e.g. Busettini et al., 1996; Masson et al., 1997), pointing out that ocular following is a good model for investigating both visual motion processing and visuomotor transformations in non-human primates. Moreover, they are obtained in monkeys without the need of extensive training. Reward is given only if the monkey attends to the center of the screen prior visual motion onset and refrains saccades during the 150 ms after motion onset. It is never delivered as a function of performance for a given set of motion stimuli. Thus, responses are free of any biases that could have been introduced from training protocols (Miles et al., 1986). Lastly, all stimuli are always fully interleaved so that a given condition cannot be predicted ensuring that early phase of ocular following is not corrupted by anticipatory pursuit responses (Kowler and Steinman, 1979).

When recorded with the scleral search coil technique in both species, these responses exhibit two striking properties: they are "machine-like" as shown by the small variability across trials and they are initiated at extremely short latencies. In non-human primates, latency is ~55 ms after target onset. In humans, it goes up only to ~85 ms. These latencies are very robust and show very little variance (<5 ms) for stimuli with high signal-to-noise ratio (Miles et al., 1986; Masson and Castet, 2002). There are two key aspects that must be noticed. First, given a shortest reaction time around 55 ms in primates, only a few cortical steps can be activated before the neuronal information flow reaches the oculomotor brainstem nuclei. In fact, latency of monkey ocular following strictly follows that of the earliest MST neurons when recorded simultaneously (Kawano et al., 1994). Thus, the temporal dynamics of early phases of ocular following must closely follow the temporal dynamics of feed-forward cortical motion processing. In line with this, a change in visual input will be reflected in the oculomotor response with the same delay, as demonstrated with double ramp experiments (Miles and Kawano, 1986). Any increase in response latency must be associated with additional delays in the processing steps encoding or decoding motion information. Second, since the earliest phase of OFRs seems to be dependent upon both speed and direction of target motion, these critical pieces of evidence must be extracted with a very short integration time. In other words, only a few spikes at the onset of neuronal responses are sufficient to convey information needed by visuomotor transformation (Osborne et al., 2004). In line with this, the time course of eye acceleration can be predicted from that of neuronal activity in areas MT and MST (Kawano et al., 1994; Takemura et al., 2001). Thus, ocular following very nicely reflects the activity at the earliest stages of the cortical motion pathway and that makes it an excellent behavioral probe of sensory processing.

Ocular following is a member of a family of short-latency smooth binocular responses. Miles and colleagues described in both monkeys and humans short-latency vergence eye movements, which are driven by binocular disparity (Busettini et al., 1996, 2001; Masson et al., 1997). These vergence eye movements depend on the activity of MST neurons encoding absolute binocular disparities for large visual scene (Takemura et al., 2001). Busettini et al. (1997) documented optic flow-driven vergence eye movements at ultra-short latency in humans. Here, the useful information is the radial, rather than planar, velocity flow field (see Miles, 1998; Miles et al., 2004 for reviews). Similar responses were subsequently observed in macaque monkeys (Kodaka et al., 2003). Very recently, Sheliga et al. (2009) reported short-latency torsional ocular following induced by rotation of the frontal images about the line of sight. Altogether, Miles and coworkers clearly demonstrated that all these short-latency ocular responses exhibit similar computational solutions in terms of linear extraction, gain control and automatic integration of motion/disparity signals. There is no room to describe these results here, but we want to pinpoint that the computational steps that will be described below are shared by many different binocular responses.

## 3. Neural bases of ocular following: cortical and sub-cortical contributions

Early work on ocular following suggested that such ultra-fast drive of the sensorimotor transformation can be explained by the feed-forward activation of template detectors such as found in area MST that extract optic flow information (Nakayama and Loomis, 1974; Duffy and Wurtz, 1991a,b). A long series of experimental work has been conducted to decipher their neural bases (see Kawano, 1999 for a review). A key argument suggesting that OFRs are cortically mediated was published only very recently: chemical lesions of cortical areas MT and MST abolish all these types of short-latency smooth eye movements (Takemura et al., 2007). Moreover, what makes ocular following an attractive model of sensorimotor transformation is that it is possible to track the neuronal flow of information at different cortical and sub-cortical stages and to correlate it with behavioral responses (Kawano, 1999).

Fig. 1 summarizes the impressive work conducted by Kawano and colleagues. For each stage, both one example of neuronal (mean firing rate) and ocular responses (mean eye velocity) are given together with a distribution of direction selectivity encountered within the neuronal population. Visual stages such as area MT/MST and DLPN are characterized by a full distribution covering all motion directions in the frontal plane (Kawano et al., 1992, 1994). On the contrary, in the ventral paraflocullus of the cerebellum (VPFL) only horizontal and vertical directions are represented indicating that contribution of extra-ocular muscles, rather than visual motion directions are represented at the level of Purkinje cells (Gomi et al., 1998). One can then follow the flow of information from visual to motor structures. Initiation of short-latency ocular following responses (~55 ms) is preceded by the onset of neuronal activity at ~40 ms in cortical areas MT and MST (Kawano et al., 1994). This early onset is consistent with the fast magnocellular input from the retino-geniculo-cortical pathway, with a relay in spiny stellate neurons of layer 4B in area V1 (Shipp and Zeki, 1989; Yabuta et al., 2001). In macaques, areas MT and MST project monosynaptically onto the dorsolateral pontine nucleus (DLPN), the main visual relay nucleus in the brainstem (Glickstein et al., 1994). Here, earliest activity was found ~10 ms before eye movement onset (Kawano et al., 1992). In contrast, neuronal activity

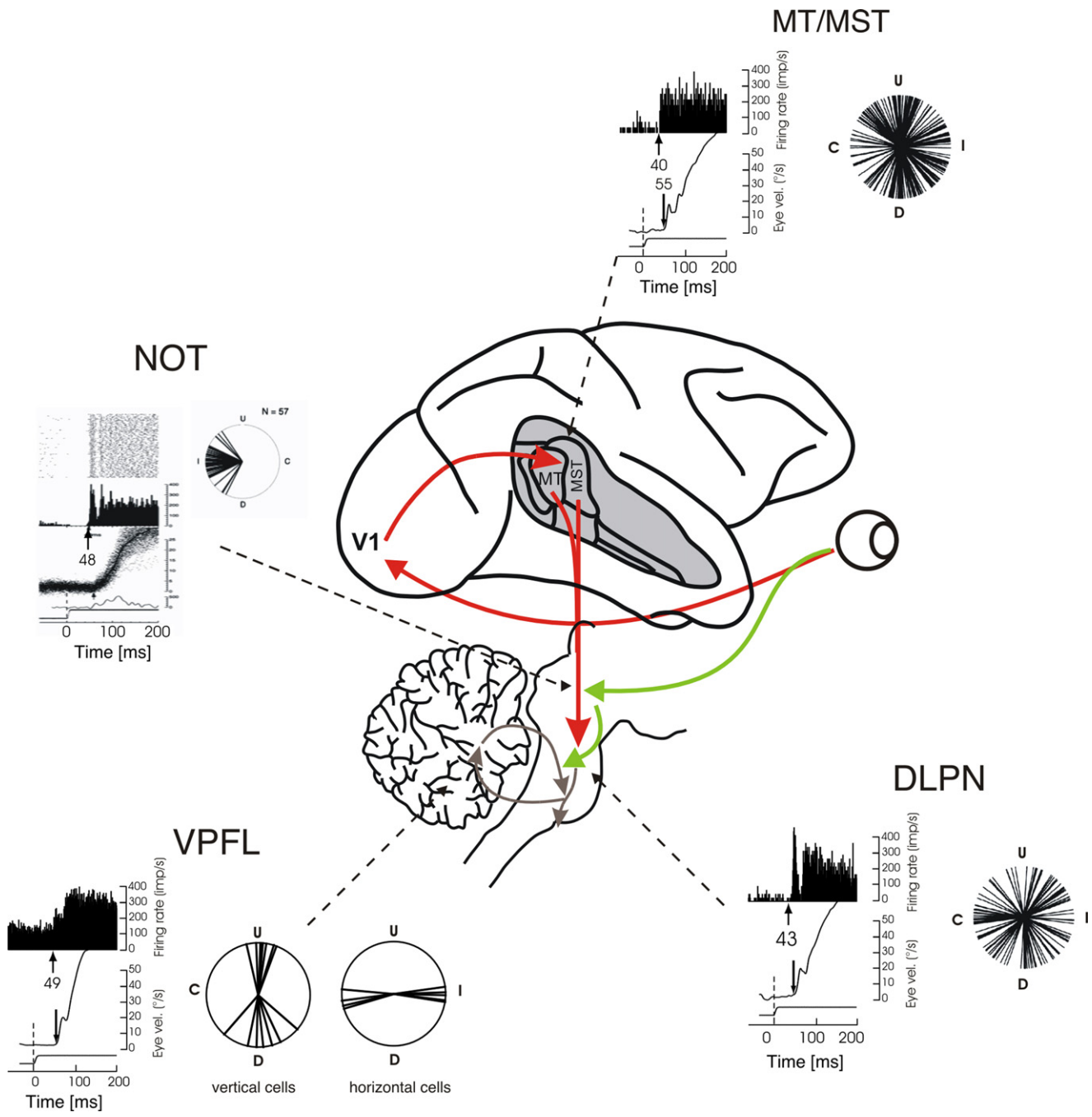


**Fig. 1.** Neural bases of ocular following in macaque monkeys. Single unit activity (mean firing rate) is shown at the level of cortical areas MT/MST, their direct projection nucleus in the brainstem, the DLPN and the oculomotor structures of the cerebellum (VPFL). Earliest neuronal latencies were respectively of about 40, 45 and 50 ms whereas eye movements onset was at ~55 ms after stimulus motion onset. Notice that earliest neuronal activity in the Nucleus of the Optic Tract, the main sub-cortical input for motion signals occurred only at ~60 ms after stimulus onset, that is after initiation of ocular following. Modified from Masson (2004) with permission from Elsevier.

in VPFL was time locked with the initiation of tracking. Kawano and colleagues suggested that visual information about motion of the visual scene is encoded in area MST and relayed via DPLN to the VPFL, which computes the motor commands driving eye muscles (Kawano, 1999). Accordingly, different signals related to either eye position or velocity, have been identified in the VPFL during ocular following (Shidara and Kawano, 1993; Gomi et al., 1998). Lastly, there are additional cerebellar afferences from the pretectal nucleus of the optic tract (NOT) that should be taken into account to render the exact wiring diagram (Inoue et al., 2000). Notice however that the earliest activity in NOT was seen only after tracking initiation, suggesting that subcortical inputs may play a role in the build-up of the response but not in triggering its earliest part.

A recurrent question regarding the neural basis of ocular following concerns the putative contribution of sub-cortical direct inputs to cortical areas along the motion pathway. Macaque area MT receives input from both the LGN and pulvinar. One might wonder whether these inputs could significantly contribute to ocular following responses. However, recent physiological evidences do not support this. First, Berman and Wurtz (2011) have dissected the signals coming from the superficial layers of the superior

colliculus via the inferior pulvinar (IP). They rejected the idea that they carry visual motion information since direction selectivity was minimal in IP neurons belonging to this ascending path. Second, Sincich et al. (2004) have provided a detailed anatomical analysis of direct geniculate input to area MT, bypassing V1. However, these geniculo-cortical neurons are mostly koniocellular and equal about 10% of the V1 population innervating MT. In macaque LGN, koniocellular cells are most often noted for carrying blue/yellow color signals (see Hendry and Reid, 2000 for a review). Their possible involvement in driving earliest responses in area MT is contradicted by the fact that color-based motion signal cannot elicit ocular following in monkeys at the ultra-short latency obtained with luminance-based motion (Guo and Benson, 1999). In brief, there is very little evidence that direct inputs to MT neurons coming from subcortical structures could play a significant role in the initiation of ocular following.

Although the information conveyed by spiking patterns of cerebellar neurons have been conducted with great details in the context of ocular following, there is still an urgent need for further research in two main directions. First, no information is available regarding the variability of neuronal responses in cortical areas MT and MST, and its impact upon the variability of the behavioral responses. Some hints are provided by the study of Lisberger and colleagues on smooth pursuit (see Lisberger, 2010 for a review). They showed that MT neuron responses to small target motions are highly correlated and that such correlation can both predict eye velocity amplitude and variance during the initial phase of monkey voluntary pursuit (Huang and Lisberger, 2009). This is important to understand how tracking eye movements are sensitive to noisy inputs and how vector averaging or other estimators of MT activity extracts speed and direction information for driving the eyes. Since sensitivity to noise could be different when using small target or large visual scene such as with ocular following, future work should further investigate the link between neuronal discharges in areas MT/MST and ocular responses. Second, very little works have been conducted to relate neuronal population activity in the primary visual cortex and ocular following responses (see Reynaud et al., 2007 for a preliminary report). Since such population activity can constrain behavioral responses in motion detection tasks for instance (Chen et al., 2006), it would be important to better understand the impact of early (V1) neuronal processing of motion information in the context of ocular following responses. We will see below that identifying the contribution of each cortical, and sub-cortical stages, into sensory, and motor, action would largely benefit from this.

Two questions remain largely open. First, it is always difficult to directly map neural mechanisms measured in macaque monkeys with putative neural pathways in human subjects. There is a large bulk of data suggesting that the human homologues of areas MT and MST are both involved in motion detection and integration (see Orban et al., 2004 for a review). It is thus highly plausible that essential cortical steps of motion processing for ocular following are similar in both species. Below, we will pinpoint striking behavioral similarities between monkey and human ocular following. These similarities, for spatial summation for instance, are a behavioral correlates of the similarities observed for spatial summation at population level in both macaque and human V1 (Nurminen et al., 2009). Still, there is an urgent need for running fMRI studies in order to correlate large-scale brain activity with ocular following responses in both monkeys and humans. Such brain imaging approach will be also important to enlarge our current view about the cortical network involved in driving ocular responses. We might find out that other cortical areas such as V2 or V4 are involved in late parts of (or delayed) responses driven by specifically designed stimuli such as plaids, barber-poles or color gratings (Masson and Castet, 2002; see Masson, 2004 for a review).

## 4. Sampling and integrating local motion: the idea of a behavioral receptive field

Tracking the cortico-subcortical sequence of neuronal activity driving ocular following is not sufficient to elucidate the neural basis of gaze stabilization. In the present article, we will extensively review the recent behavioral evidences that early linear and nonlinear visual mechanisms have a profound impact on the time course and dynamics of ocular behavior. Overall, our goal is to tease apart the contribution of each processing stage along the V1–MT–MST cascade when linking neural activity and behavioral dynamics. It is also to propose a theoretical framework for this link. We have recently proposed that the idea of a *behavioral receptive field* (bRF) can collapse these different operations into a compact representation of information processing (Barthélemy et al., 2006). Moreover, a strong isomorphism can be found between particular aspects of this bRF and what is known about neuronal receptive field (nRF) properties of direction-selective cells.

The idea that similar computations characterize behavior and physiology is not new but is at the core of the concept of perceptive field as originally proposed by Jung and Spillmann nearly 40 years ago (see Jung and Spillman, 1970; Ehrenstein et al., 2003; Neri and Levi, 2006; Spillman, 2006 for reviews). This approach aims to compare the functional properties of single neuron, neuronal populations and the whole organism. It has been demonstrated in some cases that idiosyncratic response properties of retinal or cortical neurons are propagated through subsequent stages in visual processing and can be observed in human behavior (e.g. Jung and Spillman, 1970; Ringach, 1998; Neri et al., 1999). Using reverse correlation, perceptive fields have been measured for orientation or motion direction domains (Ahumada, 1996; Neri, 2004; Solomon, 2002; Ringach, 1998; Murray et al., 2003). In many cases, there was a good match between the perceptive fields and the neuronal receptive fields of early cortical filtering. When such match is found, bottleneck information processing steps can be identified. However, this psychophysical approach suffers from severe limitations that can be overcome with ocular following.

As evidenced by the seminal work of Ahumada (1996), a perceptive field indicates the output of this final decision stage and thus aggregates properties of several neural processing units. In psychophysical tasks, the exact number of different hierarchical steps differs from one task to another. On the contrary, from the work of Kawano and colleagues, it is known that no more than 3 visual processing stages (i.e. V1, MT and MST areas) are necessary for driving ocular following in monkeys. Because all the results presented below were always collected within a very small time window (i.e. between 50 and 100 ms after stimulus onset in monkeys), behavioral receptive field reflect only transient processing along a minimalist cortical stream. Therefore, matching properties of bRF with key computations done within nRF at different stages is surely less risky than with perceptive fields.

The second limitation concerns the nonlinear properties of behavioral responses. Because different high-level factors influence the absolute psychophysical performance (Ahumada, 2002), it is most often inappropriate to establish a direct link between gain changes in perceptive fields to underlying gain changes in front-end filtering. In particular, decision criteria involved psychophysical tasks (Ahumada, 2002), as well as attentional factors (Murray et al., 2003) have a strong influence on the gain of the perceptive field. Ocular following offers a much safer alternative. Initial eye acceleration is related only to the spiking frequency of MST neurons (Takemura and Kawano, 2002). Only one final neural decision mechanism is involved since the firing rate of these MST neurons reflects the vector average of the MT population encoding

image direction and speed (Kawano et al., 1994). We will see below how such minimalist read-out allow us to map nonlinear aspects of the bRF onto a functional architecture downstream to area MST.

The third limitation is that behavior reflects the activity of population of neurons and not a single unit. This was originally stressed by Sherrington (1906) when he first coined the term of receptive field. For him, a receptive field was a functional description of an ensemble of peripheral sensors that would trigger an appropriate behavioral response (the scratch-reflex in his original description). We will see that bRF can be described by a set of functions describing the input–output transfer function. Still, each operator must be seen as a computation (e.g. gain control) performed by a population of inhibitory and excitatory neurons embedded in feedforward, feedback and recurrent connectivities. Thus, isomorphism between bRF and nRF can be used to probe such a complex functional architecture. However, we face a scaling problem. Tuning properties can be measured at behavioral and neuronal levels along the same dimension (i.e. speed and disparity) but bRF tunings encompass the full behavioral range while nRF tuning operates over a smaller sample. For instance, psychophysical (Neri et al., 1999) and ocular (Masson et al., 1997) responses to anti-correlated stimuli resemble the properties of disparity-selective neurons in macaque areas V1 (Cumming and Parker, 1997). Subsequent studies found in macaque area MST a nearly perfect fit between short-latency ocular (vergence) responses and neuronal properties at population level. However, it is only in area MST that most of the neurons encode the same range of disparities as the behavior and that both population and behavior disparity–tuning functions coincide (Takemura et al., 2001). We will show that a similar problem occurs when considering speed tuning of ocular following.

The last advantage of bRF over perceptive fields is related to the nature of smooth eye movements. Temporal dynamics of visual processing can be tracked down to the ms time scale from the time course of eye velocity profiles. In macaque monkeys, such precise timing can be used to elucidate which stage among V1, MT or MST areas determines a particular computation of the bRF. Varying spatial aspects of visual information such as visual eccentricity was a seminal strategy used by Jung and Spillman (1970) to identify which processing stage constrain the perceptive field. Other strategies such as dichoptic presentation or isoluminance can be used to narrow down the underlying brain loci. Here, we will show that temporal dynamics is a key dimension to consider when identifying the different computational elements of the bRF.

Fig. 2 illustrates our strategy to document the properties of the bRF and to link them to specific processing stages along the primate cortical motion pathway. Our objective is to reverse-engineer the near-end cortical populations from a behavioral performance. We can record reflexive tracking responses in both human and monkeys. These machine-like responses can be binned and quantified over a small 100 ms time window. Tuning functions can be reconstructed for various dimensions, as illustrated by speed tuning. Most often such behavioral tuning function coincides with single neuron tuning curves documented in area MST. These neurons have a large receptive field and integrate information from a large population of MT neurons. By presenting a motion stimulus with a given speed (grey symbol) we can probe the dynamics of a given MT subpopulation. Keeping speed constant while varying contrast, size, signal-to-noise or other aspects of the motion stimulus, we can probe the linear and nonlinear aspects of the motion integration process for this particular MT subpopulation. Because most of these computations have different contrast dynamics as well as time courses, we can identify how interactions done within V1 or MT populations can constrain the bRF. For instance, it is known that contrast gain control from surround motion is direction-selective in area MT neurons, but not in V1 (Rust et al., 2006). Step by step it becomes possible to dissect the role of each stage in the emergence of the behavior and open the door to detailed models of brain in action.

Whenever it is possible, experiments seeking for a direct comparison between neuronal dynamics and behavior shall be run. However, this is not possible in human subjects. A complete description such as the bRF helps us to elucidate the neural mechanisms underlying human behavior. Below, we will systematically compare tuning properties and temporal dynamics of human and monkey ocular following. Only the latter will be compared to neuronal data available from the literature. Still, a coherent picture emerges that highlights the detailed properties of human motion processing.

Below, we will summarize the properties of such a behavioral receptive field along different dimensions (spatio-temporal tuning, contrast, speed,. . .) that are characteristic of motion detection and integration stages. These stages are largely seen as being implemented by cortical areas V1 and MT/MST. When neuronal data are available, we will directly compare these behavioral properties with the properties of populations of V1, MT and MST neurons. Lastly, we will summarize the different linear and nonlinear operators forming this bRF and show their similarities with neuronal operators found in monkeys.

## 5. Dynamics of first stage: local motion detection

The seminal work of Rashbass (1961) demonstrated that visual motion is the primary information for smooth tracking eye movements: when a small, point-like target, is briefly stepped in one direction and then ramped into the opposite direction, smooth pursuit eye movements are initiated in the direction of the ramp (the so-called velocity error) and not the direction of the step (the so-called position error) (see also Carl and Gellman, 1987). The key question is then: how such information about speed and direction can be reconstructed?

Computing motion of a single surface is often described as a two stage process where a motion detection stage extracts local spatio-temporal changes from the images and feeds an integration stage that computes the global solution from these piecewise measurements (Nakayama, 1985; Albright and Stoner, 1995; Braddick, 1993 for reviews). In primate, areas V1 and MT/V5 are often seen as implementing detection and integration stages, respectively (Movshon et al., 1985). Several studies have investigated the contribution of the different motion detection mechanisms as well as the dynamics of motion integration to ocular following. Several recent reviews more specifically focus on the computational properties of each stage (Miles and Sheliga, 2010; Miles et al., 2004; Masson, 2004). We will summarize the most important aspects to define the linear filtering stage of the behavioral receptive field, its contrast dynamics and contextual modulations.

### 5.1. Linear motion detection: motion energy

Visual motion detection is often seen as a puzzle of different mechanisms, each of them sensing motion from different spatio-temporal cues in the image. Lu and Sperling (1995, 2001) have theorized this approach by suggesting three different systems for visual motion perception. The first-order, also called short-range (Braddick, 1974) or Fourier-based (Chubb and Sperling, 1989; Cavanagh and Mather, 1989), motion system senses direction and speed from the spatio-temporal changes in the retinal luminance pattern through a set of linear spatio-temporal filters called motion energy detectors (Adelson and Bergen, 1985; Van Santen and Sterling, 1984; Watson and Ahumada, 1985; see Derrington et al., 2004 for a recent review). These elementary detectors correspond to complex cells in primary visual

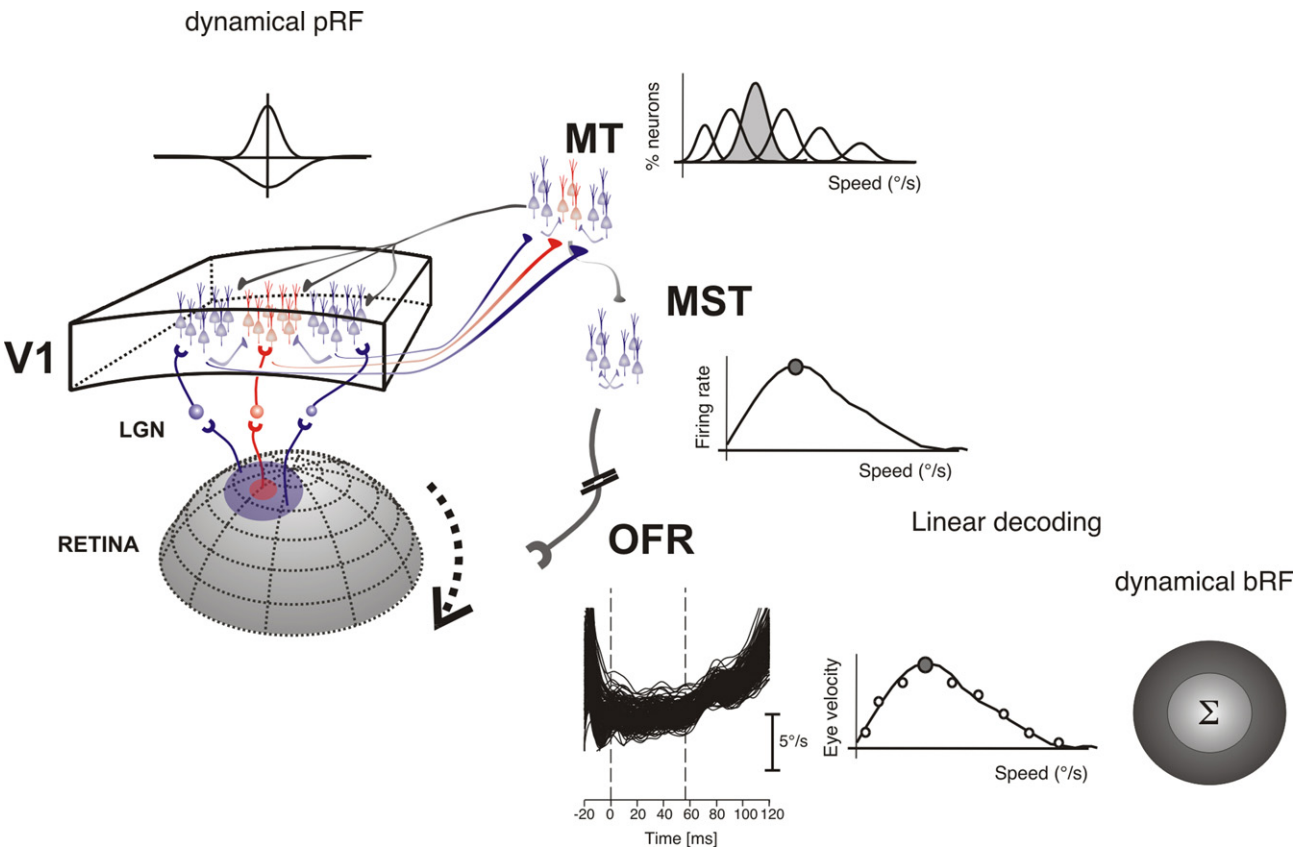


**Fig. 2.** Reconstructing population dynamics for behavioral responses. Left-side plot illustrates the hierarchy of neuronal population yielding to ocular following responses. Complex stimuli covering a large part of the retina are processed by a cascade of neuronal populations, where information is processed through feedforward, lateral and feedback interactions. Such cascade implements local, context-dependent extraction of motion information (area V1), reconstruction of target motion (MT) and effective read-out estimating the direction and speed of the global motion to be pursued, driving sub-cortical nuclei and ultimately ocular following responses. Each step can be described as a population receptive field (pRF) whose properties can be inferred from the behavioral receptive field (bRF) under some assumptions. For instance, using a constant speed that is decoded at level of area MST, we can probe the spatial interactions underlying motion processing by a specific sub-population of speed-tuned MT neurons.

cortex (Emerson et al., 1992). Under some circumstances, motion can be perceived while the spatio-temporal luminance profile remains constant. Several authors have postulated the existence of a 2nd-order or non-Fourier motion processing which sense motion from changes in contrast cues such as texture or feature elements (see Lu and Sperling, 1995, 2001). 1st- and 2nd-order motion mechanisms have very similar spatial and temporal frequencies tuning, but 2nd-order motion is slightly delayed relative to energy-based motion signals (Wilson et al., 1992). Wilson et al. (1992) have proposed that an early nonlinearity, such as half-square rectification might be sufficient to make second-order cues visible to motion energy detectors albeit at a different spatial scale (Wilson et al., 1992). Others have suggested that 2nd-order motion detectors solve the correspondence problem and therefore track significant features that are first to be extracted from the images (Cavanagh and Mather, 1989; Smith, 1994). This is however still a matter of debate as some authors have proposed that such a feature-tracking motion system (Ullman, 1979) may be dependent on our ability to attend to the features being tracked and therefore is different from the 2nd order, pre-attentive motion system (Derrington et al., 2004). Following this view, feature-tracking mechanisms are more similar to the 3rd-order, attention-based motion system postulated by Lu and Sperling (2001). This latter system is a much more sluggish mechanism that senses motion from tokens extracted from some spatio-temporal saliency maps. Thus, the first question arises: what is the respective contribution of these different motion subsystems to ocular following?

*5.1.1. Contribution of luminance-based motion energy*

Several studies have demonstrated that ocular following responses are primarily driven by luminance-based motion information. The first evidence was given by Masson et al. (2002a) in humans (Fig. 3a–c). Reversing the contrast polarity of a high-density random dot pattern (Fig. 3a) elicits motion perception in the direction opposite to the physical displacement of the pattern. Such inverted perceived motion direction was called *reversed phi motion* by Anstis (1975) and can be easily explained from a Fourier analysis of the motion stimulus: as predicted by linear system analysis, reversing contrast polarity across the single step of apparent motion flips the sign of all Fourier frequency components. In particular, the largest component of first-order motion energy is then in the direction opposite to the physical displacement. Masson et al. (2002a) found that *reversed phi* motion elicited reversed ocular following with the same latency than responses elicited by forward apparent motion (Fig. 3b). When examining the relationship between response amplitude and step size, similar tuning functions were observed albeit with opposite polarity (Fig. 3c). This strongly suggests that ocular following responses are driven by low-level motion detectors extracting motion energy through some kind of linear filtering. Similar evidence was found for another member of the family: disparity-driven vergence showed the same sign inversion when disparity was introduced between binocularly anti-correlated left and right eye images (Masson et al., 1997; Takemura et al., 2001).

The second set of evidences was provided by Miles and colleagues using the missing fundamental (*mf*) motion stimulus

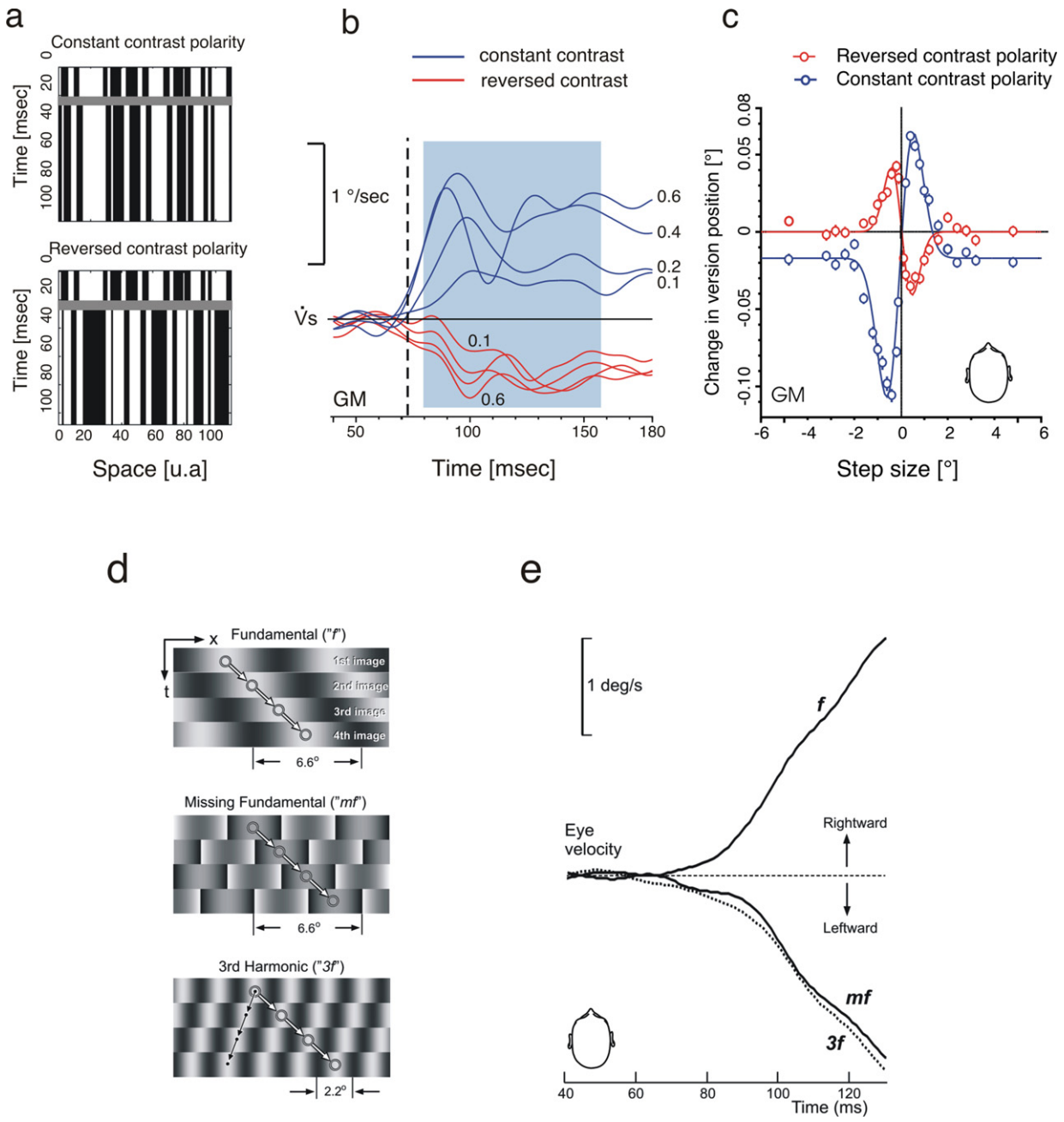


**Fig. 3.** (a,b) Forward and reversed ocular following responses. (a) Space–time diagram of the one-step apparent motion applied to a random-dot pattern (50% density, dot size ~2°). Upper and lower diagrams plot the two conditions where contrast polarity of each element is either maintained constant or reversed during the step, respectively. (b) Ocular following responses to constant (blue curves) or reversed (red curves) contrast polarity stimuli. Numbers indicate the size of the physical step (in°). Blue area indicates the time window over which the change in eye position was computed for each trial. (c) Relationships between amplitude of the response and step-size, for each condition. Continuous lines are best-fit Gabor functions indicating that similar relationships (but of opposite sign and different total amplitude) were found for both forward and reversed conditions (Masson et al., 2002a). Reprinted from Masson (2004). Elsevier. (d,e) Ocular following response to the missing fundamental stimulus. (d) Space–time diagrams of a sine-wave grating, the missing fundamental square-wave grating (*mf*) and its third harmonic (3*f*) component. (e). Initial ocular following responses obtained with a sine-wave grating (*f*) moving in the rightward direction, a mf stimulus moving in the rightward direction or a sine-wave grating at the same frequency as the 3f component of the mf stimulus albeit moving in the leftward direction. Modified after Miles and Sheliga (2010) with permission from Springer. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

illustrated in Fig. 3d (Chen et al., 2005; Sheliga et al., 2005, see Miles and Sheliga, 2010 for a review). In the *mf* motion stimulus, the fundamental frequency of a square-wave grating is removed. As pointed by Adelson (1982), the *mf* stimulus has the essential property that when set into motion by shifting its phase by ¼ wavelength steps, the $4n+1$ harmonics (like the 5th, the 9th,...) are all shifted in the *forward* direction while the $4n-1$ harmonics (like the 3rd, the 7th,...) are shifted in the *backward* direction. The amplitude of the $i$th harmonic is proportional to $1/i$, so that the major Fourier component is the 3rd harmonic (also called 3*f*). Psychophysical studies have clearly shown that in such condition, the *mf* stimulus is perceived as moving in the opposite direction to its physical translation (Adelson, 1982; Adelson and Bergen, 1985; Baro and Levinson, 1988; Brown and He, 2000). Such reversed perceived direction can be understood if motion perception is primarily driven by 1st-order motion detectors sensing the most prominent energy component from the image.

Quarter wavelength steps applied to *mf* stimuli elicit ocular following responses in humans at the usual ultra-short-latency (Fig. 3e), albeit in the backward direction (Chen et al., 2005; Sheliga et al., 2005). Fig. 3e plots the velocity profiles of responses obtained with a pure sine wave grating at the fundamental frequency (*f*) or at the spatial frequency, contrast and direction of the 3*f* component from the *mf* stimulus. Clearly, responses to *mf* and 3*f* motion stimuli were indistinguishable: ocular following was driven in the direction of the prominent frequency component of complex stimuli, as expected from a system driven by motion-energy mechanisms. Similar conclusions were reached with ocular following responses in monkeys, also using the *mf* stimulus (Miura et al., 2006). Moreover, responses driven by either *f* or *mf* stimuli scale similarly with contrast both in terms of response amplitude and latency (Sheliga et al., 2005; Miura et al., 2006).

#### 5.1.2. Lack of evidence for a contribution of second- and third-order motion

If luminance-based motion sensors drive early ocular following, what would be the contribution of the two other motion systems? Feature-tracking motion detection is believed to be extremely slow. In particular, it fails with very brief stimulus presentation (<200 ms) or very high temporal frequency stimuli because there might not be time to locate the features before they move (Derrington and Badcock, 1992). The ultra-short latency of ocular following argues against a significant contribution of feature tracking for eye movement initiation. Interestingly, human subjects can sometime perceive motion in the correct direction of the *mf* stimulus by picking up some specific features in the images, though higher level motion mechanisms. This never occurs in monkey or human ocular following. Thus, early ocular following is essentially pre-attentive and immune to contribution of third-order motion system. Coherently, investigating the initiation of voluntary tracking in human observers, Wilmer and Nakayama (2007) suggested that motion sensor might trigger the earliest phase of pursuit, while feature tracking mechanisms would be engaged much more later, that is after the first catch-up saccade most often occurring later than 200 ms after target motion onset.

If feature-tracking mechanisms have little chance to contribute to the initiation of ocular following responses, it remains possible that a 2nd-order motion mechanism does contribute. The role of motion cues based upon contrast modulation of texture elements in the image has been a matter of debate for the last two decades. Some authors have proposed the existence of a specific, non-Fourier motion system that can extract texture information through some nonlinear operation (Lu and Sperling, 1995, 2001). However, the results compiled from the psychophysical literature are highly contradictory, so that some authors have suggested that there is no specific 2nd-order motion sensors and that, instead, motion direction of contrast-modulated pattern are sensed either by regular motion energy sensors (at low contrast at least) and by feature-tracking mechanisms (at high contrast) (Derrington et al., 2004). This is consistent with the physiological observations that MT neurons are only weakly driven by contrast-modulated patterns whereas they show high responsiveness and strong direction-selectivity for luminance patterns (O'Keefe and Movshon, 1998).

What about ocular following? An early study showed, but in only one monkey, that a low-contrast texture pattern can elicit ocular following with a slightly delayed latency but with similar spatio-temporal tuning when compared with responses to luminance gratings (Benson and Guo, 1999). However, in a very recent report, the group of Kawano reinvestigated this question in both humans and monkeys. They found that pure second-order motion did not elicit short-latency ocular following (Hayashi et al., 2008). This further suggests that mostly, if not solely, first-order motion sensors trigger ocular following eye movements in primates.

### 5.2. Linear motion detection: spatio-temporal tuning

In their original studies, Miles and coworkers examined the dependence of OFRs on the spatio-temporal properties of the input. In both species, they found that OFRs were best driven by low spatial frequency stimuli (cutoff: 1 cpd) and mid-to-high (>8 Hz) temporal frequencies with optimal values around 0.3cpd and 10 Hz (Miles et al., 1986; Gellman et al., 1990). The spatial and temporal tunings of motion energy mechanisms driving ocular following has been fully characterized by several studies conducted in both humans and monkeys. Fig. 4a and d plots the normalized response amplitude again spatial frequency of grating patterns. In both human (Sheliga et al., 2005) and monkeys (Miura et al., 2006), spatial frequency tuning is best fitted by a Log-skewed Gaussian function, peaking and 0.2–0.3 cpd and of width of ~0.5 log units. Fitting functions and best-fit parameters are given in Table 1. We found similar tuning parameters in human subjects using bandpass filtered noise patterns. They remained constant over a large range of contrast (Drewes et al., 2007). Spatial frequency tuning were always measured using large field stimuli, covering a broad range of retinal eccentricities. Optimal frequency range and tuning width match that of population measurements done in visual areas MT/V5 + in humans (Singh et al., 2000; Henriksson et al., 2008) albeit the band pass tuning of ocular following function more closely fit that of V1 at large eccentricity. In monkeys, a large majority of MT similar band-pass tuning peaking at spatial frequency around 0.3 cpd (Lui et al., 2007), A direct comparison is still needed using high speed stimuli as used for ocular following to understand which neuronal mechanism shape the behavioral spatial frequency tuning function.

Fig. 4b and e plots the relationship between response amplitude and temporal frequency of a drifting sine wave grating at optimal spatial frequency. In both species, strongest responses were obtained with mid-range temporal frequency (~10 Hz). Similar low-pass temporal tuning was found with other spatial frequencies, albeit with some jittering across conditions (Miles et al., 1986; Gellman et al., 1990). Interestingly, temporal frequency cutoff (as defined by the temporal frequency yielding two-thirds of peak response) was found between 20 and 30 Hz. A similar frequency cutoff was reported for V1 neurons when presented with drifting gratings (Hawken et al., 1996; Nienborg et al., 2005). Lastly, temporal frequency tuning such as illustrated in Fig. 4e exhibits many of the key properties of V1 neurons projecting to area MT in macaque (Movshon and Newsome, 1996), suggesting that the spatio-temporal dependencies of ocular following are set at a very early stage along the cortical motion pathway. Coherently, a subpopulation of MT neurons exhibits a similar band-pass tuning peaking at high temporal frequencies (8–10 Hz). Both single neuron and ocular following tuning for temporal frequency are best fitted by a log-skewed Gaussian function (see Table 1).

Latency of ocular following is largely independent upon spatial frequency of single gratings as well as dot size of random dot patterns, in both humans and monkeys (Miles et al., 1986; Gellman et al., 1990). Latency was also weakly affected by motion speed. However, when plotting latency against temporal frequency of single sine-wave moving gratings, they found a strong decaying exponential relationship, mostly independent upon the grating spatial frequency (Fig. 4c and f). Minimum latencies were found for temporal frequencies in the 20-40 Hz range, for both humans (~85 ms) and monkeys (~55 ms). At very low temporal frequencies (<1 Hz), mean latency increased up to about 100 ms in both monkeys and humans. Overall, these results support the existence of a local mechanism that senses the temporal change in luminance to trigger ocular following (Miles et al., 1986). Interestingly, optimal temporal frequencies are different for either response amplitude (~10 Hz) or latency (~20–30 Hz) in the two species. This argues in

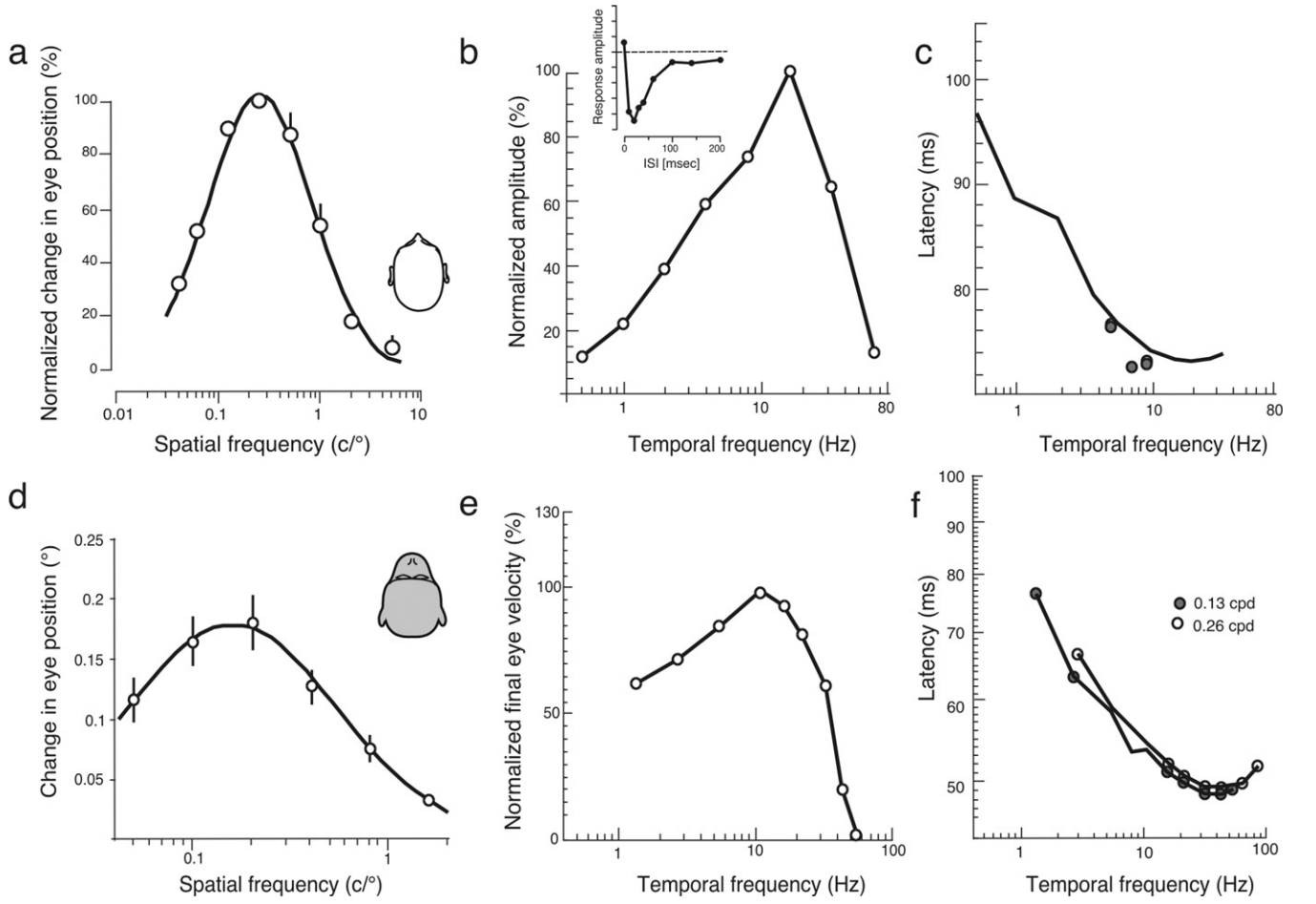


**Fig. 4.** Temporal and spatial tunings of ocular following in human (upper row) and macaque monkey (lower row). (a,d) Relationship between amplitude of the earliest ocular following and spatial frequency of a drifting grating. The smooth black curves are best-fit Gaussian functions. Reprinted from Sheliga et al. (2006a,b) and Miura et al. (2006). (b,e) Temporal frequency tuning of ocular following amplitude, obtained with pure sine-wave gratings. Inset in the human data plot illustrates the temporal impulse response measured for ocular following with two gratings presented successively after different ISI. Replotted from data published in Gellman et al. (1990) and Miles et al. (1986), respectively. (c,f) Dependency of ocular following latency upon the temporal frequency of the grating. For monkey results, the relationship is plotted for two low-spatial frequencies. Replotted from Gellman et al. (1990) and Miles et al. (1986), respectively.

favor of separate information processing for triggering responses on the one hand and controlling eye velocity on the other hand, as originally suggested by Miles et al. (1986).

### 5.3. Linear motion detection: temporal properties of motion detectors

The band-pass properties of the temporal tuning function can be explained from the temporal properties of the motion energy detectors. Recent studies using two frames movies (i.e. single steps of a vertical grating) to elicit ocular following responses found that brief Inter Stimulus Intervals (ISIs, range 10–100 ms) reversed the initial direction of these responses (Sheliga et al., 2006a; Kodaka et al., 2007). These reversals are reminiscent of the reversed perceived direction observed with brief ISIs (Pantle and Turano, 1992; Shiori and Cavanagh, 1990). They are usually attributed to the temporal dynamics of motion detection mechanisms in early visual pathways and in particular to the negative phase of the biphasic impulse response function of motion detectors (Takeuchi and De Valois, 1997; Takeuchi et al., 2001). In this schema, the polarity of the visual inputs to motion detectors is assumed to undergo reversal during the ISI, so that the neural representation of the second image is now matched to a representation of the first image which is now of opposite contrast polarity. Miles and coworkers performed the same experiment at both low (scotopic) and high (photopic) illumination conditions. As expected, higher eye velocities were obtained at high illumination and reversal were also more evident at both 40 and 100 ms ISIs. Reducing the mean illumination level to scotopic conditions resulted in smaller response and a reduced ISI window (i.e. 60–100 ms) for motion reversal. They proposed that these results correspond to the behavioral counterpart of the changes in the human modulation transfer function from band-pass to low-pass in the frequency domain and from biphasic to monophasic in the time domain when going to light-adapted to dark-adapted conditions (Kelly, 1961, 1971a,b; Roufs, 1972; Swanson et al., 1987). Since psychophysical studies have generally agreed that low-level, first-order motion mechanisms are triggered at short ISIs (<100 ms, providing that spatial distance was not too large), Miles and colleagues concluded that reversed ocular following reflects the temporal properties of low-level motion detectors. It is another example that fundamental aspects of motion processing such as biphasic temporal impulse response of early motion detectors can be probed at the level of short-latency ocular following to understand the functional properties of populations of motion sensitive neurons in cortical areas MT and MST.

### 5.4. Low-level motion detection: contrast dynamics

We have seen that ocular following is best driven by low spatial frequencies (<1 cpd) and high temporal frequencies (10–30 Hz). In

**Table 1**
A complete set of descriptive functions for modeling the behavioral receptive field underlying ocular following, together with average values of best-fit parameters. References corresponding to the original papers are given in the text.

| Computational step | Descriptive functions | | Humans | Monkeys |
|---|---|---|---|---|
| *Spatio-temporal filtering* | | | | |
| Temporal impulse response | $f(t) = (kt)^n \exp(-kt)\left[\frac{1}{n!} - \frac{B(kt)^2}{(n+2)!}\right]$ | (1) | $n = 1$<br>$k = 0.12$<br>$B = 3.5$ | NA |
| Temporal frequency tuning | $R(tf) = R_{max}\left(\exp\left[-\frac{\log((tf/tf_{opt}))^2}{2(\sigma_{tf} - \zeta\log(tf))^2}\right] - \exp\left(-\frac{1}{\zeta^2}\right)\right)$ | (2) | $tf_{opt} = 12$ c/s<br>$\zeta = 0$<br>$\sigma_{tf} = 0.46$ c/s | $tf_{opt} = 5$ c/s<br>$\zeta = 0$<br>$\sigma_{tf} = 0.55$ c/s |
| Spatial frequency tuning | $R(sf) = R_{max}\exp\left[-\frac{\log(sf/sf_{opt})^2}{2\sigma_{sf}^2}\right]$ | (3) | $s_{opt} \sim 0.3$ c/°<br>$\sigma_s \sim 0.5$ log unit | $s_{opt} \sim 0.3$ c/°<br>$\sigma_s \sim 0.5$ log unit |
| *Gain control* | | | | |
| Response latency | $\tau_c(c) = \tau_{max} + \tau_{shift}\cdot\left[\frac{c^n}{c^n + S_{50}^n}\right]$ | (4) | $n \sim 1.8$<br>$7 < S_{50} < 10$<br>$\tau_{shift} \sim 20$ ms | $n \sim 2.4$<br>$30 < S_{50} < 40$<br>$\tau_{shift} \sim 27$ ms |
| Response amplitude | $R = R_{max}\frac{C^n}{C^n + C_{50}^n}$ | (5) | $n \sim 2$<br>$5 < C_{50} < 10\%$ | $n \sim 2$<br>$10 < C_{50} < 30\%$ |
| *Spatial summation and center–surround interactions* | | | | |
| Spatial summation | $R(x) = \frac{L_c(x)}{1 + k_s L_s(x)}$ | (6) | $w_c = 10°$ | $w_c = 16°$ |
| | where $L_c(x) = \left(\frac{2}{\sqrt{\pi}}\int_0^x e^{-(y/w_c)^2}\,dy\right)^2$ | | $k_s = 6.2$ | $k_s = 12.5$ |
| | and $L_s(x) = \left(\frac{2}{\sqrt{\pi}}\int_0^x e^{-(y/w_s)^2}\,dy\right)^2$ | | $w_s = 10°$ | $w_s = 28°$ |
| *Speed tuning* | | | | |
| Speed tuning | $R(s) = R_{max}\exp\left[-\frac{\log[(s/s_{opt})]^2}{2\sigma_s^2}\right]$ | (7) | $s_{opt} \sim 40°$/s<br>$\sigma_s \sim 0.6$ | $s_{opt} \sim 41°$/s<br>$\sigma_s \sim 0.6$ |

primates, this range corresponds to the optimal tuning parameters of the magnocellular (M) component of the retino-geniculo-cortical pathway (see Shapley, 1990). Beyond area V1, this M pathway provides the main direct input to area MT (see Born and Bradley, 2005 for a review). These neurons, as well as their cortical target cells are sensitive to low spatial frequency (upper cutoff frequency, 1 cpd) and a rather large temporal tuning favoring high temporal frequencies (Shapley, 1990; Movshon and Newsome, 1996). They also manifest a preference for low contrast and have a high contrast gain as expressed by a steep contrast response function.

The dependency of ocular following upon contrast of drifting grating has been carefully investigated by several groups. Again, very similar properties were observed in both humans (Masson and Castet, 2002; Barthélemy et al., 2006, 2008; Sheliga et al., 2005) and monkeys (Miura et al., 2006; Barthélemy et al., 2010). Fig. 5a and d illustrates eye velocity profiles of ocular following for different contrast values. In humans, as contrast of a vertical drifting grating increases from 2.5 to 40%, response latencies become shorter and initial eye acceleration becomes stronger. For contrast higher than 40% no further changes were observed (Fig. 5a). Similar dynamics was found in monkeys (Fig. 5d). Fig. 5c and f plots the relationship between response amplitude and contrast. Masson and Castet (2002) have suggested to describe this ocular following contrast response function with a functional model similar to that used to describe contrast dynamics at neuronal levels for both single-cell spiking rate (Albrecht and Hamilton, 1982) and sub-threshold population activity recorded with voltage-sensitive optical imaging (Reynaud et al., 2007; Sit et al., 2010). The Naka–Rushton function (Naka and Rushton, 1966) nonlinearly links response amplitude ($R$) and contrast ($C$) (see Table 1) with only two free important parameters. Its slope sets the sensitivity of the neuronal mechanism while its half-saturation contrast determines its operating range. Similar best-fit parameters were found by the different studies. In monkeys, half-saturation values are ~5% contrast and slopes is ~2 (Miura et al., 2006). Barthélemy et al. (2010) found slightly higher half-saturation contrast as they used lower mean luminance and non optimal stimulus size (see below). In humans, mean half-saturation and slope parameters range around 8–10% and 1.8–2.2, respectively (Table 1). Barthélemy et al. (2008, 2010) investigated whether there was a change in the main parameters of the contrast response function over several successive time windows. There was a significant decrease in half-saturation contrast (from 10 to 5%) over a 100 ms time course, although major part of the reduction was seen during the first 40 ms. This indicates that contrast gain of the cortical pathway driving ocular following is set very rapidly, as found for V1 neurons (Albrecht et al., 2002). This responds to the need for a quick sensory gain control that can operate to optimize visual processing shortly after a saccade and during brief fixation periods (<200–300 ms) such as experienced during natural vision.

Fig. 5b and c illustrates the last aspect of contrast dynamics. As evidenced from the eye velocity profiles, reducing contrast from 50 to 2% resulted in a 50% increase in response latency in both species. Barthélemy et al. (2008) have shown that, similarly to V1 neuronal response (Albrecht et al., 2002, 2003), the relationship between response latency and contrast can be described using an *inverted* Naka–Rushton function (see Table 1). Best-fit parameters were very similar to those computed from the response amplitude, indicating that initial eye acceleration and response onset are linked together.

Fig. 6a compares humans and monkeys ocular following response functions (red curves) with that of neuronal populations at different cortical stages. Each curve is the best-fit function computed with the mean of the parameters distributions given by Sclar et al. (1990) for LGN, V1 and MT and Crowder et al. (2009) for MST. Clearly MT and MST curves fall in between the two behavioral functions. LGN magnocellular and V1 neurons have lower sensitivities

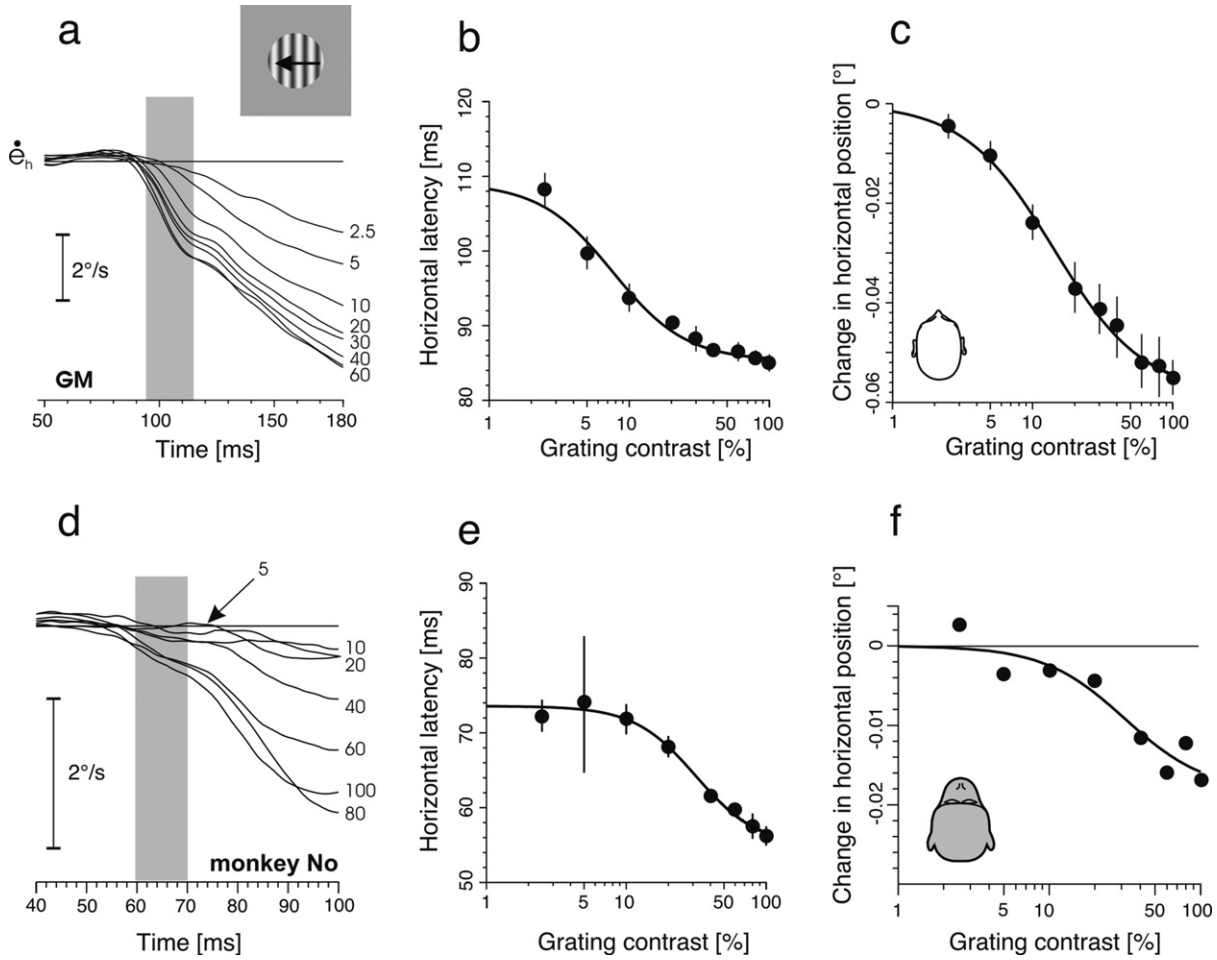


**Fig. 5.** Contrast dynamics of ocular following responses in humans (a–c) and monkeys (d–f). (a,d) Eye velocity profiles of responses to a single grating that is drifting leftward and is presented at different contrasts, as indicated by numbers. The shaded grey bar illustrates the early time window for measuring changes in eye position in each trial. (b,e) Latency of ocular following, plotted against grating contrast. Continuous line is best-fit inverted Naka–Rushton function. (c,f) Relationships between amplitude of early ocular following and grating contrast. Continuous lines are best-fits obtained with the Naka–Rushton function. Plots (a–c) are modified from Barthélemy et al. (2008). Data illustrated in (d–e) are replotted from Barthélemy et al. (2010).

because of their smaller spatial integration (Sclar et al., 1990). None of the behavioral curves match the V1 or LGN parvocellular functions, suggesting that contrast dynamics must be set at MT/MST stage.

### 5.5. Low-level motion detection: speed tuning

Initial eye acceleration is primarily sensitive to the speed of retinal motion (Miles et al., 1986). Fig. 6b illustrates the speed tuning function of monkey (grey symbols) and human (closed symbols) ocular following in response to brief motion of a large random dot pattern. These two sets of data were collected in exactly the same experimental conditions (Masson and Miles, unpublished results) for direct comparison but closely match with other datasets (e.g. Miles et al., 1986; Gellman et al., 1990; Kawano et al., 1994; Masson et al., 2002b). Both speed tunings exhibit an inverted-U shape function in log scale, peaking at ~40°/s. Data were best fitted with a log-skewed Gaussian function similar to that used for MT and MST neurons speed tuning (Nover et al., 2005) to compute tuning width and cut-off (see Table 1).

Superimposed are two sets of neuronal data. Red dotted curve plots the mean firing rate of a population of MST neurons recorded simultaneously to ocular following (Kawano et al., 1994). Clearly, speed tuning of behavior and MST population overlap, illustrating that MST neurons encode motion information for driving the eyes (Takemura and Kawano, 2002; Takemura et al., 2002). Also superimposed is the distribution of speed-tuned neurons in the macaque MT population (taken from Cheng et al., 1994, see also Maunsell and Van Essen, 1983). Such distribution fits the optimal velocity for driving MST neurons and ocular following. Black dotted lines are four theoretical MT neurons. Their log-skewed Gaussian tuning function was defined using the mean parameters obtained over a large sample by Nover et al. (2005) in awake macaques, and weighted by the population distribution of optimal speeds. From this, we can see that ocular following is driven by the population activity of MST neurons that sample speed information from MT neurons that are more narrowly tuned. Linear decoding of both MT and MST populations has been proposed for describing initial eye velocity of both voluntary (Priebe et al., 2003; Priebe and Lisberger, 2004; Lisberger, 2010) and reflexive tracking (Takemura et al., 2002). Fig. 6b illustrates the neuronal populations underlying the behavioral receptive field. As illustrated in Fig. 2, using a constant speed stimulus tap onto one specific, speed-tuned subpopulation of MT neurons.

### 5.6. Low-level motion detection: summary

We have shown above that ocular following exhibit many of the properties attributed to low-level motion detectors. These properties can be described by the first set of descriptive functions

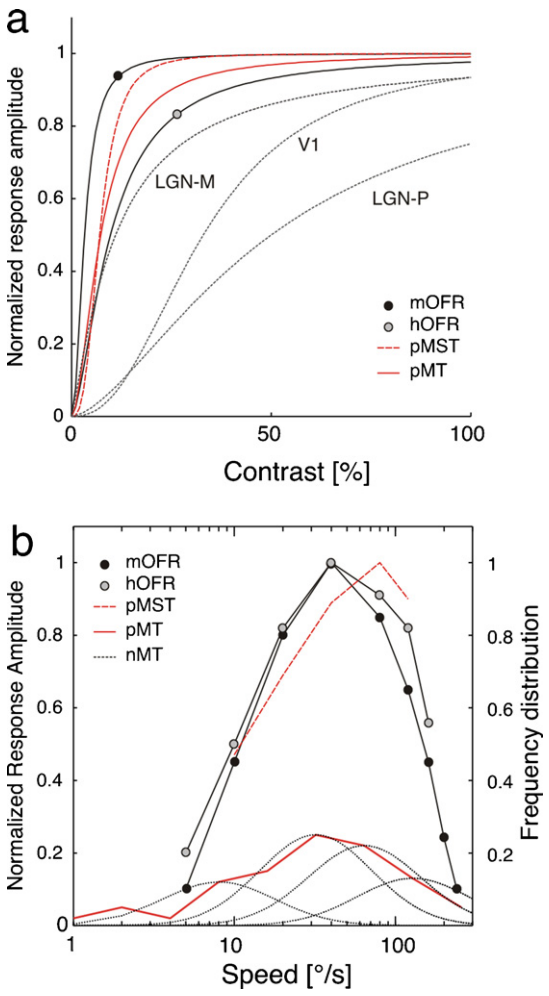


**Fig. 6.** (a) Contrast dynamics of ocular following responses in humans (hOFR) and monkeys (mOFR). Best fit Naka–Rushton functions were computed using mean parameter estimates from Barthélemy et al. (2008) and Miura et al. (2006), respectively. For comparison, these curves are plotted together with the best fit functions estimated for the population of MST neurons (Crowder et al., 2009) or MT cells (Sclar et al., 1990) as well as a population of V1, LGN magnocellular and parvocellular neurons (Sclar et al., 1990). (b) Speed tuning of human and monkey ocular following. Normalized response amplitudes of human (hOFR) and monkey (mOFR) ocular following are plotted against stimulus speed (data from Masson and Miles, unpublished). Red dotted curve plot the mean firing rate of a MST population (Kawano et al., 1994). Continuous red curve show the frequency distribution of optimal speed for a MT population (Cheng et al., 1994). Continuous grey curves are best fit tuning functions of three hypothetical MT neurons (parameters taken from Nover et al., 2005), weighted by their frequency distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

for both response amplitude and latency. These functions (spatial and temporal frequency tunings, speed tuning, contrast response, temporal impulse) are similar to that used for describing neuronal input–output transfer function. This suggests that ocular following results from a linear pooling of nonlinear units. Below, we will highlight several aspects of these nonlinear processing involved in local interactions between motion signals.

## 6. Beyond low-level motion detection: the needs for local motion integration

Local, first-order motion detectors that trigger ocular following face several pitfalls when processing natural images as, for instance, external noise, luminance and contrast fluctuations or the aperture problem. Solving these different problems requires to pool information across different spatial and temporal scales. It also requires to adaptively change response gain as a function of the local context. These mechanisms are useful to improve the estimate of local velocity. We will review the existing behavioral evidence and relate them to neural mechanisms found at different cortical stages, trying to tease apart their respective role. In particular, we will show how different types of divisive normalization act to produce optimal responses. In Section 7, we will review the impact of the neural solutions for the aperture problem for the behavioral output.

### 6.1. Combining early motion signals: linear or non-linear integration?

A single object contains different edges, features or textures that can yield to local motion signals along different directions and speeds. An immediate question is: how are these different local motions integrated to reconstruct a single and accurate estimate of object motion? The flipside of motion integration is segmentation. Integration must be selective to eliminate spurious motion belonging to background or other surfaces (Braddick, 1993). Different natural surfaces can have different global directions or speeds, different mean luminance or contrast (e.g. Balboa and Grzywacz, 2000; Frazor and Geisler, 2006) or different spatio-temporal frequencies. These low-level features can be used for automatic segmentation in order to single out the single global motion that needs to be pursued. Several studies have been conducted to decipher how different motion signals are automatically but selectively integrated for ocular following. They are rooted on earlier observations that opposite motion direction signals cancel each other (Busettini et al., 1996; Masson et al., 2001; Barthélemy et al., 2006) and that early phase of optokinetic nystagmus is driven by the vector average motion when several random dot fields moving in the same direction but with different speeds are presented simultaneously (Mestre and Masson, 1997). However, studies on voluntary and reflexive tracking have pointed out that, under some certain circumstances and timing, automatic target selection can occur so that ocular responses are driven by only one target at a point in time (Mestre and Masson, 1997; Ferrera, 2000). However, these delayed target selections must be done thought top-down mechanisms. On the contrary, ocular following studies point out the role of automatic scene segmentation based on low-level features (e.g. Busettini et al., 1996). Recent experimental work has clarified what are the different integration and competition rules as well as their biologically plausible mechanisms.

#### 6.1.1. Pooling motion information: vector averaging

Masson and Castet (2002) reinvestigated this point with moving plaids made of two orthogonal gratings of similar contrast and spatio-temporal frequencies. They found that the earliest phase of ocular following to these type I plaids are similar to the responses driven by a single oblique grating moving in the same direction and speed as the vector average prediction for plaid motion (Fig. 7a, left panel). In particular, latencies of responses to both type I plaid and single gratings were identical. Consistently, patterns made of several Gabor patches were presented with all patches having same carrier orientation/direction or the ensemble being divided into two sets of orthogonal carrier orientation/direction. Again, ocular following responses to either uni- or bi-kinetic patterns were indistinguishable (see Fig. 7a, right panel, Masson, 2004; Masson and Castet, 2002). These results suggest that ocular following is driven by a vector average computation that integrates all motion signals present in the central part of the visual field (see below) and in the plane of fixation (Masson, 2004). These results are consistent with previous psychophysical studies showing that, in the absence of special features, the visual system defaults to simple pooling over space to compute estimates of global velocity (Mingolla et al.,

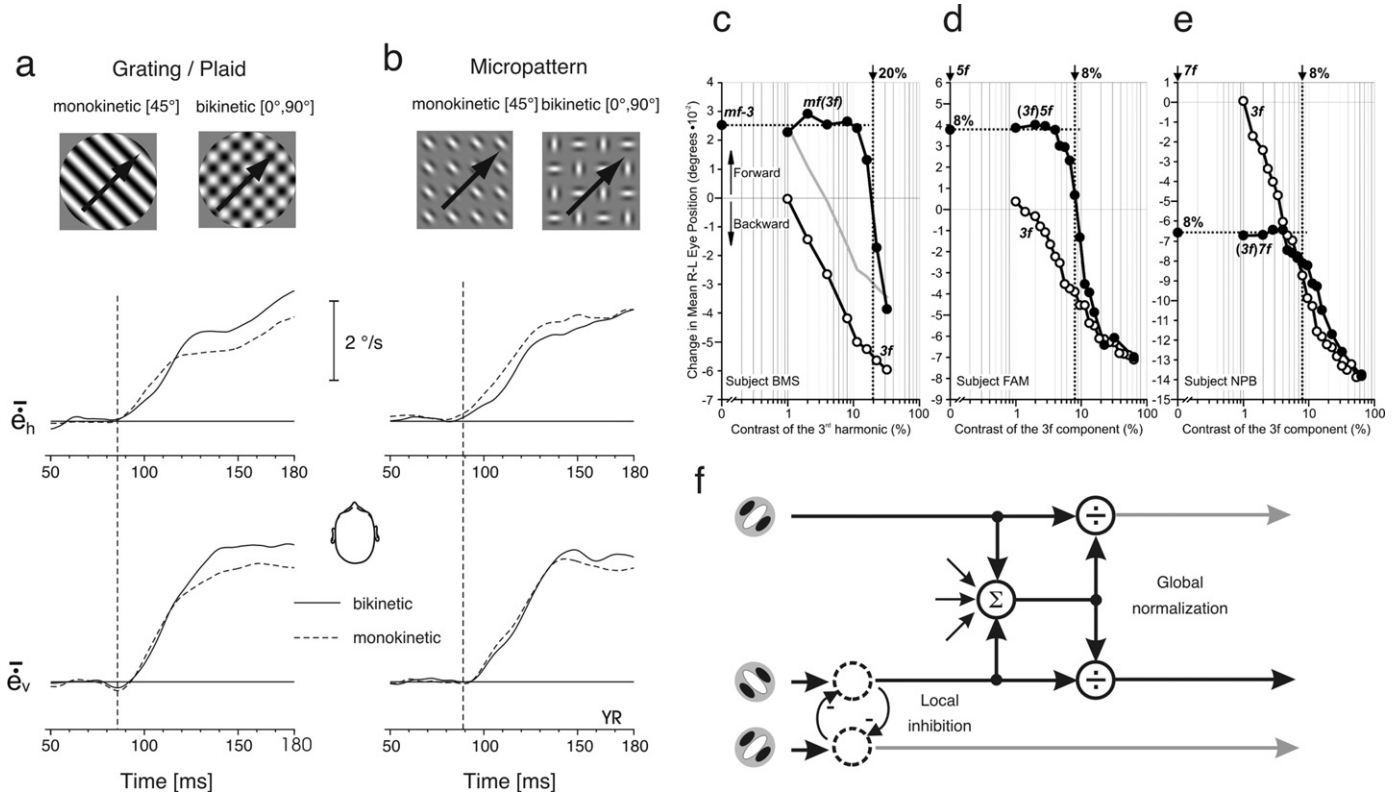


**Fig. 7.** (a,b) Ocular following responses to single vs multiple motion stimuli. (a) Mean horizontal and vertical eye velocity in response to either single grating (dotted line) and type I plaid (continuous line) whose global motion direction and speed were equivalent. (b) The same responses obtained with micropatterns made of either one type of Gabor patches or two sub-groups of Gabor with orthogonal carrier motion direction. Once again, the global motion direction and speed were identical. Reprinted from Masson (2004) with permission from Elsevier Ltd.. (c–f) Linear and non-linear interactions between competing motions. (c) Closed symbols plot (m*f*(3*f*)) the contrast response function of early ocular following responses driven by a missing fundamental stimulus where contrast of the 3*f* component was varied (from 32 to 1%) while the contrast of the other component was kept constant. For comparison, open circles plot the contrast response function obtained when varying the contrast of a pure 3*f* stimulus alone (i.e. a sine-wave grating). Point labeled m*f* − 3 indicates the responses obtained when removing the 3*f* component (i.e. contrast = 0) from the missing fundamental stimulus. Continuous grey line shows the vector average prediction from the responses to both m*f* − 3 and 3*f* stimuli. Clearly, when the contrast of the 3*f* component from the m*f*(3*f*) stimulus was reduced below 20%, no contribution of this motion stimulus was seen. Such critical value is equivalent to the contrast of the 5*f* component, now driving the responses in the forward direction. (d) Amplitude of ocular following response is plotted again the contrast of the 3*f* component when a single sine-wave of spatial frequency 3*f* is presented alone (open symbols) or together with a grating of 5*f* spatial frequency at fixed contrast (8%). The response to the later component presented alone is indicated by the single closed circle (labeled 8%) and the horizontal dotted line. Again, when the contrast of the 3*f* component dropped below the contrast of the 5*f* component, its contribution was purely eliminated. (d) Same experiment but now with a 3*f* and 7*f* components, moving in the same direction. (f) Model proposed by Sheliga et al. (2008a,b) to explain their results.

Reprinted from Miles and Sheliga (2010) with permission from Springer-Verlag.

1992; Rubin and Hochstein, 1993; Lorenceau and Zago, 1999). It is also coherent with the fact reported above that when two motion signals of opposite motion directions are located in the plane of fixation, they cancel each other (Masson et al., 2001).

These results are also consistent with observations that voluntary pursuit is also initiated in the vector average direction when presented with two targets (Ferrera and Lisberger, 1997; Lisberger and Ferrera, 1997; see Lisberger, 2010). Moreover, since earliest responses of area MT neurons to multiple motions correspond first to the vector average prediction, Recanzone and colleagues have argued that averaging occurs at this critical stage of motion processing (Recanzone et al., 1997; Recanzone and Wurtz, 1999). Altogether, these results strongly support the idea that motion signals of similar spatio-temporal contents and contrast are pooled together by the population of MT neurons underlying the behavioral receptive field.

### 6.1.2. *Pooling motion information for direction: from vector averaging to winner-take-all*

However, there is a strong need for automatic segmentation of visual scene to single out motion of the object of interest. Psychophysical studies have suggested that local motion signals can be integrated automatically only within a limited range of spatial and temporal frequencies (e.g. Adelson and Movshon, 1982; Welch and Bowne, 1992; Delicato and Derrington, 2005). Moreover, different surfaces in the visual environment can have different mean luminance or contrast values, providing other low-level segmentation cues. It is only very recently that experimental evidences have been gathered about how these low-level mechanisms are used by ocular following.

First, we found that, in a plaid pattern, moving gratings of very different spatial frequencies are not integrated together. Furthermore, when the two components of the plaid have very different contrasts, ocular following was driven in the direction of the highest contrast component (Masson and Castet, 2002). Miles and colleagues conducted a more extensive study using the *mf* stimulus already described above (Fig. 3d, see Miles and Sheliga, 2010 for a review). They reduced the stimulus to the sum of the 3rd and 5th harmonics. Overall, when the two components had similar contrast, both were effective and ocular following can be predicted by vector average of the responses to the single component motion. However, initial ocular following showed a nonlinear dependence on the relative contrast of those two sine waves (Sheliga et al., 2006b). When the two components differed in contrast by more than about an octave, ocular following was dominated by the one with the higher contrast. The second component was suppressed,

providing evidence for a winner-take-all mechanism. These results suggest that channels conveying information about the competing harmonics are mutually antagonistic so that if one harmonic has a contrast significantly greater than all others then it will tend to prevail over its competitors. What about motion signals along the same direction? Fig. 7 illustrates the interactions observed when motion stimuli were made of two sine waves of frequencies equivalent to the 3*f* and 7*f* of the *mf* stimulus. For a contrast ratio equal to or around 1, amplitude of ocular following responses was again best predicted by a vector average computation. On the contrary, for contrast ratios lower than ½ or higher than 2, the response was purely determined by highest contrast motion. The latter was interpreted as further evidence for a winner-take-all mechanism engaged when the two components have very dissimilar contrast. Overall, these experiments show that to reconstruct global motion direction, similar local motion signals are linearly combined but dissimilar ones are actively eliminated to that only one reliable solution drives tracking initiation. Contrast ratio can determine such transition from vector averaging to winner-take-all solution, but we can assume that other low level feature dimensions (orientation, spatial frequency,. . .) could be also of importance (Masson and Castet, 2002).

#### 6.1.3. Pooling motion information for speed: the competition model

Reconstructing speed offers another example for investigating integration rules. Rigid translation of a single surface will drive different channels lying along the same speed axis in the spatio-temporal space. Visual motion is extracted at different spatial scales, but we need to understand how these different bits of information are integrated to obtain the most likely speed estimate. We recently investigated this question by recording human ocular following to single or composite noise patterns. We compared responses obtained with two band-pass filtered noise pattern of different spatial frequencies presented either separately or together (Drewes et al., 2007). Notice that the two noise patterns had similar speed, direction and contrast but differed by their spatio-temporal contents. A single weighted vector average model failed to predict the responses to the composite patterns over the whole set of pairwise combinations. A winner-take-all model was also inappropriate. Overall, we found that vector average solution works when the two spatial frequencies were close one from each other by less than 1 octave. Above this, responses were stronger than predicted, this gain being larger when weak (i.e. high-spatial frequency) noise patterns were combined. To accommodate the full range of results we considered a competition model using tuned excitatory input and tuned divisive inhibition. Such feed-forward network implements a competition among the component inputs (Grossberg, 1973; Reynolds et al., 1999; Krekelberg et al., 2004) but does not necessarily lead to a single "winner" as in the winner-take-all model. Such a model is equivalent to a divisive normalization (Simoncelli and Heeger, 1996; Reynolds and Heeger, 2009) where the response R can be predicted from:

$$R = \frac{K \times E}{E + I + \sigma} \text{ with } E = \sum_i W_i^+ C_i \quad \text{and} \quad I = \sum_i W_i^- C_i$$

where $C_i$ is a spatio-temporal channel, $W_i^+$ and $W_i^-$ are 6 excitatory and inhibitory weights (given by the total number of spatial frequency tested pairwise) and $K$ is a gain factor. The competition model worked must better than a weighted vector average model. We found that excitatory and inhibitory weights were set as a function of distance between spatial frequency channels. Notice however a strong difference between these interactions and those observed when using different motion directions. When all motion components have similar speed, a supralinear interaction was found when local motions were sampled across very different spatial channels. Now here excitatory weights are given between dissimilar inputs to units tuned for the same speed, contrary to the integration rules found for motion direction (e.g. Wilson et al., 1992).

### 6.2. Local motion integration: summary

Overall, the results described above illustrate how motion information is pooled within the behavioral receptive field to drive ocular following. Automatic integration is performed when local motion signals have similar contrast or spatio-temporal contents, corresponding to the vector average solution (Masson and Castet, 2002). Notice that vector average corresponds only to a singular solution of the competition model. There is however a limited operating range for vector averaging computation. When local motion signals are too different, one must be eliminated, corresponding to automatic segmentation. This is evident when the two motions have very different contrast (contrast ratio > ½) or are given by very different directions. When local motions are of opposite directions, the competition model collapses to a winner-take-all solution since a single winner is necessary to make a decision about which direction will be used to initiate tracking. Lastly, when local motion signals are of same direction or speed, we must understand how motion information is pooled across units tuned for the same speed but that sample motion across different spatio-temporal channels (Hayashi et al., 2010). Then, the weights of these different channels can be measured using the competition model. Altogether, these results can be explained by a single competition model where weights ($W_i^+$ and $W_i^-$) are set differently along different domains: contrast, direction, speed, spatio-temporal frequencies. It shall be noticed however that when these opposite motions cannot be distinguished based on their spatio-temporal properties, additional low-level segmentation mechanism are involved such as the binocular filtering of out-of-plane motion that has been demonstrated in both humans and monkey ocular following (Busettini et al., 1996; Masson et al., 2001).

The competition model relies on tuned divisive inhibition between units, the weights of the inhibitory components being set by their relative distance and contrast. Divisive normalization can be seen as a mechanism acting to enhance sensitivity to the signal(s) of interest and lower the contribution of irrelevant signals. This model has been successfully applied to a wide range of low-level mechanisms, from direction selectivity (Heeger, 1993) to attentional modulation (Reynolds and Heeger, 2009) at both single-unit (e.g. Simoncelli and Heeger, 1996; Reynolds et al., 1999) and population levels (Sit et al., 2010). Motion integration underlying neuronal responses to plaid patterns can be also described as a cascade of tuned divisive normalizations between V1 and MT units (Rust et al., 2006). These last versions of the models involve both global and tuned components of the divisive pool acting on units tuned for spatio-temporal frequencies, direction or speed. An attractive aspect of this class of models is that it allows describing contrast dynamics and motion integration within the receptive field (see Albrecht et al., 2002 for a review) as well as from outside the classical receptive field (e.g. Schwartz and Simoncelli, 2001). It can also be expanded to understand how different motion components are integrated. Again, by varying the tuned component of the normalization model, ones can predict different solutions, from vector averaging to winner-take-all.

Since motion components can be presented either at the same location (i.e. plaids pattern, *mf* stimuli,. . .) or split among distant retinal loci, with different directions, ocular following offers an excellent probe to tease apart the contribution of global and local inhibitory interactions. In particular, interactions between motion stimuli of same or opposite directions behave differently when

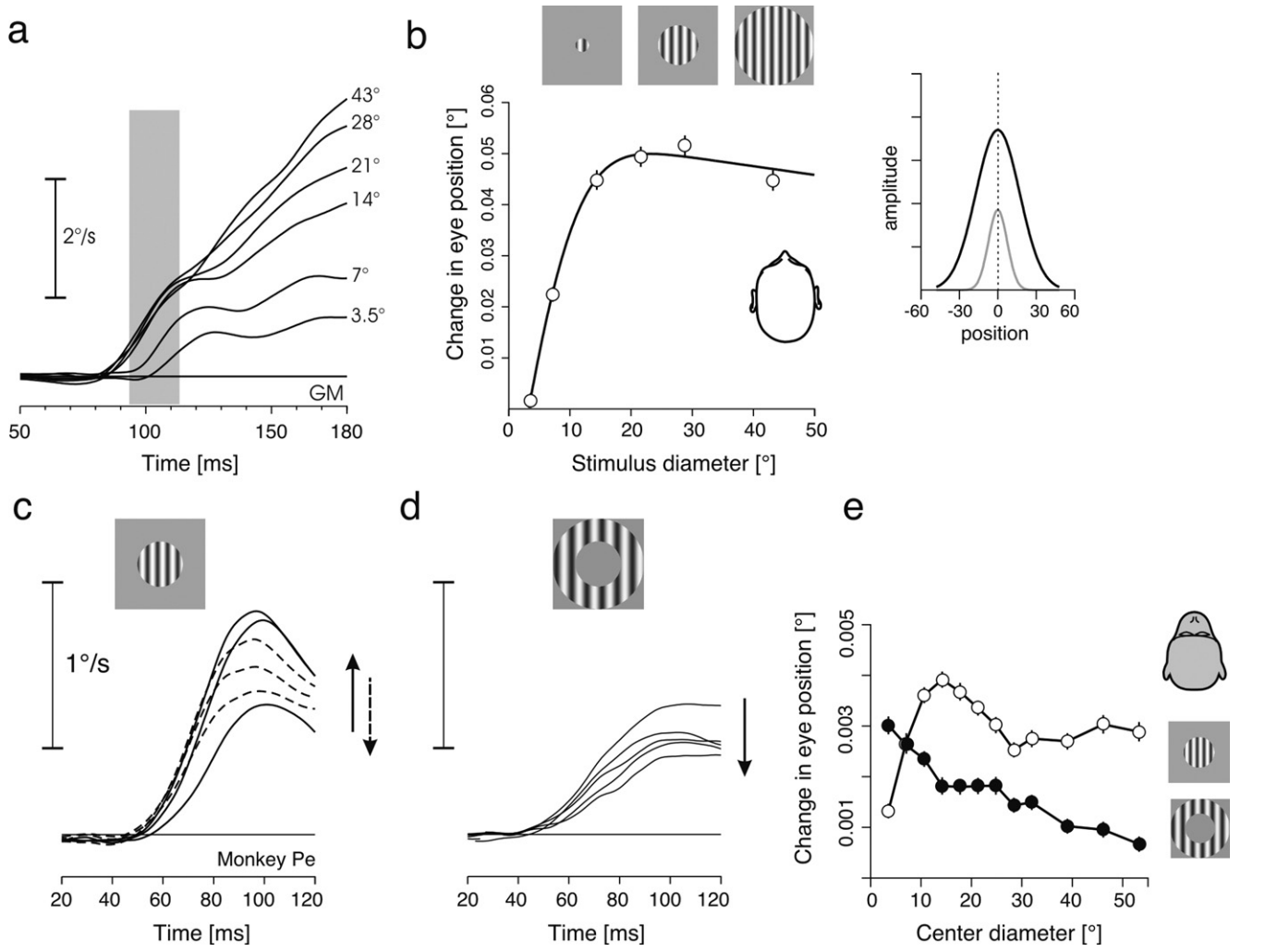


**Fig. 8.** Spatial summation for ocular following in humans (upper row) and monkeys (lower row). (a) Mean velocity profiles of ocular responses to a drifting grating presented behind an aperture of increasing diameter (from 3.5 to 43°). During the earliest part of ocular following, response amplitude rapidly grows and then saturate for diameters larger that 15°. (b) Such saturation is evident when plotting response amplitude against stimulus diameter. Continuous line is the best-fit obtained with a Difference-of-Gaussians. For low spatial frequencies, response amplitude saturates with little hyper-saturation for largest diameters. Temporal integration results in larger summation area as illustrated by the estimated Gaussian integration profiles for the earliest (grey) and latest (black) part of the ocular following responses. Modified from Barthélemy et al. (2006). (c,d) Left-hand plots illustrate mean eye velocity profiles for stimuli of expanding center diameter. Notice that in the annulus condition, this is equivalent to removing central parts of the motion stimulus and therefore probing the contribution of motion sensors of growing eccentricity. See text for explanations. (e) Relationships between change in eye position and center diameter for both disk (open symbols) and annulus (closed symbols) conditions.

different motion signals are kept separated (Sheliga et al., 2008b). Nonlinear interactions between opposite motion directions disappear when the two components are presented within narrow parallel stripes separated by more than 1°. Above this, ocular responses match the linear prediction. On the contrary, motion signals of same direction interact over much larger retinal distances. For inter-stripes gap of 8° (the largest separation tried), ocular following responses were still substantially less than predicted by the linear sum. Sheliga et al. (2008a,b) postulated the existence of two nonlinear interactions: local mutual inhibition resulting in winner-take-all effects and more global inhibition, resulting in normalization effects. Using multiple components stimuli, they found that motion inputs whose influence had been suppressed by local inhibition between opponent motions where excluded from global normalization. This suggests that local interactions occur at an earlier level than global normalization (Fig. 7d). The spatial extent of the global normalization signal is largely unknown. We can however assume that it covers a very large part of the receptive field given the facts that, (i) macaque ocular following can be driven by signals presented at eccentricities greater than 30° (see below and Fig. 8f) and (ii) that peripheral suppression of pursuit initiation in monkeys can be seen when a textured static background is set at a distance as large as 30° from the moving target (Kimmig et al., 1992).

If gain control mechanisms operate so as to selectively integrate motion signals from different parts of the visual field, it should have an impact on spatial summation, as shown for V1 neurons for instance (Cavanaugh et al., 2002a,b). Hence, we could probe the contrast dynamics of ocular following in the presence of surround motion of different directions. Several studies have been conducted to document these two aspects in both humans and monkeys: spatial summation and contextual modulation.

### 6.3. Early motion integration: local and global divisive gain control

#### 6.3.1. Nonlinear spatial summation for ocular following

Miles and colleagues originally observed that, in monkeys, amplitude of early ocular following increases as the stimulus size increases up to 40° diameter in the center of the visual field (Miles et al., 1986). However, they also found evidence that initial eye

velocity does not monotonically increase with stimulus size but show signs of saturation. It was only recently that the link was established between nonlinear spatial summation and global divisive inhibition effects. Breaking a moving vertical drifting grating into small (1° width) parallel stripes, Sheliga et al. (2008b) found that increasing the number of stripes (and this the total size of the stimulus) resulted at first in an increase in response amplitude, then a plateau and a decrease in initial eye velocity when the image filled the screen. They interpreted this as evidence for inhibitory surround input acting as a (divisive) response normalization. However, two effects were not explained. First, response latency decreases with total stimulus size. Second, changes in eye velocity were seen only in the later part of the response, that is ~40 ms after response onset. This is reminiscent to the early observation of Miles et al. (1986) that antagonistic surround motion does influence ocular following but only in the late part of the response. One potential problem with the studies of Miles and colleagues was that each competing motion elicited an oculomotor response in itself, opening the door for motor rather than visual interactions.

We documented a more direct set of evidence for surround modulation in both humans and monkeys. First, we closely re-examined spatial summation in both humans and monkeys to measure the integrative zone of the behavioral receptive field (Barthélemy et al., 2006). In humans, increasing the size of the stimulus results in a sharp increase of response amplitude, for diameters up to ~15–20°. Larger stimulus sizes did not further boost the response amplitude (Fig. 8a). Such saturation with grating diameter indicates that summation of motion information occurs within a limited integrative zone, which we defined as the central, driving part of the behavioral receptive field (Barthélemy et al., 2006). Moreover, at least for medium spatial frequency (0.4 cpd), we found little evidence for hyper-saturation indicating a weak inhibition due to peripheral inputs of same motion direction (Fig. 8b). In monkeys, we further explored the properties of the integrative zone in two different ways. First, we titrated the contribution of peripheral neurons using annuli of increasing center diameter (Fig. 8c,e) and compared this with the effect of disks of expanding size (Fig. 8d,e). Second, we varied the spatial frequency content of the stimulus to see if the spatial summation function varies with spatial frequency for ocular following as it does for single neurons (Daugman, 1985; DeAngelis et al., 1994). As shown by the two arrows in the upper-left plot, increasing stimulus size first resulted in increasing responses (continuous lines) although further expanding the grating diameter resulted in a reduction of the ocular responses (broken curves). Such nonlinear spatial summation function is best illustrated by plotting response amplitude against stimulus diameter (open symbols, Fig. 8e). Such reduction with sizes larger than optimal is often seen as a signature of surround inhibition. Notice that hyper-saturation was larger than observed in humans for similar spatial (0.4 cpd) and temporal (10 Hz) frequencies. Notice also that decrease in response amplitude was always seen as early as response initiation, but that in humans smaller stimuli resulted in longer latencies.

Such a nonlinear spatial summation with response reduction for larger than optimal sizes is classically seen as a signature of a response inhibition from surrounding motion (Allman et al., 1985). The fact that surround suppression asymptote to some less than maximal response amplitude can be explained by the distribution and size of motion detectors across the visual field. Using an annular stimulus, we documented the weight of peripheral motion detectors (Fig. 8e, closed symbols). Increasing the inner diameter of the ring exponentially reduces the response amplitude, demonstrating that the contribution of local motion detectors decreases with eccentricity. Neurons selective for higher spatial frequency have smaller receptive fields (i.e. smaller optimal size of stimulus) (Daugman, 1985) and are less represented in the periphery of the visual field since receptive field size increases linearly with eccentricity in both macaque areas V1 and MT (Dow et al., 1981; Albright and Desimone, 1987). Consistently, with higher spatial frequency gratings, we found a sharper decrease, zeroing with relatively small inner diameter. At these high spatial frequencies, no hypersaturation was observed in the spatial summation function, demonstrating that this latter is produced by peripheral motion detectors tuned for the same direction of motion as the center. This is consistent with the global normalization model postulated by Sheliga et al. (2008b).

Spatial summation is slightly, but significantly affected by contrast. In humans, we found a larger optimal size at very low contrast (inset Fig. 8b). Such contrast modulation of spatial summation has been reported at both psychophysical (Tadin et al., 2003; Tadin and Lappin, 2005) and neuronal (Cavanaugh et al., 2002a; Sceniak et al., 1999; Pack et al., 2005) levels. Moreover, surround suppression is often more pronounced at high contrast (e.g. Cavanaugh et al., 2002a; Pack et al., 2005). These results strongly suggest that surround inhibition acts as a gain control. Models postulating a divisive inhibition of the center of behavioral receptive field predict that contrast response function shall be strongly modulated by surround motion (e.g. Cavanaugh et al., 2002a; Schwartz and Simoncelli, 2001; Xing and Heeger, 2000, 2001). Such contextual modulation shall then be dependent on the relative contrast and direction of center and surround motions. We investigated the dynamics of this gain control in both humans and monkeys ocular following.

#### 6.3.2. Contextual modulation of center motion: contrast gain control

To investigate gain control mechanisms, Sheliga et al. (2006a,b, 2008b) used different grating motion moving simultaneously in opposite direction. Since the peripheral motion can drive ocular responses on its own, it remains uncertain whether a reduction in response amplitude reflects inhibitory interactions at visual or motor levels. We have investigated these inhibitory interactions using a different approach. In both human and non-human primates, we documented contrast response function of responses driven by central grating motion (20° diameter) presented either alone or surrounded by a dynamical surround. Here, the surrounding motion was a counterphase grating, i.e. the sum of two gratings with similar spatiotemporal frequencies and contrast but moving in opposite directions. One advantage of such contextual stimuli was that it did not drive any significant eye movements when presented alone, discarding the possibility that changes observed in the amplitude of center-driven grating were due to some interactions between competing motor responses (Barthélemy et al., 2006; Barthélemy and Masson, 2006).

Fig. 8a illustrates results obtained in human subjects. As can be seen from comparing eye velocity profiles, a flickering surround has a strong suppressive effect on ocular following responses to the center moving grating. However such suppressive effect was mostly seen on the later part of the responses. This is even more evident when comparing contrast response functions obtained with or without a flickering surround. In the earliest time window (light grey bar), locked with response onset, no differences were seen between the two sets of curves. It looks like as if the surround was ineffective in this early period. On the contrary, a strong difference can be observed in the later time window (dark grey bar). Adding a surround both lowered the slope of the Naka–Rushton function and increased half-saturation contrast from 10 to 30%. In other words, the surround had a strong divisive effect on the center-driven responses. Notice that the grey symbols, nearly aligned with the horizontal line at 0° amplitude response, indicate that when presented alone the flickering surround did not elicited any significant eye movements. A closer look at the dynamics of contrast gain control indicates that the effect of the surround was mostly to

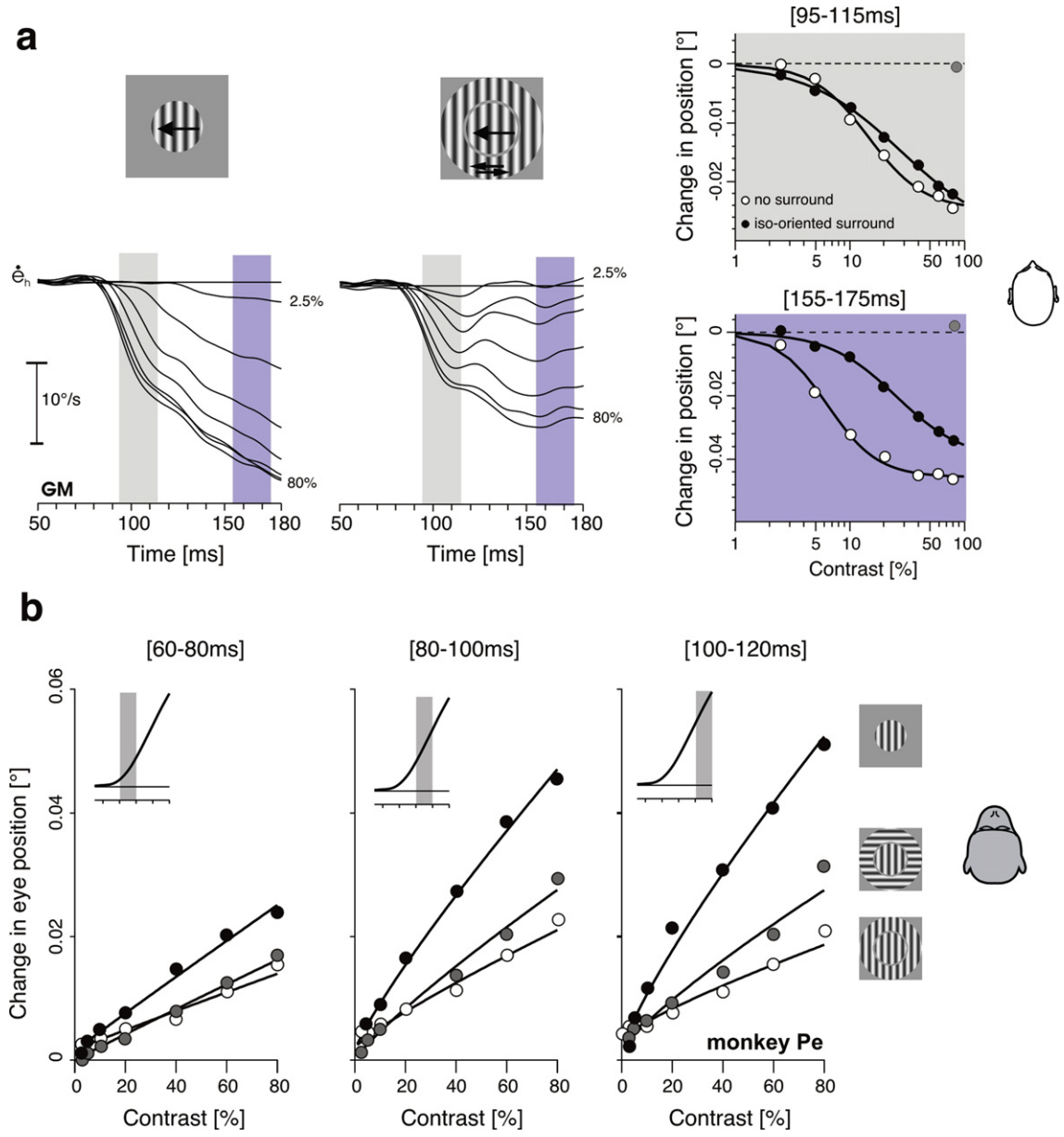


**Fig. 9.** Contrast gain control and contextual information in human (a) and macaque (b) ocular following. (a) Left-hand plots are eye velocity profiles of responses to a single grating, presented at seven different contrasts (from 2.5 to 80%). Light and dark grey bars indicate early and late time windows for quantitative analysis. Middle plots are eye velocity profiles of responses to same grating motion stimulus (contrast range: 2.5–80%) but now presented with a full contrast flickering grating in the surround. Right-hand plots illustrate the contrast response functions for ocular following to the center grating alone (open symbols) or with an iso-oriented surround (closed symbols), for the two different time windows. Reprinted from Reprinted from Barthélemy et al. (2006), J. Neurophysiol. with permission from Am. Physiol. Soc. (b) Contrast responses functions obtained with or without a flickering surround, measured at 3 different time windows after stimulus onset. (Barthelemy et al., in preparation).

clamp the contrast gain of ocular following to the level originally set at response onset, prevailing the leftward shift of the curve due to temporal integration.

How to explain that surround inhibition was delayed relative to tracking onset? One simple explanation could be that, due to the rather large diameter of our central stimulus (20°), peripheral signals were delayed. Two experiments were conducted to reject such explanation. First, we reduced the size of the center–surround stimuli to ∼6°. Second, we broke the motion stimulus into a set of nine small patches, all having a surround set to either 80 or 0% (i.e. no surround). Overall, the temporal dynamics of surround divisive inhibition was strictly similar to that reported for large motion stimuli. This result suggests that interactions were mostly local between motion detectors tuned for opposite motion direction. This results in consistent with the experiments done by the group of Miles using small bands and single gratings of opposite motion directions. Moreover, at least in humans, divisive surround effect was tuned for relative orientation/direction between center and surround motions: inhibition was strongest for iso-oriented surrounds and weakest for cross-oriented ones.

Similar results were found in macaque monkeys (Barthélemy and Masson, 2006). As shown in Fig. 8b, ocular following driven by center motion was largely suppressed by adding a flickering surround. The difference between iso- and cross-oriented surrounds was found to be smaller than observed in humans. However, the temporal dynamics of surround suppression is very similar to that observed in humans. The earliest phase (55-75 ms after motion onset) of ocular following was unaffected by a dynamical

surround. Surround motion began to have a significant, suppressive effect only ~15 ms after response onset. From then, surround inhibition grew over time. The facts that surround influence was delayed are consistent with the timing of surround inhibition found at single unit level in area V1 of macaque monkeys (Bair et al., 2003, see also Knierim and Van Essen, 1992; Lamme, 1995). Recently, we investigated this dynamics with a scaled-version of the center–surround stimuli used for ocular following (Reynaud et al., 2007). Recording population activity with voltage-sensitive dyes optical imaging, they found that surround activity suppressed the neuronal responses driven by center stimuli in area V1 of alert macaque monkeys. Such surround suppression was delayed by ~10 ms and this delay was dependent on the distance between the center and surround parts of the stimulus. Since more peripheral surround stimulus elicits neuronal responses in a more distant part of the cortex relative to center retinotopic representation, we can assume that timing of surround suppression is mostly explained by the dynamics of lateral interactions within area V1 and not by feedback modulations from higher areas such as area MT.

#### 6.3.3. *A divisive normalization for ocular following: gain control and automatic selection*

Altogether, the results presented above strongly support the idea that sensory processing for ocular following involves several nonlinearities. Changes in response amplitudes when varying stimulus size, contrast and surround suggest that these nonlinearities result from at least two divisive normalization mechanisms. A first divisive inhibition occurs between units tuned for opponent motions and operates at small spatial scales. This mechanism is tuned for motion direction, co-linear gratings producing the greatest suppression and orthogonal ones the weakest (Barthélemy et al., 2006; Barthélemy and Masson, 2006). Rust et al. (2006) have suggested that such tuned divisive normalization occurs at the level of MT neurons. If this is true, local inhibition shall be implemented in locus II rather in locus I in the general schema proposed by Sheliga et al. (2008b) and illustrated in Fig. 7f.

We, and others, found that a significant fraction of surround suppression is not tuned for direction. Such global suppression also scales with the contrast of the surround motion but emerges slightly before the tuned suppression (Barthélemy et al., 2006; Barthélemy and Masson, 2006). This global suppression would correspond to the global divisive normalization postulated by Barthélemy et al. (2006) and then Sheliga et al. (2008a,b) on different experimental grounds. Such global divisive normalization might be implemented at an early level such as area V1 as suggested by Simoncelli and Heeger (1996) and supported by the timing of surround suppression for ocular following as well as the temporal dynamics of population gain control in primary visual cortex (Reynaud et al., 2007; Sit et al., 2010). This latter aspect is critical to tease apart the contribution of these different stages. We have shown in both human and monkeys that we can track the different inhibitory mechanism at the ms time scale by looking at the time course of ocular following. Most current models using divisive normalization for motion integration are static models. Because center–surround mechanisms present different timings along the cortical motion pathway, future work will attempt to dissect the temporal architecture of this cascade.

Divisive normalization offers an integrative framework to elucidate two nonlinear properties of the behavioral receptive field: gain control and automatic selection. We have shown that the competition model implements a tuned divisive inhibition. A general property of tuned divisive normalization is that it amplifies certain motion components while damping the contribution of others. Such a mechanism could act to single out the global motion signal in different domains (spatial location, spatio-temporal properties, direction and speed). It correspond to the early suggestion of Miles et al. (1986) that ocular following is not driven by *en masse* motion but rather relies on complex front-end processing that quickly and automatically segment the object motion of interest.

## 7. Comparison with voluntary smooth pursuit

Above, we have summarized experimental data collected with ocular following eye movements, which are reflexive tracking eye movements. We must say a few words about the similarities and differences between reflexive ocular following and voluntary smooth pursuit. Steve Lisberger and his colleagues at USCF have performed a long series of experiments to investigate how local motion information is encoded by population of MT neurons to drive the initial eye acceleration of voluntary tracking (see Lisberger, 2010 for a review of their work). Overall, they found that target speed is encoded by a population of MT neurons performing local motion integration (Priebe et al., 2001), with extended spatio-temporal bandwidths that make them speed tuned (Priebe et al., 2003; Priebe and Lisberger, 2004). This is consistent with earlier report that target velocity, and in a much lesser extent target motion acceleration is represented in area MT when anesthetized monkeys are presented with small targets similar to those used for driven smooth pursuit (Lisberger and Movshon, 1999). More recently, they investigated how motion direction information is build-up in such MT neurons population. Using information theory, they proposed that behavioral performance is essentially constrained by noise level within the sensory part of the visuomotor transformation (Osborne et al., 2005) and that direction information is quickly accumulated in MT neurons so that only a few spikes are needed to correctly trigger an ocular responses with latency ~100 ms (Osborne et al., 2004, 2007). Another earlier important observation from the same group was that, when presented with multiple target motions, initial eye acceleration of voluntary pursuit in macaques is driven by the vector average solution (Lisberger and Ferrera, 1997). Similarly, direction tuning of MT neurons was found to encode the same vector average computation (Recanzone et al., 1997) at least until target selection has occurred (Recanzone and Wurtz, 1999, 2000; Wilmer and Nakayama, 2007).

Altogether, this summary points out that similar computational rules apply to both voluntary and reflexive tracking eye movements. Moreover, the two series of studies have highlighted different aspects of automatic motion integration. In particular, the work of Lisberger and colleagues favors the hypothesis that pursuit initiation is based upon a linear read-out of a population of speed tuned neurons located in area MT, implemented as a weighted vector average computation (see Lisberger, 2010). However, ocular following (Masson and Castet, 2002; Sheliga et al., 2006a,b; Drewes et al., 2007) as well as pursuit responses (Priebe and Lisberger, 2004) to compound stimuli pin point the fact that dissimilar inputs can be nonlinearly integrated, depending for instance upon their relative contrast or spatial frequency. Future work should further explore this aspect since information about speed and direction of a natural objects are available at multiple spatial scales and therefore through a set of different spatio-temporal channels (Drewes et al., 2007; Simoncini et al., 2010). Since ocular following more heavily relies on spatial integration of motion signals to active template detectors such as found in area MST, it is more probable that nonlinear pooling mechanisms are involved in triggering ocular following than smooth pursuit of a single, point-like motion stimulus. Lastly, we are still lacking trial-by-trial correlation between single unit responses in areas MT/MST and ocular following, similar to what has been done for voluntary smooth pursuit eye movements (see Lisberger, 2010 for a review).

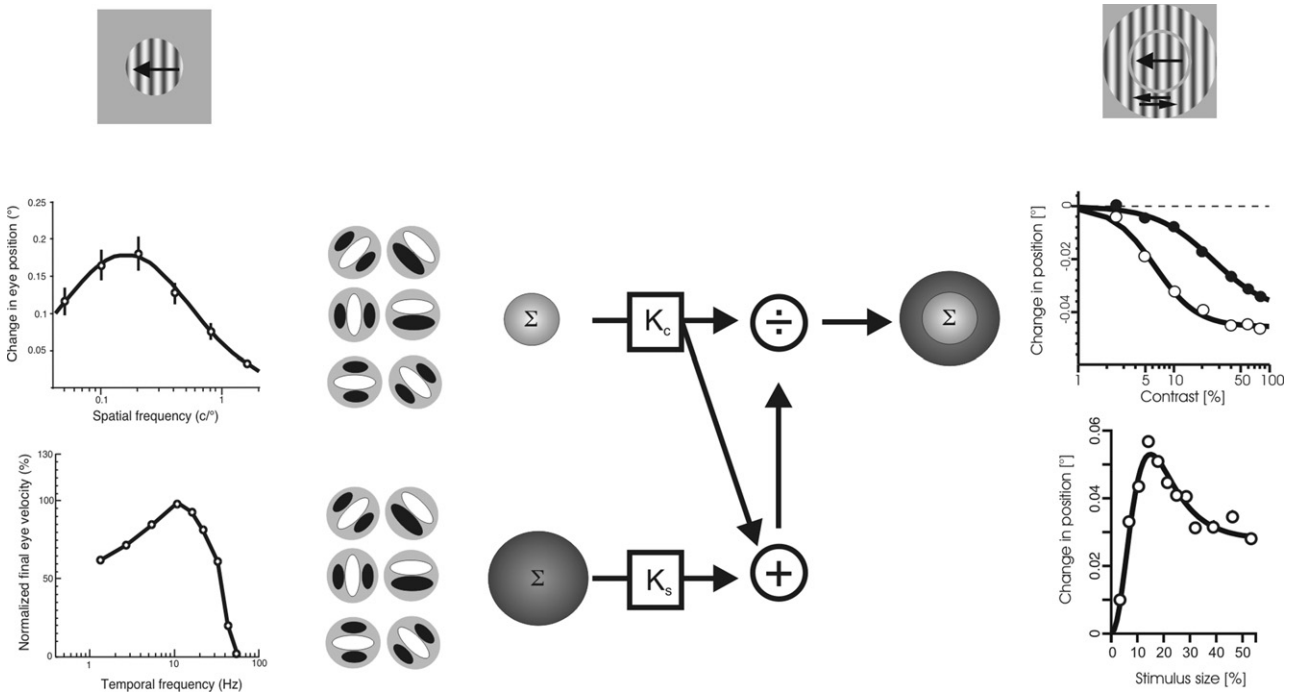


**Fig. 10.** A functional model for linear and nonlinear visual properties of ocular following. (a) Motion signals are extracted through spatio-temporal filters whose relative contribution defines the spatial and temporal tunings of the behavioral receptive field, that is the optimal window of visibility for these behavioral responses. (b) Integrative and suppressive receptive fields perform the linear integration of motion signals within the driving part of the visual field, pooling signals from different (mainly opposite) directions. The suppressive receptive field acts upon this gain control by normalizing the output of the integrative stage by a signal proportional to the global and local variances (with different direction tuning) of the filtered stimulus. (c) Such ROG model simulates several nonlinear properties of ocular following such as contrast response functions and spatial summation.

## 8. A functional description of the behavioral receptive field

All these results presented above can be captured in a descriptive model of a behavioral receptive field. Fig. 10 summarizes the essential computational properties of this behavioral receptive field. Table 1 provides a mathematical description for each step, together with mean best estimates of the parameters, for both humans and monkeys. Overall, three essential mechanisms are involved: spatio-temporal filtering with a bank of linear filters, context-dependent integration implemented as center–surround mechanisms and gain control through a cascade of tuned and untuned divisive normalizations. We will describe these three steps.

Local changes in luminance are sampled by a set of linear filters extracting the spatial and temporal profiles in a small image patch. These filters have a temporal impulse response (Sheliga et al., 2005) that can be fitted by Eq. (1), similar to both cells found in primary visual cortex and psychophysical performance (De Valois and De Valois, 1990). These spatio-temporal filters are pooled together to form the spatial and temporal envelopes of the behavioral receptive fields (Miles et al., 1986; Gellman et al., 1990). These envelopes are best fitted with log-skewed Gaussian functions (Eqs. (2) and (3)), again similar to that used previously for single neurons or population measurements (e.g. Albrecht, 1995; Henriksson et al., 2008). Interestingly, these envelopes are both band-pass tuned, indicating that ocular following samples motion over a limited window of visibility. This suggests that the population of MT/MST neurons that reconstructs object speed pools information only over a restricted subset of V1 neurons. Consistently, the tuning functions found for ocular following correspond to those observed with V1 neurons projecting to MT area (Movshon and Newsome, 1996). We obtained very similar best fitting parameters, in particular optimal temporal frequency (~5 Hz) at low spatial frequencies. These spatio-temporal properties implement a first automatic segmentation, filtering out motion from a given range of spatial and temporal scales corresponding to that of large objects whose images undergo retinal translation during head movements (Miles et al., 1986) (Fig. 10).

The behavioral receptive field captures the behavioral consequences of two nonlinear mechanisms documented at the early visual motion processing stage: a non-linear relationship between response amplitude and stimulus size and a sigmoid-like contrast response function that can be shifted along contrast axis, depending on context. We have shown above that such nonlinear relationship between response amplitude and contrast can modeled by Eq. (5) also called the Naka–Rushton function. Interestingly, the best-fit parameters found for both human and monkeys are consistent with those reported for populations of MT and MST neurons in macaque (see Fig. 6a). The correlation is even stronger between MST population and ocular responses in monkeys. This suggests strongly that dynamical properties of ocular following are indeed set at this level.

MT and MST neurons show strong center–surround interactions. These interactions have been largely investigated in primary area V1 but are less documented in these areas. In both humans and monkeys we found that surround motion affects the contrast response function in a consistent way: the Naka–Rushton function becomes less steep and half-saturation contrast increases. In monkeys, we found that on average the slope shifted from ~2 to ~1.5 and $C_{50}$ increased from 10 to 20%. More work is needed to compare these values to those observed at population level for both MT and MST areas. However, the fact that breaking the center–surround stimulus into smaller patches does not change this effect strongly suggests that context-dependent gain control mechanisms arise principally in area MT where neurons have smaller receptive fields.

Another evidence for a key role of area MT in shaping the behavioral receptive field is given by the OFR spatial summation function. Relationships between response amplitude and stimulus size exhibit a highly nonlinear shape that is best fitted with Eq. (6). At neuronal levels, center–surround interactions

underlying non-linear spatial summation and tuned/un-tuned gain control can be modeled as divisive interactions between pools of neurons covering overlapping parts of the visual field. Cavanaugh et al. (2002a,b) have proposed that a Ratio-of-Gaussian model can simulate most of these properties for V1 neurons. This RoG model implements a general divisive normalization, in particular to model the effects of grating surround on the neuron's contrast response (Chen et al., 2001). They found that $k_c$ and $k_s$ are non-linearly dependent upon contrast ($k \propto C^n$) and therefore that contrast would change the gain ratio between the two populations such that suppression was proportional to $k_c/(1+k_s)$. Such gain model is sufficient to explain changes in the summation function with lower contrast (Sceniak et al., 1999) without the need of postulating changes in spatial extent of excitatory and inhibitory fields. We applied the same logic to ocular following and found that the optimal spatial summation function can be fitted with Eq. (6). Of particular interest, motion signals are optimally integrated over the central 10–15° of the visual field. Moreover, spatial integration is modulated by a suppressive field of much larger size (Table 1). Such suppressive field could implement context-dependent gain settings as reported above (see Fig. 9). In humans, surround suppression is modulated by the contrast ratio between center and surround motion (Barthélemy et al., 2006). However, further work is needed to investigate whether the RoG model can also describe these center–surround modulations. Still, the current results support the view that spatial properties of motion integration can be seen as the product of excitatory (driving) and suppressive (modulating) mechanisms implementing both non-specific and direction-selective normalization (Barthélemy et al., 2006; Sheliga et al., 2008a,b; Miles and Sheliga, 2010).

We have seen that most of these nonlinear interactions have sluggish temporal dynamics: contrast gain rapidly changes over time and surround inhibition is delayed relative to center-driven responses. This could respectively correspond to two properties of V1 neurons: local temporal integration (Albrecht et al., 2002) and delayed input from the surround (Bair et al., 2003; Reynaud et al., 2007). Moreover, the fact that surround inhibition builds up over time indicate that the strength of the inhibitory pool grows over time, something observed at population level in macaque area V1 (Reynaud et al., 2007). In the model, this could be described by two different temporal integration mechanisms, one being local (pooling of similar motion signals) and the other being global (normalization signals). These two time scales could correspond to the different temporal dynamics of local (i.e. short-range) and global (i.e. long-range) cortico-cortical interactions.

Table 1 shows that input–output relationships between amplitude or latency of ocular following and the properties of visual motion stimuli can be collapsed into a set of descriptive functions that form the core of the spatial interactions between driving and suppressive populations of neurons tuned for motion directions. We have proposed that such a set of functions defines a *behavioral receptive field*. Moreover, it becomes possible to clarify the spatial scales at which each linear or non-linear processing occurs. However, such behavioral receptive field must be seen as a read-out of neuronal populations activity. This is further illustrated by the fact that MST neurons have themselves receptive fields with many properties that are homogeneous with what is observed at behavioral level: large, but finite integration zone, broad speed tuning, disparity-tuning peaking around 1–2° and so on and so forth. Thus a linear pooling of MST neurons can yield to a set of tuning curves, in both spatial and temporal domains, which in principle correspond to the properties of the behavioral receptive field. This approach is similar to population analysis methods that reconstruct a population receptive field from single measurements of individual neurons (see Erlhagen et al., 1999; Jancke et al., 1999).

In Section 4, we introduced the idea of the behavioral receptive field as a much more constrained framework than the original concept of perceptive fields. Table 1 illustrates the advantages of ocular following to map the functional properties of motion detection and integration stages. In its original description, the perceptive field was mostly seen as a description of human perception, trying to link phenomenology of both psychophysical performances and neuronal properties. We provide a more quantitative framework where best fitting parameters for a given functional step can be compared to the neuronal population characteristics (see Fig. 6 for instance). Moreover, ones can measure the temporal dynamics of ocular responses and compare it to time courses of neuronal activity. Such approach can be extended to 2D motion integration for which we demonstrated that time course of eye movements closely follow that of the neuronal solution of the aperture problem in area MT (see Fig. 4 in Barthélemy et al., 2010; see also Fig. 8.2, Masson et al., 2010). In the future, we will relate some of the nonlinear properties of the behavioral receptive field to the characteristics of early and late phases of ocular following when presented with plaids or barber-poles. We will then be able to test some predictions made by the behavioral receptive field model such as the similarity between in the temporal dynamics of center–surround interactions and 2D motion integration (Masson et al., 2000; Barthélemy et al., 2010), similar to some recent models of MT neurons responses to 2D motion (Tsui et al., 2010). One last advantage of the behavioral receptive field upon the perceptive field is its limited numbers of steps that make it suitable for realistic modeling using dynamical models. Future work will then target neural modeling of the bRF, focusing in particular on the spatio-temporal network properties (Perrinet and Masson, 2007).

## 9. Conclusions

In the present review, we have summarized a large bulk of results obtained with ocular following responses in human and non-human primates. We have shown that these data are a mine of information at different levels: (i) to identify what are the basic mechanisms involved in motion detection, and their timing, (ii) identify the rules of automatic motion integration and segmentation to single-out the part of the image corresponding to the object of interest, (iii) map linear and nonlinear characteristics of visual motion processing with different stages along the cortical motion pathways. We have shown how these different computational steps can be condensed into the idea of a behavioral receptive field. Such bRF describes how most relevant information is extracted to drive the eyes. Our hope is that we have convinced readers that small, reflexive eye movements offer us a delicate tool to probe cortical dynamics in both humans and monkeys.

Understanding how these non-linearities are related to decision mechanisms operating on cascaded and distributed representations of motion information is one of the future challenges. To meet this objective, we need population recordings at different spatial and temporal scales as well as correlation studies between neuronal and behavioral responses on a trial-by-trial basis. Such approaches have been pioneered for smooth pursuit (Lisberger, 2010), but ocular following offers a better window onto the temporal dynamics of context-dependent pooling mechanisms. Just as spatial orientation, ocular following could then be another example of how to link the transfer function approaches to dynamical Bayesian inference (MacNeilage et al., 2008; Perrinet and Masson, 2007).

Another challenge is to link the temporal dynamics of early motion detection and integration to that of two-dimensional (2D) motion integration. Several studies have demonstrated that linear averaging of Fourier motion is not sufficient to reconstruct the global 2D trajectory of a moving surface (e.g. Adelson and Movshon,

1982; Wilson et al., 1992). Further computational steps are necessary (e.g. Movshon et al., 1985; Wilson et al., 1992; Rust et al., 2006; Tsui et al., 2010) as well as the diffusion of information into layered, retinotopically organized maps (e.g. Grossberg et al., 2001; Tlapale et al., 2010). The last decade has unveiled that MT motion integration is a highly dynamical computation (Pack et al., 2001, 2004; Pack and Born, 2001; Smith et al., 2005). Again, the temporal dynamics of both reflexive (Masson et al., 2000; Masson and Castet, 2002) and voluntary (Masson and Stone, 2002; Wallace et al., 2005; Born et al., 2006) tracking reflects these different computational steps. Moreover, time course of ocular following strictly mimics (on a ms time scale!) that of MT neuronal populations when tested with ambiguous 2D motion patterns such as plaids or barberpoles (Barthélemy et al., 2010). Interestingly, this dynamics is also consistent with that reported above for center–surround interactions at population levels. We need to elaborate a common theoretical framework to relate these two mechanisms and therefore unveil how V1 and MT neural populations interact to compute *global* speed and direction of moving objects.

Investigating visual motion processing in the context of tracking eye movements has long been trapped between a search for objective estimates of perceptual processes and a source of noise and caveats that you should eliminate to decipher the underlying motor apparatus. After two decades of intensive works by only a few groups, tracking eye movements appear as an exquisite tool to understand in the finest details how the brain senses our world to rapidly and accurately guide our actions.

## Acknowledgments

This work was supported by the CNRS, the Agence Nationale de la Recherche (ANR-NASTATS-2005) and the European Union (FACETS, VIth Framework FET-IST Program, 2005-15879). We acknowledge the help of D. Laugier, M. Mekaouche, M. Martin, A. DeMoya and J. Baurberg for building the experimental setup and for running experiments on macaque monkeys. We thank Drs. Eric Castet, Laurent Goffart, Fréderic Chavane and Anna Montagnini for helpful comments and discussions. We greatly thank Dr. Frederick A. Miles for introducing us to the ocular following approach.

## References

Adelson, E.H., 1982. Some new motion illusions, and some old ones, analyzed in terms of their Fourier components. Invest. Ophth. Vis. Sci. 34, 144.
Adelson, E.H., Movshon, J.A., 1982. Phenomenal coherence of moving visual patterns. Nature 300, 523–525.
Adelson, E.H., Bergen, J.R., 1985. Spatiotemporal energy models for the perception of motion. J. Opt. Soc. Am. A 2, 284–299.
Ahumada, A.J., 1996. Perceptual classification images from Vernier acuity masked by noise. Perception 16, 18.
Ahumada, A.J., 2002. Classification image weights and internal noise level estimation. J. Vis. 2, 121–131.
Albrecht, D.G., 1995. Visual cortex neurons in monkey and cat: effect of contrast on the spatial and temporal phase transfer functions. Vis. Neurosci. 12, 1191–1210.
Albrecht, D.G., Hamilton, D.B., 1982. Striate cortex of monkey and cat: contrast response function. J. Neurophysiol. 48, 217–237.
Albrecht, D.G., Geisler, W.S., Frazor, R.A., Crane, A.M., 2002. Visual cortex neurons of monkeys and cats: temporal dynamics of the contrast response function. J. Neurophysiol. 88, 888–913.
Albrecht, D.G., Geisler, W.S., Crane, A.M., 2003. Nonlinear properties of visual cortex neurons: temporal dynamics, stimulus selectivity, neural performance. In: Chalupa, L.M., Werner, J.S. (Eds.), The Visual Neurosciences. MIT Press, Cambridge, pp. 747–764.
Albright, T.D., Desimone, R., 1987. Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. Exp. Brain Res. 65, 582–592.
Albright, T.D., Stoner, G.R., 1995. Visual motion perception. Proc. Nat Acad. Sci. U. S. A. 92, 2433–2440.
Allman, J., Meizin, F., McGuiness, E., 1985. Direction- and velocity specific responses from beyond the classical receptive field in the middle temporal visual area (MT). Perception 14, 105–126.
Anstis, S.M., 1975. Phi movement as a subtraction process. Vision Res. 10, 1411–1430.
Bair, W., Cavanaugh, J.R., Movshon, J.A., 2003. Time course and time–distance relationships for surround suppression in macaque V1 neurons. J. Neurosci. 23, 7690–7701.
Balboa, R.M., Grzywacz, N.M., 2000. Occlusions and their relationships with the distribution of contrast in natural images. Vision Res. 40, 2661–2689.
Baro, J.A., Levinson, E., 1988. Apparent motion can be perceived between patterns with dissimilar frequencies. Vision Res. 28, 1311–1313.
Barthélemy, F.V., Vanzetta, I., Masson, G.S., 2006. A behavioral receptive field for ocular following: dynamics of spatial summation and center–surround interactions. J. Neurophysiol. 95, 3712–3726.
Barthélemy, F.V., Masson, G.S., 2006. Spatial integration of motion for human and monkey ocular following. Soc. Neurosci Abstr., 735.5/L3.
Barthélemy, F.V., Perrinet, L.U., Castet, E., Masson, G.S., 2008. Dynamics of distributed 1D and 2D motion representation for short-latency ocular following. Vision Res. 48, 501–522.
Barthélemy, F.V., Fleuriet, J., Masson, G.S., 2010. Temporal dynamics of 2D motion integration for ocular following in macaque monkeys. J. Neurophysiol. 103, 1275–1282.
Benson, P.J., Guo, K., 1999. Stages in motion processing revealed by the ocular following response. NeuroReport 10, 3803–3807.
Berman, R.A., Wurtz, R.H., 2011. Signals conveyed in the pulvinar pathway from superior colliculus to cortical area MT. J. Neurosci. 31, 373–384.
Born, R.T., Bradley, D.C., 2005. Structure and function of visual area MT. Annu. Rev. Neurosci. 28, 157–189.
Born, R.T., Pack, C.C., Ponce, C.R., Li, S., 2006. Temporal evolution of 2-dimensional direction signals used to guide eye movements. J. Neurophysiol. 95, 284–300.
Braddick, O., 1974. A short-range process in apparent motion. Vision Res. 14, 519–527.
Braddick, O., 1993. Segmentation versus integration in visual motion processing. Trends Neurosci. 16, 263–268.
Brown, R.O., He, S., 2000. Visual motion of missing-fundamental patterns: motion-energy versus feature correspondence. Vision Res. 40, 2135–2147.
Busettini, C., Fitzgibbon, E.J., Miles, F.A., 2001. Short-latency vergence in humans. J. Neurophysiol. 85, 1129–1152.
Busettini, C., Masson, G.S., Miles, F.A., 1997. Radial optic flow induces vergence eye movements with ultra-short latencies. Nature 390, 512–515.
Busettini, C., Masson, G.S., Miles, F.A., 1996. A role for stereoscopic depth cues in the rapid visual stabilization of the eyes. Nature 380, 342–345.
Carl, J.R., Gellman, R.S., 1987. Human smooth pursuit: stimulus-dependent responses. J. Neurophysiol. 57, 1446–1463.
Cavanagh, P., Mather, G., 1989. Motion: the long and short of it. Spat. Vis. 4 (2–3), 103–129.
Cavanaugh, J.R., Bair, W., Movshon, J.A., 2002a. Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. J. Neurophysiol. 88, 2547–2556.
Cavanaugh, J.R., Bair, W., Movshon, J.A., 2002b. Nature and interactions of signals from the receptive field center and surround in macaque V1 neurons. J. Neurophysiol. 88, 2530–2546.
Chen, K.J., Sheliga, B.M., Fitzgibbon, E.J., Miles, F.A., 2005. Initial ocular following in humans depends critically on the fourier components of the motion stimulus. Ann. N. Y. Acad. Sci. 1039, 260–271.
Chen, Y., Geisler, W.S., Seidemann, E., 2006. Optimal decoding of correlated neural population responses in the primate visual cortex. Nat. Neurosci. 9 (11), 1412–1420.
Chen, C.C., Kasamatsu, P., Polat, U., Norcia, A.M., 2001. Contrast response characteristics of long-range lateral interactions in cat striate cortex. NeuroReport 12, 655–661.
Cheng, K., Hasagawa, T., Saleem, K.S., Tanaka, K., 1994. Comparison of neuronal selectivity for stimulus speed, length, and contrast in prestriate visual cortical area V4 and MT of the macaque monkey. J. Neurophysiol. 71, 2269–2280.
Chubb, C., Sperling, G., 1989. Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. Proc. Natl. Acad. Sci. U.S.A. 86, 2985–2989.
Cohen, B., Matsuo, V., Raphan, T., 1977. Quantitative analysis of the velocity characteristics of optokinetic nystagmus and optokinetic after-nystagmus. J. Physiol. (Lond.) 270, 321–344.
Crowder, N.A., Price, N.S.C., Mustari, M.J., Ibbotson, M.R., 2009. Direction and contrast tuning of macaque MSTd neurons during saccades. J. Neurophysiol. 101, 3100–3107.
Cumming, B.G., Parker, A.J., 1997. Responses of primary visual cortical neurons to binocular disparity without depth perception. Nature 389, 280–283.
Daugman, J.G., 1985. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. A 2, 1160–1169.
DeAngelis, G.C., Freeman, R.D., Ohzawa, I., 1994. Length and width tuning of neurons in cat's primary visual cortex. J. Neurophysiol. 68, 144–163.
Delicato, L.S., Derrington, A.M., 2005. Coherent motion perception fails at low contrast. Vision Res. 45, 2310–2320.
Derrington, A.M., Badcock, D.R., 1992. Two-stage analysis of the motion of 2-dimensional patterns: what is the first stage? Vision Res. 32, 691–698.
Derrington, A.M., Allen, H., Delicato, L., 2004. Visual mechanisms of motion analysis and motion perception. Annu. Rev. Psychol. 55, 181–205.
De Valois, R.L., De Valois, K.K., 1990. Spatial Vision. Oxford Psychology Series, vol. 15. Oxford University Press, London, New York.

Drewes, J., Barthélemy, F.V., Masson, G.S., 2007. Optimal speed estimation for ocular following responses in humans is based on natural scene statistics. Soc. Neurosci. Abstr., 617.3/PP5.
Dodge, R., 1903. Five types of eye movements in the horizontal meridian plane of the field of regard. Am. J. Physiol. 8, 307–329.
Dow, B.M., Snyder, A.Z., Vautin, R.G., Bauer, R., 1981. Magnification factor and receptive field size in foveal striate cortex of the monkey. Exp. Brain Res. 44, 213–228.
Duffy, C.J., Wurtz, R.H., 1991a. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of responses selectivity to large-field stimuli. J. Neurophysiol. 65, 1329–1345.
Duffy, C.J., Wurtz, R.H., 1991b. Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. J. Neurophysiol. 65, 1346–1359.
Emerson, R.C., Adelson, E.H., Bergen, J.R., 1992. Directionally-selective complex cells and the computation of motion energy in cat visual cortex. Vision Res. 32, 203–219.
Ehrenstein, W.H., Spillmann, L., Sarris, V., 2003. Gestalt issues in modern neurosciences. Axiomathes 13, 433–458.
Erlhagen, W., Bastian, A., Jancke, D., Riehle, A., Schöner, G., 1999. The distribution of neuronal population activation (DPA) as a tool to study interaction and integration in cortical representation. J. Neurosci. Meth. 94, 53–66.
Ferrera, V.P., 2000. Task-dependent modulation of the sensorimotor transformation for smooth pursuit eye movements. J. Neurophysiol. 84, 2725–2738.
Ferrera, V.P., Lisberger, S.G., 1997. The effect of a moving distractor on the initiation of smooth-pursuit eye movements. Vis. Neurosci. 15, 7472–7484.
Frazor, R.A., Geisler, W.S., 2006. Local luminance and contrast in natural images. Vision Res. 46, 1585–1598.
Gellman, R.S., Carl, J.R., Miles, F.A., 1990. Short-latency ocular following responses in man. Vis. Neurosci. 5, 107–122.
Glickstein, M., Gerrits, N., Kralj-Hans, I., Mercier, B., Stein, H., Voogd, J., 1994. Visual pontocerebellar projections in the macaque. J. Comp. Neurol. 349, 51–72.
Grossberg, S., 1973. Contour enhancement, short-term memory and constancies in reverberating neural networks. Stud. Appl. Math. 52, 213–257.
Gomi, H., Shidara, M., Takemura, A., Inoue, K., Kawano, K., Kawato, M., 1998. Temporal firing of Purkinje cells in the cerebellar ventral paraflocullus during ocular following responses in monkey. I. Simple spikes. J. Neurophysiol. 80, 832–848.
Grossberg, S., Mingolla, E., Viswanathan, L., 2001. Neural dynamics of motion integration and segmentation within and across apertures. Vision Res. 41, 2521–2553.
Guo, K., Benson, P.J., 1999. Grating and plaid chrominance motion influences the suppressed ocular following response. NeuroReport 10, 387–392.
Hawken, M.J., Shapley, R.M., Grosof, D.H., 1996. Temporal frequency selectivity in monkey visual cortex. Vis. Neurosci. 13, 477–492.
Hayashi, R., Miura, K., Tabata, H., Kawano, K., 2008. Eye movements in response to dichoptic motion: evidence for a parallel-hierarchical structure of visual motion processing in primates. J. Neurophysiol. 99, 2329–2346.
Hayashi, R., Sugita, Y., Nishida, S., Kawano, K., 2010. How motion signals are integrated across frequencies: study on motion perception and ocular following responses using multiple-slit stimuli. J. Neurophysiol. 103, 230–243.
Hendry, S.H., Reid, R.C., 2000. The koniocellular pathway in primate vision. Annu. Rev. Neurosci. 23, 127–153.
Heeger, D.J., 1993. Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. J. Neurophysiol. 70, 1885–1898.
Henriksson, L., Nurminen, L., Hyvärinen, A., Vanni, S., 2008. Spatial frequency tuning in retinotopic visual areas. J. Vision 8 (10), 1–13, 5.
Howard, I.P., Simpson, W.A., 1989. Human optokinetic nystagmus is linked to the stereoscopic system. Exp. Brain Res. 78, 309–314.
Huang, X., Lisberger, S.G., 2009. Noise correlations in cortical area MT and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements. J. Neurophysiol. 101 (6), 3012–3030.
Ilg, U.J., 2008. The role of area MT and MST in coding visual motion underlying the execution of smooth pursuit. Vision Res. 48, 2062–2069.
Ilg, U.J., 1997. Slow eye movements. Prog. Neurobiol. 53, 293–329.
Inoue, Y., Takemura, A., Kawano, M.J., Mustari, M.J., 2000. Role of the prectectal nucleus of the optic tract in short-latency ocular following responses in monkeys. Exp. Brain Res. 131, 269–281.
Jancke, D., Erlhagen, W., Dinse, H.R., Akhavan, A.C., Giese, M.A., Steinhage, A., Schöner, G., 1999. Parametric population representation of retinal location: neuronal interaction dynamics in cat primary visual cortex. J. Neurosci. 19, 9016–9028.
Jung, R., Spillman, L., 1970. Receptive-field estimation and perceptual integration in human vision. Early experience and visual information processing in perceptual and reading disorders. In: Young, F.A., Lindsley, D.B. (Eds.), National Academy of Sciences Proceedings. Washington, DC, pp. 181–197.
Kawano, K., 1999. Ocular tracking: behavior and neurophysiology. Curr. Opin. Neurobiol. 9, 467–473.
Kawano, K., Shidara, M., Watanabe, Y., Yamane, S., 1994. Neural activity in cortical area MST of alert monkey during ocular following responses. J. Neurophysiol. 71, 2305–2334.
Kawano, K., Shidara, M., Yamane, S., 1992. Neural activity in dorsolateral pontine nucleus of alert monkey during ocular following responses. J. Neurophysiol. 67, 680–703.
Keller, E.L., Heinen, S.J., 1991. Generation of smooth-pursuit eye movements: neuronal mechanisms and pathways. Neurosci. Res. 11, 79–107.
Kelly, D.H., 1961. Visual response to time-dependent stimuli. I. Amplitude sensitivity measurements. J. Opt. Soc. Am. 51, 422–429.
Kelly, D.H., 1971a. Theory of flicker and transient responses. I. Uniform fields. J. Opt. Soc. Am. 61, 537–546.
Kelly, D.H., 1971b. Theory of flicker and transient responses. II. Counterphase gratings. J. Opt. Soc. Am. 61, 632–640.
Kimmig, H.G., Miles, F.A., Schwarz, U., 1992. Effects of stationary textured background on the initiation of pursuit eye movements in monkeys. J. Neurophysiol. 68, 2147–2164.
Knierim, J.J., Van Essen, D.C., 1992. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J. Neurophysiol. 41, 1071–1095.
Kodaka, Y., Wada, Y., Kawano, K., 2003. Ocular responses to forward motion in monkeys: visual modulation at ultra-short latencies. Exp. Brain Res. 148, 541–544.
Kodaka, Y., Sheliga, B.M., FitzGibbon, E.J., Miles, F.A., 2007. The vergence eye movements induced by radial optic flow: some fundamental properties of the underlying local-motion detectors. Vision Res. 47, 2637–2660.
Kowler, E., Steinman, R.M., 1979. The effect of expectation on slow oculomotor control. II. Single target displacements. Vision Res. 19, 633–646.
Krauzlis, R.J., 2004. Recasting the smooth pursuit eye movement system. J. Neurophysiol. 91, 591–603.
Krauzlis, R.J., Stone, L.S., 1999. Tracking with the mind's eye. Trends Neurosci. 22, 544–550.
Krekelberg, B., Albright, T.D., 2004. Motion mechanisms in macaque MT. J. Neurophysiol. 93, 2908–2921.
Lamme, V.F., 1995. The neurophysiology of figure-ground segregation in primary visual cortex. J. Neurosci. 15, 1605–1615.
Leigh, R.J., Zee, D.S., 2006. The Neurology of Eye Movements, IVth edition. Oxford University Press.
Lisberger, S.G., 2010. Visual guidance of smooth-pursuit eye movements: sensation, action and what happens in between. Neuron 66, 477–491.
Lisberger, S.G., Miles, F.A., Optican, L.M., Eighmy, B.B., 1981. Optokinetic responses in monkey: underlying mechanisms and their sensitivity to long-term adaptive changes in vestibuloocular reflexes. J. Neurophysiol. 45, 869–890.
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.
Lisberger, S.G., Morris, E.J., Tychsen, L., 1987. Visual motion processing and sensory-motor integration for smooth pursuit eye movements. Annu. Rev. Neurosci. 10, 97–129.
Lorenceau, J., Zago, L., 1999. Cooperative and competitive spatial interactions in motion integration. Vis. Neurosci. 16, 755–770.
Lu, Z.L., Sperling, G., 2001. Three-systems theory of human visual motion perception: review and update. J. Opt. Soc. Am. A 18, 2331–2370.
Lu, Z.L., Sperling, G., 1995. The functional architecture of human visual motion perception. Vision Res. 35, 2697–2722.
Lui, L.L., Bourne, J.A., Rosa, M.G.P., 2007. Spatial and temporal frequency selectivity of neurons in the middle temporal visual area of new world monkeys. Eur. J. Neurosci. 25, 1780–1792.
MacNeilage, P.R., Ganesan, N., Angelaki, D., 2008. Computational approaches to spatial orientation: from transfer function to dynamic Bayesian inference. J. Neurophysiol. 100, 2981–2996.
Masson, G.S., 2004. From 1D to 2D via 3D: dynamics of surface motion segmentation for ocular tracking in primates. J. Physiol. (Paris) 98, 35–52.
Masson, G.S., Ilg, U.W., 2010. Dynamics of Biological Motion. Springer-Verlag, New-York.
Masson, G.S., Busettini, C., Miles, F.A., 1997. Vergence eye movements in response to binocular disparity without depth perception. Nature 389, 283–286.
Masson, G.S., Castet, E., 2002. Parallel motion processing for the initiation of short-latency ocular following in humans. J. Neurosci. 22, 5149–5163.
Masson, G.S., Busettini, C., Yang, D.-S., Miles, F.A., 2001. Short-latency ocular following in humans: sensitivity to binocular disparity. Vision Res. 41, 3371–3387.
Masson, G.S., Stone, L.S., 2002. From following edges to pursuing objects. J. Neurophysiol. 88, 2869–2873.
Masson, G.S., Yang, D.-S., Miles, F.A., 2002a. Reversed short-latency ocular following. Vision Res. 42, 2081–2087.
Masson, G.S., Yang, D.-S., Miles, F.A., 2002b. Version and vergence eye movements in humans: open-loop dynamics determined by monocular rather than binocular image speed. Vision Res. 32, 2853–2867.
Masson, G.S., Rybarczyk, Y., Castet, E., Mestre, D.R., 2000. Temporal dynamics of motion integration for the initiation of tracking eye movements at ultra-short latencies. Vis. Neurosci. 17 (5), 753–767.
Masson, G.S., Montagnini, A., Ilg, U.J., 2010. When the brain meets the eye: tracking object motion. In: Masson, G.S., Ilg, U.J. (Eds.), Dynamics of Visual Motion Processing. Springer-Verlag, New York, pp. 161–188.
Maunsell, J.H.R., Van Essen, D.C., 1983. Functional properties of neurons in middle temporal visual areas of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. J. Neurophysiol. 49, 1148–1167.
Mestre, D.R., Masson, G.S., 1997. Ocular responses to motion parallax stimuli: the role of perceptual and attentional factors. Vision Res. 37, 1627–1641.
Miles, F.A., 1993. The sensing of rotational and translational optic flow by the primate optokinetic system. Rev. Oculomot. Res. 5, 393–403.
Miles, F.A., 1997. Visual stabilization of the eyes in primates. Curr. Opin. Neurobiol. 7, 867–871.

Miles, F.A., 1998. The neural processing of 3-D visual information: evidence from eye movements. Eur. J. Neurosci. 10, 811–822.
Miles, F.A., Busettini, C., Masson, G.S., Yang, D.-S., 2004. Short-latency eye movements: evidence for rapid, parallel processing of optic flow. In: Vaina, L.M., Beardsley, S.A., Rushton, S. (Eds.), Optic Flow and Beyond. Kluwer Academic Press, pp. 70–107.
Miles, F.A., Kawano, K., Optican, L.M., 1986. Short-latency ocular following responses of monkey. I. Dependence on temporospatial properties of visual input. J. Neurophysiol. 56, 1321–1354.
Miles, F.A., Kawano, K., 1986. Short-latency ocular following responses of monkey. III. Plasticity. J. Neurophysiol. 56, 1381–1386.
Miles, F.A., Sheliga, B.M., 2010. Motion detection for reflexive tracking. In: Masson, G.S., Ilg, U.J. (Eds.), Dynamics of Biological Motion. Springer-Verlag, New York.
Mingolla, E., Todd, J.T., Norman, J.F., 1992. The perception of globally coherent motion. Vision Res. 32, 1015–1031.
Miura, K., Matsuura, K., Taki, M., Tabata, H., Inada, N., Kawano, K., Miles, F.A., 2006. The visual motion detectors underlying ocular following responses in monkeys. Vision Res. 46, 869–878.
Movshon, J.A., Adelson, E.H., Gizzi, M.S., Newsome, W.T., 1985. The analysis of visual moving patterns. In: Chagas, C., Gattass, R., Gross, C. (Eds.), Pattern Recognition Mechanisms. Springer-Verlag, 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.
Murasugi, C.M., Howard, I.P., Ohmi, M., 1989. Human optokinetic nystagmus: competition between stationary and moving display. Percept. Psychophys. 45, 137–144.
Murray, R.F., Sekuler, A.B., Bennett, P.J., 2003. A linear cue combination framework for understanding selective attention. J. Vision 3, 116–145.
Nakayama, K., Loomis, J.M., 1974. Optical velocity patterns, velocity-sensitive neurons and space perception: a hypothesis. Perception 3, 63–80.
Naka, K.I., Rushton, W.A., 1966. S-potentials from luminosity units in the retina of fish (Cyprinidae). J. Physiol. (Lond.) 185, 587–599.
Nakayama, K., 1985. Biological motion processing: a review. Vision Res. 25, 625–660.
Neri, P., 2004. Estimation of nonlinear psychophysical kernels. J. Vision 4, 82–91.
Neri, P., Levi, D.M., 2006. Receptive versus perceptive fields from the reverse-correlation viewpoint. Vision Res. 46, 2465–2474.
Neri, P., Parker, A.J., Blakemore, C., 1999. Probing the human stereoscopic image features with reverse correlation. Nature 401, 695–698.
Nienborg, H., Bridge, H., Parker, A.J., Cumming, B.G., 2005. Neuronal computation of disparity in V1 limits temporal resolution for detecting disparity modulation. J. Neurosci. 25, 10207–10219.
Nover, H., Anderson, C.H., DeAngelis, G.C., 2005. A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. J. Neurosci. 25, 10049–10060.
Nurminen, L., Kilpeläinen, M., Laurinen, P., Vanni, S., 2009. Area summation in human visual system: psychophysics, fMRI and modeling. J. Neurophysiol. 102, 2900–2909.
O'Keefe, L.P., Movshon, J.A., 1998. Processing of first- and second-order motion signals in cortical area MT of the macaque monkey. Vis. Neurosci. 15, 305–317.
Orban, G., Van Essen, D., Vanduffel, W., 2004. Comparative mapping of higher visual areas in monkeys and humans. Trends Cogn. Sci. 8, 315–324.
Osborne, L.C., Bialek, W., Lisberger, S.G., 2004. Time course of information about motion direction in visual area MT of macaque monkeys. J. Neurosci. 24, 3210–3222.
Osborne, L.C., Lisberger, S.G., Bialek, W., 2005. A sensory source for motor variation. Nature 437, 412–416.
Osborne, L.C., Hohl, S.S., Bialek, W., Lisberger, S.G., 2007. Time-course of precision in smooth-pursuit eye movements of monkeys. J. Neurosci. 27, 2987–2998.
Pack, C.C., Born, R.T., 2001. Temporal dynamics of a neural solution to the aperture problem in visual area MT. Nature 409, 1040–1042.
Pack, C.C., Berezovskii, V.K., Born, R.T., 2001. Dynamic properties of neurons in cortical area MT in alert and anesthetized macaque monkeys. Nature 414, 905–908.
Pack, C.C., Gartland, A.J., Born, R.T., 2004. Integration of contour and terminator signals in visual area MT of the alert macaque. J. Neurosci. 31, 3268–3280.
Pack, C.C., Hunter, J.N., Born, R.T., 2005. Contrast dependence of suppressive influences in cortical area MT of alert macaque. J. Neurophysiol. 93, 1809–1815.
Pantle, A., Turano, K., 1992. Visual resolution of motion ambiguity with periodic luminance- and contrast-domain stimuli. Vision Res. 32, 2093–2106.
Perrinet, L.U., Masson, G.S., 2007. Modeling spatial integration in the ocular following response using a probabilistic framework. J. Physiol. (Paris) 101, 46–55.
Priebe, N.J., Churchland, M.A., Lisberger, S.G., 2001. Reconstruction of target speed for the guidance of smooth pursuit eye movements. J. Neurosci. 21, 3196–3206.
Priebe, N.J., Cassanello, C.R., Lisberger, S.G., 2003. The neural representation of speed in macaque area MT/V5. J. Neurosci. 23, 5650–5661.
Priebe, N.J., Lisberger, S.G., 2004. Estimating target speed from the population response in visual area MT. J. Neurosci. 24, 1907–1916.
Rashbass, C., 1961. The relationship between saccadic and smooth tracking eye movements. J. Physiol. (Lond.) 159, 326–338.
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.
Recanzone, G.H., Wurtz, R.H., 1999. Smooth pursuit initiation and MT and MST neuronal activity under different stimulus conditions. J. Neurophysiol. 82, 1710–1727.
Recanzone, G.H., Wurtz, R.H., 2000. Effects of attention on area MT and MST neuronal activity during pursuit initiation. J. Neurophysiol. 83, 777–790.
Reynaud, A., Barthélemy, F.V., Masson, G.S., Chavane, F., 2007. Input–output transformation in the visuo-oculomotor loop: comparison of real-time optical imaging recordings in V1 to ocular following responses upon center–surround interactions. Arch. Ital. Biol. 145, 251–262.
Reynolds, J.H., Chelazzi, L., Desimone, R., 1999. Competitive mechanisms subserve attention in macaque areas V2 and V4. J. Neurosci. 19, 1736–1753.
Reynolds, J.H., Heeger, D.J., 2009. The normalization model of attention. Neuron 61, 168–185.
Ringach, D.L., 1998. Tuning of orientation detectors in human vision. Vision Res. 38, 963–972.
Roufs, J.A.J., 1972. Dynamic properties of vision. I. Experimental relationships between flicker and flash thresholds. Vision Res. 12, 261–278.
Rubin, N., Hochstein, S., 1993. Isolating the effect of one-dimensional motion signals on the perceived direction of moving two-dimensional object. Vision Res. 33, 1385–1396.
Rust, N.C., Mante, V., Simoncelli, E.P., Movshon, J.A., 2006. How MT cells analyze the motion of visual patterns. Nat. Neurosci. 9, 1421–1431.
Schwartz, O., Simoncelli, E.P., 2001. Natural signal statistics and sensory gain control. Nat. Neurosci. 4, 819–826.
Sclar, G., Maunsell, J.H.R., Lennie, P., 1990. Coding of image contrast in central visual pathways of the macaque monkey. Vision Res. 30, 1–10.
Sceniak, M.P., Ringach, D., Hawken, M.J., Shapley, R., 1999. Contrast's effect on spatial summation by macaque V1 neurons. Nat. Neurosci. 2, 733–739.
Shapley, R., 1990. Visual sensitivity and parallel pathways. Annu. Rev. Psychol. 41, 635–658.
Sheliga, B.M., Chen, K.J., FitzGibbon, E.J., Miles, F.A., 2005. Initial ocular following in humans: a response to first-order motion energy. Vision Res. 45, 3307–3321.
Sheliga, B.M., Chen, K.J., FitzGibbon, E.J., Miles, F.A., 2006a. The initial ocular following responses elicited by apparent-motion stimuli: reversal by inter-stimulus intervals. Vision Res. 46, 979–992.
Sheliga, B.M., Kodaka, Y., FitzGibbon, E.J., Miles, F.A., 2006b. Human ocular following initiated by competing image motions: evidence for a winner-take-all mechanism. Vision Res. 46, 2041–2060.
Sheliga, B.M., FitzGibbon, E.J., Miles, F.A., 2008a. Human ocular following: evidence that responses to large-field stimuli are limited by local and global inhibitory influences. Prog. Brain Res. 171, 237–243.
Sheliga, B.M., FitzGibbon, E.J., Miles, F.A., 2008b. Spatial summation properties of ocular following response (OFR): evidence for nonlinearities due to local and global inhibitory interactions. Vision Res. 48, 1758–1776.
Sheliga, B.M., FitzGibbon, E.J., Miles, F.A., 2009. The initial torsional ocular following response (fOFR): a response to motion energy in the stimulus? J. Vision 9 (12), 2.1–38.
Sherrington, 1906. The Integrative Action of the Nervous System. Charles Scribner's Sons, New York.
Shidara, M., Kawano, K., 1993. Role of Purkinje cells in the ventral paraflocullus in short-latency ocular following responses. Exp. Brain Res. 93, 185–195.
Singh, K.D., Smith, A.T., Greenlee, M.W., 2000. Spatiotemporal frequency and direction sensitivities of human visual areas measured using fMRI. Neuroimage 12, 550–564.
Shiori, S., Cavanagh, P., 1990. ISI produces reverse apparent motion. Vision Res. 30, 757–768.
Shipp, S., Zeki, S., 1989. The organisation of connections between areas V5 and V1 in macaque monkey visual cortex. Eur. J. Neurosci. 1, 308–331.
Simoncelli, E.P., Heeger, D.J., 1996. A model of neuronal responses in visual area MT. Vision Res. 36, 743–761.
Simoncini, C., Perrinet, L.U., Montagnini, A., Mamassian, P., Masson, G.S., 2010. Different pooling of information for perceptual speed discrimination and behavioural speed estimation. J. Vision. 10 (7), 837.
Sincich, L.C., Park, K.F., Wohlgemuth, M.J., Horton, J.C., 2004. Bypassing V1: a direct geniculate input to area MT. Nat. Neurosci. 7, 1123–1128.
Sit, Y.F., Chen, Y., Geisler, W.S., Miikkulainen, R., Seidemann, E., 2010. Complex dynamics of V1 population responses explained by a simple gain-control model. Neuron 64, 943–956.
Smith, A.T., 1994. Correspondance-based and energy-based detection of second-order motion in human vision. J. Opt. Soc. Am. A 11, 1940–1948.
Smith, M.A., Majaj, N.J., Movshon, J.A., 2005. Dynamics of motion signalling by neurons in macaque monkeys. Nat. Neurosci. 8, 220–228.
Solomon, J.A., 2002. Noise reveals mechanisms of detection and discrimination. J. Vision 2, 105–120.
Spillman, L., 2006. From perceptive fields to Gestalt. Prog. Brain Res. 155, 67–92.
Swanson, W.H., Ueno, T., Smith, T.C., Pokorny, J., 1987. Temporal modulation sensitivity and pulse-detection thresholds for chromatic and luminance perturbations. J. Opt. Soc. Am. A 4, 1992–2005.
Tadin, D., Lappin, J.S., 2005. Optimal size for perceiving motion decreases with contrast. Vision Res. 45, 2059–2064.
Tadin, D., Lappin, J.S., Gilroy, L.A., Blake, R., 2003. Perceptual consequences of centre–surround antagonism in visual motion processing. Nature 424, 312–315.
Takemura, A., Inoue, Y., Kawano, K., Quaia, C., Miles, F.A., 2001. Single-unit activity in cortical area MST associated with disparity-vergence eye movements. J. Neurophysiol. 85, 2245–2266.
Takemura, A., Kawano, K., 2002. Sensory-to-motor processing of the ocular following response. Neurosci. Res. 43, 201–206.
Takemura, A., Kawano, K., Quaia, C., Miles, F.A., 2002. Population coding in area MST. Ann. N. Y. Acad. Sci. 956, 284–296.

Takemura, A., Murata, Y., Kawano, K., Miles, F.A., 2007. Deficits in short-latency tracking eye movements after chemical lesions in areas MT and MST. J. Neurosci. 27, 529–541.
Takeuchi, T., De Valois, K.K., 1997. Motion-reversal reveals two motion mechanisms functioning in scotopic vision. Vision Res. 37, 745–755.
Takeuchi, T., De Valois, K.K., Motoyoshi, I., 2001. Light adaptation in motion direction judgments. J. Opt. Soc. Am. A 18, 755–764.
Tlapale, E., Masson, G.S., Kornprobst, P., 2010. Modelling the dynamics of motion integration with a new luminance-gated diffusion mechanism. Vision Res. 50 (17), 1676–1692.
Tsui, J.M., Hunter, J.N., Born, R.T., Pack, C.C., 2010. The role of V1 surround suppression in MT motion integration. J. Neurophysiol. 103, 3123–3138.
Ullman, S., 1979. The Interpretation of Visual Motion. MIT, Cambridge, 229 pp.
Van Santen, J.P., Sterling, G., 1984. Temporal covariance model of human motion perception. J. Opt. Soc. Am. A 2, 300–321.
Wallace, J., Stone, L.S., Masson, G.S., 2005. Object motion computation for the initiation of smooth pursuit eye movements in humans. J. Neurophysiol. 93, 2279–2293.
Watson, A.B., Ahumada, A.J., 1985. Model of human visual motion sensing. J. Opt. Soc. Am. A 2, 322–341.
Welch, L., Bowne, S.F., 1992. Coherence determines speed discrimination. Vision Res. 19, 425–435.
Wilmer, J.B., Nakayama, K., 2007. Two distinct visual motion mechanisms for smooth pursuit: evidence from individual differences. Neuron 54, 987–1000.
Wilson, H.R., Ferrera, V.P., Yo, C., 1992. A psychophysically motivated model for two-dimensional motion perception. Vis. Neurosci 32, 135–147.
Xing, J., Heeger, D.J., 2000. Center–surround interactions in foveal and peripheral vision. Vision Res. 40, 3065–3072.
Xing, J., Heeger, D.J., 2001. Measurement and modelling of center–surround suppression and enhancement. Vision Res. 41, 571–583.
Yabuta, N.H., Sawatari, A., Callaway, E.M., 2001. Two functional channels from primary visual cortex to dorsal cortical areas. Science 292, 297–300.