

Cortical mechanisms underlying  
low-level motion processing  
in the visual system of human  
and non-human primates

Roger Jacques Elisabeth Bours

### **Thesis committee**

#### **Thesis supervisors**

Prof. dr. ir. J.L. van Leeuwen  
Professor of Experimental Zoology  
Wageningen University

Prof. dr. R.J.A. van Wezel  
Professor of Biomedical Signals and Systems  
University of Twente

#### **Thesis co-supervisor**

Dr. ir. M.J.M. Lankheet  
Associate Professor, Experimental Zoology Group  
Wageningen University

#### **Other members**

Prof. dr. K.H. Britten, University of California, Davis, USA  
Prof. dr. F.A.J. Verstraten, Utrecht University  
Prof. dr. M. Dicke, Wageningen University  
Dr. A.J. van Doorn, Associate Professor, Delft University of Technology  
Dr. A.J. Noest, Utrecht University

This research was conducted under the auspices of the Graduate School of the Wageningen Institute of Animal Sciences.

# Cortical mechanisms underlying low-level motion processing in the visual system of human and non-human primates

Roger Jacques Elisabeth Bours

## **Thesis**

submitted in fulfillment of the requirements for the degree of doctor  
at Wageningen University  
by the authority of the Rector Magnificus  
Prof. dr. M.J. Kropff,  
in the presence of the  
Thesis Committee appointed by the Academic Board  
to be defended in public  
on Friday 29 October 2010  
at 11 AM in the Aula.

Roger J. E. Bours

Cortical mechanisms underlying low-level motion processing  
in the visual system of human and non-human primates,  
162 pages.

Thesis, Wageningen University, Wageningen, NL (2010)  
With references, with summaries in Dutch and English

ISBN 978-90-8585-794-5

**Paranimfen**

Ir. Roy Catharina Hubert Bours  
Maurice Joannes Cornelis Toepoel

*Aan mijn familie  
en vrienden*



# Preface

Knowing the research chapters for my thesis are finished, I am sitting at home watching a snow blizzard through the window. Through the falling snow I see the houses in the street with trees on both sides and cars driving by. When watching the snow flakes you can see them being carried by the wind around the houses and between the branches of the trees in numerous directions. Besides a beautiful scene it is actually an intriguing example of the power of our visual system. It enables us to follow a single snow flake getting carried by the wind, and still distinguish other snow flakes moving in various directions. You can even distinguish which branch belongs to which tree.

If you would imagine building a program or robot that could achieve all of this, then you might get a feeling of the complexity of our visual system. It is able to analyze information from the environment at a local scale (following a single snow flake), at a global scale (determine which branch belongs to which tree) or even both at the same time (determine the wind speed by following a single snowflake while it passes through the branches of a tree). Imagine a single day without being able to see motion. How would you safely move across traffic, ride your bicycle, drive your car, walk around or watch a movie. The desire to understand how our brain processes visual motion information and turn it into a percept, has been the main drive for the research in this thesis.

Our sophisticated motion processing system is the result of hundreds of millions years of evolution and has intrigued researchers for more than 2000 years. Since the 1900s motion research developed rapidly, due to technical improvements such as the introduction of the computer, improved monitor screens and more elaborated techniques. In addition to psychophysical approaches, also neurophysiological and neuroimaging techniques were used to study different aspects of the motion system, such as direction, speed, velocity, transparency, monocular and binocular vision, and rivalry.

This dissertation started out to physiologically study temporal aspects of the cortical mechanism underlying motion perception. Following a series of psychophysical experiments (Koenderink et al., 1985; Van de Grind et al., 1986; Van Doorn and Koenderink, 1982a,b; Fredericksen et al., 1993, 1994a,b,c; Morgan and Ward, 1980; Van den Berg and Van de Grind, 1989; Van Doorn and

---

Koenderink, 1984), we used similar visual stimuli (random dot patterns) to study the physiology of motion processing. We studied the temporal dynamics of neurons in the primate brain in a motion sensitive area, called area MT (see chapter 3). While interpreting our results and comparing them to findings by others, we realized that cells in the motion area might be susceptible to a motion illusion, called reverse-phi <sup>1</sup>, inherently present in the stimuli we used. This finding was the trigger for a new series of psychophysical studies. To get a grip of this unintended complication we developed a new motion stimulus that allowed us to measure temporal properties of human motion perception in more detail and without the encountered complications (chapter 4). Interestingly, a simple modification to this stimulus, enabled us to induce the reverse-phi percept and more importantly, allowed us to quantitatively measure spatial and temporal tuning properties of our motion system for reverse-phi (chapter 5). This was interesting to us, because we could now address the question whether initial motion analysis is performed independently in the ON and OFF pathway or also through combining both channels <sup>2</sup>. The results of our psychophysical studies showed that the human motion system is as sensitive for regular motion as it is for contrast-reversing motion (i.e. reverse-phi). This strongly suggested that efficient detection of correlation across ON and OFF channels occurs. The next step was to address the question how the signals from the ON and OFF channels are combined. The percept of regular motion arises through excitations in low-level motion sensors. If reverse-phi percept results from excitations, it should behave similarly as regular motion. However, if reverse-phi results from inhibition then it should in fact behave comparable to another illusion, namely the motion aftereffect <sup>3</sup>. Contrary to previous explanations, we found that reverse-phi motion does not cause excitation in the opposite direction, but rather inhibition in the direction of displacement (chapter 6). The evidence is based on a new set of perceptual phenomena showing that reverse-phi percepts in many respects behave like motion aftereffects, and not like regular motion. Motion adaptation causes reduced activity during a stationary test stimulus, which by means of directional opponency leads to motion perceived in the opposite direction. Our results show that reverse-phi motion similarly reduces the activity of low-level motion detectors. These new insights lead to several surprising prediction, e.g. whereas reverse-phi has similar properties as the motion aftereffect of regular motion, the motion aftereffect of reverse-phi behaves like regular motion (chapter 7).

---

<sup>1</sup>Reverse-phi is the phenomenon that contrast reversal with (apparent) motion leads to the perception of motion opposite to physical displacement of the stimulus.

<sup>2</sup>In our visual pathway low-level contrast information is represented in two different channels. ON-center cells signal positive contrasts and OFF-center cells signal negative contrasts.

<sup>3</sup>Motion aftereffect is the illusion that after prolonged viewing of a unidirectional moving stimulus, a stationary pattern appears to move in the opposite direction.



# Contents

<b>Preface</b>	<b>i</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motion processing by the visual system . . . . .	1
1.2 Low-level motion processing in the retino-geniculo-cortical pathway . . . . .	4
1.2.1 Preprocessing of visual information concerning motion . . . . .	5
1.2.2 Local motion detection . . . . .	6
1.2.3 Post-processing of motion information . . . . .	8
1.3 Human versus non-human primates . . . . .	9
1.4 Motion in Cartesian and Fourier space . . . . .	9
1.4.1 Apparent motion . . . . .	10
1.5 Modeling low-level motion processing . . . . .	12
1.5.1 Feature tracking and intensity-based models . . . . .	12
1.5.2 Correlation-based models for low-level motion detection . . . . .	13
1.5.3 Modeling post-processing of low-level motion . . . . .	18
1.6 Evidence for three-stage motion processing . . . . .	18
1.6.1 The ON and OFF pathway . . . . .	19
1.6.2 Transparent motion . . . . .	20
1.6.3 Motion aftereffect . . . . .	23
1.6.4 Reverse-phi . . . . .	23
1.7 Mechanisms underlying motion detection . . . . .	25
<b>2 Overview, summary and conclusions</b>	<b>29</b>

<b>3</b>	<b>Step size tuning in macaque area MT</b>	<b>35</b>
3.1	Abstract . . . . .	35
3.2	Introduction . . . . .	36
3.3	Methods . . . . .	38
3.3.1	Subjects . . . . .	38
3.3.2	Surgical preparation, neuronal recording procedures and data acquisition . . . . .	39
3.3.3	Visual stimuli . . . . .	39
3.3.4	Experimental procedure . . . . .	40
3.3.5	Data analysis . . . . .	42
3.4	Results . . . . .	43
3.4.1	Direction tuning . . . . .	43
3.4.2	Step size tuning . . . . .	45
3.4.3	Interval duration . . . . .	50
3.5	Discussion . . . . .	53
3.5.1	Step size tuning . . . . .	53
3.5.2	Response dynamics . . . . .	55
<b>4</b>	<b>Tuning for temporal interval in human apparent motion detection</b>	<b>57</b>
4.1	Abstract . . . . .	57
4.2	Introduction . . . . .	58
4.3	Methods . . . . .	61
4.4	Results . . . . .	63
4.4.1	Interactions between step size and temporal interval tuning	64
4.4.2	Dot density . . . . .	68
4.4.3	Viewing distance . . . . .	68
4.5	Discussion . . . . .	69
4.5.1	Spatio-temporal separability . . . . .	71
<b>5</b>	<b>Sensitivity to reverse-phi motion</b>	<b>73</b>
5.1	Abstract . . . . .	73
5.2	Introduction . . . . .	74
5.3	Methods . . . . .	76
5.3.1	Stimuli . . . . .	76

5.3.2	Measurement procedure . . . . .	79
5.4	Results . . . . .	80
5.4.1	Regular and reverse-phi motion . . . . .	80
5.4.2	No-phi motion . . . . .	80
5.4.3	Dot density . . . . .	85
5.5	Discussion . . . . .	86
<b>6</b>	<b>The parallel between reverse-phi and motion aftereffects</b>	<b>91</b>
6.1	Abstract . . . . .	91
6.2	Introduction . . . . .	92
6.3	Methods . . . . .	95
6.4	Results . . . . .	97
6.4.1	Experiment 1. Orthogonal reverse-phi motion . . . . .	97
6.4.2	Experiment 2: Motion nulling . . . . .	99
6.5	Discussion . . . . .	101
<b>7</b>	<b>General discussion</b>	<b>105</b>
7.1	Single-step two-frame motion stimulus . . . . .	105
7.2	Modeling motion processing . . . . .	108
7.2.1	Pre-processing: Spatial and temporal filtering . . . . .	110
7.2.2	Motion detection: Correlation and local integration stage . . . . .	111
7.2.3	Post-processing . . . . .	112
7.2.4	The mechanism underlying local motion detection . . . . .	113
7.2.5	Combining local motion signals: The role of excitatory and inhibition signals on motion percepts . . . . .	116
7.3	Concluding remark . . . . .	120
7.4	Perspective . . . . .	121
	<b>Bibliography</b>	<b>125</b>
	<b>Overzicht, samenvatting en conclusies (Dutch)</b>	<b>143</b>
	<b>Dankwoord (Dutch)</b>	<b>149</b>
	<b>List of publications</b>	<b>153</b>
	<b>Curriculum Vitae</b>	<b>155</b>



# Chapter 1

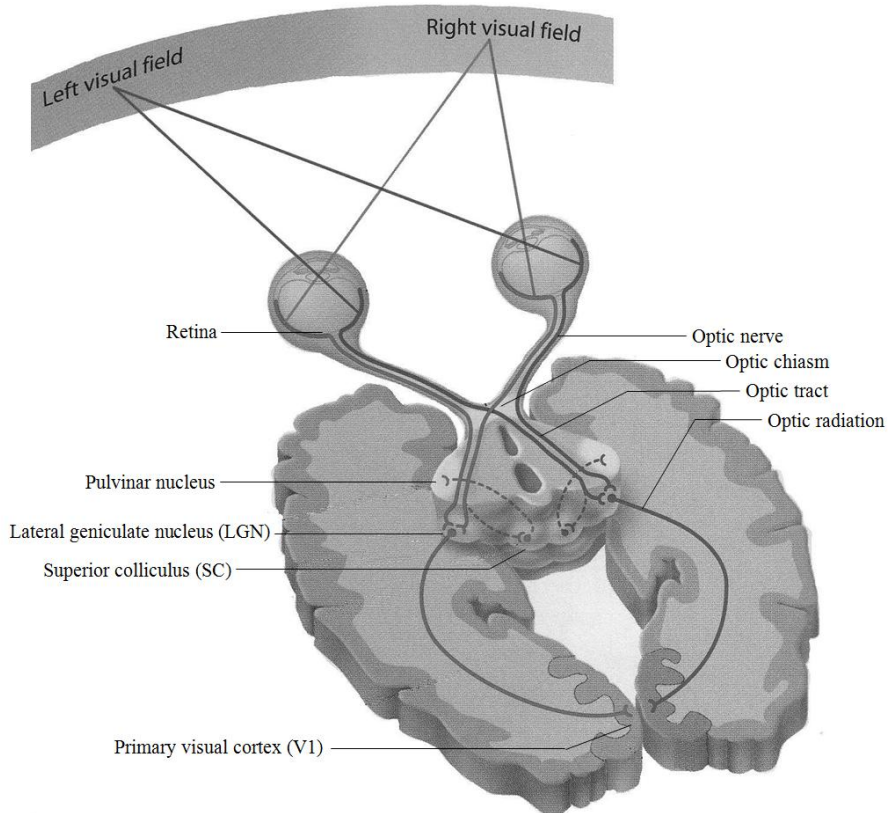
## Introduction

### 1.1 Motion processing by the visual system

Visual images are constantly projected on the retina, a thin multi-layered neuronal structure located in the back of our eye, and interpreted by our visual system. Detecting and analyzing motion in the visual scene is a major function of the visual system. The importance of motion processing for survival of visually oriented species is easily appreciated by watching nature around us. Evolutionary pressure has led to highly sophisticated motion detection and processing systems, in predators as well as in prey.

Although the visual system of primates is large and complex, it is well-organized (Stevens, 2001; Barton, 1998; Van Essen et al., 1992). The primate brain has two primary pathways for visual processing, based on where they terminate within the subcortex (figure 1.1). The retino-geniculate pathway carries visual information to the lateral geniculate nuclei (LGN) of the thalamus. This pathway provides input to cortical visual processing areas in the occipital cortex, and is also referred to as the retino-geniculo-cortical or geniculo-cortical pathway. According to the two-streams hypothesis (Mishkin and Ungerleider, 1982; Ettliger, 1990), this pathway is the basis for two major processing streams (figure 1.2). The ventral stream projects to the inferior temporal cortex and is also called the ‘what’ stream. This stream is involved in object recognition and form representation. The dorsal stream projects to the posterior parietal cortex and is involved in recognizing where objects are in space (spatial vision) and in the guidance of actions (vision for action). Accordingly, this stream is referred to as the ‘where’ and ‘how’ stream. The functional significance of this distinction remains controversial: Several studies support it (Vaina, 1994; Goodale, 1992), but others dispute it (Vaina, 1994; Franz et al., 2000, 2005). Physiological findings revealed that both streams are heavily interconnected (Farivar, 2009). Milner and

## 1.1 Motion processing by the visual system



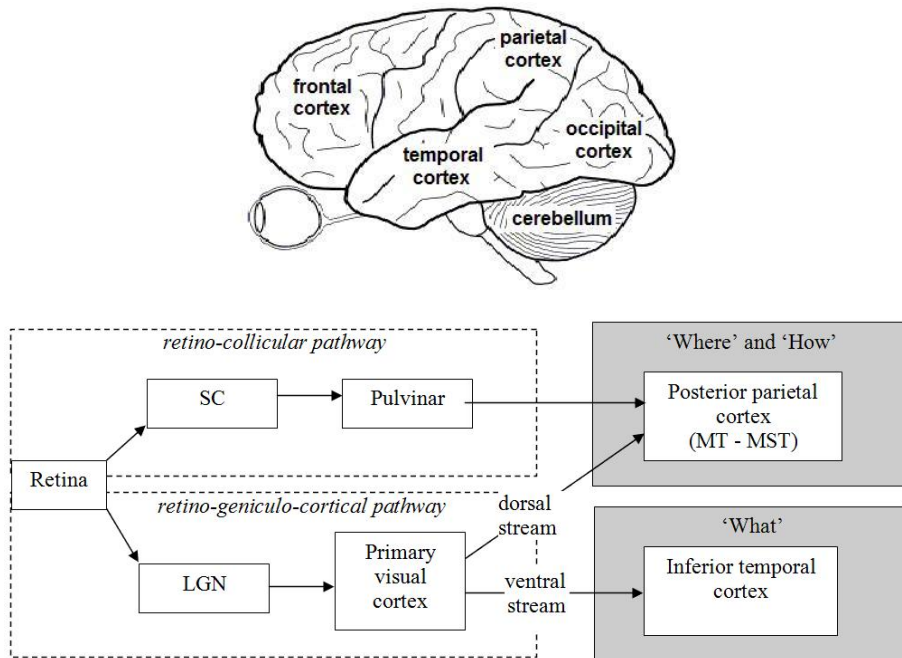
**Figure 1.1:** The primary projection pathways of the visual system based on where they terminate in the subcortex: Visual information falls on photoreceptors of the retina and is processed by several layers of retinal cells. Axons of retinal ganglion cells form the optic nerve, along which information is sent to higher visual centers. Information from the left and right visual fields stays separated and end up in the right and left hemisphere. To this end, axons carrying information from the nasal side cross over to the opposite side and combine with the axons carrying the information from the temporal side to form the optic tract. About 90% of the axons in the optic nerve project to the lateral geniculate nuclei (LGN) and subsequently to the cortex. This pathway is referred to as the retino-geniculo-cortical pathway. Almost all projections from the LGN ascend to the primary visual cortex (V1), which contains the first direction selective neurons in the retino-geniculo-cortical pathway. V1 projects to higher extra-striate cortical areas that are involved in further processing of the motion information. The remaining 10% of the axons project to the superior colliculi and pulvinar nuclei. This pathway is called the retino-collicular pathway (after figure 5.4 in Gazzaniga et al., 2002)

Goodale (2008) suggested that the differences between the two streams are best characterized by the output that the two systems serve, rather than by their visual inputs. They argue that the ventral stream is used for visual perception and the dorsal stream for ‘active vision’ (Goodale, 1992; Milner and Goodale, 1993, 2006).

Within the retino-geniculo-cortical pathway the first motion sensitive neurons are located in the primary visual cortex (V1). Interestingly, motion sensitivity already occurs at the retinal ganglion cell level (for review see Vaney et al., 2001). However, these directionally selective ganglion cells do not project to LGN but to the so-called retino-collicular pathway (figure 1.2). This pathway contains projections from the retina to the superior colliculi in the midbrain, which have connections with the pulvinar nuclei of the thalamus. This pathway projects to the posterior parietal cortex and is part of the visuomotor system. It plays an important role in eye movements and visual attention (Zihl and Von Cramon, 1979; Rodman et al., 1990), but supposedly plays no role in conscious visual perception. In this thesis, we focus on motion processing in the retino-geniculo-cortical pathway.

Where and how motion sensitivity arises in the visual pathway remained unclear until the late '50s, when several landmark studies shed a new light on the mechanisms underlying motion perception (e.a. Hubel and Wiesel, 1959, 1962, 1963; Tolhurst, 1973; Tolhurst and Movshon, 1975; Movshon et al., 1978a,b,c). These studies showed how the visual system builds a complex representation of the visual image from simple stimulus features such as orientation, color and motion. This led to a large number of studies on visual motion processing, and to the consensus that the processing of motion information is not postponed until a high, cognitive level in the brain, but is done via so-called front-end processing, i.e. at a low level.

To study visual motion processing several experimental and theoretical approaches are being used. Psychophysical studies measure the ability of a subject to perform a visual task, in order to gather information on mechanisms and limitations of the visual system. Neuroanatomy provides structural information on brain areas involved in visual processing, their cellular components and connections. Neurophysiology provides information on activity of individual neurons or groups of neurons in these areas. Neuroimaging techniques, for example functional magnetic resonance imaging (fMRI), allow non-invasive measurements of brain activation related to cognitive tasks. Finally, computational modeling is used to integrate different sources of information and to understand their relation to visual perception. In this dissertation, we have used both neurophysiological and psychophysical techniques to examine spatio-temporal properties of the primate motion system. We use computational models to address the question what the experimental results tell us about the fundamental principles underlying low-level motion detection.



**Figure 1.2:** Visual information from the retina is projected via the retino-collicular pathway and the retino-geniculo-cortical pathway. The retino-collicular pathway projects from the mid-brain to the posterior parietal cortex, which is involved in spatial vision and vision for action. The retino-geniculo-cortical pathway projects from the occipital cortex via the dorsal stream to the posterior parietal cortex and via the ventral stream to the inferior temporal cortex, which is involved in object recognition and form representation. The dorsal stream is also referred to as the ‘where’ (as in spatial vision) and ‘how’ stream (as in vision for action). The ventral stream is also called the ‘what’ stream (as in object recognition and form representation).

## 1.2 Low-level motion processing in the retino-geniculo-cortical pathway

Braddick (1974) showed that motion processing could be divided in two mechanisms, a short-range and a long-range motion mechanism. The short-range mechanism involves low-level mechanisms and acts over relatively short spatial distances. The long-range motion mechanism involves higher-order processes, e.g. object recognition, and can handle larger displacements and longer time intervals. In this thesis, the focus is on the short-range mechanism in the retino-geniculo-cortical pathway. The short-range motion pathway can be divided into three stages: (1) Processing of visual information preceding motion detection, (2) local motion detection and (3) post-processing of local motion information. Post-processing includes spatio-temporal integration of similar local motion com-



ponents, segregation of contrasting motion components, and extracting complex patterns of motion (e.g. optic flow). Each of these stages may set limits to motion processing and hence plays a fundamental role in shaping the global motion percept.

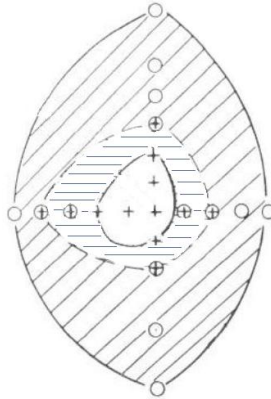
### 1.2.1 Preprocessing of visual information concerning motion

Spatio-temporal processing of visual information preceding motion detection is important because it directly affects the tuning characteristics of the motion detector stage. Preprocessing starts when visual stimuli fall on the photoreceptor cells of the retina, which translate light intensities in modulations of membrane potentials. Other types of retinal cells, namely bipolar, horizontal, amacrine and ganglion cells further process the information. The axons of the ganglion cells form the optic nerve that sends information to higher brain centers. The retinal cells perform a significant amount of spatial filtering. Information from  $\approx 260$  million photoreceptors is compressed into  $\approx 1.2$  million fibers of each optic nerve.

About 10% of the axons of the optic nerve are part of the retino-collicular pathway and connect to the superior colliculus. They play a role in e.g. controlling eye movements. About 90% of the axons of the optic nerve are part of the retinogeniculate pathway and connect to the LGN. In addition to spatial filtering, the retinal cells and the LGN are involved in temporal filtering (De Valois et al., 2000; McClurkin et al., 1991; Funke and Worgotter, 1997; Gawne, 1999).

Typically, a neuron downstream the visual pathway receives input from many cells upstream. Functional properties of a neuron can be characterized by its receptive field, the region of the retina where a stimulus alters the firing of that particular neuron. Since a neuron downstream receives input from many cells upstream, the receptive fields of those cells collectively form its receptive field. For example, the receptive field of a retinal ganglion cell is composed of all photoreceptors that provide its input. A group of ganglion cells in turn contributes to the receptive field of a cell downstream the visual pathway.

The receptive fields of vertebrate retinal ganglion cells are organized in such a way that they detect contrast rather than light intensities (Kuffler, 1953). The receptive field is divided into a central disk, the center, and a concentric ring, the surround. These regions respond with opposite polarity to light. If the center is excited by a light increase and the surround by a light decrease, it is called an ON-center cell (figure 1.3). If the center is excited by a light decrease and the surround by a light increase, it is called an OFF-center cell. The ON-center and OFF-center ganglion cells project to ON-center and OFF-center cells in the LGN (Schiller, 1992). LGN cells have receptive fields that are circularly symmetric and are therefore not directionally selective. They project to the primary visual cortex where the first directionally selective cells in retino-geniculo-cortical pathway are found. In section 1.6.1 the ON and OFF pathways are explained in more detail.

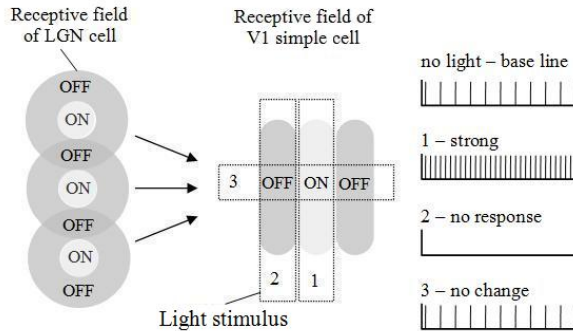


**Figure 1.3:** Drawing of receptive field of ON-center ganglion cells of a cat (redrawn from figure 6 in Kuffler, 1953). The crosses, open circles and cross-circle combination show the stimulus locations where Kuffler (1953) measured responses. The differently shaded areas give the approximate areal organization. The receptive field is divided into a central disk (center) with ON responses (cross, white area), an outer concentric ring (surround) with OFF responses (open circle, diagonally hatched) and a region in between with ON-OFF responses (cross-circle combination, horizontally hatched). In case of an OFF-center ganglion cell the center and surround properties are reversed.

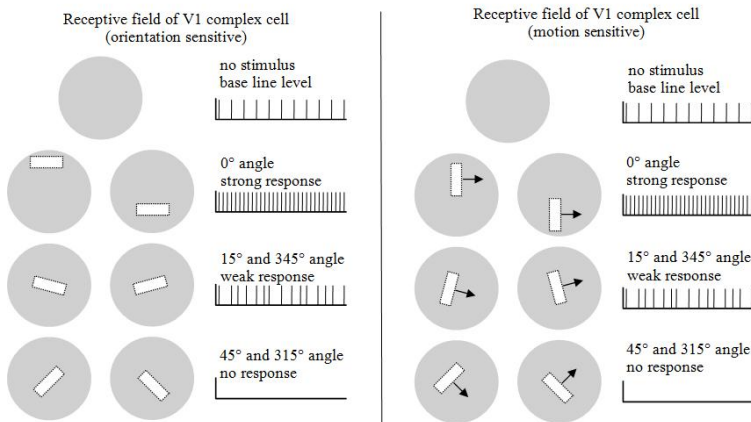
## 1.2.2 Local motion detection

To detect motion, visual information from two spatial locations has to be correlated in time. This correlation is performed by specialized front-end motion detectors. The first level in the retino-geniculo-cortical pathway where motion sensitivity arises is the primary visual cortex (Hubel and Wiesel, 1959, 1962, 1965, 1968; Zeki, 1974; Essen and Zeki, 1978; Albright, 1984; Maunsell and Van Essen, 1983; Mikami et al., 1986b). The primary visual cortex is also referred to as striate cortex or V1. Neurons in V1 can be divided in simple and complex cells (figure 1.4 and 1.5). Simple and complex cells respond primarily to oriented edges and gratings. A subpopulation of these simple and complex cells responds to movement in a certain direction (Snowden et al., 1991) and are therefore truly motion sensitive. Whereas simple cells have distinct excitatory and inhibitory regions, for complex cells these regions overlap, resulting in spatial invariance. Complex cells respond to their preferred orientation irrespective of the exact location within their receptive field. According to Hubel and Wiesel, complex cells receive input from a number of simple cells with similar orientation preference but with spatially shifted receptive fields. The summation and integration of the output of many simple cells may explain spatial invariance in complex cells. Although Hubel and Wiesel (1959, 1962, 1965) suggested that simple cells and complex cells represent two consecutive stages, later studies provided evidence that receptive fields of complex cells could also be built in parallel, by direct input from LGN ON and OFF cells. (Hoffmann and Stone, 1971; Stone et al., 1979): E.g.

Hammond and MacKay (1975, 1977) showed that some complex cells responded to visual stimuli (random dot patterns) that were not effective in driving simple cells (Hammond, 1991). Cells in area V1 project to different extrastriate areas, where further processing takes place (for extensive review on simple and complex cells see e.a. Martinez and Alonso, 2003; Skottun et al., 1991; Mechler and Ringach, 2002).



**Figure 1.4:** Simple cells in area V1 are selective for orientation. They receive input from LGN cells. LGN cells have circular, center-surround receptive fields. The receptive fields are ON-center/OFF-surround or OFF-center/ON-surround (not shown). A simple cell has excitatory and inhibitory areas that determine the preferred orientation of that cell. The response of a simple cell gradually changes with the orientation of the light stimulus which results in an orientation tuning profile.



**Figure 1.5:** Complex cells in area V1 are selective for orientation. A subpopulation of complex cells is also selective for direction. The response gradually changes with the preferred orientation or direction. Contrary to simple cells, the receptive field of complex cells has overlapping ON and OFF regions and show spatial invariance: The location of the stimulus in the receptive field is not critical.

### 1.2.3 Post-processing of motion information

Simple and complex cells in area V1 have relatively small receptive fields and therefore extract local motion information. Further processing of this information is required to construct a globally consistent motion percept. Post-processing on the one hand allows for integration of local motion signals that belong to the same moving object (or background), and on the other hand for segregation of motion components that belong to different objects. This information is then used to build more complex motion patterns (e.g. optic flow and biological motion) and finally the information is combined into a global motion percept. This process can be illustrated by a plaid stimulus, which consists of two superimposed sine-wave gratings with different orientations. Local motion detectors in V1 respond mostly to motion perpendicular to the orientations of either one of the gratings. These detectors act as component selective cells. A plaid, however, is perceived to move coherently in a direction corresponding to the vector sum of the separate components. A subpopulation of neurons in area MT responds in line with the perceived direction of motion. These cells are called pattern selective cells and integrate the information of local detectors (Movshon et al., 1986).

The output of local motion detectors in V1 is fed into the Middle Temporal (MT) area and from there to other, adjacent visual areas, e.g. Medial superior temporal (MST) area (Maunsell and Van Essen, 1983; Zeki, 1974; Albright, 1984; Mikami et al., 1986a,b; Dubner and Zeki, 1971). These areas further integrate and analyze the information of multiple motion selective cells and their receptive fields are therefore much larger than those in V1 are. Area MT consists of a large number of directionally selective neurons: They prefer motion in a particular direction (their preferred direction) and responses decrease for directions of motion shifting away from the optimal direction. Typically, the response is lower than spontaneous (i.e. inhibited) for the direction opposite to the preferred direction, which is called the anti-preferred or null direction (Albright, 1984; Britten et al., 1993; Maunsell and Van Essen, 1983; Perge et al., 2005a,b). Inhibition for responses opposite to the preferred direction is called motion opponency and presumably arises from inhibition by directionally selective cells tuned to the opposite direction. MT neurons respond vigorously to a visual pattern moving in a preferred direction, but the responses can be suppressed substantially by superimposing a second pattern moving in a non-preferred direction (Snowden et al., 1991; Bradley et al., 1995; Qian and Anderson, 1994). Heeger et al. (1999) tested for a neuronal correlate of motion opponency in the human visual cortex, using functional magnetic resonance imaging (fMRI). They found strong motion opponency in the human MT complex, but little evidence of motion opponency in primary visual cortex. A large role for area MT in generating motion opponency is in line with the physiological findings concerning the degree of motion opponency in area MT and V1 (Snowden et al., 1991; Bradley et al., 1995; Qian and Anderson, 1994).

Area MT provides direct input to area MST. Cells in dorsal MST preferentially respond to complex motion or optic flow patterns, like expansion, contraction,

clockwise and counterclockwise rotation (Sakata et al., 1985, 1994; Saito et al., 1986; Duffy and Wurtz, 1991a; Orban et al., 1992; Graziano et al., 1994; Lagae et al., 1994). Neurons in ventral MST respond to background movements behind stationary objects or object movements on a stationary background (Sugita and Tanaka, 1991; Tanaka et al., 1993).

Post-processing of local motion information in higher visual areas such as area MT and MST plays an essential role in shaping global motion percepts. Several studies have provided direct links between activity of cells in e.g. area MT and area MST and perceptual direction discrimination or complex motion discrimination (Britten et al., 1992, 1996; Krekelberg et al., 2006b) and complex motion selective cells in area MST (Heuer and Britten, 2004; Britten and Wezel, 1998, 2002; Celebrini and Newsome, 1994).

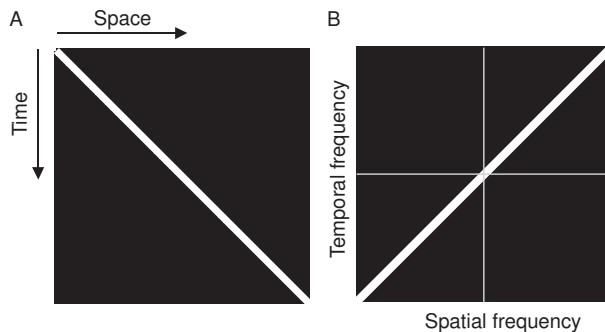
### 1.3 Human versus non-human primates

In order to understand the relationship between brain activity and motion perception one needs to know which visual areas are involved and how they respond to various motion stimuli. Much knowledge of visual areas and their physiological responses underlying motion perception comes from neuroanatomical and neurophysiological studies in monkeys. On the other hand, most knowledge of the perception of motion is inferred from human psychophysical studies. Functional MRI in humans and primates as well as physiological studies in awake behaving monkeys allow one to bridge the gap between human and monkey studies. Numerous studies on awake behaving monkeys have established a close link between neural activity of MT neurons and motion perception (Britten et al., 1992; Britten and Newsome, 1998; Krekelberg et al., 2006b) and a close link in cortical responses to visual motion between human and non-human primates (for review see e.g. Greenlee, 2000). Techniques have been developed that allow direct comparisons between non-human and human functional neuroanatomy (Dubowitz et al., 2001). These studies show that cortical organization and principles in motion processing in human and non-human primates are to a high degree similar. Still, when comparing neural activity in non-human primates with functional MRI in humans, inter-species and inter-technique variability need to be taken into account (Tootell et al., 1996; Heeger et al., 2000; Rees et al., 2000; Ungerleider and Haxby, 1994; Orban et al., 2004; Tootell et al., 1997). Evidence for several apparent species differences has been reported, e.g. in topography in area V2 (Tootell and Taylor, 1995) and in the cortical magnification factor of V1 (Serenio et al., 1995).

### 1.4 Motion in Cartesian and Fourier space

Motion in Cartesian space is represented in the space–time domain. It is also possible to describe motion in Fourier space. Motion is then represented in the

spatial and temporal frequency domain (for a more detailed explanation on motion in Cartesian and Fourier space see e.g. Watson and Pelli, 1983). Both the Cartesian and Fourier space provide complete descriptions of a motion stimulus. One can go back and forth between the two domains without losing any information. Why would one represent motion in Cartesian or Fourier space? It depends on which aspects of motion one wants to study. Some aspects are better understood in the space–time domain and others in frequency domain. Figure 1.6 represents a bar stimulus moving at constant velocity to the right in Cartesian (figure 1.6a) and Fourier space (figure 1.6b). In the space–time domain the moving bar is transformed in a straight line through the origin, with the slope corresponding to the velocity. In spatial and temporal frequency domain the same moving bar is also a straight line through the origin, with a slope of  $\frac{-1}{velocity}$ .

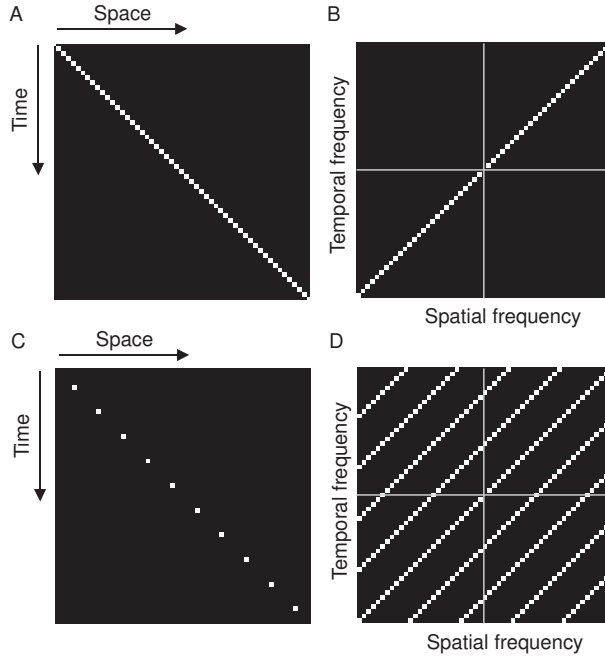


**Figure 1.6:** Representation of a white bar stimulus moving smoothly to the right in Cartesian space (A) and Fourier space (B). In the space-time domain (Cartesian space) the moving bar is transformed in a straight line to the origin. The slope corresponds to the velocity. In the spatial and temporal frequency domain (Fourier space) the velocity of the bar corresponds to  $-1/\text{slope}$ .

### 1.4.1 Apparent motion

In a discotheque with a stroboscope, that gives a light flash every second (i.e. at a frequency of 1 Hz), you perceive stills instead of continuous motion. If the time between flashes is decreased (i.e. increasing the frequency), at a critical frequency the perception of stills switches to a percept of flickering motion. Further decreasing the inter-flash interval finally results in a sensation of smooth continuous motion. The sensation of smooth motion to a physical non-continuous motion stimulus is called apparent motion. Figure 1.7 represents time-sampled version of the moving bar in 1.7. The time between samples in the space-time domain is in Fourier space transformed into a sampling frequency, i.e. the inverse of the time between samples. Sampling at fixed time intervals translates into replicas of the spectrum at intervals of the sampling frequency.

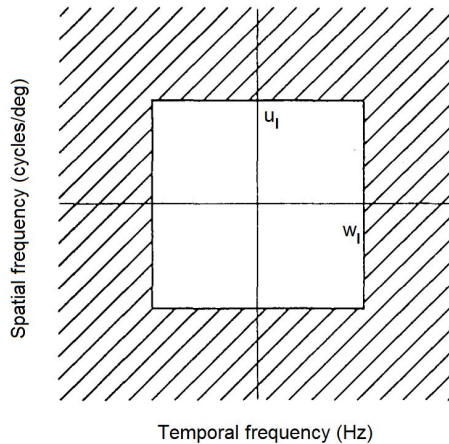
The human motion system can handle spatial and temporal frequencies within



**Figure 1.7:** Representation of two time-sampled versions of the moving bar in figure 1.6 in Cartesian space (A,C) and Fourier space (B,D). The time between samples in the space-time domain is in Fourier space transformed into the sampling frequency, i.e. the inverse of the time between samples. The time-sampling translates into replicas of the spectrum at intervals of the sampling frequency.

a certain frequency range. Spatial or temporal frequencies above a critical spatial or temporal frequency are invisible (Shade, 1956; De Lange, 1954). Moreover, spatial and temporal frequency sensitivity in human has been shown to be relatively independent of one another (Robson, 1966; Koenderink and Van Doorn, 1979). Watson and Pelli (1983) roughly characterized the limits of human visual sensitivity to spatial and temporal frequencies by a rectangular window of visibility (figure 1.8). Combinations of spatial and temporal frequencies within the window are more or less visible, and outside the window are invisible. Moreover, Watson and Pelli (1983) postulated that two stimuli will appear identical if their spectra, after passing through the window of visibility, are identical.

Limits of the human motion system can also be described in the space-time domain (Van Doorn and Koenderink, 1982a,b, 1984; Van de Grind et al., 1986, 1987; Baker and Braddick, 1985; Morgan and Ward, 1980). Morgan and Ward (1980) measured directional motion perception limits for combinations of spatial displacement and temporal delay using random dot patterns. Random dot patterns provide a stimulus containing a wide range of spatial frequencies. The displacement of the patterns and the interval between the displacement selectively



**Figure 1.8:** The window of visibility (Watson and Pelli, 1983) characterizes the limits of human visual sensitivity to spatial and temporal frequencies by a rectangular window. Within this window combinations of spatial and temporal frequencies are more or less visible. Outside this window (shaded area) combinations of spatial and temporal frequencies are invisible. The window is bounded by its limits of spatial ( $u_l$ ) and temporal resolution ( $w_l$ ) (after figure 3 in Watson et al., 1986).

activates motion detectors tuned for a particular span and delay. Morgan and Ward (1980) found that motion was only perceived when the spatial separation was less than 0.2–0.3 degrees of visual angle: the upper temporal limit varied with the state of retinal adaptation between 60 ms (under dim light) and 100 ms (under normal artificial light). Fredericksen et al. (1993, 1994a,b,c) also measured spatiotemporal limits of motion perception and found similar spatial limits, but a much higher temporal limit. In chapter 4, we will discuss this apparent discrepancy.

## 1.5 Modeling low-level motion processing

### 1.5.1 Feature tracking and intensity-based models

Several models have been developed that describe the short and long range motion processes. These models can be grouped in two classes, feature tracking models and intensity-based models. Feature tracking models describe the long-range mechanism and rely on higher order processes (Morrone and Burr, 1988; Del Viva and Morrone, 1998). They extract velocity by determining the time it takes for an object or feature to cover a certain distance. These models imply that the visual system first identifies objects or features and then tracks it in time (for review see Derrington et al., 2004). The intensity-based models describe the short-range motion mechanism and assume low level, front-end processing.



This class of models uses spatio-temporal changes in light intensity to determine direction and speed. Intensity-based models comprise gradient models and correlation models. Gradient models extract stimulus velocity independent of spatial stimulus structure (Marr and Ullman, 1981; Mather, 1990). Correlation models extract motion from spatiotemporal correlations in the stimulus.

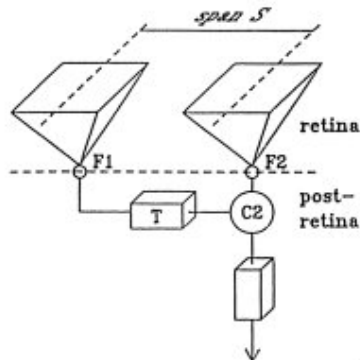
### 1.5.2 Correlation-based models for low-level motion detection

Models of motion processing have emphasized a three-stage process. First, preprocessing of motion information occurs. This stage is important since the properties of pre-filters affect the tuning characteristics of motion detectors in the next stage. Neurons at this motion detection stage detect local motion signals. Neurons at the third stage further process the local motion information to create a global motion percept. The detection of local motion requires the correlation of visual information from at least two distinct locations. In essence, intensity-based models in general, and correlation-based models in particular, use spatio-temporal correlation for local motion detection. In this section I will first introduce the bilocal correlation principle and then two commonly used models for human motion perception.

#### The bilocal correlation principle

The bilocal correlation principle describes the most basic, functional organization of a motion detection unit, tuned to motion with a specific spatial displacement and time lapse. The basic bilocal detector consists of four elements: two receptive fields (F1, F2), a delay (T), and a correlator (C) (figure 1.9). An object moving over the first receptive field (F1) triggers a response that feeds into the correlator with a specified delay T. Meanwhile, the object moves over the second receptive field (F2) and triggers a similar response. If the delayed signal from F1 and the direct signal from F2 reach the correlator (C) at the same time, it will generate an output signal. The origin of the delay in the signal is still uncertain in the literature. Possible candidates are differential temporal filtering in the preprocessing stages in the LGN and in area V1 (see section 1.2.1). The distance between the receptive fields (S or span) and the delay may vary from detector to detector, providing it with a preferred speed. In addition, the relative position (orientation) of the receptive fields creates a preferred direction of motion for each detector. This principle gives good predictions of motion processing from fly to primate (for review see Borst and Egelhaaf, 1989).

Since a flickering stimulus can already activate bilocal motion detectors tuned for different directions, the information from a single, local motion detector is ambiguous. Extracting ecologically useful information therefore requires combining multiple bilocal detectors. Setting two asymmetrical half detectors in opposite virtue of canceling each other out by subtraction, the output becomes selective



**Figure 1.9:** Bilocal motion detector. The detector is tuned to the velocity  $S/T$ , where  $S$  is the span and  $T$  is the delay. A coincidence unit ( $C$ ) receives the input from both receptive fields ( $F1$ ,  $F2$ ), one of which has an additional latency of  $T$  before reaching  $C$ . (after Van de Grind, 1988)

for direction (Borst and Egelhaaf, 1990). This direction selectivity is called motion opponency (see 1.2.3) and it is included in many low level computational theories of visual motion perception (Van Santen and Sperling, 1984, 1985; Qian and Anderson, 1994; Adelson and Bergen, 1985; Simoncelli and Heeger, 1998).

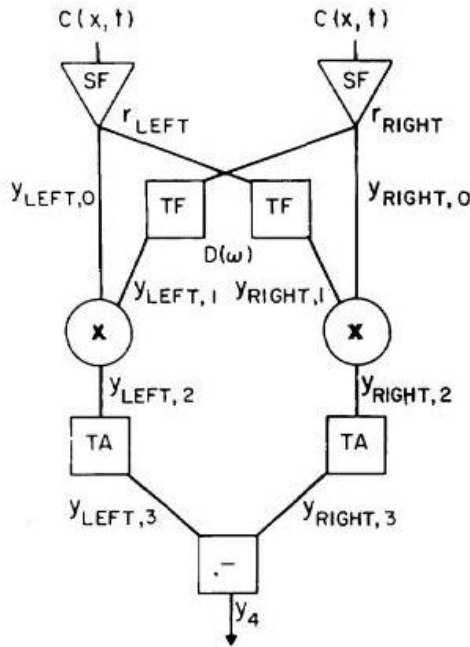
By comparing output signals of an array of detectors with various spans, delays and orientations, the motion perception system can extract visual motion information different speeds and directions. Moreover, the use of a labeled line code within a population of bilocal detection units allows for the simultaneous representation of multiple motion components, as in transparent motion perception (Van Doorn and Koenderink, 1984; Van de Grind et al., 1986; Fredericksen et al., 1993).

The bilocal correlation principle is used in multiple models to describe low-level motion detection in e.g. insects (Reichardt, 1961), rabbits (Barlow and Levick, 1965), cats (Heeger, 1993) and human and non-human primates (Van Santen and Sperling, 1984; Adelson and Bergen, 1985; Watson and Ahumada, 1985; Heeger, 1993). Next, I will introduce two commonly used correlation-based models for motion perception in human and non-human primates.

### The elaborated Reichardt detector

The elaborated Reichardt detector (Van Santen and Sperling, 1984) is a model for human motion perception (figure 1.10). This model is an elaboration on the Reichardt detector that was originally developed for insect vision (Reichardt, 1961). The elaborated Reichardt detector consists of two bilocal motion detectors (also called half detectors). One half detector correlates signals from two neighboring

receptive fields by multiplying the signal from one receptive field by a delayed version of the signal from its neighbor. The delay is a temporal filter. The output is maximal if the delay is equal to the time it takes for an image to move from one receptive field to the other. The other half detector is similar but tuned for motion in the opposite direction. By subtracting the output of both half detectors in the next stage the output becomes selective for direction.



**Figure 1.10:** (a) Elaborated Reichardt model: Visual information ( $C(x, t)$ ) falls on the receptive fields and is spatially filtered. Signals ( $y_{LEFT,0}$ ,  $r_{LEFT}$ ,  $r_{RIGHT}$  and  $y_{LEFT,0}$ ) from two neighboring spatial filters (SF) are correlated (X) by multiplying the signal from one spatial filter ( $y_{LEFT,0}$  and  $y_{RIGHT,0}$ ) by a delayed version of the signal from its neighbor ( $y_{LEFT,1}$  and  $y_{RIGHT,1}$ ). The delay is a temporal filter (TF). The output is maximal if the delay is equal to the time it takes for an image to move from one receptive field to the other. The half detectors are tuned for motion in opposite direction. The next stage subtracts (-) the output of both half detectors (TA) and the output ( $y_4$ ) becomes selective for direction (after Van Santen and Sperling, 1985).

### The motion energy detector

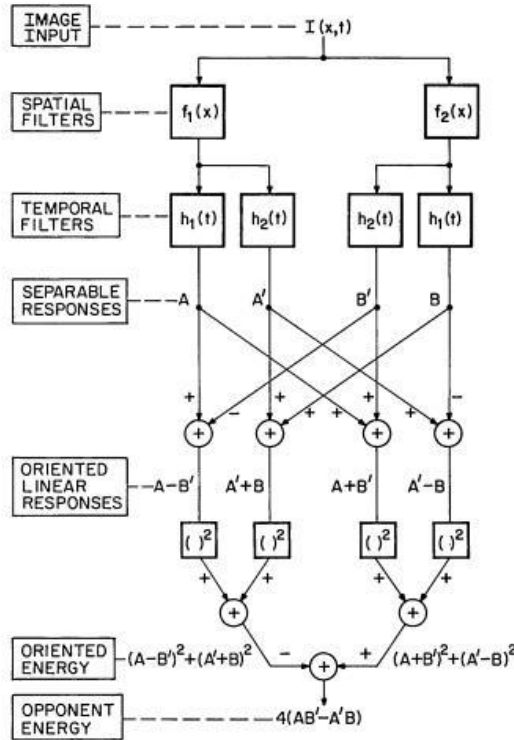
The motion energy detector (Adelson and Bergen, 1985) uses spatio-temporal filters, which are oriented in space-time and therefore match the oriented space-time

structure of moving spatial patterns (figure 1.11). This is accomplished by adding together space-time separable filters. A separable filter is determined by its spatial and temporal profile: The spatial profile has the same shape over time but its sensitivity is scaled by the temporal filter. For each direction two linear filters with a  $90^\circ$  phase difference are generated, one which is symmetric (bar-like) and one which is asymmetric (edge-like). The sum of the squares of these filter outputs is called oriented motion energy. This stage corresponds to a phase-independent unidirectional motion energy detector. Actually, such a detector measures the Fourier energy within its predefined spatiotemporal frequency window. The difference in motion energy for two opposite directions is called the opponent motion energy.

The motion energy detector is also referred to as Fourier energy detector. Motion is represented as an orientation in the space-time domain. Oriented structure in space-time produces oriented structure in Fourier space, where the coordinate bases are now spatial and temporal frequency. The Fourier spectrum of an orientation filter in space-time occupies two regions at opposite ends of a line through the origin. A sensor or receptive field, designed to sample such antisymmetric regions, is sensitive to motion. Therefore, in terms of Fourier energy, the visual system can be constructed by spatio-temporally tuned filters oriented in frequency space. Such filters allow the measurement of power in the oriented Fourier transform of the stimulus (Adelson and Bergen, 1985; Watson and Ahumada, 1985; Heeger, 1987, 1993). Watson and Ahumada (1985) showed the linear sensor fulfills these requirements. It is constructed in two stages. The first stage is an array of scalar linear sensors that are selective for 2D location, 2D spatial frequency and direction. A scalar linear sensor is an assembly of simple spatial and temporal filters. However these sensors are ambiguous with regard to the speed and direction of the image. This ambiguity is resolved in the second stage by combining the response of multiple sensors. This stage is called the vector motion sensor.

### Similarities in correlation-based models

Interestingly, correlation-based models are all highly similar, due to the flexibility of the correlation principle (Borst and Egelhaaf, 1989). Changing the spatial and temporal filter properties enables a variety of different retinal motions to be detected. Van Santen and Sperling (1985) argued that the linear, energy and bilocal correlation detectors are all similar, and that the bilocal correlator can mimic the properties of both linear and energy detector, depending on the spatial and temporal filtering properties in the preprocessing stage. Vice versa, Adelson and Bergen (1985) described a version of the Reichardt model that is formally equivalent to a version of an energy model. Based on currently available psychophysical and physiological data, it is not possible to exclude one of these models (Derrington et al., 2004). Favoring one of these models seems to depend on the stimuli of interest. Traditionally, the Reichardt-based model is adopted in studies on human perception. These studies use stimuli that are described in

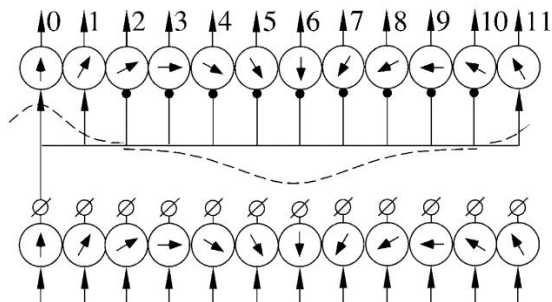


**Figure 1.11:** Motion energy model: The image input passes through two spatial filters ( $f_1$  and  $f_2$ ). Each output then passes through two temporal filters ( $h_1$  and  $h_2$ ).  $h_2$  delays the signal more than  $h_1$ . The combination of one spatial and one temporal filter is called a linear separable spatiotemporal filter that is sensitive to both leftward and rightward motion. The separable responses ( $A$ ,  $A'$ ,  $B$  and  $B'$ ) are combined (by summing or subtracting) in four oriented filters to get oriented linear responses ( $A - B'$ ,  $A' + B$ ,  $A + B'$  and  $A' - B$ ). Two of these filters are sensitive for leftward motion and two for rightward motion. The two members of one pair are about 90 degrees out of phase with each other. Local oriented motion energy is extracted by squaring the oriented linear response and summing the output of a pair. The spatio-temporally oriented energy stage is directionally selective, phase independent and also contrast polarity independent. By taking the difference between the rightward and leftward responses, the response of the motion energy stage also reflects the contrast polarity of the stimulus (after Adelson and Bergen, 1985).

terms of spatial displacement and temporal interval. The motion energy model is frequently used in physiological studies of motion processing. These models are adapted for stimuli that are described in terms of spatial and temporal frequencies.

### 1.5.3 Modeling post-processing of low-level motion

Models of motion processing have emphasized a three-stage process. In this description, the local motion detection stage is followed by a summation and integration stage. This post-processing stage comprises an opponency stage. The opponency stage is described by other models, e.g. the distribution-shift model (Mather, 1980) and Grunewald-Lankheet network model (Grunewald and Lankheet, 1996; Grunewald, 1996) with an automatic gain-control (Van de Grind et al., 2003, 2004). Figure 1.12 illustrates the Grunewald-Lankheet network model. The output is used by a global integration and segmentation stage, whose output is used for higher-level motion processing and finally combined into a motion percept.



**Figure 1.12:** Basic structure of the Grunewald-Lankheet network model with automatic gain control. The bottom row represents the direction-tuned motion sensors, which receive input from contrast-sensing neurons. The arrows in the motion sensors symbolize their tuning direction. The oblique line crossing a small circle represents the gain-control mechanisms. The motion sensors project via these gain-controls to integrator units in the opponency stage. Sensors with highly similar direction preference converge to excite an integrator neuron, whereas broad inhibition (black dots) centers around the opponent direction unit of the integrator layer. Excitatory and inhibitory connections have weights that change in Gaussian fashion with the difference in direction tuning (interrupted line). Notice that the mapping is shown for one direction. Equivalent mapping exists for all other directions (after Van de Grind et al., 2003, 2004).

## 1.6 Evidence for three-stage motion processing

Models of motion processing have emphasized a three-stage process. In the next paragraphs, I will give a short review of three motion phenomena (transparent

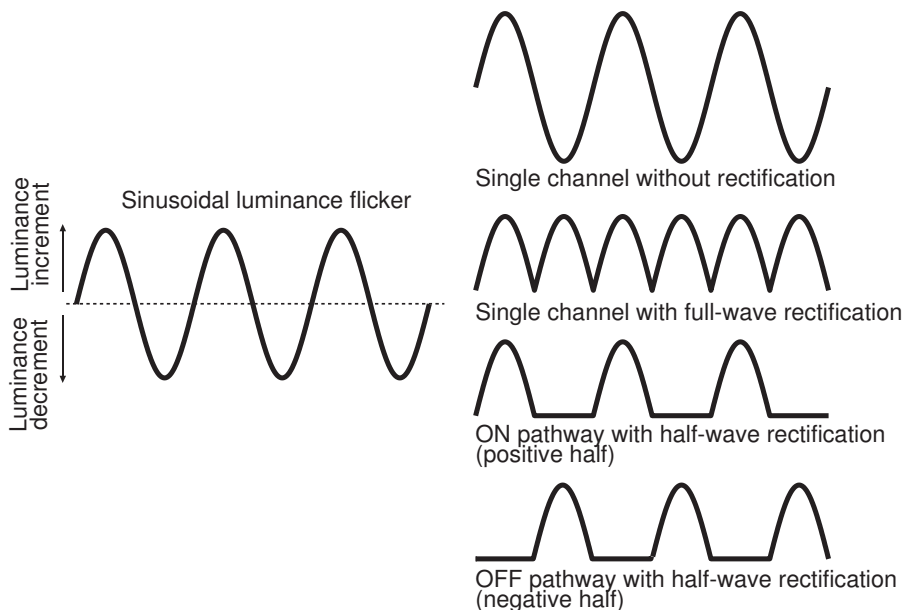
motion, the motion aftereffect and reverse-phi motion) and the role of the ON and OFF pathway, since they help to understand this process and are used in the different chapters of this thesis and in the introduction of possible mechanisms underlying motion detection in section 1.7.

### 1.6.1 The ON and OFF pathway

The ON and OFF pathways play an important role in at least the vertebrate visual pathway. Schiller et al. (1986) argued that encoding information via the ON and OFF pathway is more efficient than using a single channel (figure 1.13). The reason is that retinal ganglion cells have a low spontaneous activity and only limited information can be coded via inhibition. A higher spontaneous activity would lead to a high metabolic cost (Laughlin et al., 1998). The ON and OFF pathways first appear at the bipolar cell level. At the level of the lateral geniculate nucleus (LGN), the ON and OFF distinction is maintained within ON and OFF center cells (Edwards and Badcock, 1994). Schiller (1982, 1984, 1990, 1992) and co-workers (Schiller et al., 1986) postulated that the main functions of the ON and OFF pathways were the detection of light increments (luminance above background luminance) and decrements (luminance below background luminance), respectively, and to provide increased contrast sensitivity. They had administered 2-amino-4-phosphonobutyric acid, or APB, (Slaughter and Miller, 1981) to the primate visual system and found that APB blocks the ON pathway (from ON-bipolar cell level to the ON center cells at LGN level), but did not affect the OFF pathway. Moreover, they found that the ON-surrounds of OFF-center cells were unaffected, which indicated that the center-surround organization of these cells does not originate from interactions between ON and OFF pathways. They concluded that the two pathways converged at the level of V1-complex cells, because these cells responded to light/dark and dark/light edges under control condition, but only to light/dark edges after APB infusion.

In the visual system information of ON and OFF signals is available for the generation of direction-selective responses. The motion energy model (Adelson and Bergen, 1985) and the elaborated Reichardt model (Van Santen and Sperling, 1985) both implicitly incorporate the ON and OFF pathway in their model. Motion sensitivity is generated by the spatio-temporally oriented filters that use both ON and OFF signals. How these signals are combined depends on the filter properties. Filters used in the literature are tuned to detect same-contrast orientation in the spatio-temporal domain (Adelson and Bergen, 1985; Watson and Ahumada, 1985).

The elaborated Reichardt model assumes that correlation occurs within each channel. There are no explicit interactions between the ON and OFF signals: ON and OFF signals from one receptive field are correlated with respectively ON and OFF signals from the other receptive field. In the next stage, the output from each pathway is combined. I will discuss the implications of this scheme on the reverse-phi phenomenon in section 1.6.4.



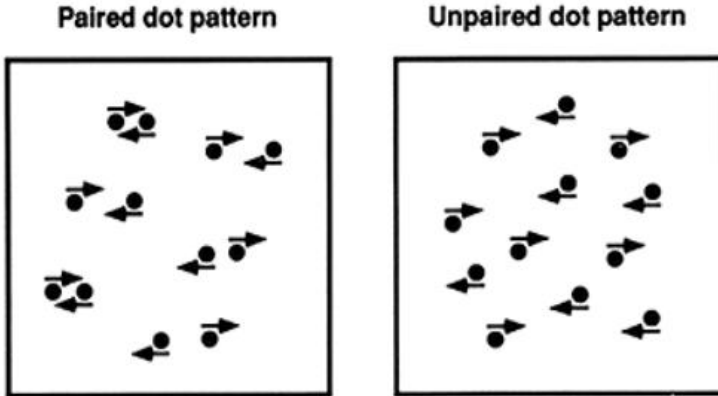
**Figure 1.13:** Schematic representation of four rectification schemes for coding luminance information in the visual pathway. Within a single channel luminance information can be encoded without rectification. This conserves the sign for luminance increments (excitation) and decrements (inhibition). Full-wave rectification allows luminance increments and decrements to pass through a single channel, although the sign is lost since both luminance increments and decrements are encoded as excitations. Half-wave rectification of luminance information in two channels preserves the sign: Luminance increments are transformed in excitations in the ON pathway and luminance decrements in excitations in the OFF pathway.

## 1.6.2 Transparent motion

Our visual system can perceive multi-vectorial motion. Multi-vectorial motion occurs when more than one motion direction and/or speed is present in our visual field. If the components are clearly separated in space or time, it is called segregation. If the components overlap, we speak of transparent motion or motion transparency.

Qian et al. (1994) reported on transparent motion percepts using paired and unpaired dot patterns (figure 1.14). Performing psychophysical experiments, they found that whenever a display had finely balanced opposing motion signals (within  $0.2^\circ$  of visual angle) in all local areas, it was perceptually nontransparent. If a display contained unpaired dots moving in opposing directions or contained locally paired motion signals in different directions that were spatially unbalanced ( $0.2^\circ$  of visual angle or more) then the display was perceptually transparent. Interestingly, if a subject concentrated on a small area in a spatially balanced dot pattern, the





**Figure 1.14:** Paired and unpaired dot patterns: Qian et al. (1994) generated paired random dot patterns that consisted of many randomly located pairs of dots. If the dot were paired with an offset less than  $0.2^\circ$  the display was perceptually nontransparent (referred to as balanced pairs). If the dots in the paired pattern had an offset of  $0.2^\circ$  or more the display was perceptually transparent (referred to as unbalanced pairs). If the paired dots were positioned independently (unpaired) the display was also perceptually transparent (after Qian et al., 1994, figure 2).

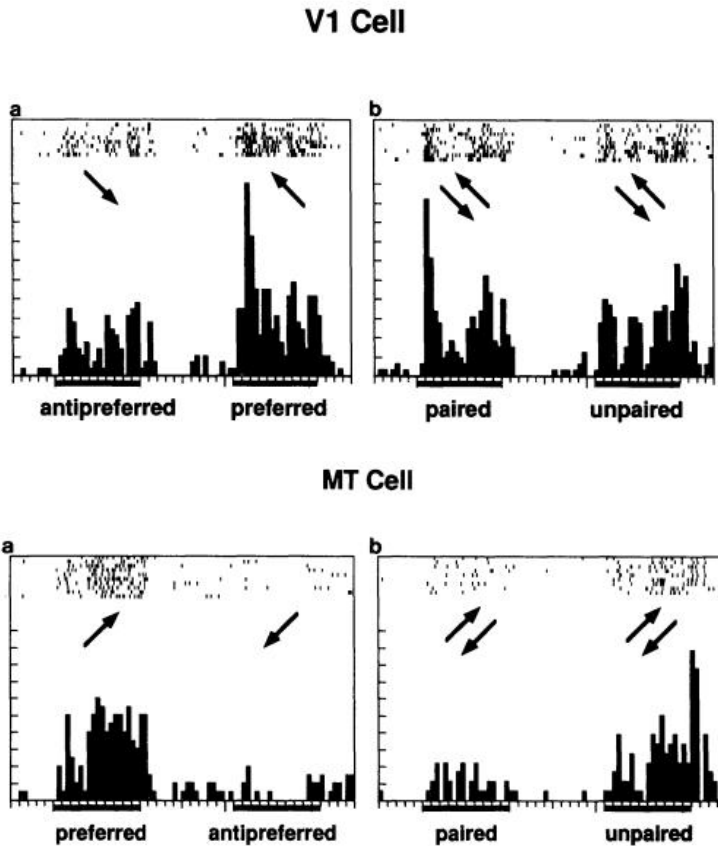
opposed motion of both dots in a pair could be easily seen. This indicates that the process of dot-pairing destroyed the global coherent motion percept.

In another study Qian and Anderson (1994) used physiological recordings from areas V1 and MT of behaving monkeys, comparing single-cell responses to the paired and the unpaired dot patterns (figure 1.14). They reported that the average V1 responses could not reliably distinguish between the finely balanced and unbalanced motion signals. On the other hand, a large fraction of MT cells responded significantly better and more vigorous to the unbalanced motion stimuli than to the balanced ones (figure 1.15).

The results of Qian and Anderson (1994) are consistent with findings by Snowden et al. (1991), who also measured the response of neurons in area MT and V1 to transparent motion. They found that area V1 cells responded to their preferred direction of movement even under transparent conditions, whereas area MT neurons were suppressed under transparent conditions.

These results are consistent with the three-stage model for motion processing: After pre-processing visual information, the second stage measures local motion and the third stage introduces suppression if different directions of motion are present at a local region of the visual field (motion opponency). Qian (1994)'s physiological data indicate that the second stage is located primarily in V1 and the third stage primarily in MT.

1.6 Evidence for three-stage motion processing



**Figure 1.15:** Results from Qian and Anderson (1994), showing the responses of a V1 and a MT cell to single sets of random dots moving in the preferred and anti-preferred direction of the neuron under study, and to paired and unpaired dot patterns with one set of dots moving in the preferred direction and the other in the anti-preferred direction. The raster in the top of each diagram represents the spike records from several trials and each small dot represents the occurrence of a spike. The histograms compiled from the rasters are shown at the bottom of the diagrams. The arrows indicate the directions of motion. Each small division in the horizontal axis represents 10 ms. The 1 s periods during which the stimuli were presented are marked by the thick black lines under the histograms. One small vertical division represents 4.0 spikes/s for the MT cell, and 7.3 spikes/s for the V1 cell (after Qian and Anderson, 1994, figure 1 and 13).

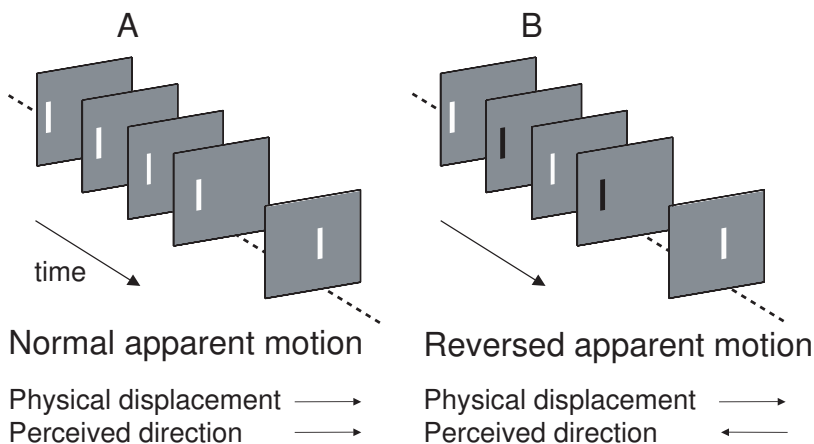
### 1.6.3 Motion aftereffect

After prolonged viewing of a unidirectional moving stimulus, a stationary pattern appears to move in the opposite direction (Addams, 1834). This phenomenon is called the motion aftereffect (MAE) or waterfall illusion and has tickled many people to study the underlying mechanisms (for a review see e.g. Mather et al., 1998). Two classical properties of the MAE that helped to understand the mechanisms underlying human motion detection were its perceived direction and its characteristic nulling behavior (Blake and Hiris, 1993; Castet et al., 2002; Lankheet and Verstraten, 1995): If one adapts to a moving stimulus in one direction, the MAE is perceived to move 180 degrees away (Verstraten et al., 1996b, 1998). However, adaptation to two superimposed (or rapidly interleaved) velocity field components leads to a uni-directional aftereffect. Its direction is opposite to the vector-average of the directions present in the adaptation condition (Mather, 1980; Verstraten et al., 1994), although it strongly depends on the speed difference of the velocity fields and the test condition (Alais et al., 2005; Verstraten et al., 1999). Moreover, Grunewald and Lankheet (1996) showed that adaptation to two superimposed random dot patterns (RDP) moving with the same speed in opposite directions led to a transparent MAE that moved in the orthogonal plane compared to the direction of the motion components during adaptation. The other characteristic of the MAE is that the illusory motion during the aftereffect can be counteracted with ‘real’ motion in the opposite direction, which is known as nulling (e.g. Blake and Hiris, 1993; Castet et al., 2002; Lankheet and Verstraten, 1995). These findings indicated that the motion system integrates motion signals across all directions. Moreover, for nulling to occur, two motion signals have to cancel each other, which has fundamental implications for any model underlying motion detection. It was suggested that the direction of the MAE of transparent motion is a resultant of the weighted summation of the component inducing vectors established with some gain control (see figure 1.12). Verstraten et al. (1994) already suggested that this gain control cannot be located in the individual motion detectors and must be situated at or after some subsequent cooperation stage of the human motion analysis system. As we will demonstrate in Chapter 6, the MAE has several features in common with another visual illusion, namely the reversed motion phenomenon also known as reverse-phi illusion.

### 1.6.4 Reverse-phi

Contrast reversal with apparent motion leads to the perception of motion opposite to physical displacement of the stimulus. Anstis (1970) first described this phenomenon, which he termed reverse-phi motion. The terminology on phi motion and apparent motion in the past century has been confusing (for review: see Steinman, 2000). In this thesis phi motion is considered as a form of apparent motion (Boring, 1942). Moreover, since contrast reversal with motion in general leads to a percept of motion in the opposite direction, we will also use the term reversed apparent motion. Figure 1.16 shows a classical cartoon illustrating the

reversed apparent motion illusion: A simple pattern of white bars on a gray background is repetitively shifted rightward. Although the physical displacement is spatially and temporally discrete, the percept is smooth, continuous rightward moving bars. This phenomenon is known as apparent motion. If the bar changes contrast (white to black and black to white) each time it is displaced, motion opposite to the physical direction of motion is perceived. Since contrast reversal leads to an opposing motion percept, this phenomenon is useful to understand fundamental principles of motion sensors and to reveal ON and OFF cell contributions to visual motion detection.



**Figure 1.16:** Illustration of normal apparent motion and reversed apparent motion of bar. A. Normal apparent motion: a white bar on a gray background is repetitively shifted rightward. Although the physical displacement is spatially and temporally discrete, the percept is smooth, continuous rightward moving bars. B. Reversed apparent motion: If the bar changes contrast (white to black and black to white) each time it is displaced, motion in the opposite direction than the physical direction is perceived, although the physical displacement is still repetitively shifted rightward.

Many studies have examined the reversed apparent motion phenomenon in psychophysical and/or physiological settings and implemented the findings in motion detection models (Adelson and Bergen, 1985; Anstis, 1970; Anstis and Rogers, 1975; Borst and Egelhaaf, 1989; Chubb and Sperling, 1989; Conway and Livingstone, 2003; Dobkins and Albright, 1994; Emerson et al., 1987; Ibbotson and Clifford, 2001; Krekelberg and Albright, 2005; Livingstone and Conway, 2003; Livingstone et al., 2001; Mo and Koch, 2003; Sato, 1989). In addition to Braddick (1974) who showed that motion processing can be divided in a short-range and a long-range motion mechanism, Chubb and Sperling (1989) showed that the reversed apparent motion illusion only involved the short-range motion system. This finding is exemplified by the classical illustration in figure 1.16. Since, the location

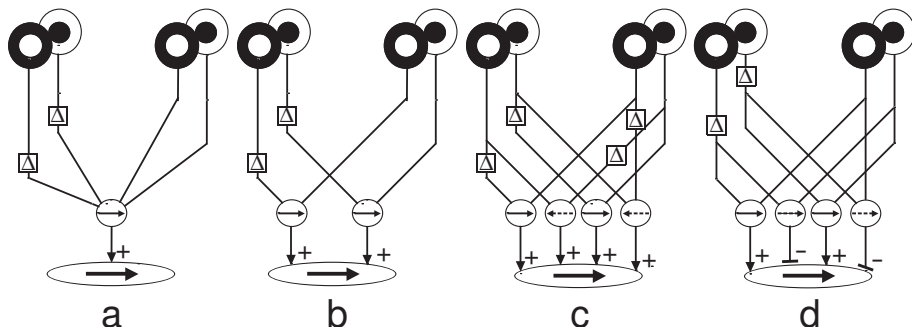
of the contrast borders behave identical under the normal and contrast reversal condition, this observation provides evidence that motion perception does not rely on higher-order processes (e.g. feature tracking). Accordingly, Sato (1989) found that the maximum spatial displacement within which motion discrimination is possible (referred to as  $D_{max}$ ) was comparable for normal and reversed apparent motion and suggested that the same short-range processes underlie normal and reversed apparent motion.

Low-level motion models that use oriented space-time filters already predict the change in perceived motion direction due to contrast reversal. These filters detect specific components of spatio-temporal Fourier or motion energy in a scene. Which frequencies of the spatio-temporal spectrum of the stimulus are detected depends on the spatio-temporal properties of the motion detectors (e.g. Adelson and Bergen, 1985; Van Santen and Sperling, 1985). This can be illustrated by space-time plots of reverse-phi motion that show energy in the reverse direction (see e.g. figure 5.1 on page 75). Although the direction of the moving stimulus is to the right, the leftward motion energy unit aligns better with the stimuli and extracts more motion energy than the right motion energy unit. These models imply that reverse-phi motion simply reflect stimulus properties. However, Mo and Koch (2003) argued that this reasoning could only partly account for the observed reverse-phi motion effect. They suggested that given the nature of reverse-phi motion, a direct interaction between the ON and OFF pathway is essential to account for the reversal of response. They proposed a double synaptic-veto mechanism to model the interaction between the ON and OFF pathway. In the next section, I will introduce this mechanism.

## 1.7 Mechanisms underlying motion detection

Regardless of their origin, intensity-based models incorporate a correlation step in which signals separated in space and time are combined and an opponent stage in which the activities of detectors with identical spatio-temporal tuning properties for opposite directions are evaluated. Current models have different configurations for combining these stages. Chubb and Sperling (1988, 1989) distinguished two ways of pre-processing before the actual motion detection stage: Half-wave and full-wave rectification (see figure 1.13).

In full-wave rectification both ON- and OFF-center receptive fields connect to the same correlator unit (figure 1.17a). The correlation unit of a Reichardt-like detector performs a multiplication operation. Since, the ON pathway is coded as a positive signal and the OFF pathway as negative, the sign of the output after the multiplication operation at the correlation step determines your percept: A positive response indicates motion in the direction of the physical displacement, whereas a negative sign in the opposite direction. However, it is debatable whether such a four-quadrant multiplication operation is achievable by a single neuron or neural network in the visual system (pro: Borst (2000), contra: Mo and Koch



**Figure 1.17:** Four possible configurations for a correlation-based mechanism underlying motion detection. (a) Full-wave rectification. The ON- and OFF receptive fields converge. The ON- and OFF-center fields connect to a single correlator unit. (b) Half-wave rectification without interaction between ON and OFF pathway. (c) and (d) Half-wave rectification with interaction between the ON- and OFF-pathway. The difference is the sign of the input from the correlator unit to the next stage.

(2003)).

In half-wave rectification the ON- and OFF pathways are separated and the ON- and OFF-center fields connect to a different correlator unit (figure 1.17b). Most correlation-based models (e.g. shunting inhibition model (Barlow and Levick, 1965), the motion energy detector (Adelson and Bergen, 1985), linear detector (Watson and Ahumada, 1985), normalization model (Heeger, 1993), Reichardt detector (Reichardt, 1961), and the elaborated Reichardt detector (Van Santen and Sperling, 1984)) are elusive on the role of ON and OFF pathways. The traditional view assumed separate ON and OFF pathways and accordingly, only the presence of same-contrast spatio-temporal correlated receptive fields (figure 1.17b). Consequently, this suggests that contrast reversal of a moving stimulus leads to activation of motion detectors tuned to motion in the opposite direction, since motion detectors are in fact spatio-temporal filters that respond to spatial and temporal frequency.

However, in 2003 Mo and Koch proposed the reversed-shunting-inhibition model, that explicitly assumes a direct interaction between ON and OFF pathways. Figure 1.17c shows a motion detector configuration equivalent to the one proposed by Mo and Koch (2003). Based on extensive simulations they suggested that the visual system should incorporate specific motion detectors sensitive to contrast reversing motion components. Accordingly, they implemented a motion detection stage that combines same-contrast and contrast-reversed spatio-temporal correlated receptive fields. In their reversed-shunting-inhibition model four types of correlations occur, namely two same-contrast correlations (ON-ON and OFF-OFF) in the preferred direction and two opposite-contrast correlations (ON-OFF and OFF-ON) in the direction opposite to the preferred direction of the same-contrast correlators. The output of these four correlators feed with a

positive sign to a summation stage. Consequently, reverse-phi motion leads to an excitation at the summation stage in the direction opposite to the physical motion, finally resulting in a perceptual directional reversal. Interestingly, the proposal of Mo and Koch raises another possible configuration (figure 1.17d). In this configuration the direction for opposite contrast combinations is the same as the same-contrast correlation but the sign is negative. In this case, a reverse-phi motion leads to an inhibition at the summation stage and by means of opponency at a higher level results in directional reversal. In chapter 6 we address the question which configuration is present in the local motion detection stage.





# Chapter 2

## Overview, summary and conclusions

This dissertation is about the mechanisms underlying low-level motion detection in the retino-geniculo-cortical pathway of human and non-human primates. This pathway plays an essential role in global motion perception. Numerous studies have been done on motion processing that led to the consensus that motion processing is done by front-end processing (e.g. Hubel and Wiesel, 1959, 1962, 1963; Tolhurst, 1973; Tolhurst and Movshon, 1975; Movshon et al., 1978a,b,c; Zeki, 1974; Essen and Zeki, 1978; Albright, 1984; Maunsell and Van Essen, 1983; Mikami et al., 1986b). Within this pathway, the first motion sensitive neurons are located in the striate or primary visual cortex. These neurons detect local motion signals and have projections to neurons in the Middle Temporal (MT) area and other extra-striate areas, e.g. the Medial Superior Temporal area. Area MT cells process the information of multiple local motion detectors and are involved in global motion detection (e.g. Movshon et al., 1986; Born and Bradley, 2005; Albright, 1984; Britten et al., 1993; Maunsell and Van Essen, 1983; Perge et al., 2005a,b).

## Chapter 3

This dissertation started out to physiologically study temporal aspects of the cortical mechanism underlying motion perception. Following a series of psychophysical (Fredericksen et al., 1993, 1994a,b,c; Van Doorn and Koenderink, 1982a,b, 1984; Van de Grind et al., 1986, 1987; Baker and Braddick, 1985; Morgan and Ward, 1980) and physiological experiments (Borghuis et al., 2003; Perge et al., 2005a,b), we used similar visual stimuli (random dot patterns) to study the physiology of

---

motion processing. Perge et al. (2005a,b) used a reverse-correlation technique (Borghuis et al., 2003) to examine biphasic temporal responses, center-surround interactions, and direction tuning of neurons in area MT. The first empirical study (chapter 3) is a natural extension into the domain of speed tuning. We used a similar technique for measuring the temporal dynamics and tuning of MT neurons for displacements of random dot stimuli. We compared responses to single displacements as a function of step size (ranging from 2.4 to 76.8 min of arc) in preferred and non-preferred directions, and at different inter stimulus intervals (ranging from 8.3 to 66.7 ms).

We found that responses were generally biphasic, with an average peak latency of about 54 ms, both in the preferred and non-preferred direction. Temporal dynamics were largely independent of step size. The most compelling finding was that although optimal step sizes were similar in preferred and non-preferred directions, step size tuning in the non-preferred direction was wider than in the preferred direction. Increasing the inter stimulus interval had remarkably little effect. Optimal step sizes and the widths of step size tuning did not vary consistently with changes in temporal interval. Response latencies and biphasic behavior were also largely independent of inter stimulus interval.

For stimulus intervals longer than one monitor frame, the stimulus remained on the monitor throughout the interval. This principle was also used in studies that psychophysically examined spatio-temporal properties of the human motion system (e.g. Fredericksen et al., 1993, 1994a,b,c). While interpreting our results and comparing them to previous psychophysical findings, we realized some shortcomings of the stimulus paradigm for measuring the temporal properties. Because we found no effect of inter-stimulus interval in our physiological measurements we wondered whether this might be due to the continuous presence of the stationary stimulus during the inter stimulus interval. Solving this question was the trigger for the study described in chapter 4.

## Chapter 4

Detection of apparent motion in random dot patterns requires correlation across time and space. In chapter 4 we introduce a new motion stimulus that allowed us to measure both the spatial and the temporal tuning properties of human motion perception in more detail. Motion processing involves three stages: Pre-filtering of visual information, local motion detection and post-processing at the next stages. Local motion detection involves a correlation step. Our stimulus allowed us to specifically study the tuning for temporal interval in the correlation step, while keeping the pre-filtering and post-integration constant. Morgan and Ward (1980) used a similar stimulus principle to measure spatial and temporal limits, but their stimulus paradigm was not suited to measure spatio-temporal tuning characteristics.

Our stimulus consisted of a sparse random dot pattern in which each dot

was presented in two frames only, separated by a specified interval. On each frame, half of the dots were refreshed and the other half was a displaced reincarnation of the pattern generated one or several frames earlier. Motion energy statistics in this stimulus do not vary from frame to frame, and the directional bias in spatio-temporal correlations is similar for different interval settings. We measured coherence thresholds for left-right direction discrimination by varying motion coherence levels in a Quest staircase procedure, as a function of both step size and interval. Results show that highest sensitivity was found for an interval of 17-42 ms, irrespective of viewing distance. The fall-off at longer intervals was comparable with the study of Morgan and Ward (1980), but were much sharper than described in other studies (Fredericksen et al., 1993, 1994a,b,c). We argue that this difference in fall-off can be explained by the different stimulus paradigms. Morgan and Ward (1980) found that based on upper and lower limits, step size and interval tuning were independent. Our results show that tuning for temporal interval was largely, but not completely, independent of step size. The optimal temporal interval slightly decreased with increasing step size. Similarly, the optimal step size decreased with increasing temporal interval.

## Chapter 5

In the physiological study in chapter 3 we encountered some cells that showed a large positive response modulation for steps in the non-preferred direction at a latency slightly larger than the optimal latency for the cell. We postulated that these neurons might be susceptible to a motion illusion, called reverse-phi, inherently present in the stimuli that we used. This finding was the trigger for a new series of psychophysical studies.

Low-level contrast information in the primary visual pathway is represented in two different channels. ON-center cells signal positive contrasts and OFF-center cells signal negative contrasts. In this study, we address the question whether initial motion analysis is performed separately in these two channels, or also through combination of signals from ON and OFF cells. We quantitatively compared motion coherence detection for regular and for reverse-phi motion stimuli. In reverse-phi motion the contrast of a pattern flips during displacements. Sensitivity may therefore be based on correlating positive and negative contrasts, whereas for regular motion it is based on correlating similar contrasts.

A simple modification to the stimulus paradigm used in chapter 4, enabled us to induce the reverse-phi percept. Visual stimuli consisted of sparse random dot patterns in which consistent motion information was either restricted to equal polarity correlations (regular motion) or to opposite polarity correlations (reverse-phi motion). Dots were black or white on a gray background. On each frame, the number of black and white dots was equal. Motion information was limited to a single combination of step size and interval.

This allowed us to quantitatively measure spatial and temporal tuning proper-

---

ties for reverse-phi motion and compare them with the spatio-temporal characteristics of regular motion. We found that tuning for step size and temporal interval was highly similar for the two types of motion. Moreover, minimal coherence thresholds for both types of motion matched quantitatively, irrespective of dot density. To test if the absence of motion information was sufficient to induce a reverse-phi percept and explain its sensitivity, we also measured sensitivity for so-called no-phi motion stimuli. In a no-phi stimulus, the contrast of displaced dots was set equal to the background luminance. Sensitivity for no-phi motion was low for stimuli containing only black or only white dots. When both dot polarities were present in the stimulus, sensitivity was completely absent. Thus, motion information based on separate contrasts was effectively canceled by a component based on different contrasts. The low sensitivity for stimuli containing only black or only white dots may be induced because the overall luminance of the stimulus is lower (black dots) or higher (white dots) than the background luminance. If both polarities were present, the overall luminance of the stimulus was equal to the background luminance. Together these results show that the human motion system is as sensitive for regular motion as it is for contrast-reversing motion. This strongly suggests that efficient detection of correlation across ON and OFF channels occurs. This postulation naturally led to the question how these signals from the ON and OFF channels are combined.

## Chapter 6

Reverse-phi motion has generally been explained with the notion that flipping contrasts actually shifts the balance of motion energy towards the opposite direction. In this sense, the reversal is trivial because any suitable motion energy detector would be optimally excited in a direction opposite to that for regular motion. This notion, however, does not explain how these two types of motion are initially detected. Mo and Koch (2003) already argued that reverse-phi involves an interaction between the ON and OFF pathway. The findings in chapter 5 (i.e. the high similarity in sensitivity for regular and reverse-phi motion and the insensitivity for no-phi stimuli) also strongly suggest that efficient detection of correlation across ON and OFF channels occurs.

In this psychophysical study, we address the question how the signals from the ON and OFF channels are combined. The first level at which the ON and OFF pathway converge is in the primary visual cortex (Slaughter and Miller, 1981). In the retino-geniculo-cortical pathway, the primary visual cortex contains the first motion sensitive cells (e.g. Hubel and Wiesel, 1959, 1962, 1965, 1968; Zeki, 1974; Essen and Zeki, 1978; Albright, 1984; Maunsell and Van Essen, 1983; Mikami et al., 1986b). Mo and Koch (2003) suggested that at this level opposite-contrast correlations lead to activation of motion detectors tuned for motion in the opposite direction. In this sense, reverse-phi motion should behave similarly as regular motion. However, an alternative is possible, namely that opposite-contrast correlations lead to reduced activity of motion detectors tuned for motion in the

same direction. In this study, we show that low-level detection of for regular and reverse-phi motion is quite different. The evidence is based on a new set of perceptual phenomena showing that reverse-phi percepts in many respects behave like motion aftereffects, and not like regular motion. Motion adaptation causes reduced activity of motion detectors during a stationary test stimulus, which by means of directional opponency in the next stage leads to a percept of motion in the opposite direction. Our results show that reverse-phi motion similarly reduces the activity of low-level motion detectors. In other words, correlations between equal contrast polarities provide positive evidence whereas correlations between opposite polarities provide negative evidence for the same motion. This mechanism is beneficial for the visual system, because weighing positive and negative evidence for the same motion at the first detection stage efficiently improves signal-to-noise ratios.

## **Chapter 7**

In the general discussion, the research chapters are integrated in a broader context. First, the principles of our new stimulus are discussed. Next, an overall model is presented that combines the essence of several models that have been used in the literature to describe the three stages of motion processing. The findings of the different chapter are linked to this model. The chapter ends with concluding remarks and future perspectives. The future perspectives contain suggestions for further research and some preliminary results that can support our conclusions.



# Step size tuning in macaque area MT

Roger J.E. Bours, Richard J.A. van Wezel and Martin J.M. Lankheet

## 3.1 Abstract

We investigated the temporal properties of directional selectivity in macaque Middle Temporal (MT) cortex. We compared responses to single displacements as a function of step size in preferred and non-preferred directions, and at different inter stimulus intervals. To study response dynamics at high temporal resolution we used a reverse correlation paradigm in which random pixel arrays were displaced according to a random sequence of step sizes. Temporal response profiles were obtained by correlating the occurrence of specific stimuli and spikes. In a single experiment we measured responses to step sizes (ranging from 2.4 to 76.8 min of arc) in the preferred and non-preferred direction, at a fixed inter stimulus interval. These measurements were then repeated for different interval values (ranging from 8.3 to 66.7 ms).

Responses were generally biphasic, with an average peak latency of about 54 ms, both in the preferred and non-preferred direction. Temporal dynamics were to a large extent independent of step size. Optimal step sizes were similar in preferred and non-preferred directions, but step size tuning in the non-preferred direction was wider than in the preferred direction. Increasing the inter stimulus interval had remarkably little effect. Optimal step sizes and the widths of step size tuning did not vary consistently with changes in interval. Response latencies and biphasic behavior were also largely independent of inter stimulus interval. We conclude that tuning for step size in MT, including temporal dynamics, does

not vary systematically with inter stimulus interval.

## 3.2 Introduction

Neurons in macaque Middle Temporal (MT) cortex play an important role in visual motion perception (Born and Bradley, 2005). Perceptual motion direction discrimination has been clearly linked to the activation of directionally selective cells in MT (Britten et al., 1992, 1996). In addition to their direction tuning, most MT neurons are also tuned to the speed of image motion (Maunsell and Van Essen, 1983; Felleman and Kaas, 1984; Rodman and Albright, 1987; Lagae et al., 1993; Liu and Newsome, 2003; Priebe et al., 2003; Nover et al., 2005). Neurophysiological recording and micro-stimulation experiments in monkeys suggest that MT neurons play a role in speed discrimination (Liu and Newsome, 2005). Recent adaptation experiments by Krekelberg et al. (2006a,b) in MT support this hypothesis. Adaptation to a moving stimulus reduced responsiveness, and sharpened speed-tuning curves, which was in line with the reduction in perceived speed and improvement of speed discrimination observed behaviorally.

Despite the presumed link between speed tuning in MT and speed perception, the representation of speed in MT remains far from clear. A basic question remains whether MT neurons are genuinely speed tuned, irrespective of the spatial and temporal content of motion stimuli. To what extent is the speed tuning curve invariant for changes in spatial and temporal parameters? The answer to this question is directly relevant for the way speed can be extracted from the population responses of MT cells, and thus for linking MT activity to perceived speed.

One obvious way to answer this question is by measuring directional selectivity as a function of both spatial and temporal frequency (Tolhurst and Movshon, 1975). With sinusoidal gratings spatial frequency tuning curves can be determined for different temporal frequencies, and vice versa. Invariant speed tuning then shows up as a co-variation of spatial frequency tuning and temporal frequency tuning. Simple cells in primary visual cortex (V1) generally do not show invariant speed tuning, but instead show separable spatial and temporal tuning profiles (Tolhurst and Movshon, 1975; Foster et al., 1985; Priebe et al., 2006). V1 complex cells, however, show different degrees of spatio-temporal covariation, in a way that makes their tuning more or less speed-invariant (Priebe et al., 2006). A minority of complex cells is genuinely speed tuned, showing the same speed preference for all spatial and temporal frequencies. Cells in MT very much resemble the properties of V1 complex cells. There is a wide variation in the degree of spatio-temporal separability and a minority shows speed tuning independent of spatial stimulus properties (Perrone and Thiele, 2001; Priebe et al., 2006).

Variations in preferred speed with spatial frequency in MT become smaller when stimuli are composed of multiple spatial frequency components (Priebe et al., 2006). Furthermore, square wave gratings yield speed tuning that is nearly



form-invariant. These results led to the suggestion that MT neurons derive form-invariant speed tuning from combinations of multiple spatial frequencies in natural scenes (Priebe et al., 2006). This suggestion assumes that all motion vision is based on low-level motion detectors that respond to a narrow range of spatial frequencies. In these detectors spatial and temporal tuning interact to generate directional selectivity. The problem of form-invariant velocity tuning is then solved at the next processing level by combining multiple motion signals. In such a scheme the emphasis is on a covariation of spatial and temporal frequency tuning for stimuli that consist of different Fourier components. This seems a viable solution and the first stages correspond fairly well to properties of V1 simple cells. Some simple cells behave like separable linear filters, whereas others show the spatio-temporal interactions required for directional selectivity. In all cases however, simple cells confound spatial and temporal frequency effects. If speed tuning is based on directionally selective simple cells then additional processing would be required to generate form invariant speed tuning.

It should be noticed though, that the covariation of spatial and temporal frequency tuning of narrowly tuned filters is not necessary for speed-selectivity. In natural scenes moving patterns mostly comprise a broad range of spatial frequencies. Purely sinusoidal patterns rarely occur (Field, 1987; Dong and Atick, 1995). For such stimuli, form-invariant velocity tuning could be based on spatio-temporal correlations of local contrast information irrespective of its frequency spectrum. Rather than generating motion sensitivity from narrowly tuned, oriented spatio-temporal filters, motion sensitivity might arise from pair wise correlations of non-directionally selective sub cortical cells (Reichardt, 1961; Adelson and Bergen, 1985; Borst and Egelhaaf, 1989). Form-invariant velocity tuning is then trivial and simply amounts to a fixed combination of spatial displacement and temporal delay, irrespective of the spatial frequency sensitivity of the input units.

If displacement rather than spatial frequency is the parameter of primary interest the use of broadband stimuli such as random dots seems appropriate, and manipulations of step size and temporal interval may resolve the essential tuning properties. Speed tuning in such a system might come about in two different ways. First, motion sensitive neurons may be narrowly tuned for both step size and temporal interval. Speed would then simply follow from the pattern of activation across a population of narrowly tuned cells. Alternatively tuning for step size and temporal interval may covary, with a preference for combinations corresponding to the same speed. Notice that this question is orthogonal to the question of spatio-temporal (in)separability pursued in sinusoidal grating experiments. Variations in step size of a random dot pattern are equivalent to different temporal frequencies for different spatial frequency components. The covariation of spatial and temporal frequencies is thus built into the stimulus, which is then used to study the width of tuning in the space and time domain.

Several groups have studied responses of MT to spatially broadband stimuli (e.g. moving random dots or bars) to characterize changes of spatial displacement tuning at different temporal delays (Mikami et al., 1986a,b; Churchland and Lis-

berger, 2001; Churchland et al., 2005). In a recent paper Pack et al. (2006) used reverse correlation in combination with sparse random dot displays to measure step size tuning at different temporal delays. They used pairs of black and white dots that were briefly flashed at a high rate. Directional sensitivity was then described by the second order interactions correlated with spike activity. Almost all MT neurons were tuned separable to spatial displacement and temporal delay. These authors, however, attributed the lack of velocity tuning to the low efficacy of their stimuli to drive MT cells.

To investigate whether MT cells are tuned to single combinations of displacement and interval, or to a range of combinations that corresponds to a single velocity, we used a different reverse correlation technique (Borghuis et al., 2003). Rather than using sparsely flashed dots and a second order analysis we used a dense random dot pattern performing a random walk. Direction and speed selectivity then follows from the correlation of steps with spike activity. We previously used this method to characterize the dynamics of direction tuning as well as the dynamics of center-surround interactions in area MT (Perge et al., 2005a,b). Here we use a similar method to study the dynamics of step size tuning for random dot motion in greater detail. The advantage of this stimulus is its high power to drive the cell, resulting in much lower noise levels in the correlation functions. If separable tuning for step size and interval would be an artifact of low contrast or low efficacy we should find clear covariations in tuning with our method. Because of the high signal to noise ratio of our measurements we were also able to measure over broader ranges of parameter variations, which would also favor inseparable tuning. Finally, our method allows for a more detailed analysis of the temporal dynamics of step size tuning, and its changes with temporal interval, step size and directional preference.

## 3.3 Methods

### 3.3.1 Subjects

We recorded motion sensitive single units in area MT in monkeys performing a fixation task. Two male rhesus monkeys (Macaca Mulatta; MA (10 kg, 5 yrs) and MS (5 kg, 3 yrs)) participated in this study. Both were experienced in fixation tasks. Group housing and experimental procedures were all in accordance with the guidelines of the Law on Animal Research of the Netherlands, European guidelines and regulations and approved by the Animal Care and Use Committee (DEC) of the Utrecht University.

### 3.3.2 Surgical preparation, neuronal recording procedures and data acquisition

To prepare monkeys for recordings they were surgically implanted with a head holding device, a golden search coil for measuring eye movements using the double induction technique (Reulen and Bakker, 1982; Malpeli, 1998), and a stainless steel recording cylinder (Crist Instrument Co, Damascus, MD) placed over a craniotomy above the left occipital lobe. The surgical procedures were performed under N<sub>2</sub>O/O<sub>2</sub> anesthesia supplemented with isoflurane. The cylinder allowed electrode access for recording in middle temporal cortex (area MT) of the left hemisphere.

A plastic grid inside the recording cylinder provided a coordinate system with guide tube support holes at 1 mm intervals (Crist Instruments Co, Damascus, MD). Guide-tubes were inserted transdurally. Parylene-insulated tungsten micro-electrodes (Microprobe Inc., Potomac, MD) were inserted through the guide tube and were driven with a micro-positioning controller (National Aperture Inc.).

Area MT was localized using both anatomical and physiological landmarks. Anatomical landmarks included recording depth and the transitions between active gray matter and ‘silent’ areas marking white matter or sulci. Physiological landmarks included brisk, direction selective responses, retinotopy, receptive field size according to eccentricity, and columnar organization for preferred direction (Zeki, 1974; Maunsell and Van Essen, 1983; Albright, 1984). In one monkey we confirmed the location of the recording site with a structural MRI scan. We have no histological confirmation of the recording sites.

Once MT was localized, activity was recorded and single units were isolated using standard extra-cellular methods. Signals were amplified (Bak Electronics Inc.), filtered (Hum Bug, Quest Scientific Instruments Inc; Krohn-Hite 3362, bandpass 0.75-2 kHz) and analyzed using a spike sorting system (Alpha Omega). Spike times were measured at 0.5 ms resolution (NI-DAQ PCI 1200, National Instruments) using a Macintosh G4 computer which was set up to analyze data on-line and store all relevant data for off-line analysis.

### 3.3.3 Visual stimuli

We used moving, binary random pixel arrays (RPAs) with 50% of the elements black and 50% white (Julesz, 1971) at high contrast (>99%) and a mean luminance of 37 cd/m<sup>2</sup> (figure 3.1A). The patterns had an unlimited dot lifetime. Stimuli were presented on a 21”CRT computer monitor (Sony Trinitron Multiscan 500 PS) at a resolution of 1024x768 pixels. The monitor, controlled by an ATI Rage graphic card in the G4 computer, was set to a refresh rate of 120 Hz. At the viewing distance of 57 cm, a monitor pixel subtended 0.04x0.04 degrees and the monitor screen subtended 39x29 degrees. Single elements in the RPA were 6x6 pixels, corresponding to 0.24x0.24 degrees. The stimulus covered an area of 16x16 degrees (400 x 400 monitor pixels), which was large enough to fully cover the

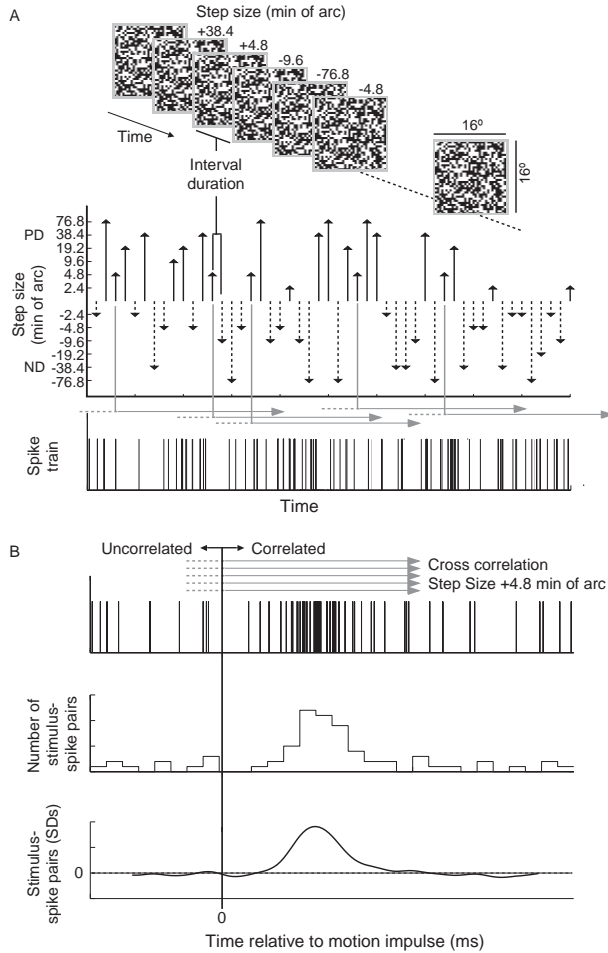
receptive fields.

Subjects sat in a custom-made primate chair and viewed the stimuli while maintaining fixation on a small red square (0.12x0.12 degrees). Stimuli were presented as long as the monkey kept fixation within a 1-degree radius fixation window. The monkey received a liquid reward for maintaining fixation each 2-3 seconds. Breaking fixation led to pausing the presentation of stimuli and lack of reward. Stimulus presentation was restarted after 300 ms of correct fixation.

#### 3.3.4 Experimental procedure

Receptive fields of isolated neurons were hand-mapped and RPAs were centered on the receptive field. We used a reverse correlation paradigm to first measure direction tuning and determine the preferred and non-preferred direction of the neuron. The procedure for measuring direction tuning has been described earlier (Borghuis et al., 2003; Perge et al., 2005a,b). In summary, a dot pattern was shifted each monitor frame (8.3 ms @ 120 Hz). The step size in the direction tuning experiments was always 7.2 min of arc. For oblique directions step sizes were rounded to the nearest pixel boundary. The direction of displacements varied from 0 to 315 degrees, in steps of 45 degrees, with 0 and 90 degrees corresponding to a leftward and upward shift, respectively. Directions were presented in a random sequence, with 1250 or 2500 repetitions for each direction. Stimulus generation, presentation and on-line analysis were performed using a motion reverse correlation method, as described in Borghuis et al. (2003). The preferred direction is defined as the direction with the highest correlation to the spike train at short latencies. The anti-preferred direction was defined as the direction 180 degrees opposite to the preferred direction.

Subsequently we studied tuning for step size in the preferred and non-preferred direction at different time intervals between displacements (figure 3.1). A single experiment consisted of 6 different step sizes in both the preferred and anti-preferred direction, presented at a fixed temporal interval, in random order. The step sizes were 1, 2, 4, 8, 16, and 32 pixels, corresponding to 2.4, 4.8, 9.6, 19.2, 38.4, and 76.8 min of arc. The same experiment was repeated for different temporal intervals between motion steps (1, 2, 4 and 8 monitor frames, corresponding to 8.3, 16.7, 33 and 67 ms). For intervals longer than a single monitor frame, the RPA remained static during intermediate frames. The total number of displacements was adjusted for longer intervals, so as to keep the duration of a single experiment constant, irrespective of stimulus interval. Depending on the stimulus interval, each step size was presented 1600 (8.3 ms), 800 (16.7 ms), 400 (33.3 ms) or 200 (66.7 ms) times.



**Figure 3.1:** The motion reverse correlation paradigm. A) A random pixel array (RPA) (50% black and 50% white pixels) is displaced in a random sequence of step sizes in the preferred and non-preferred direction (PD and ND, respectively) of a cell. In a single experiment step size varied (1, 2, 4, 8, 16, and 32 pixels, corresponding to 2.4, 4.8, 9.6, 19.2, 38.4, and 76.8 min of arc), for a constant inter stimulus interval duration (8.3 ms, corresponding to a single monitor frame, in the first experiment and 16.7, 33.3. or 66.7 ms in later experiments). Spikes were measured at 0.5 ms temporal resolution. B) Upper panel: cross correlations are performed for each stimulus separately, as illustrated for step size of 4.8 min of arc in the preferred direction. Middle panel: peri stimulus time histograms are constructed from 100 ms before to 500 ms after stimulus presentation. Lower panel: the correlation function after smoothing with a sliding Gaussian window ( $\sigma = 4\text{ms}$ ), subtracting the mean correlation function across stimuli, and dividing by the standard deviation of uncorrelated noise (100 ms preceding the motion impulse). These steps result in a correlation function indicating stimulus spike-pairs relative to the mean for all stimuli, expressed in standard deviations (SDs). For a detailed description see the methods section.

#### 3.3.5 Data analysis

Since we were interested in the temporal dynamics of step size tuning in MT neurons, we used a correlation method similar to our previous work (Borghuis et al., 2003; Perge et al., 2005a,b). On-line analysis was done by reverse correlating the occurrence of a spike with the presented stimulus sequence. Off-line analysis was done by forward correlation of motion steps with the spike train (figure 3.1B). Notice that correlating spikes to the stimulus, as reported here, is equivalent to correlating stimuli to the spikes, as reported in previous papers (Borghuis et al., 2003; Perge et al., 2005a,b). The two methods only differ in the time that serves as a reference. In forward correlation the stimuli are aligned at time zero and spike densities following the stimuli are determined. In reverse correlations the spikes are aligned at time zero and the density of stimuli before spikes is determined. In neither case the correlation functions can be interpreted as a response relative to spontaneous activity. Instead, they describe the modulation of spiking activity relative to the mean activation level in the experiment, which is set by the ensemble of stimuli. Because stimuli were presented at high temporal frequency, it is not possible to directly relate positive and negative modulations to excitation or inhibitions relative to some spontaneous activity level. The correlation functions merely describe the relative efficacy of different stimuli to generate spikes. To emphasize that our results reflect relative correlations the unit is the number of stimulus-spike pairs per second, rather than the number of spikes/s following a stimulus.

Correlation functions for each stimulus were smoothed with a sliding Gaussian window with a width ( $\sigma$ ) of 4 ms. This window-size effectively removed most of the high frequency noise, without affecting the shape of the function and its main parameters (see also Borghuis et al., 2003).

Because MT cells respond to the frame rate that we used as well as to the luminance flicker in the stimuli, the correlation functions show modulations at both the monitor frequency and at the frequency of displacements. These modulations are the same for all stimuli and were therefore not relevant for directional selectivity. To compensate for these general modulation artifacts we subtracted the mean correlation across stimuli from each correlation function. This is slightly different from the normalization used in our previous reports. Perge et al. (2005a,b) divided the correlations for each stimulus by the total number of spike-stimulus pairs. This resulted in a measure expressed as relative probabilities, with the total sum normalized to 1. Here we subtracted the mean correlation level at each point in time, without normalizing to probabilities. The advantage is that modulation amplitudes at different moments following a stimulus can be directly compared. Large amplitudes can no longer result from a small number of stimulus spike pairs, no matter how reliably they correlate. By subtracting the mean correlation, large amplitudes more directly correspond to a high number of stimulus-spike pairs.

Amplitudes of correlation functions were expressed relative to the noise level in an experiment. This noise level was determined from the standard deviation of

an uncorrelated part of the correlation functions, averaged for all stimuli. Significance of modulations in the correlation functions was obtained in a similar way. The mean number of stimulus spike pairs, and its standard deviation (across different states and time) were determined for a 100 ms window before a stimulus. Deviations from the mean in this part of the function occurred before a stimulus was presented, and therefore resulted from chance fluctuations. These deviations were normally distributed. A modulation in the correlated part was considered significant if the amplitude exceeded the level of 4 SD. This criterion closely resembles the 99% confidence interval based on noise distributions and degrees of freedom. Correlation functions were computed by custom made software written in programming language C. Final analysis was done in MATLAB (version 5.2.1. for Macintosh).

## 3.4 Results

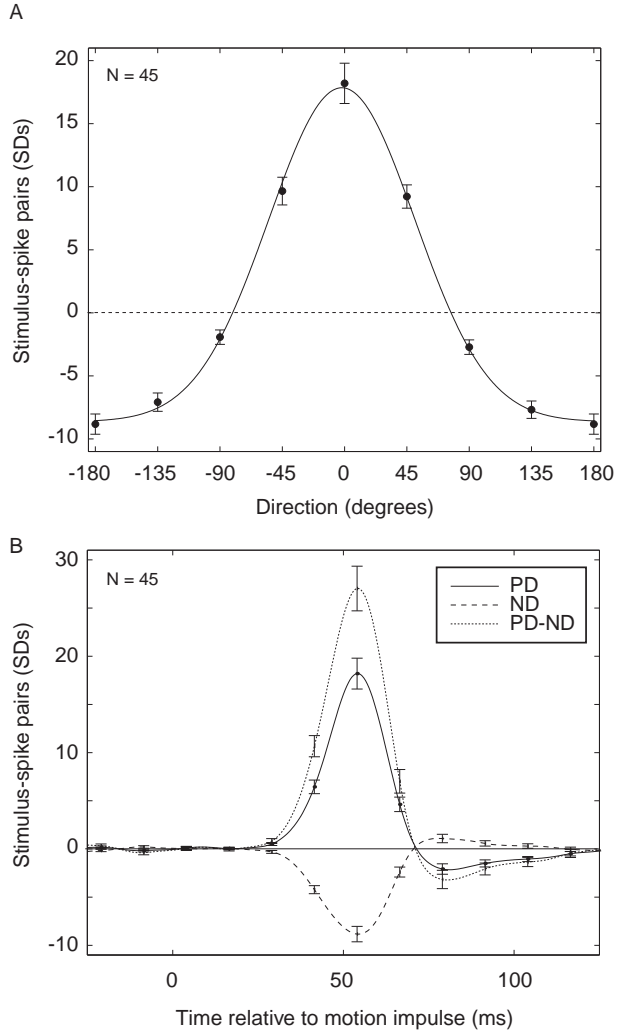
Data were collected for a total of 49 neurons in two male rhesus macaque monkeys (15 neurons in MA and 34 in MS). For each neuron we first determined its receptive field size and position, and a full direction tuning profile. In MA receptive fields were, on average 6.3 deg in diameter (SD 1.5 deg) and located at an eccentricity of 6.9 deg (SD 3.1 deg). In MS the mean eccentricity was 9.9 deg (SD 2.9 deg) and receptive fields were on average 6.9 deg in diameter (SD 2.2 deg).

### 3.4.1 Direction tuning

Figure 3.2 summarizes direction tuning properties across the whole population of MT cells. Preferred directions for all cells were aligned to 0 degrees. Response modulations for eight directions are shown at a latency of 54 ms (figure 3.2A), which corresponded to the mean peak latency of the correlation functions. Response modulations are given in units of standard deviations (SDs) for random correlations as measured in a 100 ms interval before stimulus presentations (see section 3.3.5 on page 42). On average, the maximum modulation in the preferred direction was about 18 SDs, and the reduction in the non-preferred direction was about 9 SDs. The width of the direction tuning curve was, on average 53 deg, as quantified by the SD of a Gaussian fitted to the direction profile. Based on these direction tuning measurements we selected the preferred direction; the non-preferred direction was always chosen in the opposite direction.

Figure 3.2B shows the time course of the correlation functions, averaged for all cells, for the preferred and non-preferred direction as well as their difference, which quantifies directional selectivity. In many cells the correlation functions showed the typical biphasic response as described previously (Borghuis et al., 2003; Perge et al., 2005a,b). The positive peak for the preferred direction at short latency is followed by a negative peak at a longer latency, and the response to the non-preferred direction is followed by a positive peak.

### 3.4 Results



**Figure 3.2:** Direction tuning and dynamics. Directional selectivity was measured using 8 different directions (45 degrees apart) with an interval of 8.3 ms and a step size of 0.12 degrees. The data represent the average modulation of the population of MT neurons ( $n=45$ ). The error bars represent the standard error of the mean. A) Relative efficacy of different directions to evoke spikes, at the optimal latency of 54 ms. Preferred directions were aligned at 0 degrees. The data were fitted with a Gaussian function ( $\sigma = 53$  degrees). B) The time course of the correlation functions for the preferred and non-preferred directions, as well as their difference. Error bars indicate standard deviations and for clarity are only shown at a few, equally spaced time intervals.



### 3.4.2 Step size tuning

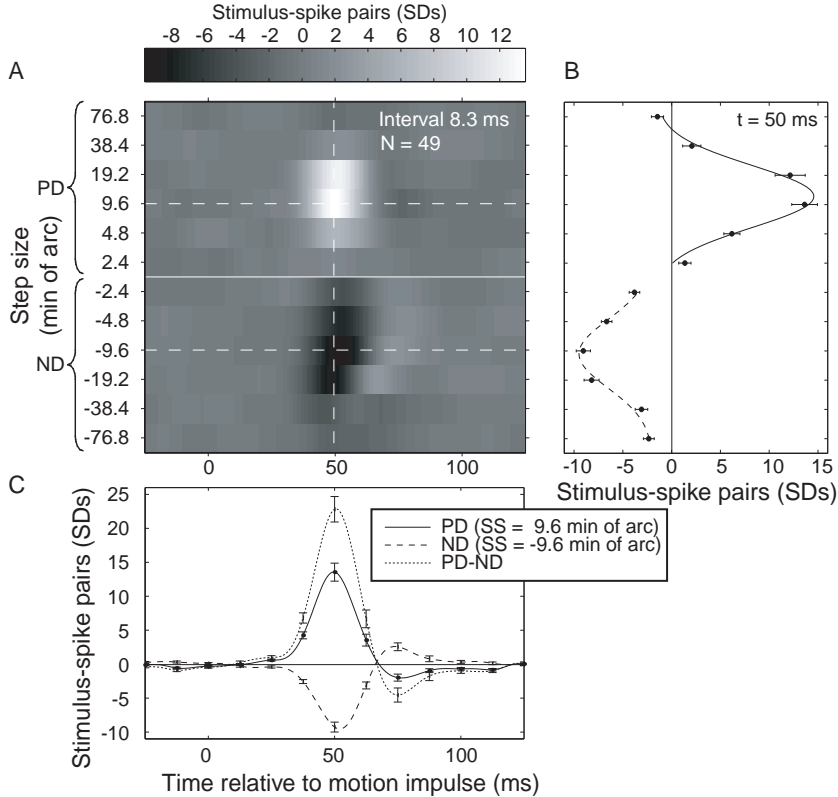
In the present study we compare the dynamics of direction tuning as a function of step size, and we show how step size tuning varies with inter stimulus interval. In a single experiment we fixed the inter stimulus interval and varied step size in both the preferred and non-preferred direction. These experiments were then repeated at different inter stimulus intervals. We will first describe the dynamics of step size tuning at the shortest inter stimulus interval (8.3 ms), and then show the changes for increasing inter stimulus intervals.

The surface plot in figure 3.3A shows the correlation functions measured with the shortest inter stimulus interval (8.3 ms), averaged for all recorded MT cells. Different step sizes are arranged along the vertical axis, with step sizes in the preferred direction increasing from the middle upwards and in the non-preferred from the middle downward. Negative numbers indicate step sizes for the non-preferred direction. Time is represented along the x-axis, relative to the presentation of stimuli (time zero). Amplitudes are given in units of standard deviations, determined over a time period of 100 ms before the stimuli. This equalizes responses in terms of signal to noise-ratios rather than mean firing rates. The latter were more variable between cells (mean  $32.0 \pm 25.0$  sp/s). Response modulations for steps in the preferred direction peaked at 13.6 SD (SEM = 1.3). The minimum value for the non-preferred direction was 9.1 SD (SEM = 0.7). All recorded cells thus showed highly significant directional selectivity at this short inter stimulus interval.

Vertical sections through the surface plot represent step size tuning at different latencies relative to the stimulus. Figure 3.3B shows an example, at a latency of 50 ms after stimulus presentation. Step size tuning functions for the preferred and non-preferred directions were separately fitted with log-Gaussian functions (Nover et al., 2005). The optimal step size for the population calculated from this fit in the preferred direction was 11.6 min of arc, and for the non-preferred direction 10.3 min of arc. The overall tuning width for the population of MT cells in the preferred direction ( $\sigma = 1.1$  log unit) was slightly smaller than in the non-preferred direction ( $\sigma = 1.2$  log unit). We will get back to this difference for the two directions in describing the results for individual cells.

Figure 3.3A also shows that at larger latencies the tuning profile is reversed. At a latency of about 77 ms the non-preferred direction clearly showed a positive modulation, and the preferred direction shows a small negative modulation. This biphasic response for preferred and non-preferred directions is more clearly shown in figure 3.3C, which corresponds to horizontal sections through the surface plot of figure 3.3A. The sections are taken at the step sizes yielding the largest positive and negative modulations. Temporal biphasic responses are more pronounced for steps in the non-preferred direction than in the preferred direction. This effect has been described in previous reverse correlation studies. Its magnitude may depend on contrast and spatial and temporal frequency content of the stimulus and other stimulus parameters (Livingstone et al., 2001; Bair and Movshon, 2004;

### 3.4 Results



**Figure 3.3:** Step size tuning for the population of MT cells ( $n=49$ ) at the shortest inter stimulus interval (8.3 ms). A) Averaged correlation functions for all step sizes. Step sizes for the non-preferred direction (ND) are indicated by negative numbers. Positive response modulations are shown in bright colors and negative modulations in dark colors, both expressed in standard deviations of uncorrelated noise. Time is represented along the x-axis, relative to displacements (time zero). B) Step size tuning at a latency of 50 ms, corresponding to the vertical, dashed line in figure 3.3A. Step size tuning functions for the preferred and non-preferred direction are separately fitted with log-Gaussian functions (Nover et al., 2005). This fit of the population average results in an optimal step size in the preferred direction of 11.6 min of arc and  $\sigma = 1.1$  log unit, and in the non-preferred direction 10.3 min of arc and  $\sigma = 1.2$  log unit C) Temporal response profiles, corresponding to the horizontal, dashed lines in figure 3.3A (9.6 min of arc) for the preferred and non-preferred direction, as a function of time. The sections are taken at the step sizes yielding the largest positive and negative modulations. The dashed line in figure 3.3C shows the difference between the preferred and non-preferred direction, which is a measure of directional selectivity. Error bars represent standard deviations and for clarity are given at a few, equally spaced time intervals.

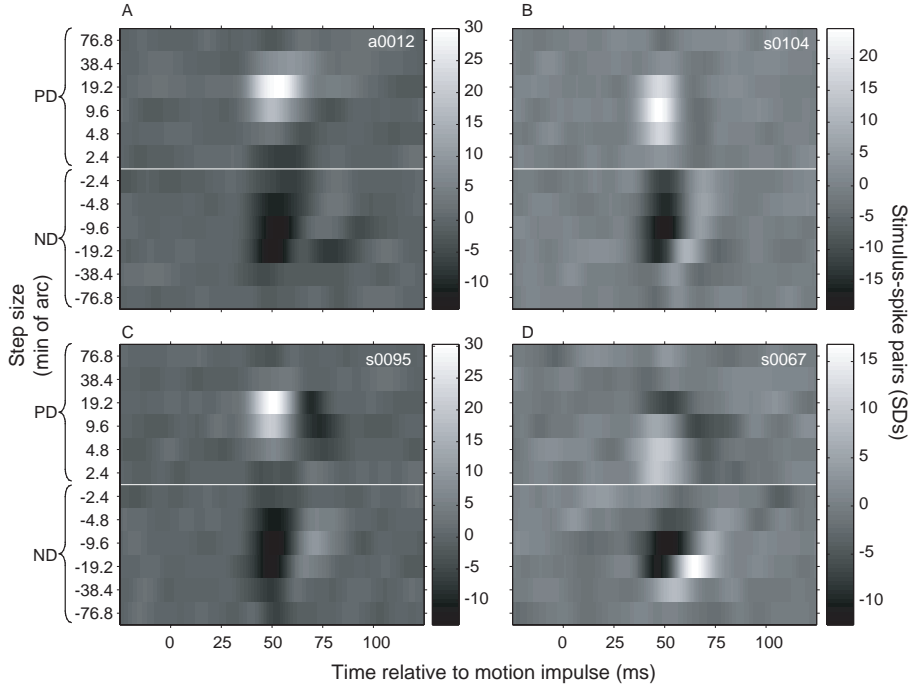
Cook and Maunsell, 2004; Perge et al., 2005a,b). Furthermore, these horizontal cross sections show that response latencies for the preferred and non-preferred direction are, on average, the same for our population of MT cells.

To provide a good impression of the variation in response properties we show results for 4 individual cells in figure 3.4 in the same format as figure 3.3A. Cells differed in their step size tuning profile as well as in their temporal dynamics. As for the cells depicted in figure 3.4 most cells were narrowly tuned for step size. This shows that the relatively larger width of the tuning profile for the population (figure 3.3A) is to a large extent the result of differences in optimal step sizes for different cells. Temporal biphasic behavior differed substantially between cells. In some cells (figure 3.4A, B) it was nearly absent, whereas others were clearly biphasic in time (figure 3.4C, D). One striking feature found for a minority of cells was a strong positive modulation at a larger latency for the non-preferred direction. It always occurred at a step size larger than the optimal step size. In some cases (8 out of 49 cells) the amplitude of this modulation was larger than the short-latency response modulation in the preferred direction. This phenomenon is clearly visible in the example cell of figure 3.4D. Another surprising feature is that some cells (figure 3.4D) also show a reversal of the sign of modulations at larger step sizes: A large step size in the preferred direction caused a negative modulation and a large step in the non-preferred direction caused a positive modulation.

We have examined several possible causes for these reversals at larger step sizes, but we have found no obvious explanation. To check whether it was caused by interactions between successively presented motion steps we performed a second order analysis of the data, in which the correlations were separated out for successive combinations of steps. If cells were responding to specific combinations of steps, e.g. two steps together corresponding to a preferred step, this would show up as a pattern of correlation for specific combinations. This was not the case. The correlations for combinations of steps did not substantially differ from what one would expect from separate steps. Thus, nonlinear interactions for successively presented stimuli did not seem to play a role. Furthermore, the inversion was not directly related to temporal biphasic behavior. Some cells were clearly temporally biphasic in the preferred and non-preferred direction, yet showed no sign of the inversion at larger step sizes (e.g. figure 3.4B). Our preliminary conclusion is that a small minority of cells seems to reverse directional selectivity for large step sizes, especially in the non-preferred direction. We will return to this issue in the discussion.

To further quantify differences in step size tuning between the preferred and non-preferred direction for the whole population we made scatter plots for the peak latency, biphasic index, optimal step size, and the width of the step size tuning (figure 3.5).

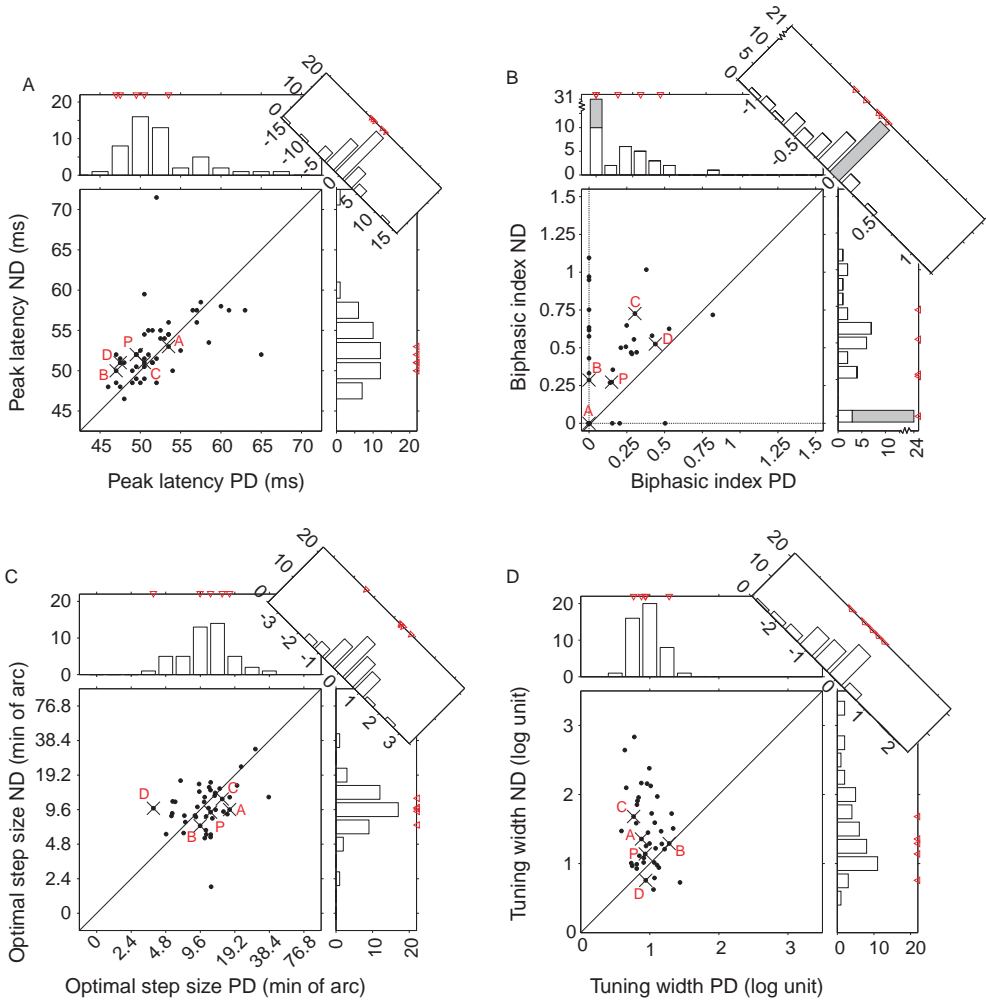
The differences in dynamics for preferred and non-preferred direction are not reflected in the peak latencies (figure 3.5A). On top and to the right of the scatter plot, the frequency distributions are plotted for the preferred (mean  $52.4 \pm 4.7$  ms) and the non-preferred direction (mean  $52.9 \pm 4.2$  ms). The diagonal line



**Figure 3.4:** Variability in step size tuning. Panels show the time course of step size tuning for an inter stimulus interval duration of 8.3 ms for four individual MT neurons. Neurons differ in biphasic behavior, sensitivity and preferred step size. Conventions are similar to those described for figure 3.3.

represents equal latencies. The histogram in the upper right corner shows the distribution of latency differences. It is centered on a mean value of  $-0.8$  ms, with a SD of only 4.3 ms. A t-test for paired samples revealed that there is no statistically significant difference in the distribution of latencies (t-test,  $p=0.59$ ). This is in contrast to a paper by Bair et al. (2002) that reported latency differences between preferred and non-preferred directions. The peak latencies were similar to those reported in other studies in MT (Raiguel et al., 1999; Perge et al., 2005a).

To quantify the degree of temporal biphasic behavior we computed for each cell a biphasic index (BI), for the optimal step size in the preferred direction as well as the non-preferred direction (figure 3.5B). The index is the ratio of negative and positive modulation amplitudes that exceed the significance threshold. For the non-preferred direction it is the ratio of positive to negative modulation amplitudes. If a modulation did not exceed the significance threshold, the ratio was set to zero and the cell was defined to behave in a monophasic manner. Low values for the biphasic index indicate the absence of the second inverted response modulation. High values indicate a relatively large second peak. 3.5B shows the distribution of BI values for the two directions, as well as their interdependence.



**Figure 3.5:** Population tuning properties. Scatter plots show response measures for the non-preferred direction as a function of that for the preferred direction. Dots represent individual MT cells. Histograms on top and to the right show the distributions for each direction separately, and the oblique histograms in the upper right corner show the distribution of their differences. Crosses (A,B,C and D in scatterplot) and triangles (histogram) indicate the location of the neurons shown in figure 3.4 and the average population response (label P in scatterplot). A) Peak latencies; B) Biphasic index; C) Optimal step size; D) Tuning width. The gray part in the bars in the histograms in figure B marks the number of neurons that were defined as monophasic in their preferred direction as well as non preferred direction (for more details see text).

### 3.4 Results

---

Of 49 neurons measured, 21 behaved monophasic, 19 neurons revealed biphasic behavior in their preferred direction, and 26 neurons showed biphasic behavior in their non-preferred direction. The median BI value for the preferred directions was 0.28 (2.5% quartile: 0.14 octaves, 97.5%: 0.82 octaves) and for the non-preferred direction it was 0.57 (2.5%: 0.14 octaves, 97.5% 1.08 octaves). Moreover, the distribution of BI for PD and ND were significantly different (t-test for paired samples,  $p < 0.001$ ).

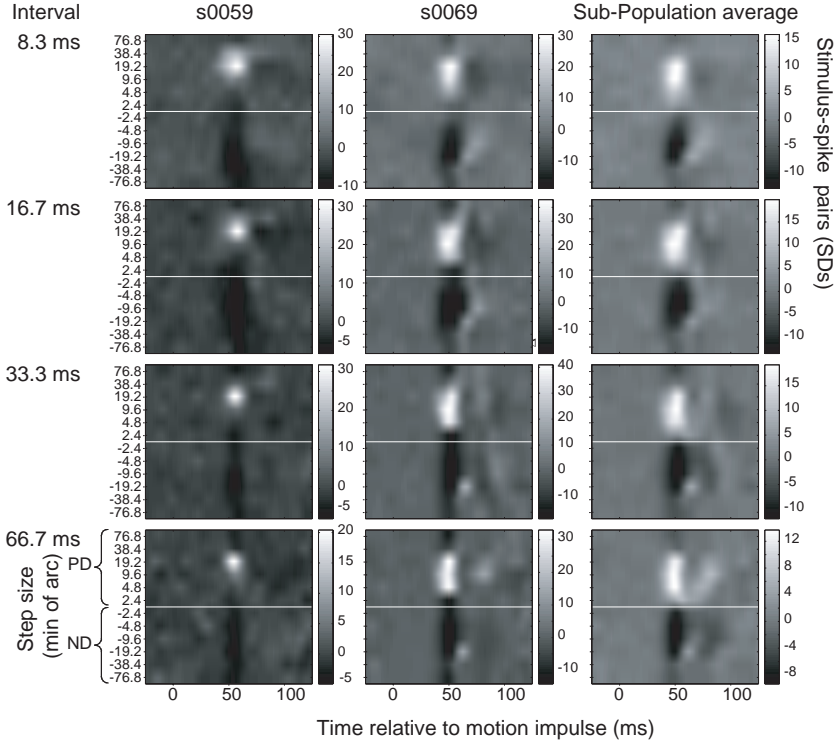
The optimal step size for the preferred direction and the non-preferred direction for a cell did not differ significantly (figure 3.5C). The optimal step sizes were estimated by fitting log-Gaussian functions to the tuning profiles for the preferred and non-preferred direction separately. A t-test for paired samples revealed that the distributions were not significantly different for the two directions (t-test,  $p=0.35$ ). The mean optimal step size for the preferred direction was 10.8 min of arc  $\pm 0.7$  octave, and for the non-preferred direction it was 9.8 min of arc  $\pm 0.7$ . The histogram in the upper right corner shows the distribution of differences, which is normally distributed around zero ( $\mu = 0.1$  min of arc,  $\sigma = 0.8$  min of arc).

As suggested by the scatter plot in figure 3.5D, the width of the step size tuning curve did differ between directions. The width of the tuning curves was given by the  $s$  parameter of the log-Gaussian fit. The median for the tuning width was 0.94 octaves (2.5% quartile: 0.62 octaves, 97.5%: 1.37 octaves) for the preferred direction and 1.36 octaves (2.5% quartile: 0.69 octaves, 97.5%: 4.09 octaves) in the non-preferred direction. A t-test for paired samples showed that the distributions were significantly different (t-test,  $p < 0.001$ ). Thus, step size tuning for the non-preferred direction is wider than for the preferred direction. The median for the difference in tuning width between preferred and non-preferred direction was -0.26 octaves (2.5% quartile: -1.07 octaves, 97.5%: 0.44 octaves).

#### 3.4.3 Interval duration

The data described so far were all measured using an interval of 8.3 ms. In this section we describe the effects of larger inter stimulus intervals. We increased the interval to 16, 33 and 66 ms and repeated the measurements. Figure 3.6 shows results for two example cells, and the average for the population of cells. It should be noticed that the different intervals were measured in different experiments. Hence, absolute response levels are not directly comparable, but we can compare spatial and temporal tuning properties. The two example cells and the population plot show that, in general, changing the inter stimulus interval had remarkably little effect.

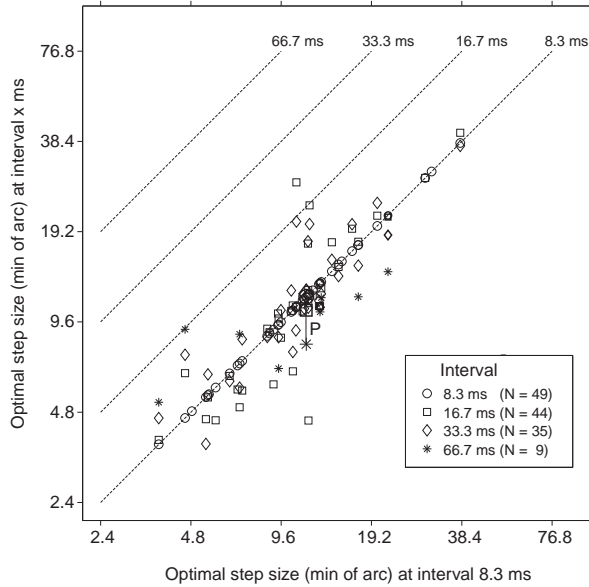
We further quantitatively analyzed the effect of inter stimulus intervals on the optimal step size, the width of the step size tuning, the peak latency and the biphasic index. Figure 3.7 shows a scatter plot of the optimal step size in the preferred direction, plotted with the longer intervals (16, 33 and 66 ms) against the shortest interval (8.3 ms). The dashed lines in the figure indicate expected



**Figure 3.6:** Step size tuning at different inter stimulus intervals. The two columns on the left show two examples for individual cells, the rightmost column shows the mean for those cells for which measurements were completed for all stimulus interval durations ( $n=9$ ). Quantitative data on step size tuning from the full set of measurements are shown in figure 3.7. Rows correspond to different inter stimulus intervals, increasing from the top downwards (8.3, 16.6, 33.3, 66.7 ms). Conventions as described for figure 3.3.

regression lines if optimal step size would be linearly proportional to inter stimulus interval. For the preferred direction the distributions for optimal step size show no consistent changes. The optimal step size does not increase systematically with increasing interval. For the non-preferred direction we found a small tendency for the optimal step size to decrease when increasing the interval from 8.3 ms to 16.6 ms (data not shown). This was opposite to what one would expect for invariable velocity tuning. The data clearly show that there is no systematic increase of optimal step size with increasing inter stimulus interval.

We also examined the differences in step size tuning width for preferred and non-preferred directions. Up to an inter stimulus interval of 33 ms there was no statistically significant change with increasing inter stimulus interval. For these intervals tuning was significantly sharper in the preferred direction than in the non-preferred direction. At the largest interval the difference in tuning widths



**Figure 3.7:** Changes of optimal step size with inter stimulus interval duration, relative to the shortest interval (8.3 ms). The dashed lines in the graph represent predictions for each inter stimulus interval duration in the case optimal step size would linearly scale with inter stimulus interval duration. The optimal step sizes at different inter stimulus intervals for the example cells and the mean response from figure 3.6 (label P) are shown enlarged.

for preferred and non-preferred direction were less and non-significant. Notice however that relatively few cells were measured with the longest inter stimulus interval ( $n=9$ ). Peak response latencies for positive modulations in the preferred direction, and negative modulations in the non-preferred direction were not significantly different at different intervals (one-way ANOVA:  $p = 0.13$ ). Similarly, the time course of directional selectivity was not systematically changed. Temporal biphasic behavior in the preferred direction was also to a large extent independent of stimulus interval (one-way ANOVA:  $p = 8.3$ ), as well as in the non-preferred direction (one-way ANOVA:  $p = 0.21$ ). Especially in the preferred direction the temporal profiles were highly similar. In the non-preferred direction the profiles were somewhat more variable, but in general there were no systematic shifts in degree of biphasic behavior. If a cell showed a positive peak at larger latencies in the non-preferred direction for a short inter stimulus interval, then it also showed up at larger intervals, at about the same point in time. This supports our finding that the response reversal in the non-preferred direction was not due to temporal interactions between subsequent stimuli. In that case we would expect large changes in dynamics with increasing interval.



## 3.5 Discussion

This study examined step size tuning in direction sensitive neurons in area MT. Using a reverse correlation method we were able to compare the dynamics of step size tuning for the preferred and non-preferred direction, and we determined how response properties changed with inter stimulus interval duration. All our data were obtained with moving random dot patterns in which we changed step size and inter stimulus interval duration. This is a fundamentally different approach from using sine wave gratings. In terms of spatio-temporal frequencies our stimuli are highly complex. Changes of speed or step size induce global shifts in temporal frequency content, for all spatial frequencies. Clearly, an explanation in terms of spatio-temporal frequencies is likewise complex, and perhaps also inconclusive. The approach based on the use of random dot patterns is however not primarily directed at disentangling the effects of separate spatio-temporal frequency combinations. Instead, it is focused on describing response properties that are more directly related to coherent motion in natural scenes. The main question in this approach is how to describe response properties independent of spatial frequency content. For studying low-level motion detection, i.e. independent of object properties, this seems a worthwhile approach.

### 3.5.1 Step size tuning

The main goal of the present study was to investigate whether this complexity translates into a simple description in terms of correlation analysis. Our main finding was that step size tuning, including its dynamics was highly similar for preferred and non-preferred direction, and for different inter stimulus intervals. We found small, though significant differences in the width of step size tuning for preferred and non-preferred directions, but optimal step sizes, and peak latencies were invariant with direction and inter-stimulus interval. Our main conclusion, therefore, is that MT neurons are invariantly tuned to step size, irrespective of inter stimulus interval duration and direction.

In terms of correlations step size tuning directly characterizes the invariant parameter for motion of a complex pattern: its displacement in a given time. Form invariant velocity tuning might thus not result from combining motion information for different spatial frequencies, but might be built in at the very first stage of directional selectivity. MT derives its input mainly from special complex cells in V1 (Movshon and Newsome, 1996). These special complex cells respond to a broad range of spatial and temporal frequencies. If V1 simple cells perform an initial form analysis, then special complex cells might perform the initial, form invariant motion correlation underlying MT directional sensitivity. This hypothesis might imply a main divergence in V1 into form related processing by simple cells, and form and position invariant processing of motion (and stereo) information in complex cells. The hypothesis that simple cells and complex cells form a divergence rather than a hierarchy is supported by several studies showing

direct input from the LGN to complex cells, rather than (only) from simple cells (see an extensive review on hierarchical and parallel input to complex cells in primary visual cortex by Martinez and Alonso, 2003).

Our data are in line with findings of Pack et al. (2006). Using sparse random dots and a second order correlation analysis they found a similar independence of step size and inter stimulus interval tuning. They suggested that the independence might have been due to the low efficacy of their stimuli to drive directionally sensitive responses, which they argued, would be expected based on the finding of Priebe et al. (2006) that velocity tuning for gratings increased with contrast. Moreover, they suggested that the noisiness of the measurements also resulted in narrow tuning curves, which therefore tended not to show any significant speed tuning related to variations in displacement and interval. Our stimuli very effectively drive MT cell responses and we found the same result. We therefore conclude that MT cells are tuned to a single combination of spatial displacement irrespective of inter stimulus interval duration.

We would like to stress, that descriptions based on Fourier analysis and correlation analysis are in no way mutually exclusive. To the contrary, Fourier analysis describes exactly the same stimulus in the frequency domain rather than the space-time domain. The two descriptions are completely equivalent. Hence, it is always possible to explain response properties based on either description. In terms of Fourier energy invariant step size tuning would be captured in a complex interaction between spatial and temporal tuning properties. Several studies employing sine wave gratings, and combinations of sine wave gratings indeed point to complex combinations of different temporal frequency components. In fact, not only energy at specific combinations of spatial and temporal frequency matter, but also their spatial phase relationships (Mechler et al., 2002; Felsen et al., 2005; Priebe et al., 2006).

Our data on variations of step size tuning with inter stimulus interval show no signs of a division in qualitatively different behaviors. This is different from earlier studies that used sine wave gratings. MT cells have been divided in pattern and component selective cells (Movshon et al., 1986; Rodman and Albright, 1989). In building network models for speed discrimination Chey et al. (1998) differentiate between speed-tuned cells and speed-sensitive cells. Speed-tuned cells respond preferentially to a limited, continuous range of speeds, as opposed to speed-sensitive cells, which vary their response with speed, but do not exhibit a preference for a particular speed. The reason for this apparent difference in diversity, measured with random dot patterns on the one hand and sine waves on the other hand, might be a diversity of low level, local receptive field profiles feeding into motion detectors. If such diversity plays a role, a random dot pattern allows for broad sampling of different inputs which assures average behavior in all cases. Gratings on the other hand selectively favor a subset of inputs, allowing for changes as stimulus parameters change. In this light it is quite evident that both random dot patterns and sine wave gratings are valuable in characterizing essential tuning properties.

### 3.5.2 Response dynamics

The second aim of this study was to characterize dynamics of motion responses, as a function of step size, direction and inter-stimulus interval. Perge et al. (2005b) previously showed that MT cells display different degrees of biphasic behavior in responses to single displacements. If one adopts the correlation function as a reasonable first order description of the motion impulse response, then biphasic behavior might explain the typical transient-sustained responses of MT cells for prolonged stimuli. Priebe et al. (2002) and Priebe and Lisberger (2002) suggested that short term adaptation caused such response transients. They adapted cells with brief (64 ms) motion presentations and measured responsiveness after different intervals. Transience was found to decline with increasing inter-stimulus interval, with a neuron specific time constant. The question thus arises whether biphasic correlation functions and short term adaptation are in fact due to the same mechanism. The preliminary findings of Perge et al. (2005b) were that these two phenomena are in fact not necessarily correlated, indicating different mechanisms. In the present study we also varied the inter-stimulus interval. If biphasic responses result from short term adaptation we would therefore expect a similar reduction in biphasic behavior with increasing interval. Clearly, this was not the case. In fact response profiles measured for different intervals were highly similar. This shows that biphasic behavior differs from adaptation and is, in fact, inherent to the response to a single motion impulse. Both mechanisms may contribute to the transient behaviour of MT cells.

In general we found remarkably little differences in dynamics for different step sizes and different directions. Positive modulations for the preferred direction and negative modulations for the non-preferred direction have similar latencies. Different step sizes, and different inter-stimulus intervals similarly had little effect on response latencies. These findings at first seem to differ from findings reported by e.g. Bair et al. (2002). They showed that most MT neurons responded faster to a transition from preferred to anti-preferred than to the opposite transition. Our data show that these differences do not primarily result from differences in response dynamics for different directions, but rather from (linear) integration of opposite responses with equal dynamics. Our data also show no signs of the large changes in dynamics with e.g. spatio-temporal structure and contrast, as reported by Bair and Movshon (2004). Obviously, changes in stimulus efficacy without changing contrast, spatial or temporal frequency do not result in changes in dynamics. Perhaps the observed changes in dynamics were not due to changes in integration at the level of directional selectivity (V1/MT), but at earlier levels. This is supported by our finding that changes in stimulus contrast in our experiment did have a significant effect on the response time course, similar to Bair and Movshon's finding (unpublished results). We did find a significant difference between the biphasic index for the preferred and non-preferred direction, but this might simply result from a nonlinear compression that attenuates negative response modulations relatively stronger than positive modulations. Such compression is inevitable due to the smaller range available for reductions of spike

frequency compared to increments.

Some cells showed a large positive response modulation for steps in the non-preferred direction. If present it generally occurred at a relatively large step size and at a latency slightly larger than the optimal latency for the cell. Typically, its presence and strength did not vary consistently with inter-stimulus interval. The reversal did not result from specific sequential effects, but was inherent to the response to individual steps. It might be related to the reverse phi phenomenon that has been extensively studied psycho-physically and physiologically. Reverse phi is the reversal of perceived direction when the contrast of a pattern reverses during displacements. Similar effects have been described in directionally selective cells in e.g. V1 and MT (Livingstone and Conway, 2003; Krekelberg and Albright, 2005). Likely the reversal we observed is related to this contrast reversed response. Although we did not reverse contrast consistently, in half of the cases pixels will have the opposite contrast. At some specific combination of step size and pixel size reverse contrast correlations may become dominant. The fact that only some cells show the effect shows that it apparently requires a fairly precise combination of spatial tuning relative to the pixel size.

The main conclusion from the present study is that response properties, including the dynamics of responses to individual motion steps are largely invariant with direction and inter-stimulus interval. Similarity of dynamics in the preferred and non-preferred direction shows that excitatory and inhibitory inputs shaping these responses are quite the same, and presumably follow similar paths. Because opponency is fairly limited in V1 (Snowden et al., 1991) directional opponency in MT must result from excitatory and inhibitory projections with equal dynamics. The fact that step size tuning was narrow and independent of inter-stimulus interval suggests that MT cells sample V1 units with similar properties. We found no indications for speed invariant covariations of step size and interval tuning.

# Chapter 4

## Tuning for temporal interval in human apparent motion detection

Roger J.E. Bours, Sanne Stuur and Martin J.M. Lankheet

*Journal of Vision* (2007b) 7(1):2 1-12

<http://journalofvision.org/7/1/2/>, doi:10.1167/7.1.2.

### 4.1 Abstract

Detection of apparent motion in random dot patterns requires correlation across time and space. It has been difficult to study the temporal requirements for the correlation step because motion detection also depends on temporal filtering preceding correlation, and on integration at the next levels. To specifically study tuning for temporal interval in the correlation step we performed an experiment in which pre-filtering and post-integration were held constant and in which we used a motion stimulus containing coherent motion for a single interval value only. The stimulus consisted of a sparse random dot pattern in which each dot was presented in two frames only, separated by a specified interval. On each frame, half of the dots was refreshed and the other half was a displaced reincarnation of the pattern generated one or several frames earlier. Motion energy statistics in such a stimulus do not vary from frame to frame, and the directional bias in spatio-temporal correlations is similar for different interval settings. We measured coherence thresholds for left-right direction discrimination by varying motion coherence levels in a Quest staircase procedure, as a function of both step size and interval. Results show that highest sensitivity was found for an interval of 17-42 ms, irrespective of viewing distance. The fall-off at longer intervals was much sharper than previously described. Tuning for temporal interval was largely, but

not completely, independent of step size. The optimal temporal interval slightly decreased with increasing step size. Similarly, the optimal step size decreased with increasing temporal interval.

## 4.2 Introduction

Two light sparks, placed closely together evoke the percept of motion when presented sequentially with a small temporal interval (Exner, 1875, 1888). Motion can be perceived for spatial offsets so small that the two sources can not be spatially resolved. It suggests that motion sensitivity is a fundamental property, based on low-level motion detectors. For motion detection in flies, Reichardt (1961) proposed such a low-level mechanism, in which signals from different receptors are correlated with a time delay for one of the signals. Such a motion detector is tuned for a specific combination of spatial span and temporal delay. Based on Reichardt's proposal, human motion detection has also been modeled by a front-end array of bi-local motion sensors, similar to the Reichardt motion detector (Reichardt, 1961; Van Santen and Sperling, 1984, 1985). Equivalent descriptions have also been developed in terms of motion-energy filtering (Adelson and Bergen, 1985; Perrone, 2004; Watson and Ahumada, 1985), or spatio-temporal gradient detection (Fennema and Thompson, 1979; Johnston et al., 1992).

Adopting the framework of bi-local motion sensors naturally divides the process of motion detection in three stages: A spatio-temporal operation generating the motion signal, which is preceded by spatio-temporal luminance filtering, and which is followed by spatial and temporal integration of motion signals. We will refer to the stage generating the motion signals as the 'correlation' stage. We use the term in a general sense, without claiming that the operations involved match those of a Reichardt detector (Emerson et al., 1992). In a physiological context, this stage would be the first level at which directional selectivity arises. In primates, in the geniculate-striate pathway directional selectivity first arises in V1 (De Valois et al., 2000; Mikami et al., 1986b; Movshon and Newsome, 1996; Perrone, 2004; Saul et al., 2005; Snowden et al., 1991). Directionally selective V1 neurons, therefore, implement the presumed correlation stage. Consequently, pre-processing would comprise all spatial and temporal filtering preceding this stage, including the retina and LGN. Most importantly, these stages transform the stimulus into a band pass-filtered image in both spatial and temporal domains. Motion signals presumably arise at multiple locations in such receptive fields which therefore also comprise a first step of spatial and temporal integration (Livingstone et al., 2001; Movshon and Newsome, 1996). Additional integration of motion signals is implemented at higher cortical stages, including area MT and MST (Britten and Heuer, 1999; Dubner and Zeki, 1971; Duffy and Wurtz, 1991a,b; Heuer and Britten, 2004; Livingstone et al., 2001).

Numerous studies used equivalent model-frameworks to psychophysically explore the spatio-temporal requirements for coherence detection in random dot dis-

plays (Fredericksen et al., 1993, 1994a,b,c; Morgan and Ward, 1980; Van de Grind et al., 1986; Van den Berg and Van de Grind, 1989; Van Doorn and Koenderink, 1982a,b, 1984). A primary aim in many studies was to examine the requirements for spatio-temporal correlation, irrespective of pre- and post-processing. Especially in the temporal domain, this has been difficult, because stimulus manipulations should neither affect the temporal frequency content, to rule out pre-filtering effects, nor the modulation of motion energy from frame to frame, to prevent temporal integration effects. Moreover, measurements of complete tuning curves, rather than just the spatial and temporal limits, require manipulation of motion coherence without affecting other relevant parameters.

In this study, we describe the results of measurements that fulfill the requirements for specifically addressing the parameters of the spatio-temporal correlation stage. Previous attempts have resulted in contrasting findings (Fredericksen et al., 1993; Morgan and Ward, 1980; Van de Grind et al., 1986; Van Doorn and Koenderink, 1982a,b), probably because one or more of the requirements were not fulfilled. Two-frame, or continuous apparent motion experiments (Baker and Braddick, 1985; Casco et al., 1989; Morgan et al., 1997; Morgan and Ward, 1980) have provided a description of maximal spatial and temporal displacement limits, but failed to provide complete tuning curves. Moreover, variations of stimulus duration and stimulus onset asynchrony may have greatly affected the total motion energy content and its temporal integration and therefore confounded correlation and temporal integration effects.

To measure complete tuning curves, Van Doorn and Koenderink (1982a,b) introduced a luminance signal-to-noise ratio (LSNR) for manipulating the motion strength in moving random pixel arrays. LSNR thresholds characterized motion sensitivity irrespective of the parameters specifying the apparent motion stimulus (i.e. step size and temporal interval). To specifically address the low-level correlation requirements for coherent motion detection, Van Doorn and Koenderink used spatial and temporal alternations of patterns moving in opposite direction. They observed, and measured, critical alternation frequencies, at which directional sensitivity was sharply reduced. For lower temporal frequencies, the two patterns could be seen to alternate in time. For high frequencies, the two patterns no longer segregated in time, but instead fused into transparency. At intermediate alternation rates motion was irresolvable. They concluded that these observations were in line with a bilocal motion detection device, with the critical alternation frequency corresponding to the preferred temporal delay of motion detectors. At this interval, they reasoned, no motion could be detected because motion sensors failed to establish the correlation. At longer intervals motion was detectable within a single interval and at shorter intervals motion was detectable through correlation over multiple alternations. Although these observations are in line with bilocal motion detection, one cannot rule out that varying the spatial or temporal frequency of alternations also affected spatial and temporal integration of motion signals.

Van Doorn and coworkers (Koenderink et al., 1985; Van de Grind et al., 1986;

Van Doorn and Koenderink, 1982a,b) reported a clear dependence of preferred step size and interval on pattern speed. At low speeds, preferred step sizes were constant and intervals varied inversely proportional with speed, whereas at high speeds intervals were constant and preferred step sizes increased with speed. This seems to contradict Morgan and Ward's (1980) finding of independent spatial and temporal limits in motion coherence detection.

Fredericksen et al. (1993) combined the manipulation of motion strength, using the luminance signal-to-noise ratio, with a single-step dot lifetime paradigm for controlling the step size and interval content of a motion stimulus. It is a continuous version of previously used two-frame motion stimuli (Baker and Braddick, 1985; Casco et al., 1989; Morgan et al., 1997; Snowden and Braddick, 1989). Although the stimulus of Fredericksen et al. (1993) was suitable for measuring spatial tuning, it was inadequate for manipulating motion interval durations without affecting the temporal frequency content of the patterns. The reason is that for intervals longer than the minimal value, patterns remained stationary during the motion interval. Consequently, the temporal frequency content varied with interval duration. Moreover, the temporal integration of motion signals was also likely affected, because displacements were presented at different temporal frequencies. In addition, and most importantly, because patterns remained stationary during the interval between motion steps, the motion comprised a wide range of temporal correlations and, hence, the stimulus failed to isolate unique combinations of step size and interval for correlation. The absence of a fall-off at large interval durations in their results, contrasts with earlier findings by Morgan and Ward (1980), and might have been due to the presence of correlation across short intervals in all of their stimuli.

Morgan and Ward's (1980) stimulus, on the other hand, effectively confined the motion energy to a single combination of step size and interval. This was achieved by showing corresponding dots for two moments only, separated by a predefined interval. Using an oscilloscope display on which dots were drawn successively, they had full control over temporal interval and were able to determine both minimal and maximal temporal intervals supporting coherence detection. Contrary to findings by Fredericksen et al. (1993) they found spatial and temporal limits to be independent of step size, and a sharp limit at large temporal intervals. Numerous methodological differences, however, prevent a direct comparison of these results to those obtained by others: Dot densities were low, stimulus durations were highly variable and measurements consisted of determining reaction times. Most importantly, rather than spatial and temporal tuning curves only the upper and lower limits were determined.

To clarify the discrepancies between previous results we used an improved version of the stimulus used by Morgan and Ward (1980), and combined it with motion coherence threshold measurements to obtain full spatial and temporal tuning curves. The stimulus was constructed by generating sparse random dot patterns on every frame of a computer monitor, and showing a shifted version of this pattern once again after a specified time interval. Therefore, each frame



contained a newly generated pattern and a displaced reincarnation of a previously shown pattern. The interval between corresponding dot patterns could be chosen freely, without affecting the number of steps per second, steps in total, and temporal frequency content. Motion information, i.e. a directional imbalance for spatio-temporal correlations present in the stimulus, was confined to a single temporal interval only and the stimulus contained constant motion energy statistics on every frame of the monitor. Moreover, measurements for different intervals were directly comparable since the level of spatio-temporal correlation at the specified motion parameters, and overall spatio-temporal energy content were invariable with interval duration.

### **4.3 Methods**

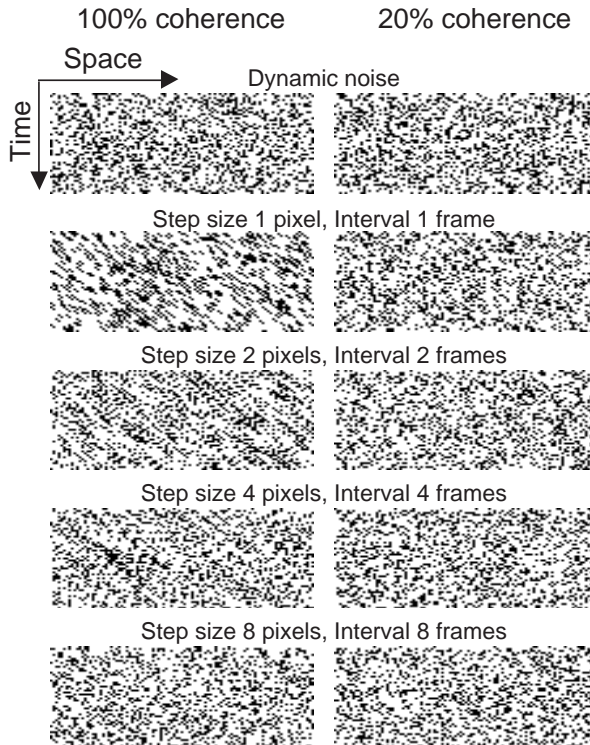
Random dot patterns were generated in realtime on a Macintosh G4 personal computer and displayed on a 19" Sony monitor (Multiscan E400) at a resolution of 800x600 pixels and a frame rate of 120 Hz. At the standard viewing distance of 125 cm one monitor pixel corresponded to  $0.02^\circ \times 0.02^\circ$ . In all experiments the patterns consisted of black dots (single monitor pixels) on a white background (mean luminance 37 cd/m<sup>2</sup>) displayed in a window of 400 x 400 pixels. The standard number of dots was 5000, yielding a dot density of 80 dots/deg<sup>2</sup>. In two control experiments, we tested the effect of different dot densities, and different viewing distance on the step size and interval tuning for an optimal interval and step size, respectively.

Consistent spatio-temporal correlations in the stimulus were limited to a single combination of step size and temporal interval. This was achieved by plotting a dot only twice, separated in time by the required temporal interval. An apparently continuous version of such a motion stimulus was constructed by newly generating 50% of the dots on each frame, and combining this pattern with a displaced version of a pattern presented one or several frames earlier. Each frame thus contained a combination of random noise (refreshed dots) and coherent motion. Step size (number of pixels) and temporal interval (number of frames) between corresponding dot patterns were varied in the experiments to measure the full spatio-temporal tuning profile.

Figure 4.1 illustrates the type of motion used in the experiments, in the form of space-time diagrams. Dot positions for a single row of the display are shown along the X-axis. The Y-axis represents time, with each line of dots corresponding to a monitor frame. The left hand column shows examples of dynamic noise and 100% coherent motion, for different combinations of step size and interval. A coherence value of 100% in these stimuli corresponds to coherent displacement of all relocated dots in combination with randomly refreshing the other half of the dots. All combinations (1/1, 2/2, 4/4 an 8/8 pixels/frames) represent the same mean velocity. Coherent motion shows up as an oriented pattern in such space-time plots, with the orientation representing the velocity. In comparison, apparent

### 4.3 Methods

motion with unlimited dot lifetime would give a noiseless, oriented pattern, with correlations across all time steps in the display. In contrast, a horizontal line in figure 4.1 correlates only with a line preceding it by the specified interval, or following it by the specified interval. There is no correlation bias across smaller or larger intervals. Importantly, the directional correlation bias in the stimulus is the same on every frame of the monitor, and does not vary with the specified temporal interval for displacements. As noted previously by Morgan and Ward (1980), apparent motion generated in this way looks surprisingly continuous, irrespective of the specified temporal interval. Observers have the impression of a rigidly moving pattern within dynamic noise, rather than a dynamically and discretely changing pattern.



**Figure 4.1:** Space-time plots for two-frame single-step motion stimuli. The left hand column shows plots for the maximal coherence value (100%), whereas the righthand plot shows the same motion step size and interval parameters for a coherence level of 20%. The top row shows dynamic noise: the pattern is randomly refreshed each time step. The additional four rows show settings for combinations of increasing step size and interval. In all four cases, the mean velocity is equal: In the top row, the pattern is displaced 1 pixel to the right with an interval of 1 frame. Step size and interval increase by a factor of two from one row to the next.

To measure sensitivity for a specific combination of step size and interval, we determined motion coherence thresholds in a two-alternative forced-choice paradigm. Coherence levels were varied by manipulating the percentage of dots taking part in the coherent motion (Britten et al., 1996). The other dots were randomly displaced in the stimulus window. The right hand column in figure 4.1 shows the effect of reducing the coherence level to 20%. Decreasing the coherence effectively reduces the directional bias in the stimulus.

Left-right discrimination thresholds were determined in a Quest staircase procedure (Watson and Pelli, 1983). A trial consisted of a single presentation of one second in which the pattern moved either to the left or to the right. Subjects indicated the perceived direction of motion by pressing the left or right arrow key on the keyboard. The staircase determined the coherence value at which observers performed at 85% correct. A single threshold measurement consisted of 40-60 trials (depending on experience of the observer), which was repeated 3-5 times to determine the mean threshold value and its standard error. All staircases were inspected, and a staircase was discarded and repeated if it had not properly converged within the maximum number of trials.

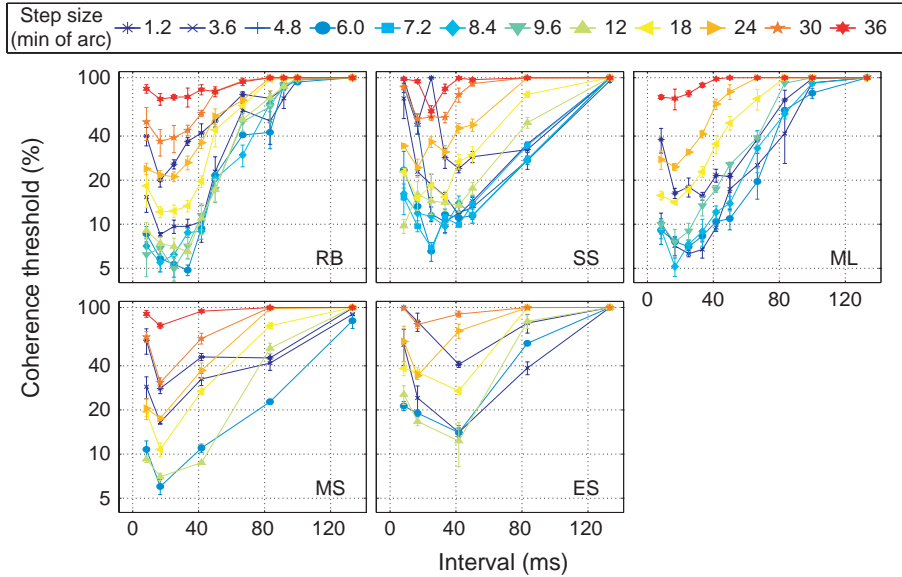
Observers viewed the stimuli binocularly, in a dark room at a viewing distance of 125cm. They were instructed to steadily fixate a small marker ( $0.08^\circ \times 0.08^\circ$ ), which was centered on the display window. Thirteen observers participated in the experiments. Three (RB, ML, SS) were experienced observers in motion and other psychophysical experiments. The other observers (MS, ES, BA, CC, EF, FR, AJ, EJ, RP, WS) were naive concerning the purpose of the experiment. Observers had full control over the pace of the experiment, starting each trial by a key press. No feedback on the correctness of answers was given. Data collection was started after subjects ran a few test staircases, to minimize any learning effects. All observers had normal or corrected-to-normal vision.

## 4.4 Results

Figure 4.2 shows coherence thresholds for various step sizes as a function of the time interval between motion steps. Each plot shows data for one subject, with step size as parameter in the graph. Coherence thresholds range from about 5%, corresponding to the highest sensitivity, to 100% (invisible). Coherence thresholds may be higher than those obtained using standard noisy motion stimuli. This is due to the substantial amount of noise that was present irrespective of coherence setting. Moreover, the motion stimulus used in this study, is more selective in triggering motion detectors that contribute to the motion percept, as biases in spatio-temporal correlation in the stimulus were limited to a single combination of step size and temporal interval. The duration of optimal intervals ranged from about 17-42 ms (2-5 frames). Sensitivity sharply dropped towards higher interval values. None of our observers could resolve an interval of 133ms (16 frames). Thresholds also increased for lower intervals, but this fall-off was less pronounced.

## 4.4 Results

All observers performed fairly well at the minimal temporal interval (8.3 ms) in our setup. At the shortest interval, our stimulus is comparable to the one used by Fredericksen et al. (1993). As expected, the perceptual behavior at short intervals is comparable to their results.



**Figure 4.2:** Coherence thresholds for different step sizes plotted as a function of temporal interval. Coherence thresholds correspond to the percentage of coherently moving dots in a Quest staircase procedure at 85% correct responses. Lower thresholds correspond to higher sensitivity, i.e. sensitivity is inversely proportional to the coherence threshold. Step sizes are given in minutes of arc and intervals in milliseconds. A single monitor pixel measured 1.2 min of arc and a single monitor frame lasted 8.3 ms. Error bars show the SEM for each threshold measurement, based on 3-5 repetitions.

### 4.4.1 Interactions between step size and temporal interval tuning

In general, temporal tuning curves had similar shapes for all step sizes. The curves seem shifted up or down, depending on the step size, but the minimum as well as the high and low interval fall-off were fairly similar. For some observers, especially ES, and to a lesser extent SS, the shapes of the curves indicate interactions between step size and interval tuning. Tuning for small step sizes tended to reach a minimum at higher interval values, whereas for larger step sizes the minimum shifted toward slightly lower intervals.

To analyze the interactions between step size and interval tuning in more detail, the data in figure 4.2 were re-plotted in the form of contour plots (left

hand column in figure 4.3). Coherence thresholds are given in colors ranging from yellow-white (low thresholds, i.e. high sensitivity) to black (high threshold, low sensitivity). Spatio-temporal interactions were quantitatively assessed by comparing a spatio-temporally separable model to models implementing several different interactions between step size tuning and temporal interval tuning. For fitting models to the data, we first converted the coherence thresholds ( $T$ ) to a sensitivity measure, in which sensitivity ( $S$ ) was defined on a logarithmic axis:

$$S = 2 - \log_{10}(T) \quad (4.1)$$

where the value 2 corresponds to the logarithm of 100% i.e. the maximum coherence threshold for an indiscriminable stimulus. Low coherence thresholds are thus converted into high sensitivity values. Logarithmic values were used in order to normalize standard deviations, at different performance levels, to more or less the same size.

Based on physiological results by Nover et al. (2005) we used log-normal functions to describe the temporal, and spatial tuning curves:

$$f(s) = \frac{1}{\sigma_s \sqrt{2\pi}} \exp\left(\frac{-\log_{10}\left(\frac{s}{\mu_s}\right)^2}{2\sigma_s^2}\right) \quad (4.2)$$

$$g(t) = \frac{1}{\sigma_T \sqrt{2\pi}} \exp\left(\frac{-\log_{10}\left(\frac{t}{\mu_T}\right)^2}{2\sigma_T^2}\right) \quad (4.3)$$

where  $f(s)$  describes tuning as function of step size ( $s$ ), and  $g(t)$  tuning as function of interval ( $t$ ).  $\mu_s$  and  $\mu_T$  are the optimal values and  $\sigma_s$  en  $\sigma_T$  represent tuning width. The separable model is given by  $h(s,t)$  and is defined by the product of spatial ( $f(s)$ ) and temporal ( $g(t)$ ) factors:

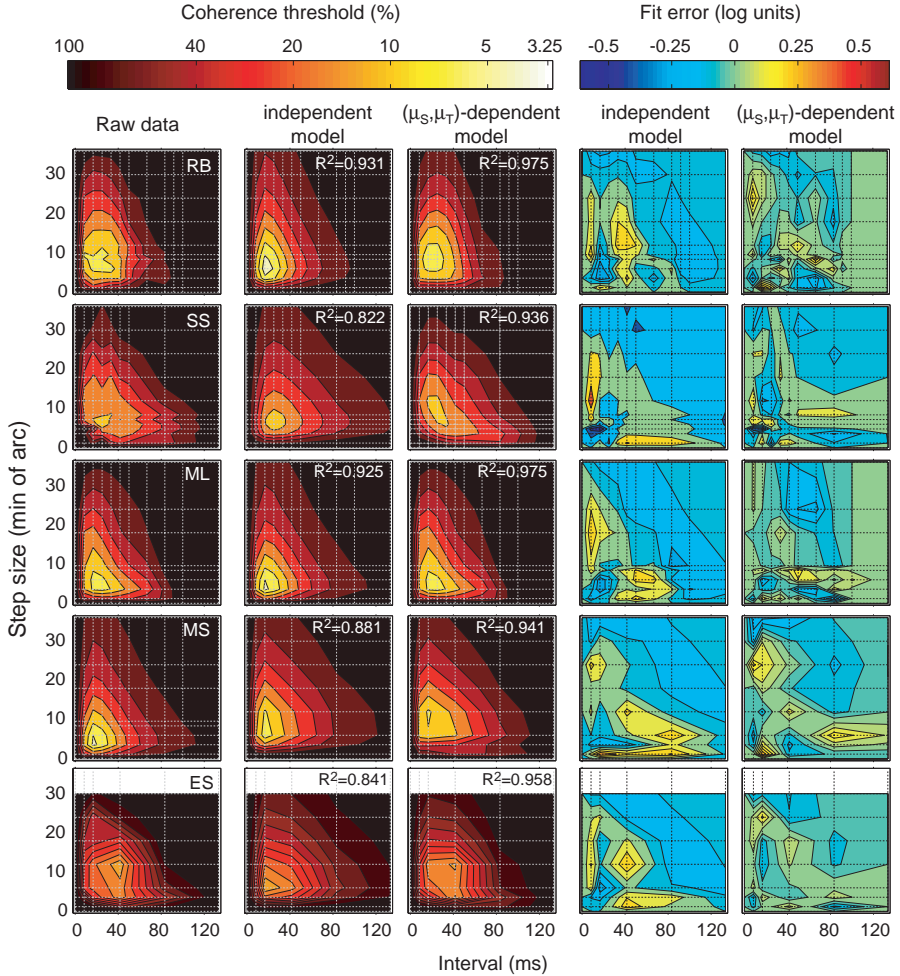
$$h(s, t) = A f(s)g(t) \quad (4.4)$$

where  $A$  represents global amplitude scaling.

To get a maximum likelihood estimate of the model parameters we used a nonlinear, least-square fitting routine based on the Gauss-Newton method. A Monte-Carlo simulation using a nonparametric bootstrapping procedure was used to assess the variability in the fitted parameters. The simulated datasets were based on the average of two randomly chosen values from the three individual measurements, for each combination of step size and interval. Subsequently, we quantified the goodness-of-fit by the  $R^2$ -value, which quantifies the fraction of the total squared error that is explained by the model. All procedures were implemented in MATLAB, using custom and standard Matlab functions.

To compare fitted tuning functions to the measured coherence thresholds, the sensitivity data were transformed back into coherence thresholds. The second column in figure 4.3 shows the resulting tuning curves for all observers. A

#### 4.4 Results



**Figure 4.3:** Contour plots of motion coherence thresholds as function of step-size and interval duration. The first column shows measured data, each row representing a different observer. The data are the same as those in figure 4.2. The second and third columns show fits of a spatio-temporally separable model, and a model including third order (cubic) variations in optimal step size and optimal interval. The fourth and fifth column show the fit errors for both models. Coherence thresholds are color-scale coded (horizontal bar above first three columns), where low detection thresholds (high sensitivity) are white/yellow and high thresholds (low sensitivity) are black. Performance below 85% correct responses at the highest coherence level (no sensitivity) was set to a coherence threshold of 100%. The fit error in the fourth and fifth column is also color scaled coded, where blue shades indicate overestimation of the coherence thresholds, and red shades indicate an underestimation. The intersections between horizontal and vertical gray or black dotted lines indicate the combinations of step size and interval that were presented (first column), or the combination used for plotting the fit and fit error. The fitting procedure is described in the text.

comparison to the measurements (first column) shows that the data can be described quite accurately with independent spatial and temporal tuning curves. ( $R^2$ -values were 0.931, 0.822, 0.925, 0.881 and 0.841 for subject RB, SS, ML, MS and ES, respectively). Oval-shaped tuning as for RB, but also asymmetric spatial tuning as for observer MS are reproduced accurately. Independent step size and temporal interval tuning, however, fails to capture any oblique effects, such as observed for SS. The fourth column in figure 4.3 quantifies the differences between experimental data and separable tuning. Differences are given as log values. The maximum difference of about 0.6 log unit for observer SS corresponds to a factor of 4.0 in sensitivity. Although deviations from independence are relatively small, the data do show a general trend. Sensitivities for combinations of large temporal intervals (16-90 ms) and small step sizes, as well as small temporal intervals and larger step sizes (6-30 min of arc) were underestimated. Combinations of large intervals and large steps, as well as small intervals and small steps, were overestimated. To assess the nature and significance of spatio-temporal interactions we extended the model to include spatio-temporal interactions. We used an F-test to compare the fits of the two models ([http://www.graphpad.com/curvefit/2\\_models\\_\\_1\\_dataset.htm](http://www.graphpad.com/curvefit/2_models__1_dataset.htm)). It provides a quantitative estimate of the significance of the increase in  $R^2$ -value, given the change in degrees of freedom due to additional parameters. The test calculates the chance (p-value) that the dataset fits the more complicated model better, if the simpler model is in fact correct. P-values are based on the F-ratio, given by

$$F = \frac{(SS_1 - SS_2)/SS_2}{(df_1 - df_2)/df_2} \quad (4.5)$$

where  $SS$  symbolizes the sum-of-squares,  $df$  the degrees of freedom, and the subscripts identify the simpler model (subscript 1) or the more complicated model (subscript 2). P-values below 0.05 were taken to indicate significant improvements of the model fits.

We tested several different dependences between spatial and temporal tuning functions. The interaction model that gave the most significant improvement of the fit included a shift of the spatial optimum ( $\mu_s$ ) with temporal interval, and a shift of the temporal optimum ( $\mu_T$ ) with step size. We compared first order (linear), second order (quadratic) and third order (cubic) shifts of spatial and temporal optima. For all observers, except ML, increasing the order resulted in significantly better fits.  $R_2$  values increased from an average value of 0.880 for no interactions to 0.922, 0.941 and 0.957 for linear, quadratic and cubic interactions respectively. Results for fits with a third order shift of optimal interval as well as optimal step size, are shown in figure 4.3 (third column). The last column in figure 4.3 quantifies the fit errors for the interaction model. In general, temporal optima decreased with increasing step sizes, although the decrement was less for larger step sizes. For observers MS, the temporal optimum was constant, irrespective of step size. Spatial optima tended to decrease with increasing temporal interval, except for ML who showed little interactions.

A variation in the width of the tuning curves did not provide significant improvements. Linearly shifting the tuning curves along the spatial or temporal axes, which is slightly different from varying the  $\mu$  parameter, did provide significant improvements, but these improvements were smaller than for the shift in  $\mu_s$  and  $\mu_T$ .

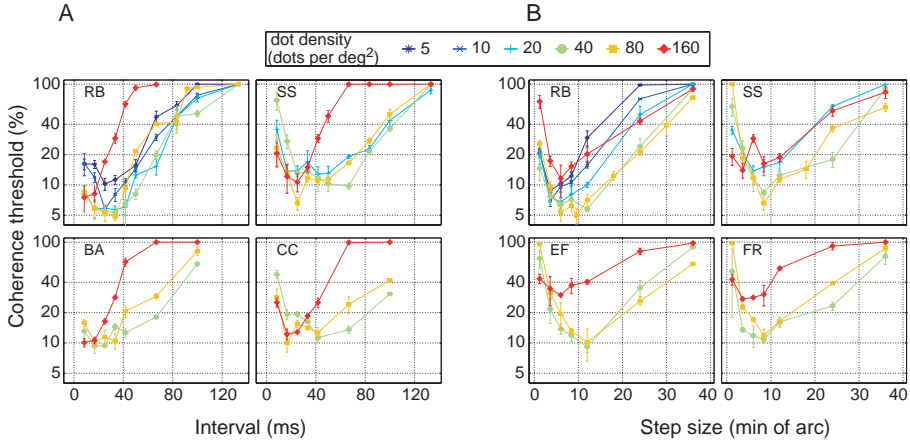
### 4.4.2 Dot density

One obvious difference between our stimuli and those of Morgan and Ward (1980) and Fredericksen et al. (1993) was dot density. In the experiments described so far, we used 80 dots/deg<sup>2</sup> per frame, at a monitor refresh rate of 120 Hz, in an 8°x8° window. Morgan and Ward (1980) used a relatively low dot density (13000 points per 728 ms, in a 2.25° x 2.25° window). A single point subtended 2.4 min of arc and was present for less than 56  $\mu$ sec. Fredericksen et al. (1993) on the other hand used much higher dot densities. Variations in dot density might therefore play a role in comparing our data to those of others. To gain insight into the effects of dot density in our stimulus we performed several control experiments at different dot densities. Figure 4.4A shows measurements of temporal tuning curves for densities ranging from 5 to 160 dots/deg<sup>2</sup>. Figure 4.4B shows similar measurements for spatial tuning curves. Temporal tuning was measured at the optimal step size of 7.2 min of arc, and spatial tuning at the optimal temporal interval of 33 ms. Measurements for 80, 40 and 20 dots/deg<sup>2</sup> did not differ substantially. A dot density of 80 dots/deg<sup>2</sup> thus seemed sufficient to reach optimal sensitivity. An increase to 160 dots/deg<sup>2</sup> dots resulted in a sharper fall off for intervals above 16.7 ms. Yet, for intervals of 16.7 ms and lower, thresholds were similar to the condition with 80 dots/deg<sup>2</sup>. A reduction to 10 and 5 dots/deg<sup>2</sup> had little effect for intervals beyond 42 ms. Threshold for intervals smaller than 42 ms were higher when compared to the condition of 80 dots/deg<sup>2</sup>. Nonetheless, the optimal interval was not affected. Figure 4.4B shows that step size tuning for 80 and 40 dots/deg<sup>2</sup> were similar. For higher and lower dot densities thresholds were raised, and optimal step sizes shifted towards lower values. In summary, we conclude that dot density may affect both spatial and temporal tuning. However, the value of 80 dots/deg<sup>2</sup> did not seem to limit performance in our experiments.

### 4.4.3 Viewing distance

Two other obvious differences between our stimuli and those of Morgan and Ward (1980) and Fredericksen et al. (1993) were dot size and viewing distance. Previous reports have shown significant effects of viewing distance on tuning for step size. Van de Grind et al. (1992) showed that this resulted in distance invariance: if step sizes were expressed in monitor pixels, i.e. object properties rather than visual angles, different viewing distances gave the same result. Changing the viewing distance affects the spatial frequency content of the stimulus, as well as the step size expressed in visual angles. Together these effects supposedly caused





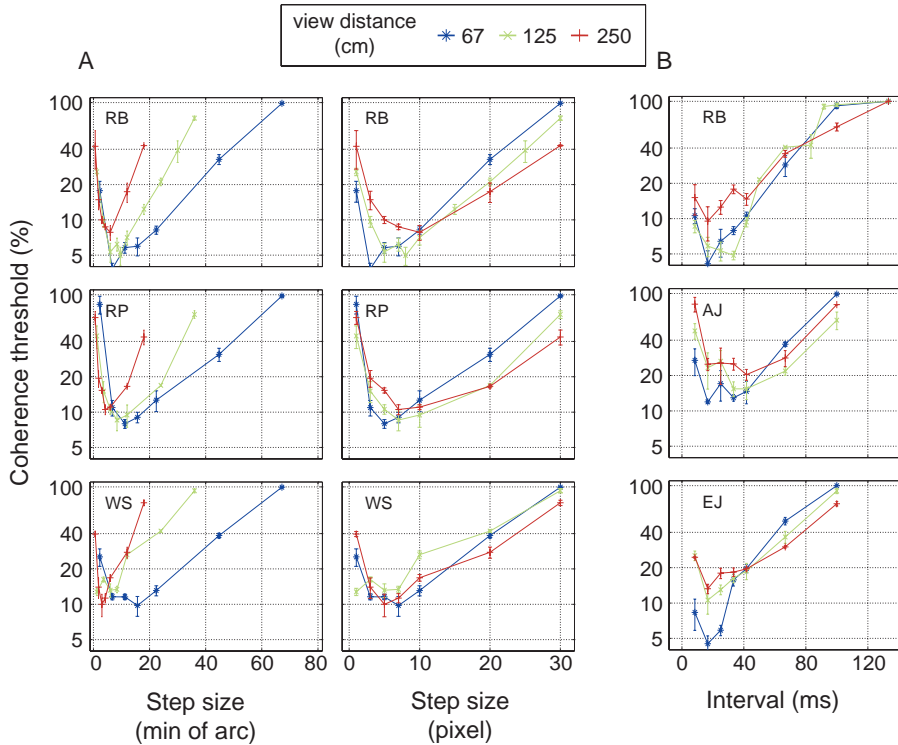
**Figure 4.4:** Performance as function of dot density: Coherence thresholds are shown as function of temporal interval (A) and step size (B) for different dot densities. Temporal tuning was measured with a step size of 7.2 min of arc and spatial tuning with an interval of 33.3 ms. Different panels show results for different observers.

distance-invariance. In a second control experiment, we checked whether this rule also holds for the stimuli used in the present set of experiments, and to what extent temporal tuning depended on viewing distance.

Figure 4.5 illustrates step size tuning measured at three different viewing distances for an interval of 33 ms. The left column shows the data with step sizes given in visual angle, the right column with step sizes expressed in monitor pixels. The latter curves nearly overlap for different viewing distances, which supports previous findings of distance-invariant coherence detection. Figure 4.5B shows temporal tuning curves measured at three different viewing distances for an optimal step size of 5 pixels. None of the observers showed a clear change in tuning curve with viewing distance: Optimal intervals as well as the fall-off at the high end seemed unaffected.

## 4.5 Discussion

We used a continuous apparent motion stimulus with two-frame, single-step, pattern lifetime to specifically investigate the requirements for correlation detection in moving random dot patterns. Variations in step size and temporal interval affected neither low-level luminance information, nor higher-level temporal integration of motion signals. Moreover, the total amount of motion energy in the stimulus was independent of step size and interval. Therefore, our results directly reflect the effect of variations in spatio-temporal correlation on low-level motion detection. By combining single-step, two-frame motion with coherence thresholds



**Figure 4.5:** Performance as a function of viewing distance: Coherence thresholds are shown as a function of step size (A) and temporal interval (B), for different viewing distances. Temporal tuning was measured with a step size of 5 pixels, which corresponds to 0.6, 1.2 and 2.4 min of arc at a viewing distance of 67, 125 and 250cm, respectively. The step size tuning was determined with an interval of 33.3 ms. Different panels show results for different observers. Stimulus conditions were the same for all viewing distances. As a result, step sizes expressed in visual angles varied. The left hand column shows step sizes expressed in visual angle, the right hand column in monitor pixels.

we were able to measure optimal values for intervals between motion steps, and the decline for longer intervals. We found optimal interval values between 17 and 42 ms, and a steep fall-off of sensitivity for larger temporal intervals. As a result, we found sharp upper limits for the temporal interval, similar to the findings by Morgan and Ward (1980).

Our data clearly differ from those presented by Fredericksen et al. (1993), who reported efficient coherence detection up to much longer intervals. One obvious reason for this discrepancy might have been a difference in mean luminance. It is well documented that both spatial and temporal properties in coherence detection strongly depend on mean luminance and the state of light adaptation (Eagle and Rogers, 1997; Lankheet et al., 2002, 1991; Morgan and Ward, 1980; Van de Grind

et al., 1987). At high luminance levels low-level (retinal) visual processing speeds up, which results in a shift towards shorter intervals in coherence detection. In the present study, we used black dots on a white background (37 cd/m<sup>2</sup>), which was practically comparable to the mean luminance in the study by Fredericksen et al. (50 cd/m<sup>2</sup>). If anything, we might have expected slightly longer optimal intervals and step sizes. The discrepancy cannot be explained by differences in dot density or viewing distance either. Our control experiments showed that the shape of the temporal tuning curves for coherence detection did not vary drastically with viewing distance. Similar to previous reports in which  $D_{max}$  measurements were found to depend on various stimulus parameters, including dot density, varying the dot density in our experiments did affect both spatial and temporal tuning. However, increasing the dot density reduced rather than increased the high temporal fall-off. We conclude, therefore, that low thresholds for long intervals in the study of Fredericksen et al. (1993) resulted from the presence of multiple correlations in their stimuli with long intervals. Because in their stimuli patterns remained stationary between displacements, at large temporal intervals correlations were also present at shorter intervals. This might very well explain why observers remained highly sensitive to long intervals in their stimulus, whereas in our stimuli sensitivity for long time intervals sharply declined.

### 4.5.1 Spatio-temporal separability

Because we measured full spatial (step size) and temporal (interval) tuning profiles our data critically test the dependence between spatial and temporal properties for coherence detection. We found that a fully separable model described the data fairly well. However, a variation of spatial tuning with temporal interval and a variation of temporal tuning with step size did provide a significantly better fit to the data. The general trend for five subjects was that the temporal optimum was inversely proportional with step size, and the spatial optimum was inversely related to interval. The reduction in fit error relative to a fully separable model was, however, relatively small.

The interaction effects were fairly subtle, which probably explains why Morgan and Ward (1980), based on upper and lower limits, found that step size and interval tuning were independent. Baker and Braddick (1985) reached the same conclusion based on two frame apparent motion experiments. Motion detection in displacements of random dots was separably dependent on the displacement from one exposure to the next, and on the time interval between exposures. They found no evidence for velocity-related dependences. Other studies reported more complex tuning properties, where spatial and temporal tuning very much depended on pattern velocity. Anderson and Burr (1985) used sinusoidal gratings to show two types of temporal tuning in motion detection, low-pass and band-pass in the range of 7-13 Hz. Temporal frequency tuning in their study very much depended on spatial frequency. Boulton and Baker (1993) used Gabor function micro-patterns to

study spatio-temporal properties for motion mechanism. Depending on stimulus onset asynchrony (SOA) they found fundamentally different behavior. At short SOAs, the mechanisms were more or less linear and could be described by Fourier motion, whereas at larger SOAs the mechanism behaved clearly nonlinear.

It should be noted, though, that step size and interval tuning for moving random pixel arrays cannot directly be compared to spatial and temporal frequencies in gratings or Gabors. Moving random pixel arrays contain a wide range of spatial frequencies, moving at the same 'speed'. Temporal frequencies vary for different spatial frequencies, and hence step size for moving patterns translates into different temporal frequencies for different spatial Fourier components. Speed-related spatio-temporal dependences for gratings therefore do not necessarily contradict independent tuning for step size and interval for random dot patterns.

Several previous studies (Koenderink et al., 1985; Van de Grind et al., 1986; Van Doorn and Koenderink, 1982a,b) also reported speed-dependent step size and interval tuning for moving random pixel arrays. They measured optimal temporal intervals using temporal alternations of two oppositely moving patterns (see introduction). Optimal step sizes were measured in a comparable experiment, using spatial alternations of two oppositely moving patterns. For wide apertures (orthogonal to the direction of motion), observers perceived both directions of motion segregated into alternating bars. For narrow apertures, the two patterns were perceived to move transparently. At intermediate aperture widths, the motion percept was greatly reduced. Since correlation detection was impossible at this critical bar width, it was interpreted as the optimal step size for motion detection. Both optimal step sizes and optimal intervals were determined as a function of pattern speed. At low speeds, optimal step sizes were constant, and optimal temporal intervals varied. At high speeds this pattern was reversed. Although these speed-related variations in optimal step size and interval seem to contradict separable step size and interval tuning, their results cannot indisputably rule out inseparable tuning for step size and interval. Drawing imaginary speed lines in the (separable) contour plots of figure 4.3 straightforwardly illustrates this notion. Low speeds have a shallow slope and cut the sensitivity surface at combinations of small step sizes and large intervals. High speeds on the other hand cut the surface for short intervals and large step sizes. Moreover, the deviations from separability that we found are compatible with the speed-dependent relationship as reported by Van de Grind et al. (1986). The important conclusion we can draw from this comparison is that the data by Van Doorn and Koenderink (1982a) and Van Doorn and Koenderink (1982b) do not necessarily imply the existence of differently tuned detectors. The same pattern could, in principle, result from separable step size and interval tuning, i.e. based on a single type of detector which is broadly tuned for step size and for interval. However, given the pattern of spatio-temporal interactions, and the regularity of fit errors in our data (figure 4.3) it does seem more likely that different combinations do play a role.

# Chapter 5

## Sensitivity to reverse-phi motion

Roger J.E. Bours, Marijn C.W. Kroes and Martin J. Lankheet

*Vision Research* (2009) 49 1-9

### 5.1 Abstract

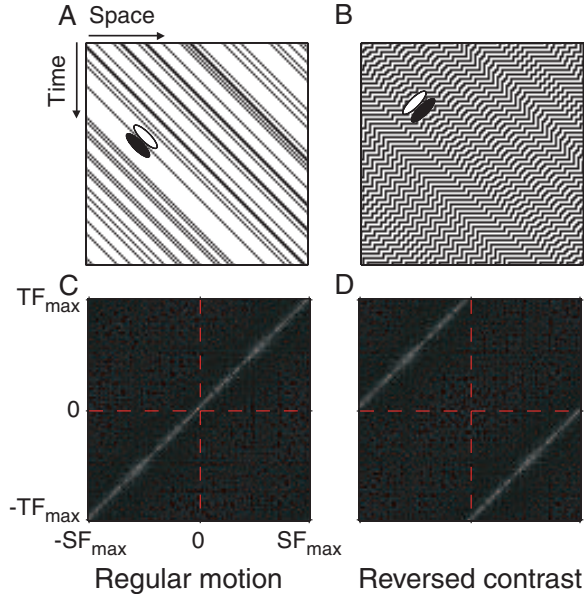
Low-level contrast information in the primary visual pathway is represented in two different channels. ON-center cells signal positive contrasts and OFF-center cells signal negative contrasts. In this study we address the question whether initial motion analysis is performed separately in these two channels, or also through combination of signals from ON and OFF cells. We quantitatively compared motion coherence detection for regular and for reverse-phi motion stimuli. In reverse-phi motion the contrast of a pattern flips during displacements. Sensitivity is therefore based on correlating positive and negative contrasts, whereas for regular motion it is based on correlating similar contrasts. We compared tuning curves for step size and temporal interval for stimuli in which motion information was limited to a single combination of step size and interval. Tuning for step size and temporal interval was highly similar for the two types of motion. Moreover, minimal coherence thresholds for both types of motion matched quantitatively, irrespective of dot density. We also measured sensitivity for so called no-phi motion stimuli, in which the contrast of displaced dots was set to zero. Sensitivity for no-phi motion was low for stimuli containing only black or only white dots. When both dot polarities were present in the stimulus, sensitivity was absent. Thus, motion information based on separate contrasts was effectively canceled by a component based on different contrasts. Together these results show equal efficiency in correlating dots of opposite contrast and of similar contrast, which strongly suggests efficient detection of correlations across ON and OFF channels.

## 5.2 Introduction

Periodic reversals of the contrast of a moving pattern cause a reversal of perceived motion direction (Anstis, 1970; Anstis and Mather, 1985; Anstis and Rogers, 1975). Classical explanations for this reversal have been based on the fact that contrast reversals simply shift the balance of motion energy to the opposite direction. The argument is easily illustrated with changes in space-time plots due to contrast reversals. Figure 5.1A gives an example of rightward apparent motion of a random dot pattern of unlimited dot lifetime. In a space-time plot this type of motion shows up as an oriented pattern. In the space-time plot of figure 5.1B the same stimulus is combined with contrast reversals between displacements of the pattern. Obviously, the main orientation in the space-time plot has shifted 90 deg, indicating a reversal of motion direction, which is what observers report when presented with this motion stimulus. Figure 5.1C and D show the Fourier transforms of these stimuli. In this representation the upper right and lower left quadrant represent motion energy for rightward motion, and the other quadrants for leftward motion. Evidently, reversing the contrast of the pattern at successive frames results in a shift of motion energy to the complementary quadrants.

However, these representations of the stimulus do not explain how detection of these two types of motion comes about, and the low-level mechanisms responsible for detecting reverse-phi motion have remained controversial. In one view, the reversal is trivial because it is embedded in the stimulus and therefore requires no specific mechanisms. The argument is that any suitable motion detector would indicate the reversal of direction (Adelson and Bergen, 1985; Krekelberg and Albright, 2005; Van Santen and Sperling, 1985). The idea is illustrated by the receptive field profiles of presumed motion energy detectors for rightward and leftward motion (figure 5.1). Obviously, the detector for rightward motion better matches the regular motion stimulus, whereas the leftward detector better matches the reverse-phi stimulus.

It should be noted though that Fourier transforms do not take into account that right from the retina contrast information is represented in two different channels. ON-center cells signal positive contrasts whereas OFF-center cells signal negative contrast. The two channels diverge at the level of retinal bipolar cells, and remain functionally segregated until convergence in primary visual cortex (Schiller, 1982, 1992; Schiller et al., 1986). Each channel provides a half-wave rectified representation of the visual image. In reverse-phi motion stimuli spatio-temporal correlations occur for dots of opposite contrast polarity, that are presumably most effectively represented in different channels. In regular motion stimuli correlations occur for dots of similar contrast. Despite the description in terms of Fourier components one can therefore ask the question whether and how information from ON and OFF cells gets combined to generate motion sensitivity. Mo and Koch (2003) studied physiologically realistic models of low-level motion detectors, and argued that sensitivity for reverse-phi cannot easily be explained without specifically combining signals from ON- and OFF-center cells. Instead,



**Figure 5.1:** Regular and reverse-phi motion of a random pattern of unlimited dot lifetime: Figures A and B show space-time plots for a coherently moving pattern in the rightward direction. Contrast reversals (B) on successive frames, i.e. along the time axis, change the main orientation. Panels C and D show Fourier energy transforms of the space-time plots. The origin of the spatio-temporal frequency coordinates is centered in the panel. The oriented receptive field profiles shown in A and B correspond to presumed motion energy detectors for rightward (A) and leftward (B) motion.

they proposed that signals from the different cell types get combined at the first level of motion detection. In their model, motion information based on correlations between two ON- or two OFF- cells for one motion direction is always complemented by similar correlations between ON and OFF cells in the opposite direction.

In this study we tested a simple prediction from the hypothesis by Mo & Koch, i.e., that motion detectors combine signals from ON and OFF cells. Because motion detectors are hypothesized to combine opposite-contrast correlations as well as same-contrast correlations, the hypothesis predicts equal sensitivity for regular motion and for reverse-phi motion. Moreover, it predicts that spatio-temporal tuning for these different types of motion should also be comparable. If on the other hand reverse-phi sensitivity would somehow be based on responses within separate channels we would expect lower sensitivity for reverse-phi motion than for regular motion. In this case sensitivity for reverse-phi motion would result from correlating incremental responses for one contrast polarity, and decremental responses for the opposite contrast polarity. These types of responses, e.g. an ON

cell excited by a white dot and inhibited by a black dot, differ both in amplitude and in time course (Chichilnisky and Kalmar, 2002, 2003), resulting in relatively weaker correlations.

To critically test whether motion detection is based on balanced input from both ON and OFF cells, we compared sensitivity as a function of step size and temporal interval for regular and reverse-phi motion. Several previous studies partly addressed this problem. Edwards and Badcock (1994) reported that no motion was perceived in contrast inverted random dot patterns, and concluded that information in ON and OFF channels remains segregated. Wehrhahn and Rapf (1992) similarly claimed that ON and OFF pathways form separate neuronal substrates for motion perception. Sato (1989), on the other hand showed that displacement limits for regular and reverse-phi motion were very similar. Most recently, Wehrhahn (2006) compared spatial and temporal limits using a two-flash apparent motion paradigm. His findings corroborate Sato's findings and in addition revealed comparable temporal limits. These studies, however, did not compare sensitivity, only detection limits. Moreover, variations of temporal interval also affected stimulus durations. As a result, temporal tuning properties for motion detection were always confounded with both changes in front-end luminance processing, and high-level temporal integration effects. The aim of this study was to measure full tuning curves for step size and for temporal interval, and to quantitatively compare sensitivity to regular and reverse-phi motion. To this end, we constructed continuous versions of regular and reverse-phi motion stimuli that were highly similar except for the type of correlation. Rather than patterns with an unlimited dot life time as shown in figure 5.1 we used two-frame (single step) dot lifetime stimuli. Regular motion stimuli contained a correlation bias for same-contrast polarities, in which correlations for opposite-contrast polarities were fully randomized. Similarly, reverse-phi motion stimuli contained a bias in correlations between opposite contrasts, with randomized correlations for similar contrasts. Variations in step size or temporal interval did not affect the spatio-temporal density of motion information, nor the duration or size of a stimulus. In combination with a variation of coherence levels these stimuli therefore allowed for a fair and quantitative comparison of tuning properties.

We found nearly equal sensitivity for regular and reverse-phi motion, and comparable step size and temporal interval tuning. Our findings therefore suggest that motion detectors effectively combine information within and across ON and OFF channels.

## 5.3 Methods

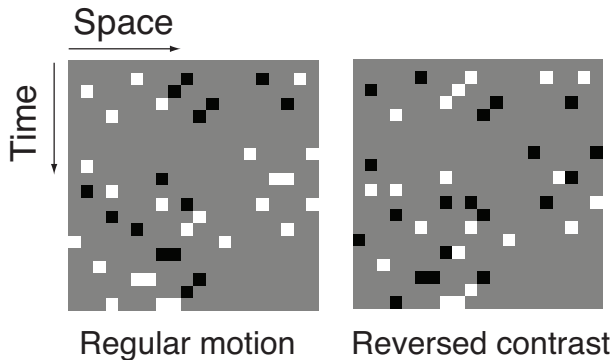
### 5.3.1 Stimuli

Visual stimuli consisted of sparse random dot patterns in which consistent motion information was either restricted to equal polarity correlations (regular motion) or



to opposite polarity correlations (reverse-phi motion). Stimuli consisted of 5000 dots ( $0.02^\circ \times 0.02^\circ$ ), presented in a window of  $8^\circ \times 8^\circ$  on a gray background of  $50 \text{ cd/m}^2$ . Half of the dots were brighter and half were darker than the mean background (contrast 96%).

Motion information was confined to a single combination of step size and temporal interval. To this end we used single-step dot lifetime stimuli similar to those described in a previous study (Bours et al., 2007b). A new set of random dots was generated and shown on each monitor frame (@120 Hz), and these dots were shown once again at a displaced location after a specified temporal interval. New and displaced dots were shown transparently. Figure 5.2 illustrates this type of motion. It shows a blow-up of a space-time diagram for a  $20 \times 20$  pixel matrix with black and white dots on a gray background. Dots are displaced once, with a step size of 2 pixels and temporal interval of 2 frames, and then randomly refreshed. Notice that, for temporal intervals longer than 1 frame the two instances of a dot are separated by frames in which newly generated dots are shown as well as displaced counterparts of previously generated dots. Interleaving correlated dots with uncorrelated dots clearly does not prevent motion detection: motion is easily perceived for temporal intervals up to at least 8 frames (66ms).



**Figure 5.2:** Space-time plot of a continuous, single-step dot lifetime motion stimulus: The diagram shows a row of 20 pixels at 20 consecutive frames. Stimuli consisted of sparse random dot patterns with 50% black and 50% white dots on a gray background. Dots were shown only twice: the first time when they are newly generated, and a second time at a displaced location after a specified temporal interval. In this example dots were displaced two pixels rightward after two frames. New dots were generated on each frame of the monitor. For regular motion the corresponding dots have the same contrast polarity, for reversed contrast a dot pair consists of a black and a white dot.

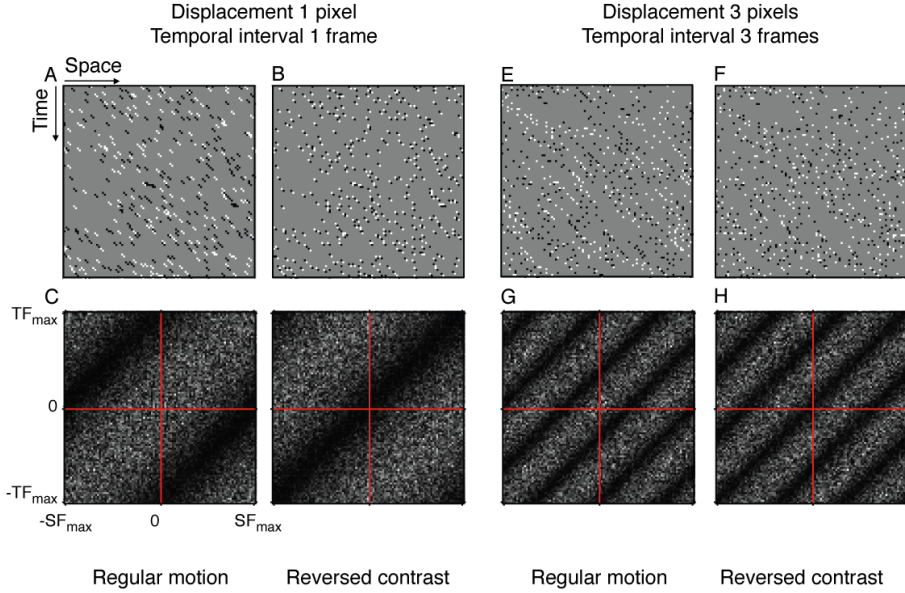
Each frame contained a combination of newly generated random dots and coherently displaced dots. This generates the same motion information on every frame of the monitor, and limits consistent correlations to the specified displacement and interval, whereas all other combinations are fully randomized. Correlations over multiple steps in the apparent motion direction, which are abundant in

unlimited lifetime stimuli, are completely randomized across motion directions. In contrast to other forms of continuous reverse-phi motion stimuli, our reverse-phi stimuli therefore contained no regular motion information at multiple time steps. Also, because dots were not visible during their temporal displacement interval, no net correlation existed for intervals shorter than the specified temporal interval. Because dots were refreshed on each frame of the monitor, irrespective of the temporal interval for displacements, variations in temporal interval did not affect the temporal frequency content of the stimulus. Changes in motion sensitivity with increasing temporal interval were, therefore, not confounded with changes in temporal frequency content or changes in temporal integration. Because each frame contained equal amounts of bright and dark dots varying the motion parameters or type of motion did not affect the temporal frequency content of the stimulus.

For regular motion, the two instances of a dot had the same polarity whereas for reverse-phi the polarity of dots was reversed upon displacement. Regular motion stimuli, therefore, contained no consistent correlations for combinations of opposite polarity, and for reverse-phi no correlations were introduced for dots of equal polarity. Notice, however, that introducing a correlation bias for same-contrast polarities induces a small imbalance for opposite-contrast polarities because presence of a same-contrast correlation automatically implies absence of an opposite-contrast correlation. In a control experiment we will investigate the relevance of this imbalance.

To measure sensitivity we varied motion coherence levels by changing the percentage of dots moving coherently. Incoherent dots were given a new random position upon displacements. At 100% coherence all reincarnated dots, i.e. half of all dots, were coherently displaced. By varying coherence levels rather than contrast, motion strength was manipulated without changing non-directional (local) contrast signals at levels before motion detection.

Figure 5.3 illustrates the two types of motion. Figures 5.3A-D illustrate motion of 1 pixel per 1 frame and Figures 3E-H show the example of 3 pixels per 3 frames. Both examples correspond to the same speed at maximum coherence level. Space-time plots reveal the noisy character of the stimuli, due to the limited dot lifetime. For regular motion, orientation is easily discerned in the space-time plots. For reverse-phi stimuli hardly any consistent orientation is perceived, yet observers could easily indicate the direction of motion in this type of stimulus. Figure 5.3C and D show the Fourier transforms of the space-time plots. In both cases the Fourier transform reveals a pattern of similar orientation. For regular motion this is a sinusoidal grating pattern with maxima on a line through the origin. For reverse-phi motion the pattern is shifted by 180 degrees, with minima on a line through the origin. For a displacement of 1 pixel per frame the spatial and temporal transforms contain a single period of a sine wave. For larger displacements the number of periods grows correspondingly. In these stimuli the total motion energy for leftward and rightward motion is nicely balanced.



**Figure 5.3:** Two-frame, single-step dot lifetime motion. Figures A and B show space-time plots for regular and reverse-phi variants of the stimulus. Figures C and D show the corresponding Fourier energy plots. The origin of the spatio-temporal frequency coordinates is centered in the panel, as indicated by the red lines. Figures E-H show the same plots for different motion parameters. A-D: 1 pixel / 1 frame; E-H: 3 pixels / 3 frames. Dot densities are comparable to those in the actual stimuli, but for clarity only a quarter (100 pixels) of the actual width of the display is shown.

### 5.3.2 Measurement procedure

We used a left-right motion direction discrimination task in a two-alternative forced-choice procedure. Coherence thresholds were determined using a Quest adaptive staircase (Watson and Pelli, 1983), converging at 85% correct performance. For regular motion a correct response corresponded to the direction of displacement, whereas for reverse-phi ‘correct’ responses corresponded to the opposite direction. Staircases consisted of 40 trials of 1 s duration, in which the direction of displacements was randomly chosen to the left or to the right. Observers started a trial by pressing a key on the keyboard and responded by pressing the left or right arrow key. No feedback was given on the correctness of responses. Mean threshold values and standard errors were based on 3-5 repetitions for each condition, and all conditions were presented in randomized order.

## 5.4 Results

### 5.4.1 Regular and reverse-phi motion

Temporal and spatial tuning curves were measured for two naive observers and for the three authors. Figure 5.4 shows results for two individual observers, and the mean for all 5 subjects. Temporal tuning curves were measured for a step size near the optimum for all observers (5 pixels, 0.1 deg, Bours et al., 2007b). Step size tuning curves were measured at three temporal intervals (1, 3 and 8 frames; 8.3, 25 and 67 ms). We found that, for the full range of step sizes and temporal intervals, thresholds for both types of motion were highly similar. Highest sensitivities, corresponding to the lowest coherence thresholds, matched and occurred at the same combination of step size and temporal interval. There was a significant tendency for lower sensitivity for reverse-phi at large temporal intervals and large step sizes, but on average this effect was relatively small. In the discussion we will come back to these differences.

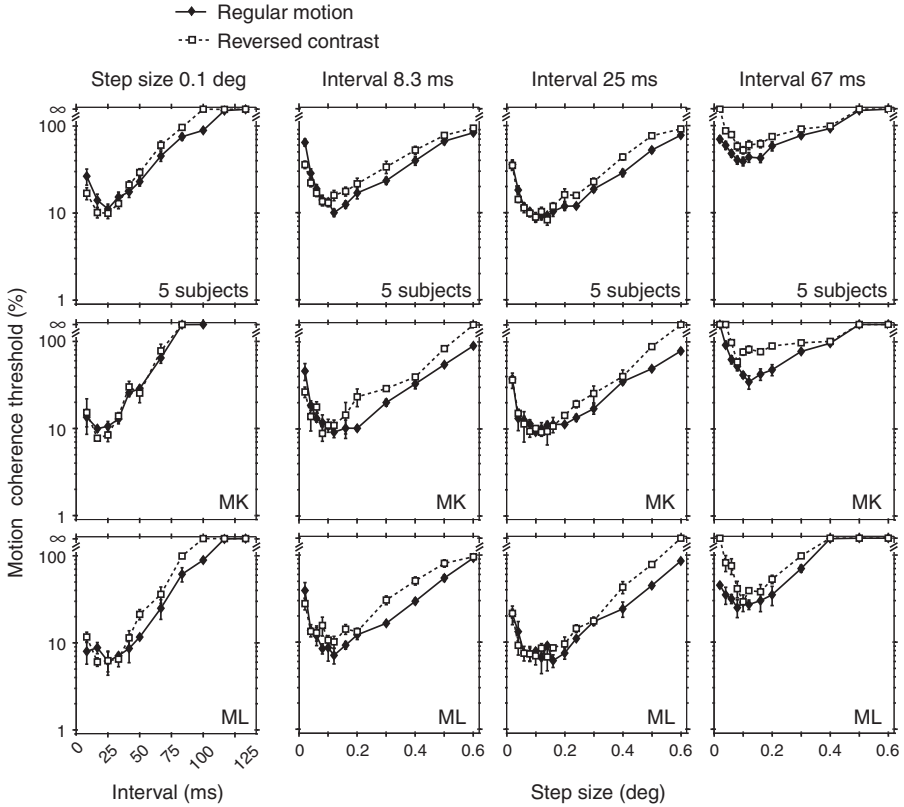
The optimal temporal interval was about 25 ms (3 frames) both for regular and for reverse-phi motion. At the lower end, sensitivity dropped fairly steeply, but the maximal frame rate of 120 Hz in our setup did not allow us to determine the lower temporal limit. The upper temporal interval limit was between 75 and 100 ms. Notice that this upper limit was measured with dynamic noise in between the first and second presentation of a dot, and therefore provides a lower bound to the upper limit. Still, this value is much higher than the limit reported by Wehrhahn (2006).

Optimal step sizes for both types of motion were about 7 min of arc. Upper displacement limits were about 36 min of arc, which is also higher than reported by Wehrhahn. Our data show a gradual decline of sensitivity for both types of motion, indicating that there is no obvious reason to propose a qualitatively different mechanism for detecting these large step sizes.

The high similarity in sensitivity and tuning properties for regular motion and reverse-phi motion indicates that motion mechanisms combine positive and negative contrasts as efficiently as similar contrasts. This strongly suggests equality of the low-level contrast responses being used in both cases, which is only possible if correlation for reversed contrast stimuli occurs across different channels.

### 5.4.2 No-phi motion

An alternative explanation for the reversal of motion direction when the contrast is periodically inverted might be that no specific correlation is required at all. In this view, flipping the contrast removes positive correlations at the specified combination of step size and temporal interval, and therefore makes the opposite displacement relatively more frequent. For example, if rightward displacements of a dot were consistently absent because the contrast inverts, displacements in



**Figure 5.4:** Temporal and spatial tuning for regular and reverse-phi motion: Left-right discrimination thresholds were determined by varying the percentage of dots moving coherently, in a staircase tracking the 85% performance level. Each condition was repeated three times. Data are shown for two individual observers (error bars: SEM,  $n=3$ ), and the average for 5 observers (error bars: SEM,  $n=5$ ). Temporal tuning curves were measured using a step size of 0.1 degrees (5 pixels) and step size tuning curves at three different temporal intervals.

the leftward direction would be relatively more frequent, due to random noise. Because motion is opponently organized, such a reduction in same-contrast correlations might tip the balance towards the opposite direction and hence induce a reversal of perceived direction. If removing same-contrast correlations were the main point, this could also explain the similarity in tuning properties for regular and reverse-phi motion, because exactly the same motion detectors would be involved.

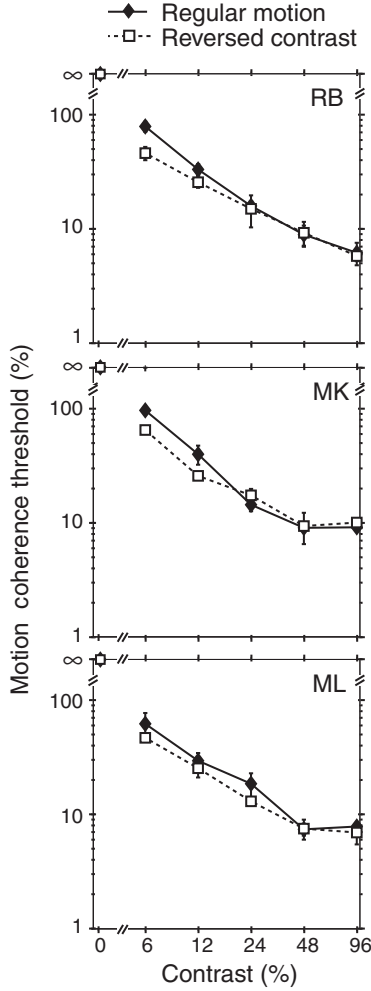
We investigated this hypothesis by creating a stimulus in which the second instance of a dot (i.e. the displaced reincarnation of a previously shown dot) was drawn in the background luminance. Notice that this is different from not drawing the dots at all. Drawing in the background color removes any dots drawn at that location and thus removes specific correlations, without introducing opposite-contrast correlations. One could call this a no-phi stimulus: the contrast is neither the same nor inverted. Such a no-phi stimulus removes specific same-contrast correlations in the same way as a reverse-phi stimulus does, but does not introduce opposite contrast correlations.

In a control experiment we measured the effect of decreasing the contrast of displaced dots, from the maximum value of 96% down to the no-phi condition of 0% contrast. We measured sensitivity for regular and for reverse-phi motion for an optimal combination of step size (5 pixels, 0.1 deg) and temporal interval (3 frames, 25 ms). The contrast of the second instance of a dot was reduced from 96% to 0%, while the contrast of the first occurrence of a dot was held constant at 96%. All other parameters were identical to those in the previous experiments. If removal of correlations from the stimulus caused direction reversals in reverse-phi stimuli, we would expect observers to be quite sensitive to no-phi motion, and results to be independent of the contrast of the second dot.

Results for this control experiment were very clear: at zero contrast none of the observers could consistently indicate the direction of this type of no-phi motion. Everyone performed at chance level, even at the maximum coherence value. Figure 5.5 shows results for different contrasts of displaced dots. Reducing the contrast from 96 to 48% had little effect on motion coherence thresholds. Further reducing the contrast, however, progressively deteriorated motion coherence detection. At zero contrast (the no-phi condition) motion was observed neither in the direction of displacement nor in the opposite direction.

We conclude that lack of correlations does not play a large role in explaining sensitivity to reverse-phi motion. If it was a significant contribution the effect of contrast for the reversed-contrast stimuli should have been very small, and the contrast curve should have been flat. Instead, we see that sensitivity falls only slightly less steeply for reversed contrast motion than for regular motion. Clearly, it is not just the absence of a consistent correlation that matters, but also the quantitative contrast mismatch between corresponding dots.

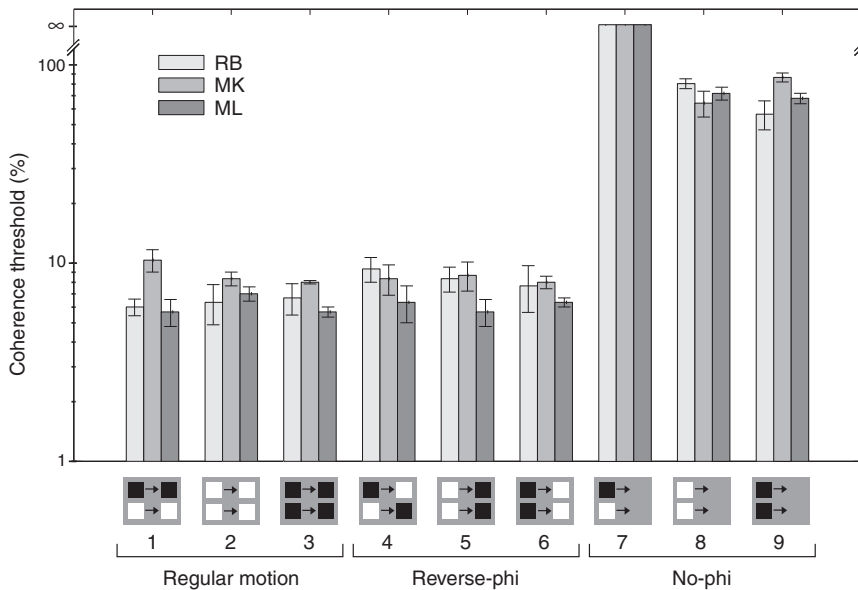
The results in figure 5.5 were obtained for stimuli containing both black and white dots, similar to the stimuli in the first experiment. We were surprised to find



**Figure 5.5:** Coherence thresholds as a function of contrast: The contrast of the second appearance of a dot was varied, while keeping the contrast of the first appearance at a fixed (maximum) level. At zero contrast the second appearance is drawn in the background luminance. Notice that this is different from not drawing the dots at all. Drawing in the background specifically removes correlations, and thus induces a directional imbalance. Random dot patterns consisted of black and white dots on a mean background. An infinite threshold indicates that performance did not reach the required 85% correct level. Error bars represent SEMs (n=3).

## 5.4 Results

a complete lack of motion sensitivity for the no-phi stimulus, because it clearly contains motion information. For both black and white dots there is an imbalance in spatio-temporal correlations. In order to further study why no motion was observed in the no-phi condition we repeated the experiment for stimuli that contained different combinations of black, white and background dots. Results are shown in figure 5.6. The first condition corresponds to the regular motion stimulus in which half of the dots were black and half were white. Conditions two and three are regular motion stimuli in which all dots are of similar contrast. Conditions 4 to 6 are the reverse-phi versions of the same stimuli. The last three conditions are three versions of a no-phi stimulus; with only white dots, only black dots, or a combination of black and white dots (i.e. the zero contrast condition from the contrast experiment). Notice that dot densities and mean luminance levels may be different for the different conditions, but they did not change during stimulus presentations, and did not affect the results significantly.



**Figure 5.6:** Motion coherence thresholds for regular, reverse-phi and for no-phi motion: Dots were either black or white at maximum contrast, or drawn in the background color (zero contrast). The cartoons below each column indicate the contrast of the first (left) and second (right) appearance of a dot. Zero contrast is indicated as the absence of a square. For each condition the upper and lower squares in the cartoon represent the contrasts of half of the dots. Coherence thresholds below 100% indicate performance above chance level (at least 85% correct). For the no-phi conditions with both black and white dots observers performed at chance level. Error bars show SEMs (n=3).

Highest sensitivities, i.e. lowest coherence thresholds were obtained for stimuli



in which the first and second instances of dots were either black or white. It did not matter very much whether they were black, white or flipped contrast. In all cases observers reached coherence thresholds of about 5-10%. For the no-phi stimuli containing only black or only white dots observers performed well above chance and reported motion in the direction opposite to the displacements. Thresholds were however between 70 and 90%, much higher than for standard reverse-phi motion. For the no-phi stimulus containing both black and white dots, i.e. the stimulus used in figure 5.5, observers performed at chance level, as indicated by an infinite threshold. These data show that lack of correlations may indeed cause a directional imbalance, which can be perceived as coherent motion. However, this effect is far too weak to explain sensitivity to reverse-phi motion. Moreover, the effect completely disappears if black and white dots are combined in a single motion stimulus.

In fact, the absence of motion sensitivity for mixed-contrast no-phi stimuli adds further support for the claim that motion detectors combine same-contrast polarities and opposite-contrast polarities with similar efficiency. If motion analysis would have been limited to similar contrasts, without cross talk between the two channels, it is difficult to explain why the motion percept vanished if black and white dots were combined in a stimulus. Absence of correlations would signal a similar motion direction for disappearing black and white dots. However, observers do not see this type of motion, presumably because the component based on separate contrasts is cancelled by a component based on correlations between black and white dots (corresponding to motion in the opposite direction). Cancellation of motion signals in a no-phi stimulus therefore suggests that opposite polarity correlations perfectly counteract the directional imbalance created for each polarity separately.

### 5.4.3 Dot density

Our results agree with combining information from ON and OFF channels into low-level motion detectors. So far, however, we did not rule out that information from positive contrasts and negative contrasts gets combined at a level before motion detection. The extent to which this might occur depends largely on the degree of low-level spatio-temporal integration of luminance information. If receptive fields of ON and OFF cells extensively integrate positive and negative contrasts, each dot polarity will have a similar but opposite effect. In this case, both dot polarities would have nearly equal efficiency in modulating responses of both ON and OFF cells. Thus, equal efficiencies for regular and reverse-phi motion might have resulted from convergence and integration of opposite polarity responses in ON and OFF channels, i.e., prior to motion detection.

To examine the contribution of low-level luminance integration we compared sensitivity for regular and reverse-phi motion as a function of dot density. Reducing dot density effectively reduces the effects of spatio-temporal luminance integration in ON and OFF cells. As a result responses in each cell type will

be more dominated by their preferred contrast polarity. Reducing dot density should therefore reveal any asymmetries. More specifically, if low-level luminance integration would play a major role, reducing dot density should cause a much larger reduction in sensitivity to reverse-phi motion than to regular motion.

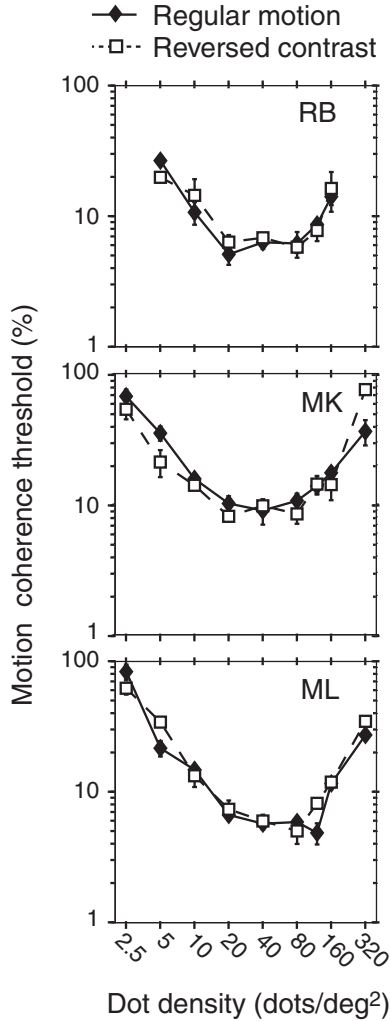
Figure 5.7 shows the results for variation of dot density. We measured sensitivity at the optimal combination of step size (0.1 deg, 5 pixels) and temporal interval (25 ms, 3 frames) for dot densities varying from 2.5 to 320 dots/deg<sup>2</sup>. Our data show no consistent signs of sensitivities diverging at lower dot densities. Observer MK shows lower thresholds for reverse-phi motion at low dot densities, which is opposite to what we would expect based on low-level luminance integration. This result makes it unlikely that positive and negative contrasts are represented equally efficient in both ON and OFF channels.

## 5.5 Discussion

The motion stimuli that we used in the present study have several advantages over those used previously. First, motion information is limited to a specific combination of step size and temporal interval, which allows for accurate and specific measurements of motion tuning properties (see also Bours et al., 2007b). Second, the stimuli allowed us to create regular motion in which no correlation bias between opposite contrasts was present, and reverse-phi motion in which no bias was present for similar-contrast correlation. For regular motion correlations between opposite contrasts were completely randomized, and for reverse-phi correlations between similar contrasts were randomized. Sensitivity and tuning for the two types of motion could therefore be compared quantitatively. Third, by using motion coherence to compare sensitivities we specifically addressed motion detection, without affecting local, low-level luminance or contrast information. Moreover, variations in motion parameters did not affect the time course of motion information: each frame contained statistically the same amount of motion information, irrespective of temporal interval. This is different from classical two-flash stimuli, in which motion information is generated at intervals corresponding to the temporal interval of displacements. Therefore, variations of inter stimulus interval always affected temporal integration of motion signals as well. In our experiments, motion information was generated on each frame of the monitor, and temporal integration of motion signals was not affected by changes in temporal interval.

Our results showed that sensitivity and tuning properties were very similar for regular and for reverse-phi motion. This observation extends the finding by Sato (1989) and Wehrhahn (2006) that displacement limits are similar for regular and reverse-phi motion. In addition to the similarity in spatial and temporal limits, our data show that absolute levels of sensitivity are nearly the same, and also the combination of step size and temporal interval at which they occur.

Our conclusion is that low-level motion detectors are equally efficient in com-



**Figure 5.7:** Sensitivity for regular and reverse-phi motion as a function of dot density: Thresholds were measured in a staircase procedure tracking the 85% correct performance level. Step size and temporal interval were fixed at optimal values (0.1 deg, 25 ms). Error bars show SEMs (n=3).

binning same polarity signals and opposite polarity signals. Similar sensitivity to regular and reverse-phi motion makes it highly unlikely that reverse-phi sensitivity is based on correlation within separate ON and OFF channels. For high-contrast stimuli, such as the ones we used, ON and OFF responses in a single channel are very different (Chichilnisky and Kalmar, 2002). For example, the response of an ON-cell to a bright dot consists of an increase in spike frequency. The response of the same cell type to a dark dot would be a reduction in spike frequency. These two types of responses differ considerably in amplitude as well as dynamics, especially at high contrast levels (Chichilnisky and Kalmar, 2002) and low dot densities. As a result our subjects should show a large difference in sensitivity if their detection of regular and reverse-phi opponent-motion were based on correlations within the same channel. We found similar sensitivities down to the lowest dot densities, and at all contrast levels. We therefore conclude that motion detection is most likely based on similar low-level responses from both ON and OFF cells, i.e. on positive contrasts exciting ON cells and negative contrasts exciting OFF cells.

Findings for the no-phi stimuli support this argument. Observers perceived motion in the so-called no-phi stimuli, but only if the stimuli contained a single contrast polarity. However, sensitivities were very low compared to the regular and reverse-phi motion stimuli. If both black and white dots were present no motion was perceived at all. This implies correlation of opposite polarities that is as efficient as correlation of similar polarities. These opposite contrast combinations generate reverse-phi motion percepts in the direction opposite to those for same-contrast combinations. The fact that the motion percept was totally absent suggests that the two motion signals were equally strong, and cancelled each other.

Our findings are at odds with many previous reports arguing against combination of ON and OFF channels in low-level motion detection (e.g. Mather et al., 1991; Wehrhahn and Rapf, 1992; Edwards and Badcock, 1994). Edwards and Badcock, for example, found no sensitivity to contrast-reversed motion in moving random dot patterns and concluded that ON and OFF channels remain segregated at the level of local motion detection. Our results, however, show that motion detection is in fact equally efficient with and without contrast reversals. Various different findings in previous experiments are likely due to differences in motion stimuli. We specifically constructed our stimuli to allow for a direct comparison of tuning properties and minimal confounds of variations in luminance integration and changes in higher-level motion integration due to changes in stimulus duration and time course of motion information. An important difference between our stimuli and those used in previous studies may be due to different contributions from second order motion mechanisms. A second order motion mechanism that discards the sign of luminance contrast (Chubb and Sperling, 1988) will respond in similar ways to regular and to reverse-phi motion. Notice that for reverse-phi stimuli the effects of first order motion and second order motion are therefore opposite, and will cancel one another. For regular motion however, the two

mechanisms yield similar signals and will add up. A small contribution from second order motion detectors could thus explain the relatively lower sensitivity for reverse-phi motion, especially because the second order system is favored by slow motion (Papathomas et al., 1996) and large displacements (Sperling, 1989). For our continuous stimuli with high degrees of dynamic noise second order mechanisms presumably play only a minor role, whereas in more classic two-frame paradigms with recognizable patterns it may play a very significant role.

The notion that sensitivity to our reverse-phi stimuli is most likely based on combining signals from ON and OFF cells does not exclude that differences in response timing may also induce illusory motion percepts. Del Viva et al. (2006), for example, recently showed that contrast-inverted Glass patterns, in which the paired dots are shown simultaneously, generate a clear motion percept. This was consistent with a small time delay in processing of positive contrasts relative to negative contrasts. The difference was on the order of a few milliseconds, which is well in line with the difference in dynamics between ON and OFF cells (Chichilnisky and Kalmar, 2002). If delay differences were compensated in the timing of bright and dark dots in the stimulus, the imbalance disappeared. This finding is not at odds with nearly similar tuning for our stimuli. Differences of a few milliseconds would cause only minor differences in temporal tuning curves because tuning for interdisplacement interval is fairly broad. The difference in response dynamics for ON and OFF cells might in fact be a second reason for the different sensitivities at large interdisplacement intervals. In this range, the interval is limiting, and an additional time delay between channels would affect sensitivity to reverse-phi motion, but not to regular motion.

Our data provide additional support for the combination of ON and OFF cell signals in generating low-level motion signals. These motion signals do not primarily result from small timing differences in responses from ON- and OFF cells, but from equal treatment of correlations within and across channels. Theoretically this makes sense, because in this way, motion detectors optimally use all available information. For a coherently moving pattern motion information is not only contained in the presence of positive correlations within separate channels, but also in the absence of correlations across ON and OFF channels. In other words, correlations between different contrasts signal the absence of motion. Combining positive and negative evidence for one direction of motion at an early level is an efficient way to improve the signal to noise ratio. Our data support the hypothesis that motion detectors use both types of information. Reverse-phi motion is then simply explained as a side effect of the second type of motion information.

This notion is in line with the general assumption in motion energy models that motion energy extraction is based on full, non rectified contrast signals. At the same time, similar sensitivity and tuning properties for reverse-phi and regular motion force one to consider how low level contrast signals are used in motion detection. A description in terms of Fourier energy does not take into account the fact that contrast is represented in separate ON and OFF channels. Moreover, it is not trivial to explain sensitivity for our stimuli based on spatio-

temporal frequency content. Global integration over different quadrants obviously does not suffice, since the total Fourier energy for leftward and for rightward motion is typically nicely balanced in our stimuli, both for regular and for reverse-phi stimuli. Figure 5.3 shows that, especially for larger step sizes and temporal intervals complementary quadrants contain equal amounts of Fourier energy. The main difference between regular and reverse-phi motion is a phase shift of the pattern. In both cases, however the total energy for leftward and rightward motion is similar. Simple linear summation of energy for the two directions therefore cannot explain sensitivity for our stimuli. To explain our findings one has to assume more complex, nonlinear combinations of different Fourier components. This conclusion is in line with findings by Krekelberg and Albright (2005). Based on single cell recordings in area MT of macaque monkeys they also concluded that linear combinations of motion energy components could not explain responses to more complex stimuli. Their data revealed complexities in the integration of motion components that are not normally incorporated in motion energy models.

Rather than specifying the complex integration of Fourier components, our experiments specifically addressed the way low-level contrast signals are used to generate directional selectivity. Our findings strongly support the notion by Mo and Koch that sensitivity for reverse-phi requires direct correlations between displaced ON and OFF cells.

Notice that the requirement for combining signals from ON and OFF cells does not specify exactly how such correlations cause a direction reversal. In a related paper we showed that reverse-phi percepts in many respects behave like motion aftereffects. This suggests that, similar to aftereffects, low-level responses to reverse-phi motion consist of low-level inhibitions. Rather than exciting detectors tuned for motion opposite to the direction of displacements, opposite contrast correlations generate inhibitions for the direction of displacements (Bours et al., 2007a). Because the motion system is opponently organized such a low level inhibition translates into an excitation into the opposite direction at a higher level (Grunewald and Lankheet, 1996). We therefore propose that reverse-phi motion results from direct combinations of ON and OFF cell signals that cause inhibition at the detection level, and due to disinhibition cause excitations at higher levels, and motion percepts for opposite directions.

# Chapter 6

## The parallel between reverse-phi and motion aftereffects

Roger J.E. Bours, Marijn C.W. Kroes and Martin J.M. Lankheet

*Journal of Vision* (2007a) 7(11):8 1-10

<http://journalofvision.org/7/11/8/>, doi:10.1167/7.11.8.

### 6.1 Abstract

Periodically flipping the contrast of a moving pattern causes a reversal of the perceived direction of motion. This direction reversal, known as reverse-phi motion, has generally been explained with the notion that flipping contrasts actually shifted the balance of motion energy towards the opposite direction. In this sense the reversal is trivial because any suitable motion energy detector would be optimally excited in a direction opposite to that for regular motion. This notion, however, does not address the question how these two types of motion are initially detected. Here we show several perceptual phenomena indicating that low-level detection of the two types of motion is quite different. Reverse-phi motion percepts in many respects behave more like motion aftereffects than like regular motion. Motion adaptation causes reduced activity during a stationary test stimulus, which by means of directional opponency leads to motion perceived in the opposite direction. Our findings suggest that reverse-phi motion similarly reduces the activity of low-level motion detectors.

## 6.2 Introduction

Reverse-phi motion is the reversal of perceived direction when the contrast of a moving pattern is inverted during displacements (Anstis, 1970; Anstis and Mather, 1985; Anstis and Rogers, 1975). It is one of the most compelling motion phenomena, similar in strength to the well-known motion aftereffect. Unlike the common account for motion aftereffects, which result from reduced activity after prolonged motion adaptation, the explanation for reverse-phi motion has remained controversial.

The classical account points out that contrast reversals shift the balance of Fourier energy towards the opposite direction, which any suitable motion detector (Adelson and Bergen, 1985; Krekelberg and Albright, 2005; Van Santen and Sperling, 1985) would pick up. In this sense the reversal is trivial and merely reflects stimulus properties rather than characteristics of mechanisms underlying motion detection. The notion that contrast reversals shift motion energy to the opposite direction, however, does not explain how, and at what level the direction reversal originates. Here we investigate whether properties of reverse-phi percepts can provide insight into the way how low-level motion mechanisms are organized.

In reverse-phi stimuli, motion information consists of correlations between contrast increments and decrements. Low-level contrast information is separated in two channels: ON-center cells in the retina and LGN signal positive contrasts and OFF-center cells signal negative contrasts. Mo and Koch (2003) studied models of low-level motion detectors, and concluded that sensitivity for reverse-phi motion cannot easily be explained by the assumption of a strictly separate processing of signals from ON-center and OFF-center cells. Instead, they propose that information from ON-center and OFF-center cells gets combined at the first level of motion detection. In this paper we address the question how correlations across different contrasts lead to perceptual direction reversals.

Mo and Koch (2003) proposed a model in which interactions between ON and OFF cells generate excitatory signals for detectors tuned to the direction opposite to the displacement. Reverse-phi stimuli thus excite detectors tuned to the opposite direction of regular motion. Notice, however, that a similar result could also be obtained if reverse-phi motion does not excite detectors for the direction opposite to the displacement, but instead inhibits detectors tuned to the ‘same’ direction, i.e. in the direction of displacement. Because motion is opponently organized, such response reductions become perceptually equivalent to response increments in the opposite direction. Under the assumption that reverse-phi results from combining information from ON and OFF cells, two different configurations are possible, as illustrated in figure 6.1. The diagrams show how local ON and OFF cell signals are combined in generating directional selectivity. The circles with solid or dashed arrows represent directionally selective operators for the different combinations of local ON and OFF cell signals, which combine into a single directionally selective cell (for instance in V1). Directionality is determined by the type of operator (combination of either same cell type or of different cell

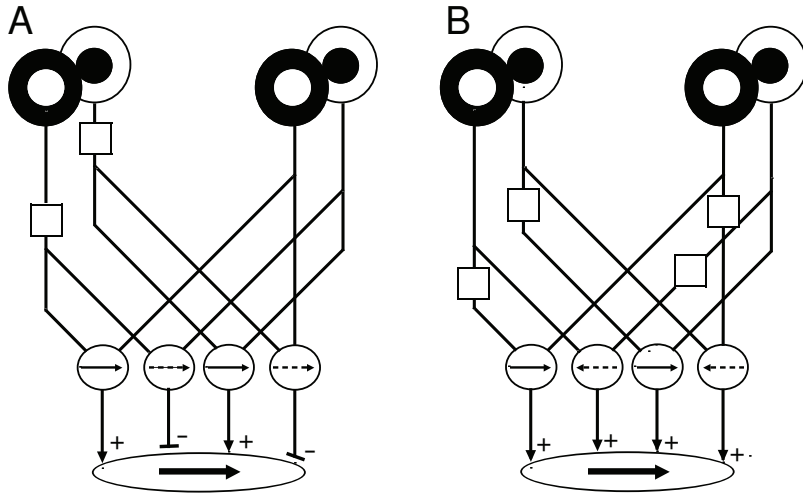


types), and by the position of the delay (one side delayed or the opposite side). This initial stage of motion detection is then followed by an integration stage (not shown) which implements integration across space and directions and which also implements directional opponency. Figure 6.1B is functionally equivalent to the scheme proposed by Mo and Koch. In this case a detector tuned to rightward motion would sum excitations from operators tuned to regular rightward motion, and contrast reversed leftward motion. Similarly, rightward reversed contrast motion would excite an operator tuned to leftward motion. In this way, both regular and reverse-phi motion leads to excitations of low-level motion detectors. Figure 6.1A shows the alternative, in which contrast reversed motion in the same direction causes an inhibition. In this case regular motion causes excitations and reverse-phi motion causes inhibitions of detectors tuned to the same direction. If this initial motion detection stage is followed by opponent directional interactions it also explains direction reversals for reverse-phi motion. Please notice that scheme A is also equivalent to first subtracting ON and OFF cell responses, and then performing a spatio-temporal correlation. As long as such combined signals are not translated into a rectifying spike train before correlation this would yield similar results. Both schemes A and B utilize the full contrast signal, as suggested in motion energy models. They differ however in the order of the operations. In scheme A directions remain initially the same but the sign inverts, in scheme B the sign remains the same but directions reverse.

In combination with opponent directional interactions as found in area MT, both schemes result in reversals of perceived direction. The two types of responses are, however, not equivalent. In fact excitations and inhibitions at the lower detection level behave quite differently, especially under transparent motion conditions. This has been extensively documented for motion aftereffects. Aftereffects result from reduced activity after prolonged adaptation, which by means of disinhibition presumably induces perceived motion in the opposite direction (Krekelberg et al., 2006b; Petersen et al., 1985; Tootell et al., 1995; He et al., 1998; Taylor et al., 2000).

Directional tuning for aftereffects, however, differs from that for direct motion vision. Two populations of dots moving in different directions are easily seen to move transparently, yet, the aftereffect of two such components is mostly unidirectional, opposite to the vector sum of the components (Mather, 1980; Alais et al., 2005; Verstraten et al., 1994). For example, adaptation to two motion vectors, one to the upper left and one to the upper right, causes a single aftereffect in the downward direction. The differences in directional tuning are easily explained if excitatory interactions are narrowly tuned, and inhibitory interactions more broadly tuned (Grunewald and Lankheet, 1996). Narrow tuning for excitations supports transparency for multiple, regular motion components. However, reduced activity in the adapted directions leads, through broad disinhibition, to a single, wide distribution during the aftereffect, which is perceived as a single, unified motion component.

Here we ask the question whether reverse-phi motion behaves like motion af-



**Figure 6.1:** Combining signals from local, ON and OFF cells into motion detectors. A minimal requirement for motion sensitivity is the combination of two signals separated in space and time. In each of the two locations we have an ON-center cell and an OFF-center cell. One can pair each local unit with a delay to one of the two units in the other pair. This provides two signals based on similar contrasts (indicated by the solid arrows), and two based on opposite contrasts (dashed arrows). The two configurations differ in the directionality and sign of the opposite contrast combinations. In Fig. A, i.e. the scheme we propose, the four directions are the same, but the sign is inverted. The alternative (B) is to keep the sign positive, but reverse directions for opposite contrast combinations. Squares represent a time delay. In scheme A reverse-phi motion leads to an inhibition of the low-level motion detector, which by means of opponency at a higher level (not shown) results in a direction reversal.

tereffects, or like real motion stimuli. To answer this question we study directional interactions between two reverse-phi motion components shown transparently. If reverse-phi motion results from reduced activity we would expect it to behave like motion aftereffects. However, if it results from increased activity it should behave like direct motion percepts.

In pilot experiments we tested this prediction with random dot patterns containing two motion components, one moving to the upper left and one to the upper right. Without contrast reversals during displacements the dots were perceived as two sheets moving transparently. The two motion components remained clearly segregated. However, when the contrast of each dot was inverted during displacements, the percept consisted of a single vector, in the downward direction. This suggests that reverse-phi motion, like motion aftereffects, is based on a reduction of activity in the same direction, rather than an increase in the opposite direction. It suggests that the scheme as presented in figure 6.1A is the most likely candidate for combining low level contrast information. We further tested

and corroborated this hypothesis in two quantitative psychophysical experiments.

## 6.3 Methods

To specifically study directional interactions for either regular or reverse-phi stimuli, we developed a stimulus in which motion information was limited to either same-polarity correlations (regular motion) or to opposite polarity correlations (reverse-phi motion). To this end we used a single-step dot lifetime stimulus similar to that used in a previous paper (Bours et al., 2007b). Dot patterns were randomly refreshed after each displacement, thus removing a correlation bias over multiple time steps.

All motion stimuli consisted of dynamic, sparse random dot patterns presented on a gray background of 50 cd/m<sup>2</sup>. Half of the dots was brighter and half was darker than the mean background (contrast 96%). A single pattern consisted of 2500 dots, displayed in an 8° x 8° window on a computer monitor (120 Hz frame rate). Stimuli consisted of two such patterns shown transparently. Dots (0.02° x 0.02°) were dynamically refreshed on every frame of the monitor, but re-appeared once again at a displaced location, and after a specified interval. Refreshed dots and displaced dots were shown on every frame of the monitor. Thus, each frame contained statistically the same amount of coherent motion information, irrespective of step size or temporal interval.

For regular motion the contrast of a dot remained the same upon displacement; white remained white and black remained black. For reverse-phi motion the contrasts were reversed; white became black and black became white. Stimuli therefore contained equal numbers of black and white dots on each frame of the monitor. Because both regular and reverse-phi stimuli consisted of equal numbers of bright and dark dots random correlations were the same for both types of stimulus. The only difference is that regular motion stimuli contain a directional bias for same-polarity correlations and reverse-phi for opposite polarity correlations.

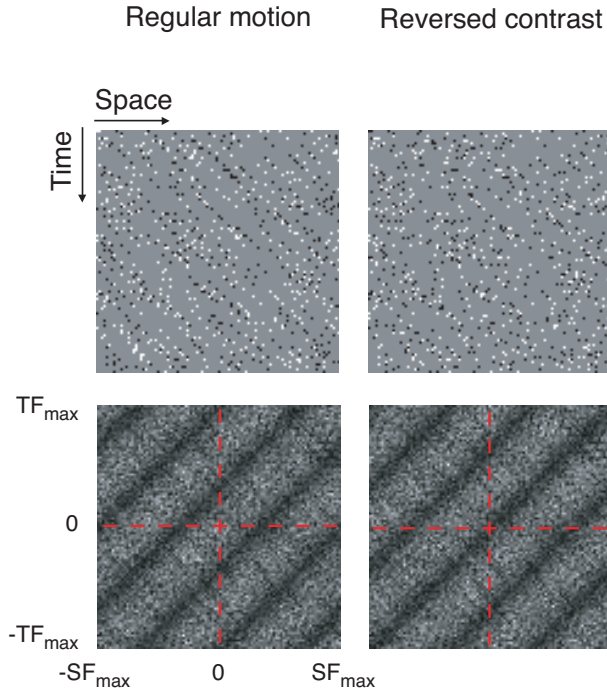
Motion coherence levels were varied by varying the percentage of dots moving coherently. Incoherent dots were displaced to a new random position. At 100% coherence all reincarnated dots, i.e. half of all dots, were coherently displaced.

Figure 6.2 shows space-time plots for both types of motion. The examples show a single pattern displaced rightward 3 pixels every 3 frames. Coherent motion corresponds to an oriented pattern in such space-time plots. Without a contrast reversal one easily perceives the orientation in the space time plot. The right-hand column in figure 6.2 shows the same motion, combined with reversals of contrast on each displacement. The only consistent correlation is the correspondence between dots of opposite contrast polarity. Spatio-temporal correlations for dots of similar polarity are maximally randomized across directions. Contrast reversals in this stimulus completely abolish the overall orientation in the space-time plot (Burr and Ross, 2006). Yet, observers are very sensitive to this type of motion, which is perceived in the direction opposite to the displace-

ments. Despite the clear differences in Fourier transforms as shown in figure 6.2 observers are nearly equally sensitive to the two types of motion.

The advantage of these stimuli is that they optimally isolate both types of motion information, and randomize unintended motion components. This is also true for two such patterns combined transparently, because the patterns are uncorrelated and contain equal numbers of dark and bright dots (see figure 6.3).

One might argue that the reverse-phi type of motion could be detected as an imbalance for same-contrast correlations since spatio-temporal correlation for opposite polarities automatically removes same-polarity correlations at the same space-time offset. The quantitative similarity in sensitivity for both types of motion, as shown for example in experiment 2 strongly argues against a significant contribution from this effect. We will return to this issue in the discussion.



**Figure 6.2:** Space-time plots, and corresponding Fourier transforms for a single component of regular and reverse-phi motion. Stimuli consisted of a set of dots dynamically refreshed on every frame of the monitor. Each dot re-appeared once again at a displaced location, and after a specified interval. The contrast between the first and second instance of a dot was either held constant (regular motion) or inverted (reverse-phi). To reduce the three dimensional (XYT) stimulus space to two dimensions we plotted a single row of (horizontal) pixels as a function of time (along the vertical axis). Spatial displacement was set to 3 pixels and temporal interval to 3 frames. Fourier transforms were averaged over 100 frames.

## 6.4 Results

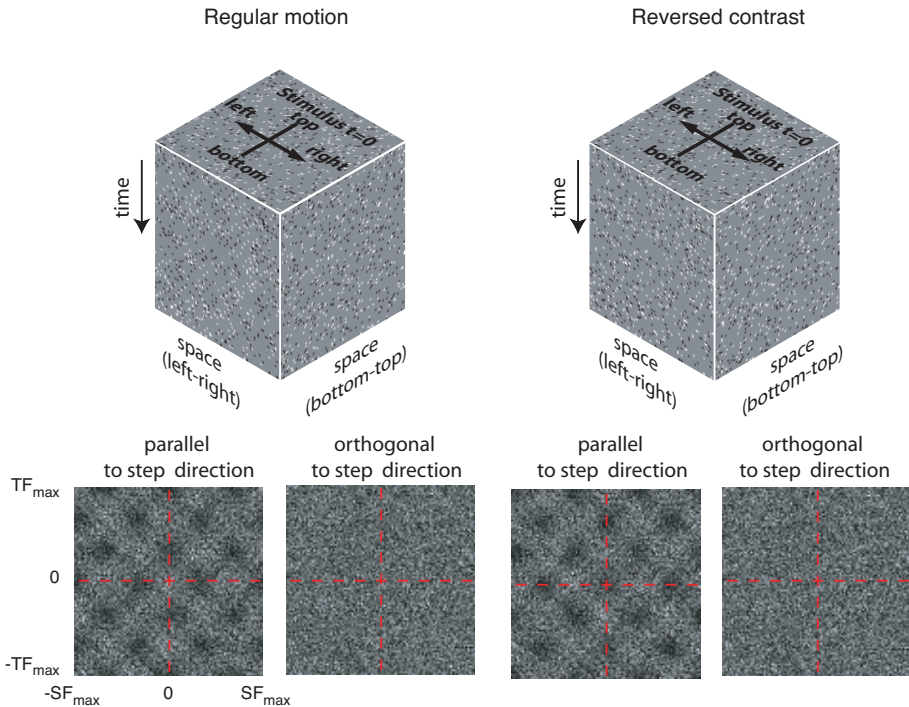
### 6.4.1 Experiment 1. Orthogonal reverse-phi motion

The experiment is based on the orthogonal motion aftereffect illusion. This illusion is perceived after adaptation to two patterns moving in opposite directions. In this case the directional interactions as previously described cancel the imbalance along the axis of stimulation, yet induce perceived motion in the orthogonal directions (Grunewald and Lankheet, 1996). If reverse-phi percepts similarly result from response reductions, then observers should also perceive two oppositely directed reverse-phi components to move along an orthogonal orientation. For example, transparent reverse-phi motion to the left and to the right should be seen to move upward and downward. Motion detectors in both stimulus directions now become inhibited, and orthogonal directions would be disinhibited. These two peaks in opposite directions, orthogonal to the displacements should then segregate like two regular, opponent motion components.

We showed observers transparent motion consisting of two opposite motion vectors, which were either two regular motion components, or two reverse-phi components. The two components could be oriented along the rightward or leftward oblique. Figure 6.3 shows space-time plots and corresponding Fourier transforms of the two types of transparent motion. Combining two patterns moving in opposite direction has no effect on motion energy in the orthogonal directions.

The task for the observer was to indicate the orientation along which motion was perceived. The experiment was constructed to minimize any perceptual difference between regular and reverse-phi motion. The only difference was the contrast reversal between the first and the second occurrence of a dot for reverse-phi stimuli, whereas for regular motion the contrast polarity for each dot remained the same. All other stimulus parameters were identical for regular and reverse-phi motion. The step size was 5 pixels (0.1 deg) and the temporal interval was 3 frames (25 ms). Both parameters were in the optimal range as determined in separate experiments. Psychometric curves were measured using a method of constant stimuli. In a single experiment we randomly interleaved regular motion and reverse-phi motion, at different strengths of the motion signal, i.e. different coherence values. Each stimulus was repeated 10 times in a single experiment, and experiments were repeated 3 times. Trials lasted 1 s, after which observers indicated the orientation along which they perceived motion (left or right oblique). For both regular motion and reverse-phi, ‘correct’ performance corresponded to the axis along which dots were displaced. Psychometric curves and 95% confidence intervals were obtained by maximum likelihood fits of cumulative Gaussians, and bootstrapping methods (Wichmann and Hill, 2001a,b).

Data for 5 subjects are shown in figure 6.4. Transparent motion of two regular components (diamonds in figure 6.4) was readily seen along the correct orientation. Performance was at chance level for low signal strengths, and increased to nearly 100% correct performance for the highest signal strengths. The combi-



**Figure 6.3:** Space-time plots of transparent motion stimuli and corresponding Fourier transforms: The left-hand side shows an example of transparent regular motion consisting of a leftward and rightward moving pattern. Space-time plots and Fourier plots are shown for the direction parallel to displacements, and orthogonal to displacements. The right-hand side shows the same analysis for two reverse-phi patterns moving in opposite directions. Combining two opponent motion stimuli clearly does not introduce motion energy orthogonal to the displacements. Fourier transforms were averaged over 100 frames.

nation of two similar, transparent reverse-phi components (squares in figure 6.4) gave a reversed result. In this case observers consistently saw the motion along an orientation orthogonal to the displacements. They judged the orientation as if it were two real motion vectors orthogonal to the actual displacements. Naive observers generally reported afterwards that they were unable to distinguish trials with regular motion from trials with reverse-phi motion.

Observers were only slightly less consistent in their answers for the reverse-phi condition. Absolute slopes of the psychometric curves were only a factor of 2.1 smaller. At maximum signal strengths observers scored on average 93% correct for regular motion, and 74% incorrect for reverse-phi motion. The orthogonal effect is much stronger than observed for motion aftereffects (Grunewald and Lankheet, 1996). This probably results from the fact that for reverse-phi the percept is directly stimulus driven, whereas for aftereffects it starts fading immediately after

adaptation.

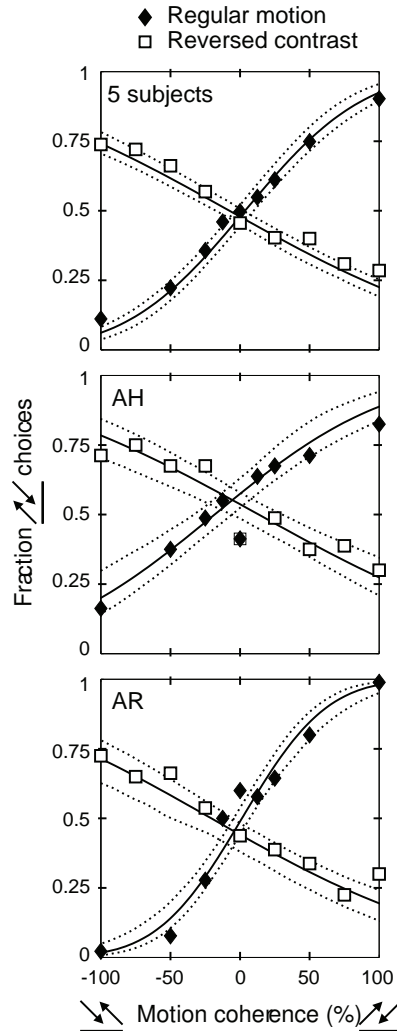
## 6.4.2 Experiment 2: Motion nulling

The orthogonal percept for transparent reverse-phi motion, as shown in the previous experiment reveal that reverse-phi percepts behave like motion aftereffects rather than like regular motion percepts. It suggests that reverse-phi similarly causes inhibitions at the front-end level of motion detection. Directional interactions and disinhibition supposedly transform inhibitions into relative excitations that correspond to the motion percept. In a second experiment we constructed a more direct test to study the sign of the low-level response. It is based on the finding that motion aftereffects can be nulled with real, regular motion (Blake and Hiris, 1993; Lankheet and Verstraten, 1995). Adding real motion to a test stimulus opposite to a perceived aftereffect does not result in transparent motion, but cancels the motion percept altogether. The reason is that the aftereffect results from a response reduction. Adding real motion simply restores the reduced activity levels to their unadapted state, canceling the directional imbalance. Nulling thus implies cancellation of positive and negative responses (Van de Grind et al., 2003). If reverse-phi percepts also arise from inhibitions we would predict similar nulling. If on the other hand reverse-phi results from excitations in the opposite direction, the two presumed excitations should induce a transparent motion percept.

Stimuli consisted of one regularly moving pattern and a second, reverse-phi pattern. In a single experiment the direction and coherence level (signal strength) for the contrast reversed pattern were held constant. Negative coherence values corresponded to (perceived) leftward motion, positive to rightward motion. The coherence level of the regularly moving pattern was varied in a method of constant stimuli. The spatial displacement for both patterns was 5 pixels (0.1 deg) and the temporal interval was 3 frames (25 ms). Trials lasted 1s and each stimulus was repeated at least 30 times. In a two-alternative forced-choice experiment, observers indicated the direction of motion which could either be to the left or to the right. Psychometric curves and 95% confidence intervals were obtained by maximum likelihood fits of cumulative Gaussians, and bootstrapping methods (Wichmann and Hill, 2001a,b).

Observers reported that they did not perceive the two patterns moving transparently, but instead saw a single direction of motion. This was qualitatively different from combining two regular motion components, which always resulted in two components perceived to move transparently. Quantitative results are shown in figure 6.5. Figure 6.5A shows full psychometric curves measured for a single combination of step size and temporal interval. The two panels on the left show data for individual observers. The right-hand panel shows results pooled for 7 observers (two authors, 5 naive). Adding reverse-phi motion simply shifted the psychometric curve leftward or rightward. Slopes for different amounts of added reversed-contrast motion were not significantly different. In other words, the

## 6.4 Results



**Figure 6.4:** Two opponent reverse-phi motion components are perceived along an orientation orthogonal to their displacement. Two regular (diamonds), or two reverse-phi motion components (squares) were shown transparently along the left or right oblique directions. Observers indicated the orientation of perceived motion, as a function of motion coherence. Negative coherence values corresponded to the upper left oblique and positive values to the upper right oblique. We used a two-alternative forced-choice paradigm with 1 s trial duration. All stimuli were interleaved in a single experiment and were repeated at least 30 times. Results shown are for a step size of 0.1 deg and a temporal interval of 25 ms. The contrast between the first and second instance of a dot was either held constant (regular motion) or reversed (reverse-phi). The two lower panels show data for individual, naive observers. The top panel shows results pooled for 5 subjects, including two authors.



presence of the reversed-contrast motion component caused a large bias, without affecting direction discriminability. This is clearly different from detecting regular motion in the presence of another regular motion component, which only has a minor effect (Edwards and Nishida, 1999; Lindsey and Todd, 1998; Verstraten et al., 1996a). The coherence level for the reverse-phi stimulus required to fully cancel the regular motion component was only slightly lower than the coherence level of regular motion. A reverse-phi component of 50% was, on average, nulled with a regular component of 45 %, and at 100% it was on average 87%.

Figure 6.5B shows that nearly identical results were obtained for other combinations of step size and temporal interval. Data for different observers were highly consistent.

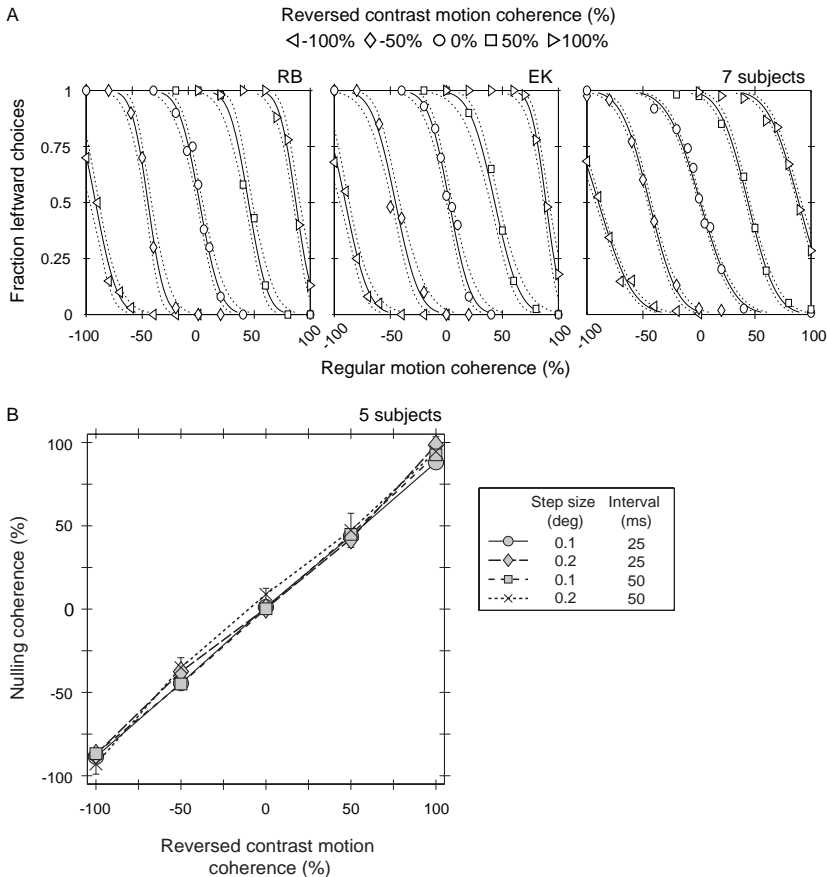
## 6.5 Discussion

Explanations for reverse-phi motion so far were based on excitation of motion detectors tuned to the direction opposite to the displacement. Clearly, at a high level of analysis this must be the case. It has been shown that the level of activity in monkey area MT (middle temporal), an area dedicated to motion analysis, corresponds to the monkeys motion percept (Britten et al., 1993; Newsome et al., 1989, 1990; Shadlen et al., 1996). Microstimulation in MT may bias the monkeys perceptual choice depending on the tuning properties of stimulated neurons (Salzman et al., 1992). Krekelberg and Albright (2005) have shown that MT activity for reverse-phi stimuli also corresponds to the monkey's percept. At the level of MT reverse-phi causes activation of cells tuned to the direction opposite to the displacement.

These findings, however, do not reveal how the two types of motion are initially detected. The initial stage of motion detection in primates is located in primary visual cortex (V1), where simple and complex cells derive directional selectivity from non-directionally sensitive geniculate cells. Contrast information in the retina and LGN is divided across two different channels: ON-cells most effectively signal positive contrasts, whereas OFF-cells signal negative contrasts. High sensitivity for reverse-phi stimuli suggest that low-level motion detectors specifically combine information from ON- and OFF-cells (Mo and Koch, 2003). These authors proposed a model in which such opposite contrast correlations excite detectors tuned to the direction opposite to the displacement (figure 6.1B). However, here we show evidence that the alternative, as shown in figure 6.1A, is more plausible. Motion signals based on correlations between ON and OFF channels probably cause inhibition of motion detectors tuned to the direction of displacement.

The critical observation to support this conclusion is the similarity between motion aftereffects and reverse-phi motion. Directional interactions differ between direct motion percepts and motion aftereffects, due to differences in directional tuning width for excitations and inhibitions. Low-level excitations remain nar-

## 6.5 Discussion



**Figure 6.5:** Nulling reverse-phi motion with regular motion: Stimuli consisted of one regularly moving pattern and a second, contrast-reversed pattern. In a single experiment the coherence level (signal strength) for the contrast-reversed pattern was held constant, as indicated in the figure. Negative coherence values corresponded to (perceived) leftward motion, positive to rightward motion. The coherence level of the regularly moving pattern was variable. In a two-alternative forced-choice experiment, observers indicated the direction of motion which could either be to the left or to the right. Figure A shows examples of psychometric curves, and their 95% confidence intervals, measured with a step size of 5 pixels (0.1 deg) and temporal interval of 3 frames (25ms). The two panels on the left show data for individual observers. The right-hand panel shows results pooled for 7 observers (two authors, five naive). Figure B shows shifts of the psychometric curves for different combinations of step size and temporal interval, averaged for five observers (two authors, three naive). Error bars represent the standard error of the mean across five observers.

rowly tuned and as a result multiple components remain segregated. Response reductions after adaptation, however, induce the opposite percept through disinhibition (Van de Grind et al., 2003, 2004). If inhibitions are more broadly tuned this leads to broad directional tuning for aftereffects (Grunewald and Lankheet, 1996). It explains the orthogonal motion aftereffect and the fact that adaptation to multiple motion components mostly results in a single, unified motion aftereffect. We propose that exactly the same interactions may be responsible for the behavior of reverse-phi percepts. This would imply that reverse-phi, like motion adaptation, causes a response reduction at the detection level, which through disinhibition transforms into a motion percept in the opposite direction.

The proposal is further supported by results for the nulling experiment. Whereas two regular motion patterns are readily seen to move in opposite directions, one regular and one reverse-phi component always canceled each other. This shows that reverse-phi is fundamentally different from regular motion in the opposite direction. The behavior is similar to that observed for motion aftereffects, but the effect is much stronger. Motion aftereffects can also be nulled with real motion but the motion strength required is much lower (Blake and Hiris, 1993; Lankheet and Verstraten, 1995; Van de Grind et al., 2003). The fact that nearly equal motion strengths were required to cancel the percept shows that cancellation was not due to a relatively small sensitivity for reverse-phi motion. In fact, in pilot experiments we found observers to be equally sensitive to the two types of motion. Motion coherence thresholds measured for the same motion settings as in the nulling experiments did not differ significantly between regular and reverse-phi motion.

Similarity of motion strengths required for nulling, and quantitative similarity of motion sensitivity for regular and reverse-phi motion also rules out that the lack of same-contrast correlations in the reverse-phi stimulus can account for the present findings. If the main effect of reversing the contrast would be removing same-contrast correlations at the specific spatio-temporal offset, we should expect large differences in sensitivity. In this case regular motion would result in a large bias in one specific direction whereas reverse-phi would correspond to incoherent noise in all directions, except one. Among other things, such a broad distribution of directions would largely cancel because most components would be balanced by their opponent counterpart. Equal sensitivity for regular and for reverse-phi motion thus strongly suggests that opposite-polarity correlations are actually used for motion detection.

The most likely explanation for the nulling results is that regular and reverse-phi motion cause excitations and inhibitions of similar strength, which are added at the very first level of motion detection.

Clearly, this proposal does not deny the observation that the difference between regular motion and reverse-phi motion is embedded in the stimulus. Nor does it deny that ideal, linear processing of Fourier components could predict the reversal (Adelson and Bergen, 1985; Krekelberg and Albright, 2005; Edwards and Nishida, 2004). It rather addresses the level at which the reversal takes place.

Our experiments show that we can differentiate between different schemes (figure 6.1) of combining positive and negative contrast signals. We propose that, at low levels of motion detection (V1), contrast reversals reverse the sign of the response, not the direction. At the next level of global direction integration (area MT) this then gives rise to a direction reversal.

Once we realize that contrast reversals may lead to perceived direction reversals through inhibitions rather than excitations we may also appreciate the reason for this specific sensitivity. Combining opposite contrasts with a negative sign makes optimal use of all available information. If spatio-temporal correlations between pairs of ON-cells or pairs of OFF-cells signal coherent motion from one receptive field to the other, then correlations between ON and OFF cells signal the absence of coherent motion. Thus, correlations between equal contrast polarities provide positive evidence whereas correlations between opposite polarities provide negative evidence for the same motion. Weighing positive and negative evidence for the same motion at the first detection stage efficiently improves signal-to-noise ratios.

It is interesting to notice that the space-time representation of our motion stimuli, as presented in figure 6.2 is similar to Glass patterns that have been used to study orientation sensitivity. Results for orientation detection are, however less consistent, and aftereffects and contrast reversals may yield different results. Burr and Ross (2006) showed that contrast reversals for Glass patterns can counteract the effect of equal-polarity correlations, but do not result in orientation percepts. This would suggest inhibition without opponency. However, Clifford and Weston (2005) used a nulling procedure for adaptation to Glass patterns similar to the procedure we used for motion. They found clear orthogonal aftereffects, indicating opponency between orthogonal orientations. Dakin (1997) also reported induction of orthogonal percepts from contrast-reversed Glass patterns. These different findings show an essential difference between orientation analysis and motion analysis. For orientation detection aftereffects and reversed contrast stimuli may yield different results. For motion, however, aftereffects and reverse-phi percepts behave in similar ways.

Our hypothesis is in line with neurophysiological results in the primate and feline visual system. Directional selectivity, the corner stone of motion sensitivity, arises in primary visual cortex, especially in complex cells. This is also the level where information from ON and OFF pathways merges (Schiller, 1992; Schiller et al., 1986). V1 complex cells in both cat (Emerson et al., 1992) and macaque (Livingstone and Conway, 2003) respond with relative inhibitions to reversed contrast correlations. At this level, there is relatively little opponency (Snowden et al., 1991). At the level of area MT, however, negative modulations have been transformed into positive modulations, and perceived motion corresponds to excitations (Krekelberg and Albright, 2005). Thus, it seems likely that directional interactions translating inhibitions to excitations (Grunewald and Lankheet, 1996) are effectuated in the projection of V1 to MT.

# Chapter 7

## General discussion

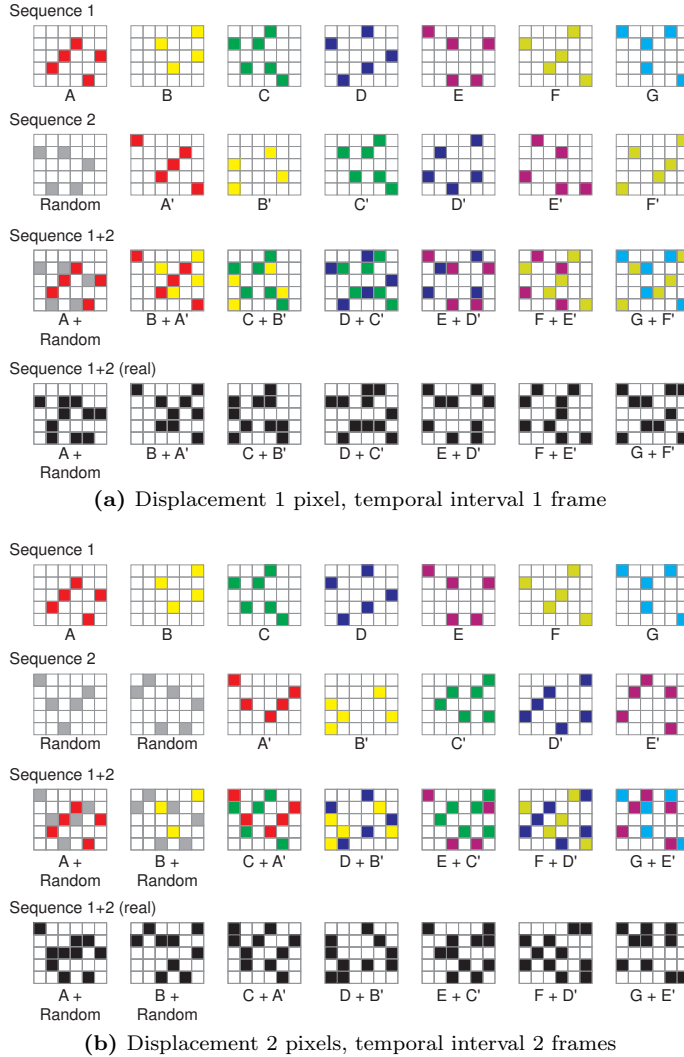
The main goal of this thesis was to study motion detection mechanisms in the primate geniculate-cortical pathway. Processing of visual information in this pathway is essential for conscious motion perception. To understand the mechanism underlying the motion detection stage, it is essential to understand the function and the physiological properties of the whole chain of visual processing. Motion pathways have been studied in numerous experimental studies, the results of which have been captured in several motion detection models. In essence, these models divide motion detection in the geniculate-cortical pathway in three stages: Local motion signals are generated by low-level detectors that perform the required spatio-temporal correlation. The detection stage, which is located in the primary visual cortex, is preceded by a pre-processing stage that involves spatial and temporal filtering of luminance information and is followed by a post-processing stage that involves spatial and temporal integration and segregation of motion signals. In this chapter, I will present an overview that combines models for the motion detection stage and for the post-processing stage. First, I will discuss the stimuli that we used in this thesis and those used in previous studies to study the spatio-temporal properties of the motion detection system in more detail. Next, I address the different stages in the overview model in figure 7.2 and discuss how it incorporates findings from psychophysical, electro-physiological and neuro-imaging studies. Finally, I will summarize our main conclusions, present several resulting predictions and show preliminary results testing these predictions.

### 7.1 Single-step two-frame motion stimulus

A large number of studies on spatio-temporal requirements for the motion correlation step (e.g. Baker and Braddick, 1985; Casco et al., 1989; Morgan et al.,

1997; Snowden and Braddick, 1989; Koenderink et al., 1985; Van de Grind et al., 1986; Van Doorn and Koenderink, 1982a,b; Fredericksen et al., 1993, 1994a,b,c; Morgan and Ward, 1980; Van den Berg and Van de Grind, 1989; Van Doorn and Koenderink, 1984) has resulted in contrasting estimates of spatio-temporal limits, sensitivity and space-time separability (see e.g. introduction of chapter 4). To a large extent, these contrasting findings can be explained by differences in the visual stimuli that have been used.

Morgan and Ward (1980) used a random pixel paradigm that effectively confined motion energy to a single combination of spatial displacement and interval duration. However, rather than spatial and temporal tuning curves only the upper and lower limits were determined. Moreover, due to numerous methodological aspects (e.g. low dot densities, highly variable stimulus durations and reaction time measurements for assessing sensitivity) a direct comparison of these results to those obtained by others is difficult. Van Doorn and co-workers (Van Doorn and Koenderink, 1982a,b; Van de Grind et al., 1986; Koenderink et al., 1985) measured optimal step sizes and temporal intervals, using spatial and temporal alternations of two random pixel patterns moving in opposite direction. By changing the spatial and temporal frequency of alternations they determined critical frequencies at which direction discrimination was sharply reduced. For frequencies below the critical frequency, the two patterns could be seen to alternate in time or space. For high frequencies, the two patterns no longer segregated, but instead fused into a single transparent display. The temporal intervals and spatial displacements corresponding to the critical frequency, were taken as best estimates of preferred step size and delay of motion detectors involved. A disadvantage of this technique is that one cannot rule out that varying the spatial or temporal frequency of alternations also affects spatial and temporal integration of motion signals (Huk and Shadlen, 2005). Fredericksen et al. (1993) used a single-step dot lifetime paradigm for controlling the step size and interval content of a motion stimulus. It is a continuous version of previously used two-frame motion stimuli (Baker and Braddick, 1985; Casco et al., 1989; Morgan et al., 1997; Snowden and Braddick, 1989). Their paradigm is suited for measuring step size tuning, but is less appropriate for measuring the temporal aspects of motion detection. The reason is that for long time intervals, patterns remained stationary between displacements. As a result, for large temporal intervals, correlations were not limited to the intended interval but were also present at shorter intervals. Consequently, the net motion information and energy differed from frame to frame and it is unclear which interval provided the largest contribution to motion detection. Our physiology data, in which we used comparable stimuli, reveal the effect of inter-stimulus interval on the step size tuning in MT cells (see section 3.4.3). Our results for neurons in area MT clearly show that variations in inter-stimulus interval had very little effect on the step size tuning and on the dynamics of responses to individual motion steps. Psychophysical tuning for temporal interval therefore most likely did not result from limitations in the correlation step but presumably from temporal filtering before or after this step.



**Figure 7.1:** Principle of the single-step two-frame-dot-lifetime stimulus illustrated for a displacement of 1 pixel every frame (A) and for a displacement of 2 pixels every second frame (B). In this example, a dot corresponds to 1 pixel. The stimuli were created by superimposing two sequences of sparse Random Pixel Arrays. The first sequence (sequence 1) consists of randomly refreshed sRPAs. The second sequence resembles the first sequence, but the pixel arrays are shifted in time and space, i.e. a pixel in sequence 1 reappears in sequence 2 after a certain time interval and with a predetermined spatial displacement to the left or the right. This method allows us to precisely control the spatio-temporal parameters of the stimulus (i.e. displacement and time interval), and keep the motion information and overall motion energy constant, irrespective of the displacement and temporal interval.

In this thesis we developed a motion stimulus to specifically examine the requirements for spatio-temporal correlation, irrespective of pre- and post-processing. This retro-innovative stimulus is an improved version of the stimulus used by Morgan and Ward (1980) to examine spatio-temporal limits of the human motion detection system. The stimulus consisted of sparse random dot arrays that contained spatio-temporal correlations at a single spatial displacement and single temporal interval. Each dot was presented in two frames only, separated by a specified interval. On each frame, half of the dots was refreshed and the other half was a displaced reincarnation of the pattern generated one or several frames earlier. Figure 7.1 illustrates the principle of the stimulus. In our stimulus, the motion information is the same on every frame irrespective of the spatial displacement or temporal interval. This prevents unwanted temporal integration effects. Moreover, the overall spatio-temporal energy content in the stimulus is constant, irrespective of the displacement or interval, which minimizes pre-filtering effects.

By combining the single-step two-frame motion stimulus with coherence thresholds, we measured complete spatio-temporal tuning curves. At first, we used black dots on a white background to measure spatio-temporal tuning of regular motion (see chapter 4). A simple adjustment to the stimulus allowed us to study the spatio-temporal properties for reverse-phi motion. To this end, we used a gray background with 50% of the stimulus dots brighter and 50% darker than the background. By simply reversing the contrast of the dots belonging to the displaced reincarnation, the stimulus changes from a regular motion stimulus into a reverse-phi motion stimulus, in which observers see motion in the direction opposite to the displacements of the dots.

## 7.2 Modeling motion processing

The overview in figure 7.2 emphasizes a three-stage model for motion processing. The overview combines a correlation-based model to describe the local motion detection stage and a network model for directional integration. The network model describes how the local motion signals combine into a global motion signal that finally results in a motion percept. The correlation-based model incorporates interactions between same-contrast signals (ON-ON and OFF-OFF), as well as cross-combinations of contrast (ON-OFF and OFF-ON). To emphasize these combinations we included a specific halfway motion detector for contrast reversing motion (i.e. spatio-temporal correlation of opposite contrast polarities). The output of these local motion sensors are combined and integrated at a motion opponency stage. The motion opponency stage integrates outputs over the whole population of motion detectors. The global motion integration stage can be described by an elaborated distribution-shift model (Grunewald, 1996; Grunewald and Lankheet, 1996) with a gain-control (Van de Grind et al., 2003, 2004).



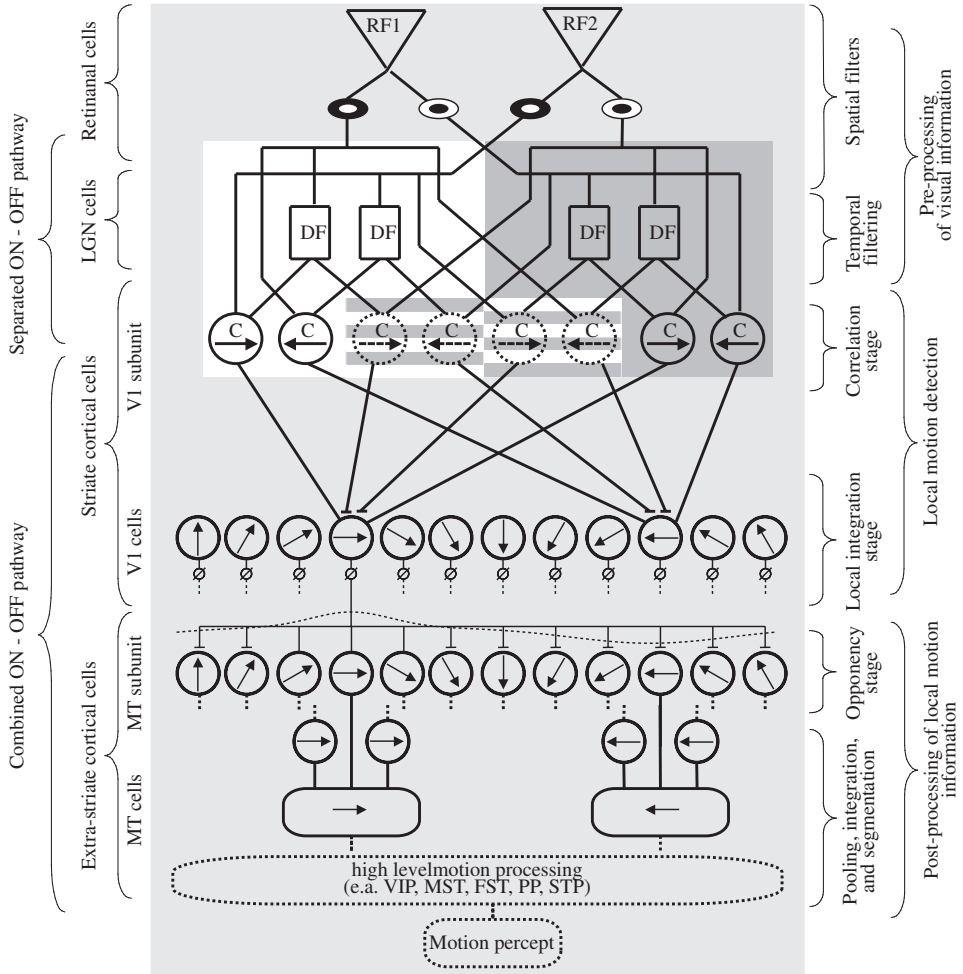


Figure 7.2: See caption 7.2 on page 110.

**Figure 7.2:** Motion processing is functionally divided in three stages: Pre-processing, local motion detection and post-processing. Pre-processing of visual information involves spatial and temporal filtering by retinal and LGN cells. Opposite contrast polarities are processed in separate ON and OFF channels. Local motion detection is performed by striate cortical cells. Four types of halfway detectors exist: Two same-contrast (ON-ON and OFF-OFF) units and two opposite-contrast (ON-OFF, OFF-ON) units. These sub-units correlate two input signals based on the delay-and-compare principle. Same-contrast correlations lead to excitations of local integrator units. Notice that the correlation and integration can be performed by a single neuron. Opposite-contrast correlation inhibits corresponding local integrator neurons. There is no interaction between correlator-units or between integrator neurons. The output has strong directionality but is non-opponent. The output of the local motion detector is evaluated at the opponency stage via a network model (Grunewald, 1996; Grunewald and Lankheet, 1996) that also incorporates an automatic gain-control principle (Van de Grind et al., 2003, 2004). Depending on the preferred direction of the detector, the detector excites opponent motion detectors that have highly similar preferred direction and inhibits others. The degree of excitation and inhibition depends in a weight factor that can be described by two Gaussian functions: Excitation is described by a sharp tuning centered around the preferred direction of the signaling integrator unit; inhibition is described by a broad tuning centered around its null direction (i.e. opposite to the preferred direction). The output of a motion opponent detector is strongly directional and motion opponent. The next stage involves pooling of detectors with similar directional preferences. The motion opponency and pooling stage are presumably located in area MT. Further processing enables detection of more complex motion components and is performed by other extra-striate areas.

### 7.2.1 Pre-processing: Spatial and temporal filtering

During the pre-processing stage visual information is spatially and temporally filtered. The properties of these filters affect the tuning characteristics at the next stages. The retina consists of photoreceptors that detect luminance changes in the visual field. These photoreceptors feed into bipolar cells. The bipolar cells consist of ON-center-OFF-surround and in OFF-center-ON-surround cells, referred to as ON-cells and OFF-cells, respectively. The ON-cells respond to contrast increments and are inhibited by contrast decrements. The OFF-cells respond in the opposite manner. These bipolar cells are the beginning of separate ON- and OFF-channels (Nelson et al., 1978), which stay separated in the LGN and V1 simple cells (Ibbotson and Clifford, 2001; Emerson et al., 1987, 1992; Livingstone et al., 2001; Schiller, 1984, 1992). Separated ON and OFF channels is the most efficient way to preserve full contrast information at minimal metabolic costs (Schiller et al., 1986; Laughlin et al., 1998). Slaughter and Miller (1981) found that 2-amino-4-phosphonobutyric acid (APB) resulted in the ON bipolar, amacrine and ganglion cells to become insensitive to light, while the OFF pathway was unaffected. This technique was later used to extensively examine the primate visual system (Schiller, 1982, 1984, 1990, 1992; Schiller et al., 1986). A fascinating

finding was that V1-complex cells that normally respond to both dark/light and light/dark edges of a moving bar were only sensitive to light/dark edge following APB. This suggests that ON and OFF pathways converge at the level of V1-complex cells. At the level of area MT and at higher levels, ON and OFF pathways are merged into a single pathway (Zeki, 1974; Tanaka et al., 1989).

A correlation type of motion detector requires a temporal asymmetry in its inputs. (Borst and Egelhaaf, 1989). Reichardt (1961) proposed a temporal delay filter for generating this asymmetry. An alternative was proposed by Adelson and Bergen (1985). They built motion detectors from combinations of spatio-temporally separable filters. Temporal asymmetry in this case is based on a phase difference between these filters. To extract spatiotemporal motion energy, filters can be chosen as quadrature whose outputs are squared and summed. In this way, one can derive a phase-independent-motion energy response from the outputs of two linear filters that are sensitive to motion in the same direction but with sensitivities 90 deg out of phase.

## **7.2.2 Motion detection: Correlation and local integration stage**

Instead of the motion energy approach by Adelson and Bergen (1985), which is based on combinations of linear space-time filters, the motion detection stage can also be modeled by an elaborated Reichardt detector. Two half-way motion detectors correlate spatially separated inputs with a temporal delay. The correlation involves a multiplication operation. Although different at first sight, the outcomes of the motion energy model and the Reichardt detector are mathematically equivalent. Based on numerous studies it seems likely that the low-level motion detection stage, whether a correlation or motion energy operation, is located in area V1 (Movshon and Newsome, 1996; Livingstone et al., 2001; Rodman and Albright, 1989; Girard et al., 1992).

In our overview, the local motion detection stage consists of a correlation and integration stage. The correlator stage feeds into the local integrator stage. The integration stage emphasizes the role of the ON and OFF pathway, which converge at the level of V1-complex cells (Livingstone and Conway, 2003; Schiller, 1990, 1992): Livingstone and Conway (2003) mapped the substructure of direction-selective fields in macaque V1 neurons using same-contrast and inverting-contrast dot-pairs. Their results clearly showed a reversal of the interaction pattern for inverting contrast dot-pairs compared to the same-contrast pairs (Figure 7 in Livingstone and Conway (2003) shows a V1 complex cell for the four contrast combinations).

We propose that in V1 a local integrator unit receives four types of input from V1 subunits in the correlation stage (reason is clarified in section 7.2.4): Two excitatory inputs from the same-contrast correlation units and two inhibitory inputs from the inverted-contrast correlation units. Notice, that at this stage

no directional opponency is present. The lack of directional opponency in area V1 is supported by a physiological experiment of Snowden et al. (1991), who measured the response of area MT and V1 neurons to transparent motion. They found that V1 cells responded to their preferred direction of movement even under transparent condition, whereas area MT cells were suppressed under transparent conditions. This suggests that motion opponency arises in the projection and integration of information from area V1 to area MT.

Interestingly, the study of Livingstone and Conway (2003) seems to suggest that some opponent interaction is present at the level of V1. In their study, they computed sequential interaction maps for all contrast combinations using a reverse correlation technique in a paradigm in which two dots were presented in each frame, one white and one black, on a gray background. Spikes were then correlated with the difference in position of pairs of stimuli, relative to the previous frame. In addition to an excitatory region, they also found a suppressed region. Since the two regions corresponded to opposite directions of motion, this might indicate opponent interactions. Notice, however that Livingstone and Conway's facilitation and suppression do not necessarily imply excitation and inhibition relative to spontaneous activity. Although they used an accurate definition for facilitatory and suppressive interactions, this terminology may therefore be confusing. Livingstone and Conway use facilitation and suppression in a relative setting, i.e. facilitation and suppression are defined according to a baseline that corresponds to an average response of a neuron, and is not related to its spontaneous activity. Therefore, one can only conclude that a suppressive region corresponds to a relatively low spiking probability, not that such a region reveals an underlying excitatory and inhibitory mechanisms. As mentioned before, these maps do clearly reveal the reversal of the interaction pattern for inverting-contrast dot pairs, and thus strongly suggest a convergence of the ON and OFF pathway at the level of complex V1-cells.

### 7.2.3 Post-processing

#### **Global motion processing: Opponency, pooling, integration and segmentation**

The next stages in our model concern global motion detection. This involves integration of multiple local motion signals and is most likely performed by areas in the extra-striate cortex, e.g. area MT and MST (Britten and Heuer, 1999; Livingstone et al., 2001; Maunsell and Van Essen, 1983; Movshon and Newsome, 1996; Raiguel et al., 1999; Tanaka et al., 1993). Several studies have found directional interactions (motion opponency) at the level of area MT (Britten et al., 1993; Heeger et al., 1999; Snowden et al., 1991; Van Wezel and Britten, 2002a,b). Moreover, a close relationship between neuronal activity in the extra-striate cortex and behavioral judgments of motion has been observed (Newsome et al., 1989; Salzman et al., 1990, 1992; Britten et al., 1992, 1996). Exactly how the popula-

tion response of area V1 is read out by area MT for different tasks is still unclear (Weiss et al., 2002; Born and Bradley, 2005).

Here, we propose the automatic gain control model (Van de Grind et al., 2003), that was originally developed to quantify several aspects of motion adaptation, like the duration of a motion aftereffect, its strength and its nulling behaviour. This model elaborates the MAE-network model by Grunewald and Lankheet (1996), by adding an automatic gain control stage. In essence, direction-tuned neurons from the local motion detection stage excite integrator neurons in the opponency stage that have a similar directional preference, whereas broad inhibition centers around the opponent direction unit of the integrator layer (Van de Grind et al., 2003). The strength of inhibition and excitation depends on the differences in direction tuning, according to a Gauss profile. Snowden et al. (1991) found that V1 cells responded to their preferred direction of movement even under transparent conditions, whereas area MT cells were suppressed in the transparent condition. This suggests absent or weak opponent interactions in area V1 but clear opponency in area MT. In chapter 3 we compared the dynamics of direction tuning of MT neurons as a function of step size in the preferred and null directions. These results allow one to extend the gain-control model in the spatio-temporal domain. Optimal step sizes in the preferred direction and null direction are similar, whereas the tuning in the null direction is broad and the tuning in the preferred direction sharp.

Perceptual direction discrimination and complex motion discrimination experiments have been linked to the activation of direction selective cells in area MT (Britten et al., 1992, 1996) and complex motion selective cells in area MST (Heuer and Britten, 2004; Britten and Wezel, 1998, 2002; Celebrini and Newsome, 1994). In line with such a correspondence between psychophysics and MT physiology, we found that spatio-temporal tuning of humans is similar to the optimal step size tuning for MT neurons. We therefore assume that the output of area MT (and MST) is combined to construct a motion percept. To allow for the representation of multiple motion components simultaneously it seems likely that different tuning properties among subpopulations of motion detectors are represented using a labeled-line code (Van Doorn and Koenderink, 1984; Van de Grind et al., 1986; Fredericksen et al., 1993).

#### **7.2.4 The mechanism underlying local motion detection**

Reverse phi motion differs from regular motion in that correlated information is based on opposite contrast signals, which presumably are represented most effectively in separate ON and OFF channels. Regular motion is based on same-contrast signals, which are most effectively represented in either one of the two channels. By limiting motion information to specific combinations of step size and temporal interval, we measured spatio-temporal tuning properties of regular and reverse-phi motion. The results revealed that low-level motion detectors are equally efficient in combining same polarity signals and opposite polarity signals.

This observation extends the finding by Sato (1989) that displacements limits are similar for regular and reverse-phi motion. Using coherence thresholds, we showed that not only the limits are similar, but also maximum sensitivities and the combination of spatial and temporal parameters at which they occur. The pivotal question in explaining reverse-phi motion therefore is how low-level ON and OFF signals feed into motion detectors.

### **Motion detection without interaction between ON and OFF signals**

One postulation is that only same-contrast (ON-ON or OFF-OFF) spatio-temporally oriented motion detectors exist and that tuning properties simply follow from the Fourier energy distribution of the stimulus and the spatio-temporal distribution of detectors in the visual system (Adelson and Bergen, 1985; Krekelberg and Albright, 2005; Van Santen and Sperling, 1985). In this view, the reversal is trivial because reverse-phi simply reflects stimulus properties and consequently reverse-phi does not provide information on motion detection mechanisms. This explanation is illustrated in figure 5.1 (page 75). The reversal of contrast changes the orientation in the space-time domain. The subsequent stage would simply consist of vector averaging over all velocities (i.e. directions and speeds) and result in a global motion percept. Notice that this requires equally efficient encoding of contrast increments and decrements. Since increments and decrements are represented in separate channels, this requires combining the two signals in a single channel.

In chapter 5 we found a remarkable similarity between spatial and temporal tuning properties of normal and reversed apparent motion. Can this be explained by simple stimulus properties? This is unlikely, since a reversal of contrast not only changes direction of the motion information or energy, but also changes the spatio-temporal characteristics of the stimulus. Accordingly, the detectors activated by the reverse-phi not only differ in direction, but also in their spatio-temporal properties. In theory, the similarity in the spatio-temporal tuning could arise at a higher-level where vector averaging over all velocities takes place. Intuitively, this explanation is not satisfactory since it is unlikely that the sensitivity would be unaffected in all cases. Therefore, the postulation that reverse-phi solely arises from stimulus properties is insufficient to explain the high degree of similarity in our data. Since reversing the contrast also affects spatio-temporal stimulus properties we should have found different tuning properties, which we did not find.

One could argue that the sensitivity for reverse-phi motion is induced because contrast reversal simply removes the same-contrast correlation at the specified combination of step size and temporal interval. Consequently, the correlation in the opposite direction occurs relatively more often. Due to motion opponency, motion in the opposite direction would be perceived. We tested this hypothesis in the no-phi experiments in chapter 5. In a no-phi stimulus, the contrast of displaced dots was set to background luminance. Sensitivity for no-phi motion was low for

stimuli containing only black or only white dots. When both dot polarities were present in the stimulus, sensitivity was absent. Thus, motion information based on separate contrasts was effectively canceled by a component based on different contrasts. The low sensitivity for stimuli containing only black or only white dots may be induced because the overall luminance of the stimulus is lower (black dots) or higher (white dots) than the background luminance. In the case that both polarities were present, the overall luminance of the stimulus was equal to the background luminance. The results of the no-phi experiments (see section 5.4.2 on page 80) clearly showed that observers could not consistently detect the motion direction. Accordingly, we can conclude that the lack of correlation is insufficient to explain the similarity in sensitivity found in our experiments in chapter 5.

### **Motion detection with interaction between ON and OFF signals**

Since it seems unlikely that the similarity in sensitivity for regular and reverse-phi motion can be achieved by correlation within separate ON and OFF channels, we proposed that the sensitivity for reverse-phi is the result of specific interaction of the ON and OFF pathway. Notice, however, that this reasoning does not rule out illusory motion percepts that arise merely because of stimulus properties or by temporal asymmetries in ON and OFF responses within one channel (Del Viva et al., 2006).

This explanation, as shown in the overview (figure 7.2), emphasizes the role of an interaction between the ON and OFF pathway and requires detectors specifically tuned for reversed motion. This reasoning assumes that for every same-contrast motion detector, a reversed-contrast detector exists with identical spatio-temporal properties. This principle was already acknowledged by Mo and Koch (2003) and integrated in their double synaptic-veto model: They implemented a detector for contrast reversing motion that generates excitatory signals for the direction opposite to the displacement. Since, both detectors have the same spatio-temporal properties, it is likely that the coherence thresholds are also comparable. In other words, opposite-contrast correlations provide *evidence for* the presence of motion in the direction *opposite* of the physical movement. On the other hand, if one assumes an opponent directional interaction following low-level detection, inhibition in the direction of displacement would result in a similar percept. In this sense, opposite-contrast correlations provide *evidence against* the presence of motion in the *same* direction of the physical movement. This may seem trivial, but this proposition has a major implication for the percept of motion aftereffects induced by adaptation to reversed-apparent motion. Figure 7.4 illustrates how the excitatory and inhibitory signals at the motion detection stage have different effects on the motion percept. In section 7.2.5 we address this topic in more detail. In chapter 6 we have provided evidence for the latter proposal: Low-level motion detection implements an inhibition step for correlations based on opposite contrasts. Thus, a same-contrast correlation provides positive evidence of motion

in the direction of the displacement and opposite-contrast correlation is used as evidence against motion in the direction of the displacement. Integration of the correlation signals results in non-opponent, but strongly directional responses.

### 7.2.5 Combining local motion signals: The role of excitatory and inhibition signals on motion percepts

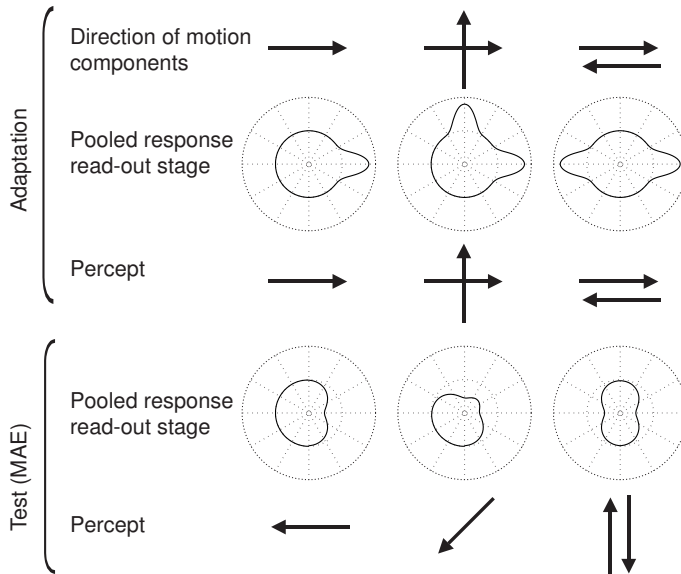
In section 7.2.4 we argued that the inhibition step is implemented in the low-level processing stream of a correlation-type motion detector: Same-contrast correlation provides positive evidence of motion in the direction of the displacement and opposite-contrast correlation is used as evidence against motion in the direction of the displacement. Mo and Koch (2003) proposed that the sign of a reverse-contrast motion detector was positive, but reversed the directions for opposite contrast combinations (see also figure 6.1, page 94). To distinguish between the two alternative propositions, we have investigated the interaction between motion aftereffects and the reverse-phi motion and incorporated the findings in our model. Our model implements an automatic gain control stage. This model predicts motion aftereffects due to adaptation to one or multiple velocity fields, as well as nulling properties of the motion aftereffects. The essence is that the motion percept corresponds to the properties of the relatively more active subpopulation of motion sensitive neurons.

#### Motion aftereffect

Numerous studies have characterized and tried to explain the motion aftereffect (e.g. Heeger et al., 1999; Lankheet and Verstraten, 1995; Petersen et al., 1985; Priebe et al., 2002; Priebe and Lisberger, 2002; Van Wezel and Britten, 2002a,b; van Wezel et al., 1994; Verstraten et al., 1994, 1996a,b). A motion aftereffect results from the reduced activity of motion detectors during a test stimulus, following prolonged adaptation. Motion opponency converts reduced activity at a lower level due to adaptation in one direction to increased activity at a higher level in the opposite direction. This results in a global motion percept opposite to the direction of adaptation. Figure 7.3 illustrates how adaptation to transparent motion leads the motion aftereffect percept. After adaptation to a transparent motion, a uni-directional aftereffect moving opposite to the vector sum of the components of the transparent motion is perceived. For example, adaptation to two motion vectors, one to the upper left and one to the upper right, causes a motion aftereffect with a single motion component in the downward direction. The difference in tuning width for excitation and inhibition nicely explains the apparent contradiction between segregation of multiple, transparent motion components in direct motion precepts, and integration into a unified, single motion aftereffect: Excitations during the direct motion percept remain narrowly tuned, and may support transparency. Reduced activity in the adapted direction leads, through broad inhibition, to a single, wide distribution during the aftereffect, which is



perceived as a single, unified motion component. Interestingly, Grunewald and Lankheet (1996) reported that after adaptation to transparently presented patterns that moved in opposite direction, an orthogonal MAE arises. The mechanism that leads to a global read-out is well described by the automatic gain-control model by Van de Grind et al. (2003). This model explains how two populations of dots moving in different directions are readily perceived as two transparently moving sheets, but also how a uni-directional percept arises after adaptation to these stimuli (see also figure 7.2: The model is implemented in the post-processing stage).



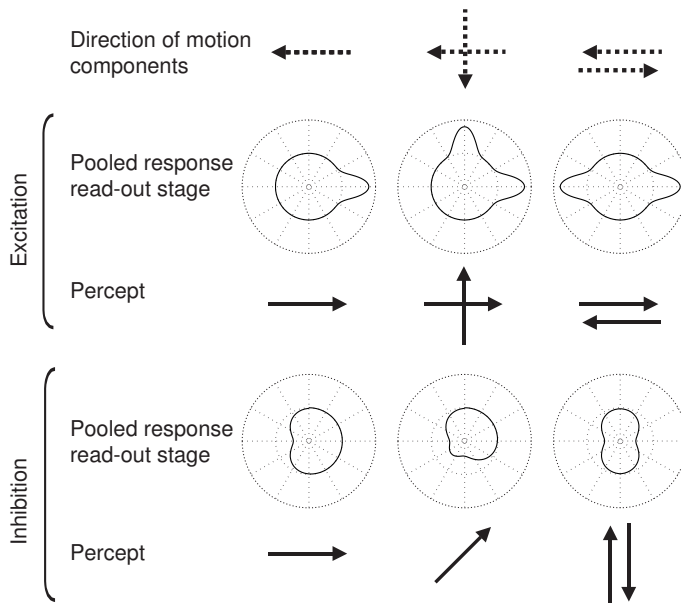
**Figure 7.3:** The pooled response of the read-out stage (polar plot) to three adaption conditions determines the percept. All direction sensors are equally active. During adaptation, the pooled response changes as result from the motion components (arrow) in the adaptation stimulus. This results in a narrow excitation in the pooled response. The motion aftereffect follows the relative higher activity in the pooled response. During the test period, no net motion is present. The more broadly tuned inhibitory response results from fatigue of the motion sensors. The percept follows the relative higher activity in the pooled response. See text for more detailed explanation.

## Reverse-phi

The mechanism underlying local motion detection makes clear predictions for the percept of reverse-phi motion stimuli consisting of single and multiple velocity fields. Figure 7.4 illustrates how the sign (excitation versus inhibition) of the output of opposite-contrast detectors affects the percept for reverse-phi stimuli.

## 7.2 Modeling motion processing

The prediction of the percept that arises after presenting transparent reverse-phi especially forms a critical test for the different configurations of ON-OFF correlations. If reversed-contrast detectors excite global integrator units, a transparently moving contrast-reversing motion under an angle of 90 deg should be perceived as transparent motion. Inhibition however should result in a unidirectional percept. If the patterns move in opposite direction, the excitatory configuration predicts transparent motion in the same plan, whereas the inhibitory configuration predicts transparent motion in the orthogonal plane.



**Figure 7.4:** The pooled response at the read-out stage to motion components of a reverse-phi stimulus (dotted arrow) depends on the sign of the output of reversed-contrast motion sensors. Top: A detector for contrast-reversing motion generates excitatory signals for the direction opposite to the displacement (see configuration A in figure 6.1). Excitatory projections result due to motion opponency in an increased activity in the pooled response at the read-out stage. The percept follows the relatively higher activity in the pooled response. Two reversed-contrast motion components result in a transparent motion percept in the direction opposite to the physical motion components. Bottom: A detector for contrast reversing motion generates inhibitory signals for the direction of the displacement (see configuration B in figure 6.1). Inhibitory projections (bottom) result in a decreased activity in the pooled response at the read-out stage. For components at an angle of 90 deg, the percept is unidirectional, opposite to the vector sum of the components. For components at an angle of 180 deg, the percept is two-directional, but orthogonal. See text for more detailed explanation.

## Nulling characteristics of motion aftereffects and reversed apparent motion

The finding that reversed apparent motion phenomenon can be perceptually nulled if a certain fraction of signal dots of a regular moving pattern move in the same direction as the physical displacement of the contrast reversing signal dot, has major implications for motion detection models. The nullification property may arise via low-level mechanisms (local motion detectors) or at a higher-level. Blake and Hiris (1993) reported on an analogue phenomenon, where a dynamic MAE can be perceptually nulled if a certain percentage of signal dots move coherently opposite to the MAE direction. Here, the explanation is twofold. Firstly, it assumes that adaptation leads to fatigue in neurons at a low-level stage. The resulting higher-level read-out stage, the ultimate cause of MAEs, as pointed out by Van de Grind et al. (2003), converts this signal into motion in the opposite direction, as predicted by a gain-control model (Grunewald, 1996; Grunewald and Lankheet, 1996; Van de Grind et al., 2003). Simplified, this model considers (using vector-averaging) only the relatively more active neuronal populations to determine the motion percept. Supplying real motion during the test phase, moving in the same direction as the adaptation stimulus can counteract this process. Consequently, the population of fatigued neurons is stimulated. If this occurs in a balanced manner, the activity is equal for all directions, and accordingly, no net motion is perceived. If we apply the same reasoning to our finding that reverse-phi motion can be nulled by real motion, we must assume that the two signals have opposite polarity, one excitatory and one inhibitory. The positive signal is induced by the real motion component, whereas the negative signal is derived from the contrast-reversing motion. Mo and Koch (2003)'s model uses inversed-contrast motion detection as positive evidence for motion in the opposite direction to the physical displacement, since it results in a positive (excitatory) signal for that direction. Interestingly, if one assumes a gain-control model to underlie the higher-order read-out stage, their model predicts a transparent motion percept, whereas our data clearly falsified that prediction. However, Qian et al. (1994) showed that the percept of motion transparency induced by two superimposed sets of dots regularly moving in opposite directions is determined by the distance between the dots of both pairs. If the distance between paired dots is small ( $<0.2$  degrees of visual angle), subjects do not perceive transparency. One could argue that this is the reason that our subjects reported no net motion or only uni-directional motion. However, our dot density is within the 'safe zone' (i.e.  $\geq 0.2$  degrees of visual angle. This is supported by findings in chapter 6: Superimposing two regular motion patterns are perceived transparently.

The finding of orthogonal motion percepts for two transparently presented reverse-phi patterns moving in opposite directions has major implications for any motion detection model. Grunewald and Lankheet (1996) reported on an analogue phenomenon involving motion adaptation. They showed that after adaptation to transparently presented patterns that moved in opposite direction, an orthogonal MAE arises. Their experiment was motivated by the MAE-network model. Their

model predicted that adaptation to simultaneous motion in opposite direction would lead to an orthogonal MAE. Moreover, they emphasized that broadly tuned inhibitory interactions are likely to be important in the integration and segregation of motion signals. Our results for reverse-phi motion support this reasoning. Notice that in addition to the latter experiment and with the assumption that a gain-control model underlies the read-out stage, Mo and Koch (2003)'s model predicts for the superimposed reversed motion condition a transparent motion percept, whereas the orthogonal motion percept clearly falsifies that prediction.

## 7.3 Concluding remark

The physiological study showed that response properties of neurons in area MT and the dynamics of their responses to individual motion steps are largely invariant with direction and inter-stimulus interval. Similarity of the temporal dynamics in the preferred and non-preferred direction shows that excitatory and inhibitory inputs shaping these responses are quite the same, and presumably follow similar paths. Because opponency is limited in V1 (Snowden et al., 1991) directional opponency in MT must result from excitatory and inhibitory projections with equal dynamics. The fact that step size tuning was narrow and independent of inter-stimulus interval suggests that MT cells sample V1 units with similar properties. Moreover, we found small, though significant differences in the width of step size tuning for preferred and non-preferred directions. This suggests that the properties of the wiring for preferred and null direction are different. We found no indications for speed invariant covariations of step size and interval tuning.

We used a new stimulus paradigm to study spatio-temporal properties of motion detection for regular and reverse-phi motion. We found similar sensitivity for both motion types, no sensitivity for no-phi motion and evidence that ON and OFF signals are combined. Why would a visual system have opposite-contrast detectors? We suggest that opposite-contrast detectors help to improve signal-to-noise ratios. Spatio-temporal correlations between pairs of ON-cells or pairs of OFF -cells signal coherent motion from one receptive field to the other. Opposite-contrast detectors on the other hand contain the information that a particular motion is absent. In the configuration of Mo and Koch (2003) this information is then used as positive evidence for motion in the opposite direction. In this sense, the absence of motion is equivalent to motion in the other direction. This configuration leads to activation of two oppositely-tuned motion detectors at the local integration stage. The information is integrated in later stages. We argue, on the other hand, that the correlations of opposite polarities provide negative evidence for the same motion. In this configuration, the two signals combine on a single motion detector at the local integration stage. Logically this makes sense, since all available motion information is used. Moreover, weighing positive and negative evidence for the same motion at the first detection stage efficiently improves signal to noise ratios. Our experiments support this theory. Moreover, this view unites reverse-phi, regular motion and their aftereffects in a single model for

global motion detection.

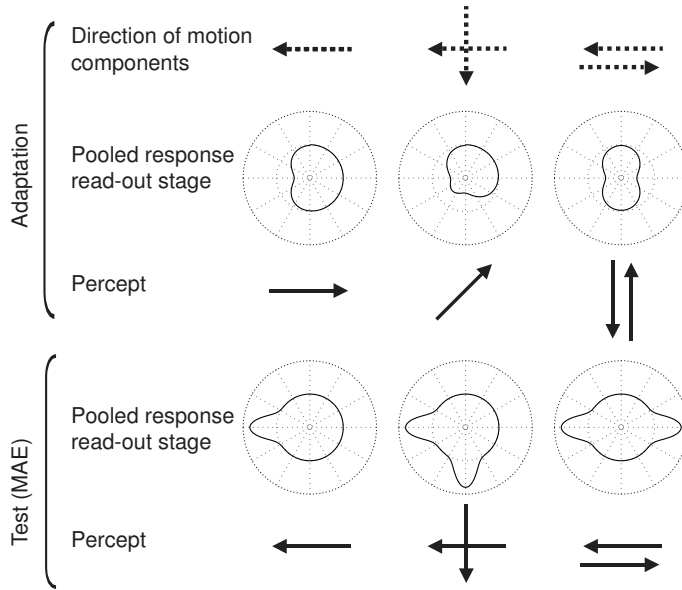
## 7.4 Perspective

Our model (figure 7.2) has strong predictions on the properties of the MAEs of regular and reverse-phi motion. We have shown that reverse-phi has similar properties as the MAE of regular motion. Interestingly, the model predicts that MAE of reverse-phi motion has the same properties as regular motion. For example, after adaption to transparent reversed motion moving under an angle of 90 degrees, we predict that during the adaption period an uni-directional motion percept arises for motion in the opposite direction as the vector-average of the components (see figure 7.5). In the subsequent test period, the observer will perceive transparent motion. In fact, this motion aftereffect for reverse-phi motion has all properties of regular motion. In a pilot experiment, we tested this prediction and indeed perceived a transparent motion aftereffect after adapting to a transparent reverse-phi stimulus.

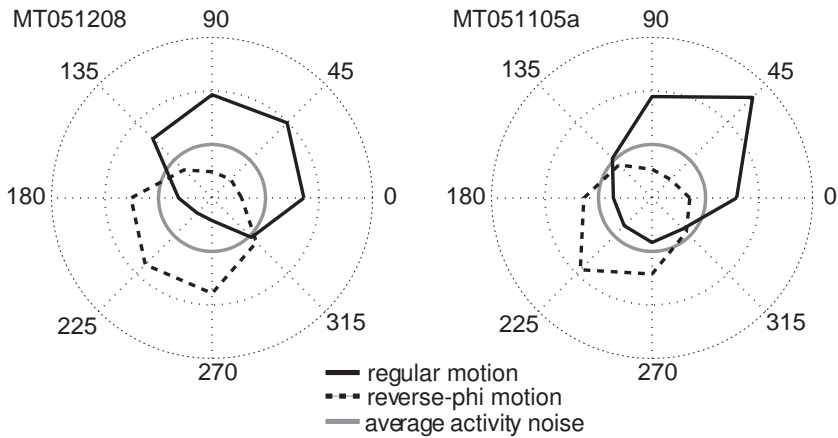
In a physiological study, we measured direction tuning of MT neurons for regular and reverse-phi motion using a stimulus similar to those in the psychophysical reverse-phi experiments. Preliminary results show that the stimuli drive MT neurons well. Figure 7.6 shows two example cells. The direction tuning is reversed for reverse-phi in comparison to the regular motion.

Our proposition that opposite-contrast correlations provide evidence against the presence of motion in the same direction seems also supported by an fMRI study. We used similar stimuli as in our reverse-phi experiments to study global brain activation in V1, V2 and MT. Preliminary results show that reverse-phi motion leads to lower activity in V1, which suggest inhibitory responses.

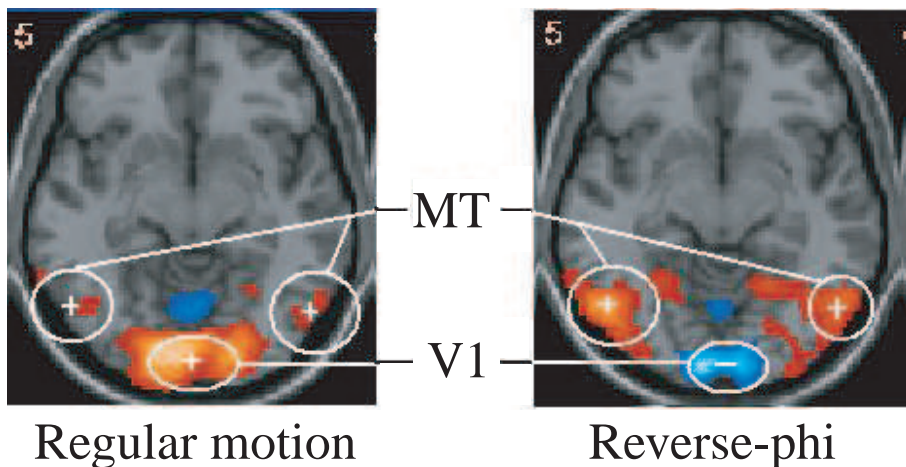
## 7.4 Perspective



**Figure 7.5:** Motion adaptation to reverse-phi: The illustrations are based on a detector for contrast reversing motion that generates inhibitory signals for the same direction as the displacement: The pooled response of the read-out stage (polar plot) to three adaptation conditions determines the percept to reverse-phi motion. During adaptation, the pooled response changes as result from the motion components (dotted arrow) in the adaptation stimulus. This results in a broad inhibition in the pooled response due to motion opponency. The percept follows the relative higher activity in the pooled response. During the test period, no net motion is presented. The narrowly tuned excitatory response results from motion opponency. The percept follows the relative higher activity in the pooled response. See text for more detailed explanation.



**Figure 7.6:** Directional tuning for two MT neurons to regular (black straight line) and reverse-phi (black dotted line) plotted in polar plot. The response is showed relative to spontaneous response of the neuron (gray straight line). The preferred direction of reverse-phi motion (225 deg) is opposite to the preferred direction of regular motion (45 deg).



**Figure 7.7:** Preliminary fMRI data: The net activation of motion sensitive neurons in V1 is increased during regular motion, and decreased during reverse-phi motion, relative to dynamic noise. The net activation of motion sensitive neurons in MT+ is increased during both regular and reverse-phi motion, relative to dynamic noise.





# Bibliography

- R Addams. An account of a peculiar optical phaenomenon seen after having looked at a moving body. *London and Edinburgh Philosophical Magazine and Journal of Science*, 5:373–374, 1834.
- EH Adelson and JR Bergen. Spatio-temporal energy models for the perception of motion. *J Opt Soc Am A*, 2:284–299, 1985.
- D Alais, FA Verstraten, and DC Burr. The motion aftereffect of transparent motion two temporal channels account for perceived direction. *Vision Res*, 45(4):403–412, 2005.
- TD Albright. Direction and orientation selectivity of neurons in visual area mt of the macaque. *J Neurophysiol*, 52:1106–1130, 1984.
- SJ Anderson and DC Burr. Spatial and temporal selectivity of the human motion detection system. *Vision Res*, 25(8):1147–1154, 1985.
- SM Anstis. Phi movement as a subtraction process. *Vision Res*, 10(12):1411–1430, 1970.
- SM Anstis and G Mather. Effects of luminance and contrast on direction of ambiguous apparent motion. *Perception*, 14(2):167–179, 1985.
- SM Anstis and BJ Rogers. Illusory reversal of visual depth and movement during changes of contrast. *Vision Res*, 15:957–961, 1975.
- W Bair and JA Movshon. Adaptive temporal integration of motion in direction-selective neurons in macaque visual cortex. *J Neurosci*, 24:7305–7323, 2004.
- W Bair, JR Cavanaugh, MA Smith, and JA Movshon. The timing of response onset and offset in macaque visual neurons. *J Neurosci*, 22:3189–3205, 2002.
- CL Jr Baker and OJ Braddick. Temporal properties of the short-range process in apparent motion. *Perception*, 14(2):181–192, 1985.

## BIBLIOGRAPHY

---

- HB Barlow and WR Levick. The mechanism of directionally selective units in rabbit's retina. *J Physiol*, 178(3):477–504, 1965.
- RA Barton. Visual specialization and brain evolution in primates. *Proc Biol Sci*, 265(1409):1933–1937, 1998.
- R Blake and E Hiris. Another means for measuring the motion aftereffect. *Vision Res*, 33(11):1589–1592, 1993.
- BG Borghuis, JA Perge, I Vajda, RJ Van Wezel, WA Van de Grind, and MJ Lankheet. The motion reverse correlation mrc method a linear systems approach in the motion domain. *J Neurosci Methods*, 123:153–166, 2003.
- EG Boring. *Sensation and Perception in the History of Experimental Psychology*. D. Appleton-Century, New York, 1942.
- RT Born and DC Bradley. Structure and function of visual area mt. *Annu Rev Neurosci*, 28:157–189, 2005.
- A Borst. Models of motion detection. *Nat Neurosci*, 3:1168, 2000.
- A Borst and M Egelhaaf. Principles of visual motion detection. *Trends Neurosci*, 12:297–306, 1989.
- A Borst and M Egelhaaf. Direction selectivity of blowfly motion-sensitive neurons is computed in a two-stage process. *Proc Natl Acad Sci*, 87:9363–9367, 1990.
- JC Boulton and CL Jr Baker. Dependence on stimulus onset asynchrony in apparent motion evidence for two mechanisms. *Vision Res*, 33(14):2013–2019, 1993.
- RJE Bours, MCW Kroes, and MJM Lankheet. The parallel between reverse-phi and motion aftereffects. *J Vision*, 7(11(8)):1–10, 2007a.
- RJE Bours, S Stuur, and MJM Lankheet. Tuning for temporal interval in human apparent motion detection. *J Vision*, 7(1(2)):1–12, 2007b.
- RJE Bours, MCW Kroes, and MJM Lankheet. Sensitivity for reverse-phi motion. *Vision Res*, 49(1):1–9, 2009.
- O Braddick. A short-range process in apparent motion. *Vision Res*, 14(7):519–527, 1974.
- DC Bradley, N Qian, and RA Andersen. Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature*, 373:609–611, 1995.
- KH Britten and HW Heuer. Spatial summation in the receptive fields of mt neurons. *J Neurosci*, 19(12):5074–5084, 1999.
- KH Britten and WT Newsome. Tuning bandwidths for near-threshold stimuli in area mt. *J Neurosci*, 80:762–770, 1998.

- KH Britten and RJA Wezel. Electrical microstimulation of cortical area mst biases heading perception in monkeys. *Nature Neurosci*, 1:1–5, 1998.
- KH Britten and RJA Wezel. Area mst and heading perception in macaque monkeys. *Cereb Cortex*, 12:692–701, 2002.
- KH Britten, MN Shadlen, WT Newsome, and JA Movshon. The analysis of visual motion a comparison of neuronal and psychophysical performance. *J Neurosci*, 12:4745–4765, 1992.
- KH Britten, MN Shadlen, WT Newsome, and JA Movshon. Responses of neurons in macaque mt to stochastic motion signals. *Vis Neurosci*, 10(6):1157–1169, 1993.
- KH Britten, WT Newsome, S Shadlen, MN Celebrini, and JA Movshon. A relationship between behavioral choice and the visual responses of neurons in macaque mt. *Vis Neurosci*, 13(1):87–100, 1996.
- D Burr and J Ross. The effects of opposite-polarity dipoles on the detection of glass patterns. *Vision Res*, 46(6-7):1139–1144, 2006.
- C Casco, MJ Morgan, and RM Ward. Spatial properties of mechanisms for detection of moving dot targets in dynamic visual noise. *Perception*, 18(3):285–291, 1989.
- E Castet, DRT Keeble, and FAJ Verstraten. Nulling the motion aftereffect with dynamic random-dot stimuli: Limitations and implications. *J Vis*, 2(4):302–311, 2002.
- S Celebrini and NT Newsome. Neuronal and psychophysical sensitivity to motion signals in extrastriate area mst of the macaque monkey. *J Neurosci*, 14:4109–4124, 1994.
- J Chey, S Grossberg, and E Mingolla. Neural dynamics of motion processing and speed discrimination. *Vision Res*, 38:2769–2786, 1998.
- EJ Chichilnisky and RS Kalmar. Functional asymmetries in on and off ganglion cells of primate retina. *J Neurosci*, 22(7):2737–2747, 2002.
- EJ Chichilnisky and RS Kalmar. Temporal resolution of ensemble visual motion signals in primate retina. *J Neurosci*, 23(17):6681–6689, 2003.
- C Chubb and G Sperling. Drift-balanced random stimuli a general basis for studying non-fourier motion perception. *J Opt Soc Am A*, 5:1988–2007, 1988.
- C Chubb and G Sperling. Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proc Natl Acad Sci USA*, 86:2985–2989, 1989.

## BIBLIOGRAPHY

---

- MM Churchland and SG Lisberger. Shifts in the population response in the middle temporal visual area parallel perceptual and motor illusions produced by apparent motion. *J Neurosci*, 21:9387–9402, 2001.
- MM Churchland, NJ Priebe, and SG Lisberger. Comparison of the spatial limits on direction selectivity in visual areas mt and v1. *J Neurophysiol*, 93:1235–1245, 2005.
- CW Clifford and E Weston. Aftereffect of adaptation to glass patterns. *Vision Res*, 45(11):1355–1363, 2005.
- BR Conway and MS Livingstone. Space-time maps and two-bar interactions of different classes of direction-selective cells in macaque v-1. *J Neurophysiol*, 89:2726–2742, 2003.
- EP Cook and JH Maunsell. Attentional modulation of motion integration of individual neurons in the middle temporal visual area. *J Neurosci*, 24:7964–7977, 2004.
- SC Dakin. Glass patterns some contrast effects re-evaluated. *Perception*, 26(3):253–268, 1997.
- H De Lange. Relationship between critical flicker frequency and a set of low frequency characteristics of the eye. *J Opt Soc Am*, 44:380–389, 1954.
- RL De Valois, NP Cottaris, LE Mahon, SD Elfar, and JA Wilson. Spatial and temporal receptive fields of geniculate and cortical cells and directional selectivity. *Vision Res*, 40(27):3685–3702, 2000.
- MM Del Viva and MC Morrone. Motion analysis by feature tracking. *Vision Res*, 38(22):3633–3653, 1998.
- MM Del Viva, M Gori, and DC Burr. Powerful motion illusion caused by temporal asymmetries in on and off visual pathways. *J Neurophysiol*, 95(6):3928–3932, 2006.
- AM Derrington, HA Allen, and LS Delicato. Visual mechanisms of motion analysis and motion perception. *Annu Rev Psychol*, 55:181–205, 2004.
- KR Dobkins and TD Albright. What happens if it changes color when it moves?: the nature of chromatic input to macaque visual area mt. *J Neurosci*, 14:4854–4870, 1994.
- DW Dong and JJ Atick. Statistics of natural time varying images. *Netw Comput Neural Syst*, 6(3):345–358, 1995.
- R Dubner and SM Zeki. Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Res*, 35(2):528–532, 1971.

- DJ Dubowitz, DY Chen, JA Atkinson, M Scadeng, A Martinez, MB Andersen, RA Andersen, and WG Bradley Jr. Direct comparison of visual cortex activation in human and non-human primates using functional magnetic resonance imaging. *J Neurosci Methods*, 107:71–80, 2001.
- CJ Duffy and RH Wurtz. Sensitivity of mst neurons to optic flow stimuli i a continuum of response selectivity to large-field stimuli. *J Neurophysiol*, 65(6): 1329–1345, 1991a.
- CJ Duffy and RH Wurtz. Sensitivity of mst neurons to optic flow stimuli ii mechanisms of response selectivity revealed by small-field stimuli. *J Neurophysiol*, 65(6):1346–1359, 1991b.
- RA Eagle and BJ Rogers. Effects of dot density patch size and contrast on the upper spatial limit for direction discrimination in random-dot kinematograms. *Vision Res*, 37(15):2091–2102, 1997.
- M Edwards and DR Badcock. Global motion perception interaction of the on and off pathways. *Vision Res*, 34(21):2849–2858, 1994.
- M Edwards and S Nishida. Global-motion detection with transparent-motion signals. *Vision Res*, 39(13):2239–2249, 1999.
- M Edwards and S Nishida. Contrast-reversing global-motion stimuli reveal local interactions between first- and second-order motion signals. *Vision Res*, 44(16): 1941–1950, 2004.
- RC Emerson, MC Citron, WJ Vaugh, and SA Klein. Nonlinear directionally selective subunits in complex cells of cat striate cortex. *J Neurophysiol*, 58(1): 33–65, 1987.
- RC Emerson, JR Bergen, and EH Adelson. Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Res*, 32(2): 203–218, 1992.
- DC Essen and SM Zeki. The topographic organization of rhesus monkey prestriate cortex. *J Physiol*, 277:193–226, 1978.
- G Ettliger. Object vision and spatial vision: the neuropsychological evidence for the distinction. *Cortex*, 26:319–341, 1990.
- S Exner. Experimentelle untersuchung der einfachsten psychischen processe. *Pfluger's Arch Physiol*, 11:403–432, 1875.
- S Exner. Einige beobachtungen über bewegungsnachbilder. *Centralblatt für Physiologie*, 1:135–140, 1888.
- R Farivar. Dorsal-ventral integration in object recognition. *Brain Res Rev*, 61(2): 144–153, 2009.

## BIBLIOGRAPHY

---

- DJ Felleman and JH Kaas. Receptive-field properties of neurons in middle temporal visual area mt of owl monkeys. *J Neurophysiol*, 52:488–513, 1984.
- G Felsen, J Touryan, F Han, and Y Dan. Cortical sensitivity to visual features in natural scenes. *PLoS Biol*, 3:e342, 2005.
- C Fennema and W Thompson. Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, 9:301–315, 1979.
- DJ Field. Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Am A*, 4:2379–2394, 1987.
- KH Foster, JP Gaska, M Nagler, and DA Pollen. Spatial and temporal frequency selectivity of neurones in visual cortical areas v1 and v2 of the macaque monkey. *J Physiol*, 365:331–363, 1985.
- VH Franz, KR Gegenfurtner, HH Bülthoff, and M Fahle. Grasping visual illusions: no evidence for a dissociation between perception and action. *Psychol Sci*, 11(1):20–25, 2000.
- VH Franz, F Scharnowski, and KR Gegenfurtner. Illusion effects on grasping are temporally constant not dynamic. *J Exp Psychol Hum Percept Perform*, 31(6):1359–1378, 2005.
- RE Fredericksen, FA Verstraten, and WA Van de Grind. Spatio-temporal characteristics of human motion perception. *Vision Res*, 33(9):1193–1205, 1993.
- RE Fredericksen, FA Verstraten, and WA Van de Grind. An analysis of the temporal integration mechanism in human motion perception. *Vision Res*, 34(23):3153–3170, 1994a.
- RE Fredericksen, FA Verstraten, and WA Van de Grind. Spatial summation and its interaction with the temporal integration mechanism in human motion perception. *Vision Res*, 34(23):3171–3188, 1994b.
- RE Fredericksen, FA Verstraten, and WA Van de Grind. Temporal integration of random dot apparent motion information in human central vision. *Vision Res*, 34(4):461–476, 1994c.
- K Funke and F Worgotter. On the significance of temporally structured activity in the dorsal lateral geniculate nucleus (lgn). *Prog Neurobiol*, 52:67–119, 1997.
- TJ Gawne. Temporal coding as a means of information transfer in the primate visual. *Crit Rev Neurobiol*, 13:83–101, 1999.
- MS Gazzaniga, RB Ivry, and GR Mangun. *Perception and encoding*. W.W. Norton & Company, Inc, 2 edition, 2002.
- P Girard, PA Salin, and J Bullier. Response selectivity of neurons in area mt of the macaque monkey during reversible inactivation of area v1. *J Neurophysiol*, 67(6):1437–1446, 1992.

- AD Goodale, MA and Milner. Separate visual pathways for perception and action. *Trends Neurosci*, 15:20–25, 1992.
- MSA Graziano, RA Andersen, and R Snowden. Tuning of mst neurons to spiral stimuli. *J Neurosci*, 14:54–67, 1994.
- MW Greenlee. Human cortical areas underlying the perception of optic flow: brain imaging studies. *Int Rev Neurobiol*, 44:269–292, 2000.
- A Grunewald. A model of transparent motion and non-transparent motion after-effects. In DS Touretzky, MC Mozer, and ME Hasselmo, editors, *Advances in Neural Information Processing Systems*, pages 837–843. MIT Press, Cambridge, MA, 1996.
- A Grunewald and MJ Lankheet. Orthogonal motion after-effect illusion predicted by a model of cortical motion processing. *Nature*, 384(6607):358–360, 1996.
- P Hammond. On the response of simple and complex cells to random dot patterns: a reply to skottun, grosf and de valois. *Vision Res*, 31:47–50, 1991.
- P Hammond and DM MacKay. Response of cat visual cortical cells to kinetic contours and static noise. *J Physiol (Lond)*, 252:43–44, 1975.
- P Hammond and DM MacKay. Differential responsiveness of simple and complex cells in cat striate cortex to visual texture. *Exp Brain Res*, 30:275–296, 1977.
- S He, ER Cohen, and X Hu. Close correlation between activity in brain area mt/v5 and the perception of a visual motion aftereffect. *Curr Biol*, 8:1215–1218, 1998.
- DJ Heeger. Model for the extraction of image flow. *J Opt Soc Am*, 4:1455–1471, 1987.
- DJ Heeger. Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *J Neurophysiol*, 70:1885–1898, 1993.
- DJ Heeger, GM Boynton, JB Demb, E Seidemann, and WT Newsome. Motion opponency in visual cortex. *J Neurosci*, 19(16):7162–7174, 1999.
- DJ Heeger, AC Huk, WS Geisler, and DG Albrecht. Spikes versus bold: what does neuroimaging tell us about neuronal activity? *Nat Neurosci*, 3:631–633, 2000.
- HW Heuer and KH Britten. Optic flow signals in extrastriate area mst comparison of perceptual and neuronal sensitivity. *J Neurophysiol*, 91(3):1314–1326, 2004.
- KP Hoffmann and J Stone. Conduction velocity of afferents to cat visual cortex: a correlation with cortical receptive field properties. *Brain Res*, 32:460–466, 1971.
- DH Hubel and TN Wiesel. Receptive fields of single neurons in the cat’s striate cortex. *J Physiol*, 148:574–591, 1959.

## BIBLIOGRAPHY

---

- DH Hubel and TN Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, 160(1):106–154, 1962.
- DH Hubel and TN Wiesel. Shape and arrangement of columns in cat's striate cortex. *J Physiol*, 165:559–568, 1963.
- DH Hubel and TN Wiesel. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J Physiol*, 28:229–289, 1965.
- DH Hubel and TN Wiesel. Receptive fields and functional architecture of monkey striate cortex. *J Physiol*, 195(1):215–243, 1968.
- A Huk and MN Shadlen. Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J Neurosci*, 24(45):10420–10436, 2005.
- MR. Ibbotson and CWG Clifford. Interactions between on- and off-signals in directional motion detectors feeding the nucleus of the optic tract of the wallaby. *J Neurophysiol*, 86:997–1005, 2001.
- A Johnston, PW McOwan, and H Buxton. A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proc Biol Sci*, 250(1329):297–306, 1992.
- B Julesz. *Foundations of Cyclopean Perception*. Univ of Chicago Press, Chicago, IL, 1971.
- JJ Koenderink and AJ Van Doorn. Spatiotemporal contrast detection threshold surface is bimodal. *Optics Letters*, 4:32–34, 1979.
- JJ Koenderink, AJ Van Doorn, and WA Van de Grind. Spatial and temporal parameters of motion detection in the peripheral visual field. *J Opt Soc Am A*, 2(2):252–259, 1985.
- B Krekelberg and TD Albright. Motion mechanisms in macaque mt. *J Neurophysiol*, 93(5):2908–2921, 2005.
- B Krekelberg, GM Boynton, and RJ Van Wezel. Adaptation from single cells to bold signals. *Trends Neurosci*, 29(5):250–256, 2006a.
- B Krekelberg, RJ Van Wezel, and TD Albright. Adaptation in macaque mt reduces perceived speed and improves speed discrimination. *J Neurophysiol*, 95:255–270, 2006b.
- SW Kuffler. Discharge patterns and functional organization of mammalian retina. *J Neurophysiol*, 16:37–68, 1953.
- L Lagae, S Raviguel, and GA Orban. Speed and direction selectivity of macaque middle temporal neurons. *J Neurophysiol*, 69:19–39, 1993.



- L Lagae, H Maes, S Raiguel, D Xiao, and GA Orban. Responses of macaque sts neurons to optic flow components: a comparison of mt and mst. *J Neurophysiol*, 71:1597–1626, 1994.
- MJ Lankheet and FA Verstraten. Attentional modulation of adaptation to two-component transparent motion. *Vision Res*, 35(10):1401–1412, 1995.
- MJ Lankheet, RJ Van Wezel, and WA Van de Grind. Light adaptation and frequency transfer properties of cat horizontal cells. *Vision Res*, 31(7-8):1129–1142, 1991.
- MJ Lankheet, AJ Van Doorn, and WA Van de Grind. Spatio-temporal tuning of motion coherence detection at different luminance levels. *Vision Res*, 42(1): 65–73, 2002.
- SB Laughlin, RR De Ruyter van Steveninck, and JC Anderson. The metabolic cost of neural information. *Nat Neurosci*, 1:36–41, 1998.
- DT Lindsey and JT Todd. Opponent motion interactions in the perception of transparent motion. *Percept Psychophys*, 60(4):558–574, 1998.
- J Liu and WT Newsome. Functional organization of speed tuned neurons in visual area mt. *J Neurophysiol*, 89:246–256, 2003.
- J Liu and WT Newsome. Correlation between speed perception and neural activity in the middle temporal visual area. *J Neurosci*, 25:711–722, 2005.
- MS Livingstone and BR Conway. Substructure of direction-selective receptive fields in macaque v1. *J Neurophysiol*, 89(5):2743–2759, 2003.
- MS Livingstone, CC Pack, and RT Born. Two-dimensional substructure of mt receptive fields. *Neuron*, 30(3):781–793, 2001.
- JG Malpeli. Measuring eye position with the double magnetic induction method. *J Neurosci Methods*, 86:55–61, 1998.
- D Marr and S Ullman. Directional selectivity and its use in early visual processing. *Proc R Soc Lond B Biol Sci*, 211(1183):151–180, 1981.
- LM Martinez and AA Alonso. Complex receptive fields in primary visual cortex. *Neuroscientist*, 9:317–331, 2003.
- G Mather. The movement aftereffect and a distribution-shift model for coding the direction of visual movement. *Perception*, 9(4):379–392, 1980.
- G Mather. Computational modelling of motion detectors: responses to two-frame displays. *Spat Vis*, 5(1):1–14, 1990.
- G Mather, B Moulden, and A O’Halloran. Polarity specific adaptation to motion in the human visual system. *Vision Res*, 31(6):1013–1019, 1991.

## BIBLIOGRAPHY

---

- G Mather, F Verstraten, and S Anstis. *The motion aftereffect: A modern perspective*. MIT Press, Cambridge, Mass, 1998.
- JHR Maunsell and DC Van Essen. Functional properties of neurons in the middle temporal visual area mt of the macaque monkey i selectivity for stimulus direction speed and orientation. *J Neurophysiol*, 49:1127–1147, 1983.
- JW McClurkin, TJ Gawne, BJ Richmond, and LM Optican. Lateral geniculate neurons in behaving primates. ii. encoding of visual information in the temporal shape of the response. *J Neurophysiol*, 66:794–808, 1991.
- F Mechler and DL Ringach. On the classification of simple and complex cells. *Vision Res*, 40:1017–1033, 2002.
- F Mechler, DS Reich, and JD Victor. Detection and discrimination of relative spatial phase by v1 neurons. *J Neurosci*, 22:6129–6157, 2002.
- A Mikami, WT Newsome, and RH Wurtz. Motion selectivity in macaque visual cortex i mechanisms of direction and speed selectivity in extrastriate area mt. *J Neurophysiol*, 55:1308–1327, 1986a.
- A Mikami, WT Newsome, and RH Wurtz. Motion selectivity in macaque visual cortex ii spatio-temporal range of directional interactions in mt and v1. *J Neurophysiol*, 55:1328–1339, 1986b.
- AD Milner and MA Goodale. Visual pathways to perception and action. *Prog Brain Res*, 95:317–337, 1993.
- AD Milner and MA Goodale. *The visual brain in action*. Oxford University Press, Oxford, 2 edition, 2006.
- AD Milner and MA Goodale. Two visual systems re-viewed. *Neuropsychologia*, 95:774–785, 2008.
- M Mishkin and LG Ungerleider. Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behav Brain Res*, 9(1): 57–77, 1982.
- CH Mo and C Koch. Modeling reverse-phi motion-selective neurons in cortex double synaptic-veto mechanism. *Neural Comput*, 15(4):735–759, 2003.
- MJ Morgan and R Ward. Conditions for motion flow in dynamic visual noise. *Vision Res*, 20(5):431–435, 1980.
- MJ Morgan, R Perry, and M Fahle. The spatial limit for motion detection in noise depends on element size not on spatial frequency. *Vision Res*, 37(6):729–736, 1997.
- MC Morrone and DC Burr. Feature detection in human vision: a phase-dependent energy model. *Proc R Soc Lond B Biol Sci*, 235(1280):221–245, 1988.

- JA Movshon and WT Newsome. Visual response properties of striate cortical neurons projecting to area mt in macaque monkeys. *J Neurosci*, 16(23):7733–7741, 1996.
- JA Movshon, ID Thompson, and DJ Tolhurst. Spatial summation in the receptive fields of simple cells in the cat’s striate cortex. *J Physiol*, 283:53–57, 1978a.
- JA Movshon, ID Thompson, and DJ Tolhurst. Receptive field organization of complex cells in the cat’s striate cortex. *J Physiol*, 283:79–99, 1978b.
- JA Movshon, ID Thompson, and DJ Tolhurst. Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat’s visual cortex. *J Physiol*, 283:101–120, 1978c.
- JA Movshon, EH Adelson, MS Gizzi, and WT Newsome. The analysis of visual moving patterns. *Exp Brain Res*, 11:117–152, 1986.
- R Nelson, EV Famiglietti Jr, and H Kolb. Intracellular staining reveals different levels of stratification for on- and off-center ganglion cells in cat retina. *J Neurophysiol*, 41:472–483, 1978.
- WT Newsome, KH Britten, and JA Movshon. Neuronal correlates of a perceptual decision. *Nature*, 341(6237):52–54, 1989.
- WT Newsome, KH Britten, CD Salzman, and JA Movshon. Neuronal mechanisms of motion perception. *Cold Spring Harb Symp Quant Biol*, 55:697–705, 1990.
- H Nover, CH Anderson, and GC DeAngelis. A logarithmic scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *J Neurosci*, 25(43):10049–10060, 2005.
- GA Orban, L Lagae, A Verri, S Raiguel, D Xiao, H Maes, and V Torre. First-order analysis of optical flow in monkey brain. *Proc Natl Acad Sci U S A*, 89(7):2595–2599, 1992.
- GA Orban, D Van Essen, and W Vaduffel. Comparative mapping of higher visual areas in monkeys and humans. *Trends Cogn Sci*, 8(7):315–324, 2004.
- CC Pack, BR Conway, RT Born, and MS Livingstone. Spatiotemporal structure of nonlinear subunits in macaque visual cortex. *J Neurosci*, 26:893–907, 2006.
- TV Papathomas, A Gorea, and C Chubb. Precise assessment of the effective mean luminance of texture patches an approach based on reverse-phi apparent motion. *Vision Res*, 36:3775–3784, 1996.
- JA Perge, BG Borghuis, RJE Bours, MJM Lankheet, and RJA Van Wezel. Dynamics of directional selectivity in mt receptive field centre and surround. *Eur J Neurosci*, 22:2049–2058, 2005a.

## BIBLIOGRAPHY

---

- JA Perge, BG Borghuis, RJE Bours, MJM Lankheet, and RJA Van Wezel. Temporal dynamics of direction tuning in motion-sensitive macaque area mt. *J Neurophysiol*, 93:2104–2116, 2005b.
- JA Perrone. A visual motion sensor based on the properties of v1 and mt neurons. *Vision Res*, 44(15):1733–1755, 2004.
- JA Perrone and A Thiele. Speed skills measuring the visual speed analyzing properties of primate mt neurons. *Nat Neurosci*, 4:526–532, 2001.
- SE Petersen, JF Baker, and JM Allman. Direction-specific adaptation in area mt of the owl monkey. *Brain Res*, 346:146–150, 1985.
- NJ Priebe and SG Lisberger. Constraints on the source of short-term motion adaptation in macaque area mt ii tuning of neural circuit mechanisms. *J Neurophysiol*, 88:370–382, 2002.
- NJ Priebe, MM Churchland, and SG Lisberger. Constraints on the source of short-term motion adaptation in macaque area mt i the role of input and intrinsic mechanisms. *J Neurophysiol*, 88:354–369, 2002.
- NJ Priebe, CR Cassanello, and SG Lisberger. The neural representation of speed in macaque area mt/v5. *J Neurosci*, 23:5650–5661, 2003.
- NJ Priebe, SG Lisberger, and JA Movshon. Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *J Neurosci*, 26:2941–2950, 2006.
- N Qian and RC Anderson. Transparent motion perception as detection of unbalanced motion signals ii: Physiology. *J Neurosci*, 14:7367–7380, 1994.
- N Qian, RC Anderson, and EH Adelson. Transparent motion perception as detection of unbalanced motion signals i: Psychophysics. *J Neurosci*, 14:7357–7366, 1994.
- SE Raiguel, DK Xiao, VL Marcar, and GA Orban. Response latency of macaque area mt/v5 neurons and its relationship to stimulus parameters. *J Neurophysiol*, 82:1944–1956, 1999.
- G Rees, K Friston, and C Koch. A direct quantitative relationship between the functional properties of human and macaque v5. *Nat Neurophysiol*, 3:716–723, 2000.
- W Reichardt. Autocorrelation a principle for the evaluation of sensory information by the central nervous system. In Rosenblith, editor, *Sensory communication*, pages 303–317. New York Wiley, 1961.
- JP Reulen and L Bakker. The measurement of eye movement using double magnetic induction. *IEEE Trans Biomed Eng*, 29:740–744, 1982.

- JG Robson. Spatial and temporal contrast sensitivity functions of the visual system. *J Opt Soc Am*, 56:97–108, 1966.
- HR Rodman and TD Albright. Coding of visual stimulus velocity in area mt of the macaque. *Vision Res*, 27:2035–2048, 1987.
- HR Rodman and TD Albright. Single-unit analysis of pattern-motion selective properties in the middle temporal visual area mt. *Exp Brain Res*, 75:53–64, 1989.
- HR Rodman, CG Gross, and TD Albright. Afferent basis of visual response properties in area mt of the macaque. ii. effects of superior colliculus removal. *J Neurosci*, 10:1154–1164, 1990.
- H Saito, M Yukie, K Tanaka, K Hikosaka, Y Fukada, and E Iwai. Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J Neurosci*, 6:145–157, 1986.
- H Sakata, H Shibusutani, K Kawano, and TL Harrington. Neural mechanisms of space vision in the parietal association cortex of the monkey. *Vision Res*, 25(3):453–463, 1985.
- H Sakata, H Shibusutani, Y Ito, K Tsurugai, S Mine, and M Kusunoki. Functional properties of rotation-sensitive neurons in the posterior parietal association cortex of the monkey. *Exp Brain Res*, 101:183–202, 1994.
- CD Salzman, KH Britten, and WT Newsome. Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, 346(6280):174–177, 1990.
- CD Salzman, CM Murasugi, KH Britten, and WT Newsome. Microstimulation in visual area mt effects on direction discrimination performance. *J Neurosci*, 12(6):2331–2355, 1992.
- T Sato. Reversed apparent motion with random dot patterns. *Vision Res*, 29(12):1749–1758, 1989.
- AB Saul, PL Carras, and AL Humphrey. Temporal properties of inputs to direction-selective neurons in monkey v1. *J Neurophysiol*, 94(1):282–294, 2005.
- PH Schiller. Central connections of the retinal on and off pathways. *Nature*, 297(5867):580–583, 1982.
- PH Schiller. The connections of the retinal on and off pathways to the lateral geniculate nucleus of the monkey. *Vision Res*, 24(9):926–932, 1984.
- PH Schiller. The on and off channels in the visual system. In B Cohen and I Bodis-Wollner, editors, *Vision and the brain: The organization of the central visual system*, pages 35–41. Raven Press, New York, 1990.

## BIBLIOGRAPHY

---

- PH Schiller. The on and off channels of the visual system. *Trends Neurosci*, 15 (3):86–92, 1992.
- PH Schiller, JH Sandell, and JH Maunsell. Functions of the on and off channels of the visual system. *Nature*, 322(6082):824–825, 1986.
- MI Sereno, AM Dale, JB Reppas, KK Kwong, JW Belliveau, TJ Brady, BR Rosen, and RBH Tootell. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268:889–893, 1995.
- OH Shade. Optical and photoelectric analog of the eye. *J Opt Soc Am*, 46:721–739, 1956.
- MN Shadlen, KH Britten, WT Newsome, and JA Movshon. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci*, 16(4):1486–1510, 1996.
- EP Simoncelli and DJ Heeger. A model of neural responses in visual area mt. *Vision Res*, 38:743–761, 1998.
- BC Skottun, RL De Valois, DH Grosop, JA Movshon, DG Albrecht, and AB Bonds. Classifying simple and complex cells on the basis of response modulation. *Vision Res*, 31:1079–1086, 1991.
- M Slaughter and R Miller. 2-amino-4-phosphonobutyric acid: a new pharmacological tool for retina research. *Science*, 211:182–185, 1981.
- RJ Snowden and OJ Braddick. The combination of motion signals over time. *Vision Res*, 29(11):1621–1630, 1989.
- RJ Snowden, S Treue, RG Erickson, and RA Andersen. The response of area mt and v1 neurons to transparent motion. *J Neurosci*, 11(9):2768–2785, 1991.
- G Sperling. Three stages and two systems of visual processing. *Spatial Vision*, 4: 183–207, 1989.
- RM Steinman. Phi is not beta, and why wertheimer’s discovery launched the gestalt revolution. *Vision Res*, 40(17):2257–2264, 2000.
- CF Stevens. An evolutionary scaling law for the primate visual system and its basis in cortical function. *Nature*, 411:193–195, 2001.
- J Stone, B Dreher, and A Leventhal. Hierarchical and parallel mechanisms in the organization of visual cortex. *Brain Res*, 180:345–394, 1979.
- Y Sugita and K Tanaka. Occlusion-related cue used for analysis of motion in the primate visual cortex. *Neuroreport*, 2:751–754, 1991.
- K Tanaka, Y Fukada, and H Saito. Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol*, 62:642–656, 1989.

- K Tanaka, Y Sugita, M Moriya, and H Saito. Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. *J Neurophysiol*, 69:128–142, 1993.
- JG Taylor, N Schmitz, K Ziemons, ML Grosse-Ruyken, O Gruber, HW Mueller-Gaertner, and NJ Shah. The network of brain areas involved in the motion aftereffect. *NeuroImage*, 11:257–270, 2000.
- DJ Tolhurst. Separate channels for the analysis of the shape and movement of a moving visual stimulus. *J. Physiol*, 231:385–402, 1973.
- DJ Tolhurst and JA Movshon. Spatial and temporal contrast sensitivity of striate cortex neurons. *Nature*, 176:87–100, 1975.
- RBH Tootell and JB Taylor. Anatomical evidence for mt and additional cortical visual areas in humans. *Cereb Cortex*, 5:39–55, 1995.
- RBH Tootell, JB Reppas, AM Dale, RB Look, MI Sereno, R Malach, TJ Brady, and BR Rosen. Visual motion aftereffect in human cortical area mt revealed by functional magnetic resonance imaging. *Nature*, 375:139–141, 1995.
- RBH Tootell, AM Dale, MI Sereno, and R Malach. New images from human visual cortex. *Trends Neurosci*, 375:481–489, 1996.
- RBH Tootell, J Mendola, NK Hadjikhani, PJ Ledden, AK Liu, JB Reppas, MI Sereno, and AM Dale. Functional analysis of v3a and related areas in human visual cortex. *J Neurosci*, 17(18):7060–7078, 1997.
- LG Ungerleider and JV Haxby. 'what' and 'where' in the human brain. *Curr Opin Neurobiol*, 4(2):157–165, 1994.
- LM Vaina. Functional segregation of color and motion processing in the human visual cortex: clinical evidence. *Cereb Cortex*, 4(5):555–572, 1994.
- WA Van de Grind. The possible structure and role of neuronal smart mechanisms in vision. *Cognit Systems*, 2:163–180, 1988.
- WA Van de Grind, JJ Koenderink, and AJ Van Doorn. The distribution of human motion detector properties in the monocular visual field. *Vision Res*, 26(5):797–810, 1986.
- WA Van de Grind, JJ Koenderink, and AJ Van Doorn. Influence of contrast on foveal and peripheral detection of coherent motion in moving random-dot patterns. *J Opt Soc Am A*, 4(8):1643–1652, 1987.
- WA Van de Grind, JJ Koenderink, and AJ Van Doorn. Viewing-distance invariance of movement detection. *Exp Brain Res*, 91(1):135–150, 1992.
- WA Van de Grind, MJ Lankheet, and R Tao. A gain-control model relating nulling results to the duration of dynamic motion aftereffects. *Vision Res*, 43(2):117–133, 2003.

## BIBLIOGRAPHY

---

- WA Van de Grind, MJ Van der Smagt, and FA Verstraten. Storage for free a surprising property of a simple gain-control model of motion aftereffects. *Vision Res*, 44(19):2269–2284, 2004.
- AV Van den Berg and WA Van de Grind. Reaction times to motion onset and motion detection thresholds reflect the properties of bilocal motion detectors. *Vision Res*, 29(9):1261–1266, 1989.
- AJ Van Doorn and JJ Koenderink. Temporal properties of the visual detectability of moving spatial white noise. *Exp Brain Res*, 45(1-2):179–188, 1982a.
- AJ Van Doorn and JJ Koenderink. Spatial properties of the visual detectability of moving spatial white noise. *Exp Brain Res*, 45(1-2):189–195, 1982b.
- AJ Van Doorn and JJ Koenderink. Spatiotemporal integration in the detection of coherent motion. *Vision Res*, 24(1):47–53, 1984.
- DC Van Essen, CH Anderson, and DJ Felleman. Information processing in the primate visual system: An integrated systems perspective. *Science*, 255(5043):419–423, 1992.
- JP Van Santen and G Sperling. Temporal covariance model of human motion perception. *J Opt Soc Am A*, 1(5):451–473, 1984.
- JP Van Santen and G Sperling. Elaborated reichardt detectors. *J Opt Soc Am A*, 2(2):300–321, 1985.
- RJ Van Wezel and KH Britten. Motion adaptation in area mt. *J Neurophysiol*, 88(6):3469–3476, 2002a.
- RJ Van Wezel and KH Britten. Multiple uses of visual motion. the case for stability in sensory cortex. *Neuroscience*, 111(4):739–759, 2002b.
- RJ van Wezel, FA Verstraten, RE Fredericksen, and WA Van de Grind. Spatial integration in coherent motion detection and in the movement aftereffect. *Perception*, 23(10):1189–1195, 1994.
- DI Vaney, S He, WR Taylor, and WR Levick. Direction-selective ganglion cells in the retina. In JM Zanker and J Zeil, editors, *Motion Vision - Computational, Neural, and Ecological Constraints*, pages 11–56. Springer Verlag, Berlin Heidelberg New York, 2001.
- FA Verstraten, RE Fredericksen, and WA Van de Grind. Movement aftereffect of bi-vectorial transparent motion. *Vision Res*, 34(3):349–358, 1994.
- FA Verstraten, RE Fredericksen, RJ Van Wezel, JC Boulton, and WA Van de Grind. Directional motion sensitivity under transparent motion conditions. *Vision Res*, 36(15):2333–2336, 1996a.



- FA Verstraten, RE Fredericksen, RJ Van Wezel, MJ Lankheet, and WA Van de Grind. Recovery from adaptation for dynamic and static motion aftereffects: evidence for two mechanisms. *Vision Res*, 36(3):2333–2336, 1996b.
- FA Verstraten, MJ Van der Smagt, and WA Van de Grind. Aftereffect of high-speed motion. *Perception*, 28(11):1397–1411, 1998.
- FA Verstraten, MJ Van der Smagt, RE Fredericksen, and WA Van de Grind. Integration after adaptation to transparent motion: static and dynamic test patterns result in different aftereffect directions. *Vision Res*, 39(4):803–810, 1999.
- AB Watson and AJ Jr Ahumada. Model of human visual-motion sensing. *J Opt Soc Am A*, 2(2):322–341, 1985.
- AB Watson and DG Pelli. Quest a bayesian adaptive psychometric method. *Percept Psychophys*, 33(2):113–120, 1983.
- AB Watson, AJ Jr Ahumada, and J Farrell. Window of visibility: psychophysical theory of fidelity in time-sampled visual motion displays. *J Opt Soc Am A*, 3(3):300–307, 1986.
- C Wehrhahn. Reversed phi revisited. *J Vision*, 6(10):1018–1025, 2006.
- C Wehrhahn and D Rapf. On- and off-pathways form separate neural substrates for motion perception psychophysical evidence. *J Neurosci*, 12(6):2247–2250, 1992.
- Y Weiss, EP Simoncelli, and EH Adelson. Motion illusions as optimal percepts. *Nat Neurosci*, 6(5):598–604, 2002.
- FA Wichmann and NJ Hill. The psychometric function i fitting sampling and goodness of fit. *Percept Psychophys*, 63(8):1293–1313, 2001a.
- FA Wichmann and NJ Hill. The psychometric function ii bootstrap-based confidence intervals and sampling. *Percept Psychophys*, 63(8):1314–1329, 2001b.
- SM Zeki. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J Physiol*, 236:549–573, 1974.
- J Zihl and D Von Cramon. Collicular function in human vision. *Exp Brain Res*, 35(3):419–424, 1979.



# Overzicht, samenvatting en conclusies (Dutch)

Dit proefschrift gaat over de corticale mechanismen die ten grondslag liggen aan de bewegingsdetectie in de retino-geniculo-corticale route, die onderdeel is van het visuele systeem in het brein van mensen en andere primaten. Deze route speelt een essentiële rol in de verwerking van visuele bewegingsinformatie en de tot standkoming van het globale bewegingspercept. De afgelopen decennia's zijn er talrijke studies uitgevoerd met betrekking tot de verwerking van beweging. Deze onderzoeken hebben geleid tot de consensus dat de detectie van beweging via de retino-geniculo-corticale route in een vroeg stadium binnen deze keten plaatsvindt (o.a. Hubel and Wiesel, 1959, 1962, 1963; Tolhurst, 1973; Tolhurst and Movshon, 1975; Movshon et al., 1978a,b,c; Zeki, 1974; Essen and Zeki, 1978; Albright, 1984; Maunsell and Van Essen, 1983; Mikami et al., 1986b). De eerste bewegingsgevoelige neuronen (zenuwen) in deze route bevinden zich in de primaire visuele cortex: Dit gebied wordt ook striate cortex of kortweg V1 genoemd. Deze neuronen detecteren lokale beweging en hebben verbindingen met neuronen in de medio-temporaalkwab (midden van de slaapkwab; MT) en andere extra-striate gebieden, o.a. naar de medio-superiore temporaalkwab (midden-boven van de slaapkwab; MST). MT cellen verwerken de informatie van meerdere lokale bewegingsdetectoren uit V1 en zijn betrokken bij de globale bewegingsdetectie (o.a. Movshon et al., 1986; Born and Bradley, 2005; Albright, 1984; Britten et al., 1993; Maunsell and Van Essen, 1983; Perge et al., 2005a,b). Bewegingsdetectie in ons visuele systeem kan onderverdeeld worden in drie fases: (1) Het voorverwerken van visuele informatie, (2) lokale bewegingsdetectie, (3) het naverwerken van lokale bewegingsinformatie. De bevindingen in dit proefschrift zijn relevant voor een beter begrip over de werking van het visuele systeem van primaten. Zij geven een nieuw inzicht in de mechanismen van lokale bewegingsdetectie en de integratie tot een globaal bewegingspercept door het visuele systeem.

---

## Hoofdstuk 3

In hoofdstuk 3 hebben we de temporele karakteristieken van het corticale mechanisme dat ten grondslag ligt aan het waarnemen van beweging bestudeerd. Naar aanleiding van een reeks psychofysische (Fredericksen et al., 1993, 1994a,b,c; Van Doorn and Koenderink, 1982a,b, 1984; Van de Grind et al., 1986, 1987; Baker and Braddick, 1985; Morgan and Ward, 1980) en fysiologische experimenten (Borghuis et al., 2003; Perge et al., 2005a,b), hebben we vergelijkbare visuele stimuli (random (willekeurig) dot patronen) gebruikt om de fysiologie van de verwerking van beweging te bestuderen. Perge et al. (2005a,b) maakten gebruik van een omgekeerde-correlatie (reverse-correlation) techniek (Borghuis et al., 2003) om de richtingsgevoeligheid van MT cellen in de tijd en de ruimtelijk interactie tussen centrum en omgeving te bestuderen. De studie in dit hoofdstuk is een logische uitbreiding naar het domein van snelheid. We hebben een vergelijkbare techniek gebruikt om de temporele dynamiek en tuning van MT neuronen voor verplaatsingen van random dot stimuli te meten. Hiervoor hebben we de respons van MT cellen op bewegende random dot patronen gemeten als functie van de stapgrootte (de verplaatsing van een dot; variërend met een visuele hoek van 2,4 tot 76,8 boogminuten) in de voorkeurs- en niet-voorkeursrichting van het onderzochte neuron en als functie van verschillende tijdsintervallen (tijd tussen een verplaatsing; variërend van 8,3 tot 66,7 ms) van de stimulus (prikkel).

De resultaten laten zien dat de de respons van de cellen over het algemeen bifasisch is, met een gemiddelde piek-latentie van ongeveer 54 ms, zowel in de voorkeurs- en niet-voorkeursrichting. De temporele dynamieken waren voor een groot deel onafhankelijk van de stapgrootte. De meest opvallende conclusie is dat, hoewel de optimale stapgroottes in de voorkeurs- en niet-voorkeursrichtingen vergelijkbaar zijn, de breedte van de stapgrootte tuning in de niet-voorkeursrichting breder is dan in de voorkeursrichting.

Het vergroten van het tijdsinterval tijdens een verplaatsing van het patroon had opvallend weinig effect. Optimale stapgroottes, en de breedte van de stapgrootte tuning werden niet beïnvloed door veranderingen in het temporele interval. Latentietijden en het bifasische gedrag waren ook grotendeels onafhankelijk van het inter-stimulus interval.

Het stimulus paradigma in dit onderzoek was gebaseerd op de principes van eerdere psychofysische studies die de spatio-temporele karakteristieken van het menselijk bewegingssysteem onderzochten (o.a. Fredericksen et al., 1993, 1994a,b,c). Voor een inter-stimulus interval met tussenpozen langer dan een monitor frame, werd hierbij het random dot patroon statisch weergegeven. In essentie betekent dit dat de spatio-temporele impuls voor tussenpozen langer dan een frame altijd een frame is. Dit paradigma heeft beperkingen bij het meten van de temporele eigenschappen, doordat de temporele integratie van stimuli met verschillende inter-stimulus intervallen verschilt. Deze constatering was de aanleiding voor de studie beschreven in hoofdstuk 4.

## Hoofdstuk 4

Detectie van schijnbare beweging (apparent motion) in random dot patronen vereist correlatie in tijd en ruimte. In hoofdstuk 4 introduceren we een nieuwe bewegingsstimulus die het mogelijk maakt om niet alleen de spatiele (ruimtelijke), maar ook de tuning van temporele eigenschappen van de menselijke bewegingsperceptie in detail te meten. De verwerking van beweging in de retino-geniculocorticale route kan worden onderverdeeld in drie fases: pre-filtering van visuele informatie, lokale bewegingsdetectie en post-processing in de volgende fasen. Lokale bewegingsdetectie impliceert een correlatie stap, namelijk het in verband brengen van twee (of meerdere) posities in de tijd. Onze stimulus maakt het mogelijk om specifiek de tuning voor het temporele interval in de correlatie stap te onderzoeken, en daarbij de pre-filtering en post-integratie constant te houden. Morgan and Ward (1980) hebben een soortgelijk principe gebruikt om spatio-temporele limieten in het bewegingsysteem te meten. Hun stimulus paradigma is echter niet geschikt om de eigenschappen van de spatio-temporele tuning te meten.

Onze stimulus bestaat uit een random dot patroon waarin elke dot in twee monitorframes, gescheiden door een bepaald tijdsinterval, wordt getoond. In elk frame, wordt de helft van de dots vernieuwd en de andere helft is een verplaatste reïncarnatie van het patroon dat één of meerdere frames eerder is gegenereerd. Bewegingsenergie-statistieken in deze stimulus variëren niet tussen frames, en de richtingsbias in de spatio-temporele correlaties is constant ongeacht de inter-stimulus interval. In deze studie hebben we de bewegingsgevoeligheid van proefpersonen gemeten in een links-rechts bewegingsdiscriminatie experiment als functie van de stapgrootte en inter-stimulus interval waarbij de coherentie (samenhang tussen dots) werd gevarieerd via een Quest-staircase (stapsgewijs) procedure. De resultaten laten zien dat proefpersonen het meest gevoelig zijn voor stimuli met een inter-stimulus interval tussen 17 en 42 ms. De temporele limieten waarbinnen de discriminatie taak door de proefpersonen goed kon worden uitgevoerd, komen overeen met de studie van Morgan and Ward (1980). De maximale temporele limiet was kleiner dan beschreven in andere studies (Fredericksen et al., 1993, 1994a,b,c). We concluderen dat dit verschil in temporele gevoeligheid kan worden verklaard door verschillen in de stimulus paradigma's. We hebben ook gekeken naar de afhankelijkheid tussen stapgroottes en inter-stimulus intervallen. Morgan and Ward (1980) hebben spatio-temporele limieten gemeten. Op basis van interpolatie van hun gegevens, kwamen zij tot de conclusie dat stapgrootte en interval tuning onafhankelijk waren. Onze resultaten laten echter zien dat de tuning voor stapgrootte en temporeel interval een beperkte mate van afhankelijkheid kent: Het optimale temporele interval neemt af met toenemende stapgrootte en de optimale stapgrootte neemt af als het temporele interval groter wordt.

---

## Hoofdstuk 5

In de fysiologische studie in hoofdstuk 3 troffen we een aantal cellen aan die een grote positieve responsmodulatie hebben voor stappen in de niet-voorkeursrichting en een latency iets groter dan de optimale latentie van de onderzochte cel. Een mogelijke verklaring hiervoor is een bewegingsillusie, reverse-phi. Reverse-phi is de illusie dat een waarnemer een bewegend object dat in een bepaalde richting beweegt, in de tegenovergestelde richting ziet bewegen. Deze illusie kan worden opgewekt door het contrast van het bewegend object en zijn omgeving om te keren. In random dot patronen treedt deze contrast omkering (onbedoeld) ook op. Dit besef was de trigger voor een nieuwe reeks van psychofysische studies.

Contrast informatie is in de retino-geniculo-corticale route vertegenwoordigd in twee verschillende kanalen: ‘AAN’ (ON-center) cellen verwerken positieve contrasten en ‘UIT’ (OFF-center) cellen negatieve contrasten. In deze studie gaan we in op de vraag of de analyse van locale beweging afzonderlijk wordt uitgevoerd in deze twee kanalen (geen interactie), of ook door het combineren van signalen van ON en OFF cellen (wel interactie). Hiervoor hebben we de gevoeligheid om beweging te detecteren voor normale en reverse-phi bewegingsstimuli kwantitatief vergeleken. In een reverse-phi stimulus keert het contrast van een patroon om tijdens verplaatsingen. Voor een reverse-phi stimulus kan de gevoeligheid worden afgeleid uit de correlatie tussen entiteiten met een positief en negatief contrast, terwijl voor normale beweging deze is gebaseerd op de correlatie van entiteiten met vergelijkbare contrasten.

Een eenvoudige aanpassing in het stimulus paradigma uit hoofdstuk 4 stelde ons in staat om een reverse-phi percept op te wekken, de spatio-temporele eigenschappen te onderzoeken en te vergelijken met de spatio-temporele eigenschappen van normale beweging. Visuele stimuli bestonden uit random dot patronen waarin consistente bewegingsinformatie was beperkt tot correlaties tussen vergelijkbare contrastpolariteit (normale beweging) of tot tegenovergestelde contrastpolariteiten (reverse-phi beweging). Dots waren zwart of wit op een grijze achtergrond. Bewegingsinformatie werd beperkt tot een enkele combinatie van stapgrootte en inter-stimulus interval. Het aantal zwarte en witte dots was voor elk frame gelijk, ongeacht het stimulustype (reverse-phi of normale beweging), de stapgrootte en het temporele interval. De gevoeligheid hebben we gemeten in een links-rechts bewegingsdiscriminatie experiment als functie van de stapgrootte en inter-stimulus interval waarbij de coherentie (samenhang tussen dots) werd gevarieerd via een Quest-staircase (stapsgewijs) procedure.

Dit paradigma maakte het mogelijk om de eigenschappen van de spatiele en temporele tuning voor reverse-phi kwantitatief te meten en te vergelijken met de eigenschappen van normale beweging. De resultaten laten zien dat de tuning voor stapgrootte en temporele interval voor beide bewegingstypen sterk overeenkomt, ongeacht de dotdichtheid van de stimulus. Om te testen of de afwezigheid van bewegingsinformatie voldoende was om een reverse-phi percept te induceren en zo de gevoeligheid ervan uit te leggen, hebben we ook de gevoeligheid gemeten

voor deze zogenoemde no-phi bewegingsstimuli. In een no-phi stimulus wordt het contrast van de verplaatste dots ingesteld op de achtergrond luminantie. Indien de stimulus alleen zwarte of witte dots bevatte, was de gevoeligheid voor de no-phi stimulus laag. Wanneer beide contrastpolariteiten in de stimulus aanwezig was, kon geen beweging worden waargenomen (ongevoelig). De no-phi experimenten laten zien dat het ontbreken van bewegingsinformatie de gevoeligheid uit de reverse-phi experimenten niet kan verklaren. Deze studie laat duidelijk zijn dat het bewegingssysteem even gevoelig is voor normale beweging en reverse-phi beweging. Deze bevindingen suggereren dat efficiënte detectie via correlatie tussen ON en OFF kanalen optreedt. In het volgende hoofdstuk gaan we in op de vraag hoe de signalen van de ON en OFF kanalen worden gecombineerd.

## Hoofdstuk 6

Reverse-phi beweging wordt doorgaans verklaard doordat het omkeren van contrast in opeenvolgende monitorframes de bewegingsenergie in de stimulus verschuift naar de tegenovergestelde richting. In deze verklaring is de verandering van het bewegingspercept triviaal, omdat elke geschikte bewegingsenergie-detector gevoelig voor de energie in de tegengestelde richting wordt gestimuleerd. Deze redenering verklaart echter niet welk mechanisme ten grondslag ligt aan de detectie van normale en reverse-phi beweging. Daarnaast geeft deze redenering geen verklaring voor de grote overeenkomst in de spatio-temporele tuning en gevoeligheid voor normale en reverse-phi beweging en de ongevoeligheid voor no-phi stimuli (hoofdstuk 5).

In dit hoofdstuk gaan we in op de vraag hoe de signalen uit de ON en OFF kanalen gecombineerd kunnen worden. Het eerste niveau waarop de ON en OFF kanalen samenkomen is in de primaire visuele cortex (Slaughter and Miller, 1981). In de retino-geniculo-corticale route, zijn de eerste bewegingsgevoelige cellen ge-localiseerd in de primaire visuele cortex (o.a. Hubel and Wiesel, 1959, 1962, 1965, 1968; Zeki, 1974; Essen and Zeki, 1978; Albright, 1984; Maunsell and Van Essen, 1983; Mikami et al., 1986b). Mo and Koch (2003) suggereerden dat op dit niveau correlaties tussen tegengestelde contrastpolariteiten voorkomen die vervolgens leiden tot activering van bewegingsdetectoren die gevoelig zijn voor beweging in de tegenovergestelde richting. Deze redenering betekent dat reverse-phi beweging zich hetzelfde zal gedragen als normale beweging. Echter, een alternatieve verklaring is mogelijk, namelijk dat correlaties met tegengestelde contrastpolariteiten leiden tot een verminderde activiteit van bewegingsdetectoren gevoelig voor beweging in dezelfde richting dan de fysieke beweging.

In deze studie tonen we aan dat deze alternatieve hypothese een betere verklaring geeft. Het bewijs is gebaseerd op een nieuwe reeks van perceptuele fenomenen waaruit blijkt dat reverse-phi percepten zich in veel opzichten gedragen als bewegingsnaeffecten, en niet, zoals voorgesteld door Mo and Koch (2003), als normale beweging. Bewegingsadaptatie leidt tot verminderde activiteit van

---

de geactiveerde bewegingsdetectoren. De verminderde activiteit kan zichtbaar gemaakt worden door het aanbieden van een statische test stimulus. De verminderde activiteit van de betreffende bewegingsdetectoren leidt in de volgende fase, als gevolg van bewegingsopponentie, tot de waarneming van beweging in de tegenovergestelde richting (het bewegingsnaeffect). Onze resultaten tonen aan dat reverse-phi beweging zich gedraagt als een bewegingsnaeffect en de activiteit van bewegingsdetectoren vermindert. Correlaties tussen gelijke contrastpolariteiten verhogen de activiteit en worden gebruikt als ‘positief’ bewijs dat een bewegingscomponent aanwezig is. Correlaties tussen tegengestelde polariteiten verlagen de activiteit en leveren ‘negatief’ bewijs dat een bewegingscomponent afwezig is. Een dergelijk mechanisme is voordelig voor het visuele systeem, omdat het afwegen van positief en negatief bewijs van bewegingscomponenten op het laagste niveau van bewegingsdetectie de signaal-ruis verhouding vergroot en daardoor leidt tot een efficiënter systeem voor bewegingsdetectie.

## Hoofdstuk 7

In de algemene discussie worden de bevindingen uit de eerdere hoofdstukken geïntegreerd en in een bredere context geplaatst. Als eerste ga ik in op de principes van onze nieuwe stimuli. Vervolgens presenteer ik een model dat de essenties meeneemt van modellen uit de literatuur die de drie fases van de verwerking van beweging beschrijven. De bevindingen uit dit proefschrift zijn verwerkt in dit model en de onderbouwing ervan. Het hoofdstuk eindigt met een slotbeschouwing en toekomstperspectieven. De toekomstperspectieven bevatten suggesties voor verder onderzoek en enkele voorlopige resultaten die onze conclusies ondersteunen.



# Dankwoord (Dutch)

En hier, het meest gelezen hoofdstuk van een proefschrift. Een promotie doe je niet alleen. Hierbij wil ik dan ook iedereen bedanken die me op welke wijze dan ook heeft geholpen. Een aantal mensen wil ik hier in het bijzonder bedanken.

Allereerst mijn directe begeleider, Martin. Een ware onderzoeker en mijn mentor in mijn biologische carrière. Eerst tijdens mijn doctoraal stage en daarna tijdens dit promotietraject. Ik dank Martin voor zijn vertrouwen, zijn enthousiasme, de discussies en stimulerende gesprekken. Ik heb heel veel geleerd in deze periode. En verder natuurlijke de gesprekken naast het werk, de mooie verhalen over mountainbiken en valpartijen.

Als tweede... Richard van Wezel. Ik wil je bedanken voor je enthousiasme en het stimuleren om verder te kijken dan het NEST. Met dank aan jou ben ik op veel plekken buiten Utrecht geweest. Het begon al met de retraite op Schiermonnikoog waar ik als pre-AIO mee mocht om kennis te maken met de Helmholtz groep. Vervolgens ben ik, dankzij jou en je connecties, naar de Summerschool in Wakoshi in Japan geweest en heb daar aansluitend een korte stage in het lab van Keiji Tanaka gevolgd. Een zeer mooie en leerzame ervaring. Vervolgens heb je me in contact gebracht met Ken, alwaar ik in zijn lab een half jaar heb kunnen werken en ervaringen op doen. Thanks to Ken and the members of the Britten lab for the opportunity to work with interesting people and to gain a lot of experience on research, GO, roadkill bbq's, local beers and bars (Al Bundy style).

Naast mijn directe begeleiding wil ik ook Wim van de Grind bedanken. Een wandelende encyclopedie, vol passie, humor en altijd een kritische noot. Verder wil ik Bert van den Berg en Johan van Leeuwen bedanken voor hun kritische commentaar tijdens dit traject en op de hoofdstukken in dit proefschrift.

Dank aan mijn collega's. Mijn kamergenoot in het NEST János Perge, een stimulerende discussiepartner, tot in de late uurtjes, bovenop het NEST en altijd in voor gezelligheid. Jeanette Lortije mijn lotgenote in Siberië. Michiel Tolboom voor de gesprekken over wetenschap, de klimpartijen en het overhalen om de 250km door het Limburgse heuvelland te fietsen. Lonneke Eeuwes voor de verhalen over meervallen en de zwempartijen in die halve ballon in Rijnsweerd. André Noest voor mooie anekdotes, verhalen over geologie en het wijzen op de schitterende gesteentes in de vloeren van het Hermitage in Sint Petersburg. Verder

---

Jacob Stoickov Duijnhouwer, Erik Aarnoutse, Rob Peters, Frank Brettschneider, Jan Brascamp, Casper Erkelens, de studenten Sanne Stuur, Martijn Agterberg, Marijn Kroes, en de 'nieuwe' garde Anna Oleksiak en Chris Klink voor discussie, het geduld, de conferenties, de gezelligheid en het altijd vinden van een reden voor het houden van een bbq: Midwinter, zomer of artikel, het maakt niet uit.

Daarnaast de leden van de AIO raad, Tanja Nijboer, Jeroen van Bortel, Gaston Hilhuysen, Krista Overvliet, Rolf Houben, voor de interessante vergaderingen, de trips (o.a. naar het Frankrijk in Amsterdam) en het gezellige borrelen.

En natuurlijk de mensen achter de schermen die op verschillende manieren een bijdrage hebben geleverd aan dit onderzoek, altijd klaar staan voor een gezellig gesprek, mooie anekdotes te vertellen hebben over het toen en nu, komen met een mop en/of een wijsheid: Henk Westland en Max, Hans Borgeld, Gerard Stoker en Ed van der Veen, Rob van Weerden en Rob Loos en Miriam van Hattum. En speciaal wil ik Theo Stuivenberg en Annemiek van der Velde hier bedanken: NEST, pizza, verrassende 50 jarige jubileums, motoren, meisje van Prinsenbeek, Viva La France, eigengemaakte saté, de beste haring en Ketel-1, Joris, Suzi's, Cagiva's en BMW's... Wanneer gaan we weer een keer naar Hulsberg? Ik heb weer tijd ;-)

Verder een woord van dank aan mijn unitmanager Janneke Meijer en teamleiders Joan van der Heijden en Els Gilijamse voor de steun, het begrip en de ruimte die jullie gaven om naast mijn 'bijbaantje' door te gaan met het afronden van de promotie. En uiteraard, de collega-vrienden met wie samenwerken in een team een plezier is, delen van kennis in een boshut en België zeker en vast tot gezelligheid leidt, bijhouden van wind- en golfvoorspellingen een dagvullend onderwerp kan zijn, whisky afspraken nog nagekomen moeten worden, en limburgse gemoedelijkheid gedeeld kan worden: Carel, Nick, Wilco, Lennart, Tim en Erik (die in zijn eentje meer kilometers rijdt dan heel de afdeling samen).

Promoveren was een leven op het NEST en het Kruyt. De uitvalbasis was op de Enny in Rijnsweerd, Zé Country Club is still Alive: Lennart, Esther en Thomas, mijn buuv Marise en Kees, Ellen, Philip, Roel, Maarten en Lieke, Gerben, Roxane en Tycho, Houkes, Marian en Mike, Martine, Martin, de oude garde Harmke, Saskia, Han en Kaatje en de frivole Bram, Mo, Jimmy, Mini en Maxi. Hartelijk dank voor een heel gezellige tijd (understatement); in vogelvlucht: Achtung die Kurve, to boldly go where no man has gone before, De Hoogte in Eesveen, Monday poker night, het A tot Z bij de Cambridge bar, weiden, schapen, een leren broek en een sloot bij terugkomst.

Sport is een perfecte manier om het hoofd te legen en ruimte te maken voor nieuwe ideeën. Ik wil de wie-de-sportschoen-passe-trekke-hem-aan mensen van Hellas Triatlon bedanken, voor ontspanning en gezelligheid. In het bijzonder de mensen die avontuur niet schuwen: adventure Dirk en Hilde, Jacosminalap, Vera, Edith, Jan en May, Marcel, no-waxer Wouter von Snapps. Daarnaast Suus en Kees, filosoof van de club, Lennart Link en personal motivator en fietskoerier

Ronald.

En ook buiten Rijnsweerd zijn er vrienden die hebben geholpen (wellicht onbewust) bij het afronden van het proefschrift: Het Al-gooien-ze-mijn-ballen-voor-de-trein-mijn-laatste-woorden-zullen-zijn-BACO-BACO!-team: Benidorm bastard Baco-Ed, Benji alias Highlander alias Sherif\* en Lady Bernadet, Bas, Zorro-Ed<sup>2</sup> en Elaine, MTB-Emiel en Stralende Kristel, Uitsmijter Sander en Anneke, Johnny Boy, Coert, Miek en Mark, Jes en Roel, en Jelte. Allemaal hartelijk dank voor de ontspanning, het filosoferen onder de sterren en gewoon de (h)eerlijke gezelligheid.

En dan wil ik bedanken Barbra die me heeft laten zien dat je heel veel kan bereiken als je het wilt en dan ook gewoon doet. Heel veel succes met HEART en al je plannen!

En bovenal wil ik bedanken voor hun extra support en steuners door dik en dun: mijn ouders Lou en Paula, Gerrien, mijn lief met een gouden hart, enthousiast en altijd klaarstaand (tenminste als er geen fiets in de buurt is). Mijn grote broer Roy en lieve schoonzus Manon en mijn altijd lachende neefje Kasper. Paranimf Maurice en Wendy, Thomas, Kiara, Ruben\* en natuurlijk Ruby. Onvoorwaardelijk staan jullie klaar, weten me te motiveren, af te leiden en rust te brengen als het nodig is.

Iedereen heel erg bedankt!



# List of publications

## PUBLICATIONS

E Sander, R Bours, I Duijnste and B Van der Zwaan. Experimental effects of an organic matter pulse and oxygen depletion on a benthic foraminiferal shelf community. *Journal of Foraminiferal Research*, 35(3):177-197, 2005

JA Perge, BG Borghuis, RJE Bours, MJM Lankheet and RJA van Wezel. Temporal Dynamics of Direction Tuning in Motion-Sensitive Macaque Area MT. *Neurophysiol*, 93: 2104-2116, 2005

JA Perge, BG Borghuis, RJ Bours, MJ Lankheet, RJ van Wezel. Dynamics of directional selectivity in MT receptive field centre and surround. *Eur J Neurosci*, 22(8):2049-58, 2005

RJ Bours, S Stuur and MJ Lankheet. Tuning for temporal interval in human apparent motion detection. *J Vis*, 8;7(1):2, 2007

RJ Bours, MC Kroes and MJ Lankheet. The parallel between reverse-phi and motion aftereffects. *J Vis*, 24;7(11):8.1-10, 2007

RJ Bours, MC Kroes and MJ Lankheet. Sensitivity for reverse-phi motion. *Vision Res*, 49(1):1-9, 2009 (Epub 2008 Nov 5)

## ABSTRACTS

RJE Bours, RJA van Wezel and MJM Lankheet. Step size tuning in macaque area MT. 619.8. Washington, DC: Society for Neuroscience, 2005

RJE Bours, S Stuur and MJM Lankheet. Tuning for temporal interval in human apparent motion detection. *Perception*, 34 ECVP Abstract Supplement 2005

RJE Bours, M Kroes, MJM Lankheet. Spatiotemporal properties of apparent-motion detection. *Perception*, 35 ECVP Abstract Supplement 2006

---

P R E S E N T A T I O N S

RJE Bours. Spatio-temporal tuning in macaque area MT. Proceedings of the third Helmholtz Retreat 2005, Bergen, The Netherlands

RJE Bours, M Tolboom, J Duijnhouwer. Visuele illusies. Wetenweek Faculteit Betawetenschappen 'Natuurkrachten aan het werk' 2005, Utrecht, The Netherlands

RJE Bours. Speed tuning of neurons in macaque MT. Proceedings of the second Helmholtz Retreat 2003, Bergen, The Netherlands

RJE Bours, RJA van Wezel, MJM Lankheet and K Britten. The effect of visual motion adaptation on directional motion sensitivity of MT-neurons and motion discrimination in primates, USC Annual Neuroscience Retreat 2002, Bodega Marine Laboratory, California, USA

# Curriculum Vitae



## P E R S O N A L I A

<b>Family name</b>	Bours
<b>First name</b>	Roger
<b>Middle names</b>	Jacques Elisabeth
<b>Nationality</b>	Dutch
<b>Date of birth</b>	February 23rd 1978
<b>Place of birth</b>	Munstergeleen
<b>E-mail</b>	rogerindavis@hotmail.com

---

## W O R K      E X P E R I E N C E

- 2010-** *ASR Insurance*, Utrecht, The Netherlands  
IS officer
- 2007-** *Sogeti Nederland B.V.*, Vianen, The Netherlands  
**2010** Software-Engineer and Information analyst  
at Enterprise Software Engineering Division,  
Requirements and design
- 2001-** *University Utrecht*, Utrecht, The Netherlands  
**2006** PhD-student at Dept. Functional Neurobiology, Faculty of Biology,  
Helmholtz Institute  
Project: 'Motion and speed processing by the primate visual system'
- 2004-** *Helmholtz Institute*, Utrecht, The Netherlands  
**2006** Member of the PHD-counsel
- 2002** *Center for Neuroscience*, University of California, Davis, USA  
Project: 'Motion adaptation in Macaque monkey'  
Supervisor: Prof. K. Britten
- 2000** *University Utrecht*, Graduation project at Neurobiology, The Netherlands  
Project: 'Spatio-temporal properties of visual speed perception in cats'  
Supervisor: dr. ir. M.J.M. Lankheet
- 2001** *University Utrecht*, Graduation project at Neurobiology, The Netherlands  
Project: 'Experimental effects of an organic matter pulse and  
oxygen depletion on a benthic foraminiferal shelf community'  
Supervisor: dr. S. Ernst
-

---

## EDUCATION AND TRAINING

- 2001-2010** See *Training and Supervision Plan on page 157*
- 1996-2001** Examen doctorale scientiae biologiae (Master of science in Biology)  
University Utrecht, Faculty of Biology
- 1990-1996** Pre-university education ('VWO') at 'Serviam lyceum', Sittard  
(Dutch, English, Latin, Economics, Mathematics, Physics, Chemistry,  
& Biology)
- 

## HONORS

- 2010** ICT Trots award, ASR, The Netherlands
- 2001** Travel grant and support for project at Center for Neuroscience at the  
University of California in Davis, USA, by Stichting Dondersfonds, Utrecht
- 2001** Full financial support for attending RIKEN-BSI summer program, Japan
- 1997** Excellent-Tracé, University Utrecht, Faculty of Biology
- 

## OTHER ACTIVITIES

- 2006** Member of 1st organizing committee of 'Krommerijn Adventure Race  
Utrecht'
- 2004-2005** Member of 1st organizing committee of 'Krommerijn Triathlon Utrecht'  
of the Hellas Triathlon division (Utrecht)
-



## TRAINING AND SUPERVISION PLAN

**The Basic Package**

Bio-ethics, Utrecht, The Netherlands (NL)	2001
Presentation skills, Sogeti, Vianen, NL	2007
Behaviour in teams, Sogeti - Lumage, Vianen, NL	2008
Styles of behaviour, Sogeti - Lumage, Vianen, NL	2008
Effective influencing behaviour, Sogeti - Mulder Projectmanagement, Vianen, NL	2009

*Subtotal Basic Package*    3 credits<sup>1</sup>

**Scientific Exposure***International conferences*

Neuroscience, Washington, USA (5 days)	2005
European Conference for Visual Perception, St Petersburg, Russia (5 days)	2006

*Seminars and workshops*

First Helmholtz Retreat, Schiermonnikoog, NL (3 days)	2001
USC Annual Neuroscience Retreat 2002, Bodega, California, USA (3 days)	2002
Second Helmholtz Retreat, Bergen, NL (3 days)	2003
Third Helmholtz Retreat, Bergen, NL (3 days)	2005
Attendance of the Helmholtz lectures of guest researchers	2001-2004
Attendance of the PhD student days ('AIO-days'), (two or three times a year at different locations in The Netherlands)	2001-2006
AIO Helmholtz-day, Utrecht, NL	2004
Dag van de perceptie, TNO, Soest, NL	2006

*Poster presentations*

Step size tuning in macaque area MT, Neuroscience	2005
Tuning for temporal interval in human motion detection, ECVF	2005
Spatiotemporal properties of apparent motion detection, ECVF	2006

*Oral presentations*

Presentation RIKEN Brain Science Institute, Wako-shi, Japan	2001
The effect of visual motion adaptation on directional motion sensitivity of MT-neurons and motion discrimination in primates, Bodega, USA (Retreat USC Neuroscience)	2002
Speed tuning of neurons in macaque MT, Bergen, NL (HH retreat)	2003
Speed tuning of neurons in macaque MT, Utrecht, NL (AIO day)	2004
Spatio-temporal tuning in macaque area MT, Bergen, NL (HH retreat)	2005

*Subtotal Scientific Exposure*    20,6 credits<sup>1</sup>

**In-Depth Studies**

Summer Program at RIKEN Brain Science Institute, Wako-shi, Japan	2001
Literature discussion group (Friday-lunch club of Helmholtz School)	2001-2006
Electrophysiology (Faculty of Biology, Utrecht, NL)	2002
Linear system analysis, Neuroscience, Erasmus University, Rotterdam	2004

*Subtotal In-Depth Studies*    13,5 credits<sup>1</sup>

---

### Statutory Courses

Course on Laboratory Animal Science; ex art 9 of the Law on Animal Experiments (Faculty of veterinary medicine, Utrecht)	2001
<i>Subtotal Statutory Courses</i>	<i>4 credits<sup>1</sup></i>

### Professional Skills Support Courses

Efficient supervision of students, Faculty of Biology, Utrecht	2002
English writing for publication, James Boswell Institute, Utrecht	2002
Business course, High performance teaming, Ohio University, Athens, USA	2007
Functional design basic/experienced, Sogeti, Vianen	2007-2008
Information analysis basic/experienced, Sogeti, Vianen	2007-2009
Interview techniques, Sogeti, Vianen	2008
Management fundamentals, Sogeti, Vianen	2009
Organizing workshops: Metaplan® Fundamentals, Sogeti, Vianen	2009
Advisory report and project proposition, Sogeti - Hendrikx Van der Spek, Vianen	2009
Branche management (Bedrijfskunde Brancheleer), Sogeti, Vianen	2009
Measuring and estimating project costs, Sogeti, Vianen	2010
<i>Subtotal Professional Skills Support Courses</i>	<i>13,3 credits<sup>1</sup></i>

### Research Skills Training

2 weeks exchange, Keiji Tanaka Lab for Cognitive Brain Mapping, RIKEN Brain Science Institute, Wako-shi, Japan	2001
6 months exchange, Ken Britten Lab, UC Davis, California, USA	2002
<i>Subtotal Research Skills Training</i>	<i>2,5 credits<sup>1</sup></i>

### Didactic Skills Training

#### *Supervising practicals*

Assistance at course Neurosciences, 2004	2004
Assistance at course Neurosciences, 2005	2005

#### *Supervising thesis*

Supervision Roxane Weijenberg	2003
Supervision Martijn Agterberg	2004
Supervision Sanne Stuur	2005
Supervision Marijn Kroes	2006
Supervision Helene Veenstra	2006

*Subtotal Didactic Skills Training* 16 credits<sup>1</sup>

### Management Skills Training

Member of Helmholtz AIO council	2004-2006
Organising AIO day	2005

*Subtotal Management Skills Training* 8 credits<sup>1</sup>

**Total ECTS credits<sup>1</sup> 80,9**

Herewith the WIAS Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of WIAS.

<sup>1</sup>One European Credit Transfer System (ECTS) credit equals a study load of approximately 28 hours.

---

# Propositions

1. Knowing that something is not moving in a particular direction is as valuable for motion perception as knowing that it is moving in that direction.
2. Reverse-phi motion perception and motion aftereffects arise through similar principles of low-level motion processing. As a consequence, reverse-phi motion perception has similar properties as the motion aftereffect of regular motion perception, and vice versa.
3. Exaggerative emphasis on efficiency in health-care organizations and universities undermines their effectiveness.
4. Erupting volcanoes help to solve the Carbon Dioxide (CO<sub>2</sub>) problem in multiple ways.
5. If you turn your work into a hobby, you need to find new work.
6. The improvement of the understanding of reality requires imaginative thoughts.

Propositions belonging to the thesis, entitled

**“Cortical mechanisms underlying low-level motion processing  
in the visual system of human and non-human primates”**

Roger J.E. Bours, Wageningen, 29 October 2010.

# Stellingen

1. Weten dat iets niet in een bepaalde richting beweegt is even waardevol voor het bewegingszien als weten dat het wel in die richting beweegt.
2. Reverse-phi en bewegingsnaeffecten zijn gebaseerd op dezelfde principes in het low-level bewegingszien. Hieruit volgt dat reverse-phi dezelfde eigenschappen heeft als het bewegingsnaeffect van gewone beweging en vice versa.
3. Overdreven nadruk op efficiëntie in zorginstellingen en universiteiten ondermijnt hun effectiviteit
4. Het uitbarsten van vulkanen helpt op meerdere manieren het CO2 probleem op te lossen.
5. Als je van je werk je hobby hebt gemaakt, moet je nieuw werk zoeken.
6. Het begrijpen van de werkelijkheid begint met fantasierijke gedachten.

Stellingen bij proefschrift, getiteld

**“Cortical mechanisms underlying low-level motion processing  
in the visual system of human and non-human primates”**

Roger J.E. Bours, Wageningen, 29 Oktober 2010.

The research described in this thesis was supported by the Helmholtz Institute, Utrecht University and by the Innovational Research Incentives Scheme (VIDI) of the Netherlands Organization for Scientific Research (NWO).

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

*I don't know what I want.  
But I know what I don't want.*

“Harley Davidson and the Marlboro Man” (1991)